A multicentre Spanish study for multivariate prediction of perioperative in-hospital cerebrovascular accident after coronary bypass surgery: the PACK2 score†

Fernando Hornero*a, Elio Martinb, Rafael Rodríguezb, Manel Castelláa, Carlos Porrasd, Bernat Romere, Luis Marotof and Enrique Pérez De La Sotag, on behalf of the Working Group on Arrhythmia Surgery and Cardiac Pacing of the Spanish Society for Cardiovascular and Thoracic Surgery (SECTCV)

a Servicio de Cirugía Cardiaca, Instituto Cardiovascular, Hospital General Universitario, Valencia, Spain
b Servicio de Cirugía Cardiaca, Área del cor, Hospital Vall d’Hebrón, Barcelona, Spain
c Servicio de Cirugía Cardiaca, Hospital Clinic i Provincial, Barcelona, Spain
d Servicio de Cirugía Cardiaca, Hospital Clínico Universitario Virgen de la Victoria, Málaga, Spain
e Servicio de Cirugía Cardiaca, Hospital Germans Trias i Pujol, Badalona, Spain
f Servicio de Cirugía Cardiaca, Hospital Clínico San Carlos, Madrid, Spain
g Servicio de Cirugía Cardiaca, Hospital 12 de Octubre, Madrid, Spain

* Corresponding author. Universitat de València, Servicio de Cirugía Cardiaca, Instituto Cardiovascular, Hospital General Universitario, Avda Tres Cruces 2, 46014 Valencia, Spain. Tel: +34-961972000; fax: +34-961972163; e-mail: hornero_fer@gva.es (F. Hornero).

Received 12 September 2012; received in revised form 11 December 2012; accepted 28 December 2012

Abstract

OBJECTIVES: To develop a multivariate predictive risk score of perioperative in-hospital stroke after coronary artery bypass grafting (CABG) surgery.

METHOD: A total of 26,347 patients were enrolled from 21 Spanish hospital databases. Logistic regression analysis was used to predict the risk of perioperative stroke (ictus or transient ischaemic attack). The predictive scale was developed from a training set of data and validated by an independent test set, both selected randomly. The assessment of the accuracy of prediction was related to the area under the ROC curve. The variables considered were: preoperative (age, gender, diabetes mellitus, arterial hypertension, previous stroke, cardiac failure and/or left ventricular ejection fraction <40%, non-elective priority of surgery, extracardiac arteriopathy, chronic kidney failure and/or creatininemia ≥2 mg/dl and atrial fibrillation) and intraoperative (on/off-pump).

RESULTS: Global perioperative stroke incidence was 1.38%. Non-elective priority of surgery (priority; OR = 2.32), vascular disease (arteriopathy; OR = 1.37), cardiac failure (cardiac; OR = 3.64) and chronic kidney failure (kidney; OR = 6.78) were found to be independent risk factors for perioperative stroke in uni- and multivariate models in the training set of data; P < 0.0001; AUC = 0.77, 95% CI 0.73–0.82. The PACK2 stroke CABG score was established with 1 point for each item, except for chronic kidney failure with 2 points (range 0–5 points); AUC = 0.76, 95% CI 0.72–0.80. In patients with PACK2 score ≥2 points, off-pump reduced perioperative stroke incidence by 2.3% when compared with on-pump CABG.

CONCLUSIONS: PACK2 risk scale shows good predictive accuracy in the data analysed and could be useful in clinical practice for decision making and patient selection.

