Validation of the Intensive Care National Audit and Research Centre Scoring System in a UK Adult Cardiac Surgery Population

Priyadharshanan Ariyaratnam, BM, BSc (Hons), MRCS,* Mahmoud Loubani, MD, FRCS, FECTS,* James Biddulph, FRCA,† Julie Moore, BSc,† Neil Richards,* Mubarak Chaudhry, FRCS,* Vincent Hong, FRCA,† Mark Haworth, MRCA,† and Anantha Ananthasayanam, FRCA†

Objective: **The Intensive Care National Audit and Research Centre (ICNARC) scoring system was conceived in 2007, utilizing 12 physiologic variables taken from the first 24 hours of adult admissions to the general intensive care unit (ICU) to predict in-hospital mortality. The authors aimed to evaluate the ICNARC score in predicting mortality in cardiac surgical patients compared to established cardiac risk models such as logistic EuroSCORE as well as to the Acute Physiology and Chronic Health Evaluation (APACHE) II.**

Design: **Retrospective analysis of data collected prospectively.**

Setting: **Single-center study in a cardiac intensive care in a regional cardiothoracic center.**

Participants: **Patients undergoing cardiac surgery between January 2010 and June 2012.**

Methods: **A total of 1,646 patients were scored preoperatively using the logistic EuroSCORE and postoperatively using ICNARC and APACHE II. Data for comparison of scoring systems are presented as area under the receiver operating characteristic curve.**

Measurements and Main Results: **The mean age at surgery was 67 years ± 10.1. The mortality from all cardiac surgery was 3.2%. The mean logistic EuroSCORE was 7.31 ± 10.13, the mean ICNARC score was 13.42 ± 5.055, while the mean APACHE II score was 6.32 ± 7.731. The c-indices for logistic EuroSCORE, ICNARC, and APACHE II were 0.801, 0.847 and 0.648, respectively.**

Conclusion: **The authors have, for the first time, validated the ICNARC score as a useful predictor of postoperative mortality in adult cardiac surgical patients. This could have implications for postoperative management, focusing the utilization of resources as well as a method to measure and compare performance in the cardiothoracic ICU.**

KEY WORDS: cardiac surgery, intensive care, ICNARC, risk stratification, EuroSCORE, APACHE II

PREOPERATIVE SCORING SYSTEMS FOR cardiac surgery, such as the logistic European System for Cardiac Operative Risk Evaluation (EuroSCORE) in Europe and the Society of Thoracic Surgeons (STS) score in North America, regularly are created, validated, and modified.[1,2] Although intended to be broadly applicable, quite often, their accuracy in predicting a particular institution's actual mortality falls short because of varying patient characteristics, surgical techniques, and postoperative care.[3,4] Despite this, they are still routinely employed, not only for informed patient consent, but also as a means to evaluate and compare an institution's performance.[5]

However, because mortality after cardiac surgery also may be determined by surgical and anesthetic techniques not apportioned to these preoperative scoring system, the prognosis for a patient immediately postoperatively may be quite different from that in their preoperative assessment.

In the UK, general intensive care units (GICU) use a variety of scoring systems discerned from parameters evident in the first 24 hours of admission. These range from long-established scores such as the Acute Physiology and Chronic Health Evaluation (APACHE) II score[6,7] to the newer Intensive Care National Audit & Research Centre (ICNARC) scoring system.[8,9]

ICNARC was devised in 2007 and, like APACHE II, focuses on physiologic variables measured during the first 24 hours of admission to the GICU. However, such scoring systems have not been used routinely to predict mortality in cardiac patients. Postoperative cardiac patients are admitted electively to the ICU, having had a significant controlled iatrogenic insult inflicted on them during their procedure; unlike GICU patients who have entered there quite unintentionally, which may minimize the usefulness of these physiologic scores for cardiac patients.

The team involved in the care of cardiac patients including surgeons, intensivists, and allied professionals might be in a better position to identify patients at risk following surgery by considering cardiac surgical patients as GICU patients and thus applying risk stratification along the GICU scoring models.

Hence, the authors sought to evaluate the role of the ICNARC scoring model in predicting mortality in postoperative cardiac surgical patients.

METHODS

The authors prospectively collected data on 1,646 consecutive adult (>18 years old) patients undergoing cardiac surgery between January 2010 and June 2012 at the authors' institution. Preoperative data were collected by cardiac surgeons and ICU admission data were collected by intensive care specialists. Permission for this project were obtained from the Institutional Review Board.

The APACHE II variables, logistic EuroSCORE variables, and ICNARC variables are shown in Table 1. The respective definitions and calculations are described previously in the literature.[10–12]

The ICNARC score is taken during the first 24 hours of admission. Most variables relate to the maximum or minimum value of the parameter; for example, serum sodium refers to the 'highest' serum sodium recorded and not the average. Other variables, such as Glasgow Coma Score (GCS), refer not only to the GCS itself, but also are dependent on whether the patient

*From the *Department of Cardiothoracic Surgery, and †Department of Cardiothoracic Anaesthesia Castle Hill Hospital, Cottingham, UK.*

Address reprint request to Priyadharshanan Ariyaratnam, BM, BSc (Hons), MRCS, Department of Cardiothoracic Surgery, Castle Hill Hospital, Cottingham, UK. E-mail: priyadariyaratnam@yahoo.co.uk