Keywords: Coronary • Stroke • Off-pump

INTRODUCTION

Although mortality during cardiac surgery has decreased in recent decades [1] morbidity has increased, mainly because patients undergoing cardiac surgery are older and have an increased vulnerable comorbid condition [2]. Stroke remains a severe complication after coronary artery bypass grafting (CABG) procedures, and its prevention is particularly important as the patients affected spend prolonged stays in intensive care and have a markedly higher mortality rate. The risk of stroke in CABG surgery is not homogeneous, and different clinical and echocardiographic features have been identified to help in stratifying it. Several prediction models have been developed to preoperatively identify this subgroup of patients with an increased risk of perioperative stroke; however, none has been extensively applied in routine clinical practice [3]. The NNECDSG stroke risk scale [4] was based on 33,062 consecutive patients undergoing isolated CABG surgery in Northern New England between 1992 and 2001, and showed good predictive accuracy with an area under

†Presented at the 26th Annual Meeting of the European Association for Cardio-Thoracic Surgery, Barcelona, Spain, 27–31 October 2012.
the receiver operating characteristic curve (c-statistic) of 0.70 [95% confidence interval (CI), 0.67–0.72]. However, it is clear that an ideal, generally applicable, simple and widely accepted stroke risk assessment tool does not exist and that each one has its own limitations.

The purpose of this study was to define a clinical grading scale to assess perioperative stroke risk in the CABG population, using the predictive preoperative criteria of stroke outcome in a rapid and efficient procedure carried out at the time of surgery.

**METHODS**

**Study setting, patient cohort and data collection**

We designed a retrospective, multicentre, observational study on consecutive adult (≥18 years of age) patients undergoing CABG surgery as a single procedure, recruited from 21 Spanish National Health System hospitals (see Appendix 1). The patients were obtained from hospital administrative databases in reverse chronological order, starting on 31 December 2011. In order to meet the inclusion criteria, we only considered CABG as a single surgical procedure. Patients undergoing >1 relevant procedures during the study period were only considered for analysis in the first surgical procedure. All patients who underwent CABG with associated surgical procedures were excluded. Using standardized case report forms, we collected selective preoperative data, which were entered in a computer database. All the clinical variables collected had previously been shown to have a significant impact on perioperative stroke risk, in accordance with the NNECDSG prediction model [4].

**Variables**

Perioperative stroke was defined as any new temporary or permanent, focal or global neurological defect, within 30 days after surgery, or later if still in hospital, in accordance with the published guidelines [5]. Temporary stroke included transient ischaemic attack, defined as a fully reversible neurological deficit lasting <24 h. Prolonged reversible ischaemic neurological deficits were defined as events lasting >24 h and <3 weeks. All the stroke outcomes included in this study were diagnosed by a neurologist, and in most cases, brain CT-scan or MRI was used for lesion assessment. The study excluded patients with diffuse postoperative brain encephalopathy presented as delirium, confusion, prolonged alteration in mental status and agitation in the immediate postoperative period, which could be related to circulatory bypass time and might also reflect longer exposure to anaesthesia [6].

The perioperative stroke risk variables included the 13 variables proposed by the NNECDSG prediction model, clustered into nine major factors: age, female gender, diabetes mellitus, arterial hypertension, vascular disease (including all extracardiac vascular disease manifestations such as intermittent claudication, amputation, absent pedal pulse and/or lower extremity ulcers; previous surgery and/or percutaneous intervention on the abdominal, thoracic aorta, lower extremity and supra-aortic vessels), previous stroke events (ictus with/without residual neurological defect or transient ischaemic attack), chronic renal failure (known glomerular filtration impairment <60 ml/min/1.73 m², patients requiring dialysis and/or preoperative creatininemia ≥2 mg/dl), preoperative cardiac failure (left ventricular ejection fraction <40% by echocardiography or ventriculography and/or preoperative NYHA functional class III–IV), and non-elective priority of surgery (urgency: operation required within 24 h to minimize the chance of further clinical deterioration; or emergency: there should be no delay in providing surgical intervention). The presence of preoperative atrial fibrillation was also considered. The NNECDSG score was calculated for each of the patients and was used to validate our multicentre study. In its assessment, the vascular disease variable was computed by the combination of the above definition plus the presence of a previous stroke event as defined in [4]. The intraoperative variable collected was the performance of CABG surgery under the off-pump beating heart (OPCAB) or cardioplegic cardiopulmonary bypass (CPBCAB) technique.

**Statistical analysis**

The continuous variables were presented as mean ± standard deviation (SD) and categorical variables were shown as a percentage (%). All group comparisons were unpaired. Continuous variables were compared using the Student’s t-test, and categorical variables were compared using χ² analysis or Fisher’s exact test, as appropriate. Age was considered both as a categorical clustered and continuous variable. A two-tailed value of P < 0.05 was considered as statistically significant.

The entire data set was divided randomly into a ‘training’ set used to develop the clinical prediction rule and a ‘test’ set to assess the performance of the rule when applied to data different from that from which it was developed. We first performed a univariate analysis, including preoperative variables and OPCAB/CPBCAB performance. Considering perioperative neurological outcome, variables with a significant level of P < 0.05 in the univariate analysis were included in a stepwise multivariate logistic regression model, to determine the independent perioperative stroke predictors. A predictive score model was developed including former independent risk factors, weighting their strength of association with neurological outcome to produce a simple and intuitive model. The predictive accuracy of the multivariate model, NNECDSG and the score scheme developed was assessed by means of the area under the receiver operating curve (c-statistic). The Hosmer–Lemeshow (H–L) goodness-of-fit test was also applied to calibrate these models.

CHAIDS (Chi² Automatic Interaction Detector) analysis, a statistical multi-way tree algorithm, was used to build profiles to help us better identify groups of perioperative stroke risk, integrating the relationships between the predictive score system and OPCAB/CPBCAB performance.

A Statistical Stressing Test to test the predictive score system was developed by randomly selecting 50 groups of 1000 patients from the entire data set. The c-statistic was calculated for each individual analysis of the groups.

The Statistical Package for the Social Sciences (IBM® SPSS® Statistic version 19.0; SPSS, Inc., Chicago, IL, USA) was used for the statistical analysis. The study’s protocol was approved by the Human Subjects Review Committee of each of the participating institutions.

**RESULTS**

**Patient cohort**

A total of 28,296 patients were included in the study. Single-institution accrual was variable (range 157–3060 patients).
Of these, 1949 (6.89%) patients had missing preoperative variables and were excluded from the analysis. Complete information was available for 26,347 (93.11%) patients. The data set was divided randomly into a ‘training’ set (n = 13,096, 49.7%) and a ‘test’ set (n = 13,246, 50.27%) to develop and check the predictive score, respectively.

The overall incidence of perioperative stroke was 1.38%, with values ranging from 0.3 to 2.5% among the participating centres. The characteristics of the patients studied are displayed in Table 1. The NNECDSG score showed very good discriminatory accuracy in predicting perioperative stroke in our cohort (c-statistic = 0.69 (0.67–0.72), P < 0.0001; H–L test: χ² = 8.32, P = 0.40).

Identification of independent preoperative risk factors of stroke

Univariate analysis was carried out on the ‘training’ group (Table 2). Off-pump surgery was not related to a lower rate of perioperative stroke (1.31% CPBCAB vs 1.38% OPCAB, P > 0.05). The multivariate analysis showed cardiac failure (OR = 3.64, P < 0.0001), non-elective surgery (OR = 2.32, P < 0.0001), vascular disease (OR = 1.37, P = 0.007) and chronic kidney failure (OR = 6.78, P < 0.0001) as independent risk factors for perioperative development of a stroke after CABG surgery (Table 3). The H–L goodness-of-fit test of the multivariate model showed good

---

### Table 1: Baseline characteristics of global ‘training’ and ‘check’ groups

<table>
<thead>
<tr>
<th>Group</th>
<th>'Training' group (n = 13,096)</th>
<th>'Check' group (n = 13,246)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55–59</td>
<td>64.88 ± 9.92</td>
<td>64.83 ± 10.01</td>
<td>0.46</td>
</tr>
<tr>
<td>60–64</td>
<td>3230 (12.26%)</td>
<td>1715 (12.99%)</td>
<td>0.74</td>
</tr>
<tr>
<td>60–64</td>
<td>4228 (16.05%)</td>
<td>2131 (16.27%)</td>
<td>0.34</td>
</tr>
<tr>
<td>65–69</td>
<td>4998 (18.74%)</td>
<td>2492 (19.03%)</td>
<td>0.25</td>
</tr>
<tr>
<td>70–74</td>
<td>5106 (19.38%)</td>
<td>2536 (19.36%)</td>
<td>0.94</td>
</tr>
<tr>
<td>75–79</td>
<td>3605 (13.68%)</td>
<td>1818 (13.88%)</td>
<td>0.36</td>
</tr>
<tr>
<td>≥80</td>
<td>995 (3.78%)</td>
<td>537 (4.05%)</td>
<td>0.08</td>
</tr>
</tbody>
</table>