1053-0770/2601-0001$36.00/0

http://dx.doi.org/10.1053/j.jvca.2014.09.013

Table 1. Variables Present in Each Scoring System

Variables in the EuroSCORE I (with Maximum Weighting)	Variables in APACHE II (with Maximum Weighting)	Variables in ICNARC (with Maximum Weighting)
Age (1)	Age (6)	Heart rate (14)
Sex (1)	Temperature (4)	Systolic blood pressure (16)
Chronic obstructive pulmonary disease (1)	Mean arterial pressure (4)	Temperature (12)
Extracardiac arteriopathy (2)	Arterial pH (4)	Respiratory rate (5)
Neurologic dysfunction (2)	Heart rate (4)	Pao2/F_IO_2 ratio (8)
Previous cardiac surgery (3)	Respiratory rate (4)	Arterial pH (4)
Creatinine > 200 μmol/L (2)	Serum sodium (4)	Serum urea (5)
Active endocarditis (3)	Serum potassium (4)	Serum creatinine (4)
Critical preoperative state (3)	Serum creatinine (4)	Serum sodium (8)
Unstable angina (2)	Hematocrit (4)	Urine output (7)
LV dysfunction (3)	White blood cell count (4)	White blood cell count (6)
Recent myocardial infarction (2)	Oxygenation (4)	Glasgow coma score (11)
Pulmonary hypertension (2)	Glasgow coma scale (12)	
Emergency (2)	Chronic disease (5)	
Other than isolated coronary artery bypass grafting (2)		
Surgery on the thoracic aorta (3)		
Post-infarct septal rupture (4)		

Abbreviation: APACHE, Acute Physiology and Chronic Health Evaluation; EuroSCORE, European System for Cardiac Operative Risk Evaluation; ICNARC, Intensive Care National Audit and Research Centre.

is sedated or paralyzed for the duration of the first 24 hours. The full breakdown of the ICNARC scoring system previously has been described in detail.[12] APACHE II also was taken in the first 24 hours of admission to the intensive care unit from cardiac operating rooms.

The primary outcome measured was perioperative mortality (defined as in-hospital death occurring during the same admission as the surgery or 30-day mortality if patients were discharged within 30 days of the surgery).

The performance of the 3 risk models was compared in this group of patients by measuring each model's discrimination (c-statistic) and calibration using Hosmer-Lemeshow goodness-of-fit test. Receiver operating characteristic curves were used to generate a c-statistic (area under the curve). C-statistic values closer to 1.0 indicate better discrimination by the model. To compare the scoring systems, the authors used the Hanley and McNeil method of comparing the areas under the receiver operating curves derived from the same case.[13]

Linear regression was used to assess whether there were any preoperative factors that may influence the ICNARC score. The following variables were examined: Sex, diabetes mellitus, chronic obstructive pulmonary disease, renal failure, New York Heart Association (NYHA) status, Canadian Cardiovascular Society status, and left ventricular ejection fraction.

Multivariate logistic regression was used to assess the impact of each of the variables in the ICNARC and APACHE II scores on perioperative mortality to discern which of the variables were most important in the cardiac surgery cohort.

Finally, the authors looked at whether a higher EuroSCORE, ICNARC, and APACHE II scoring systems had any effect on the time spent by the patient in cardiac intensive care.

Statistical analysis was performed using IBM SPSS Statistics 20.

RESULTS

The demographic details of the patients, as well as the operative and postoperative details, are shown in Table 2. The authors' perioperative mortality was 3.2%.

The intraoperative characteristics, which may influence mortality outcomes, are included in Table 3. Only blood loss in the first 24 postoperative hours including the operating room loss and blood transfusion were important determinants of survival. Figure 1 demonstrates the receiver operating characteristic graphs for EuroSCORE, APACHE II, and ICNARC.

Table 4 gives the c-indices for each scoring system. This is a measure of discrimination for the predictive test. ICNARC has a higher c-index compared to the other 2 scores. The difference between ICNARC and APACHE II area under the receiver operating curve (AUC) was statistically significant ($p < 0.0001$). The same was true between EuroSCORE and APACHE II ($p < 0.005$). There was no statistical difference between the AUCs for ICNARC and EuroSCORE. This demonstrated that although the ICNARC and EuroSCORE have a clear discriminatory advantage over APACHE II, there is less difference between the EuroSCORE and ICNARC in predicting perioperative mortality.

Table 5 gives the results of the Hosmer Lemeshow test, which is a measure of calibration of the predictive test. The ICNARC score demonstrates good calibration for validation.

It was found from multivariate analysis that New York Heart Association stage IV status ($p < 0.005$), preoperative renal failure ($p < 0.005$), and poor ejection fraction ($p < 0.005$) were the only preoperative factors related to a higher ICNARC score.

From univariate analysis of the ICNARC and APACHE II scores, only blood pressure, respiratory rate, total urine output, serum sodium concentration, lactate concentration, and arterial blood gas pH were found to be significant predictors of mortality. Multivariate analysis revealed that from these, only blood pressure, total urine output and lactate were the only significant predictors. The results of this analysis are shown in Table 6.

Regarding length of intensive care unit stay, all 3 scoring systems significantly predicted the duration of stay on the unit. Table 7 demonstrates the results of the correlation.