### Table 2: Univariate analysis in terms of perioperative neurological outcome performed in ‘training’ group

<table>
<thead>
<tr>
<th>Variable</th>
<th>No postoperative stroke (n = 12,925)</th>
<th>Postoperative stroke (n = 174)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>64.92 ± 9.82</td>
<td>65.16 ± 9.66</td>
<td>0.75</td>
</tr>
<tr>
<td>55–59</td>
<td>1594 (12.33%)</td>
<td>21 (12.07%)</td>
<td>0.92</td>
</tr>
<tr>
<td>60–64</td>
<td>2096 (16.22%)</td>
<td>35 (20.11%)</td>
<td>0.18</td>
</tr>
<tr>
<td>65–69</td>
<td>2462 (19.05%)</td>
<td>30 (17.24%)</td>
<td>0.62</td>
</tr>
<tr>
<td>70–74</td>
<td>2509 (19.41%)</td>
<td>27 (15.52%)</td>
<td>0.21</td>
</tr>
<tr>
<td>75–79</td>
<td>1792 (13.86%)</td>
<td>26 (14.94%)</td>
<td>0.67</td>
</tr>
<tr>
<td>≥80</td>
<td>449 (3.47%)</td>
<td>9 (5.17%)</td>
<td>0.22</td>
</tr>
<tr>
<td>Female</td>
<td>2297 (17.77%)</td>
<td>26 (14.94%)</td>
<td>0.37</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>4774 (36.64%)</td>
<td>82 (47.13%)</td>
<td>0.007</td>
</tr>
<tr>
<td>Arterial hypertension</td>
<td>7433 (57.51%)</td>
<td>120 (68.97%)</td>
<td>0.003</td>
</tr>
<tr>
<td>Cardiac failure</td>
<td>614 (4.75%)</td>
<td>16 (9.19%)</td>
<td>0.01</td>
</tr>
<tr>
<td>Non-elective surgery</td>
<td>1919 (14.85%)</td>
<td>77 (44.25%)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Vascular disease</td>
<td>1648 (12.75%)</td>
<td>50 (28.74%)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Chronic kidney failure</td>
<td>2046 (15.83%)</td>
<td>54 (31.03%)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Preoperative atrial fibrillation</td>
<td>1052 (8.14%)</td>
<td>74 (42.53%)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Off-pump coronary bypass surgery</td>
<td>534 (4.13%)</td>
<td>10 (5.75%)</td>
<td>0.25</td>
</tr>
<tr>
<td>Postoperative stroke</td>
<td>5193 (40.18%)</td>
<td>72 (41.38%)</td>
<td>0.76</td>
</tr>
</tbody>
</table>
Development and validation of predictive PACK2 score

An outcome risk-stratification scale was developed from the logistic regression model. Table 4 indicates the specific punctuation assignments used in calculating the PACK2 score.

In the 'training' group, predictive accuracy for the NNECDSG scheme was c-statistic = 0.70 (95% CI 0.67–0.74), P < 0.0001; H–L test: χ² = 8.11, P = 0.42; and for PACK2 score c-statistic = 0.77 (95% CI 0.73–0.81), P < 0.0001; H–L test: χ² = 0.16, P = 0.69.

In the 'test' group, predictive accuracy for the NNECDSG scheme was c-statistic = 0.69 (95% CI 0.65–0.73), P < 0.0001; H–L test: χ² = 4.13, P = 0.85; and for PACK2 score c-statistic = 0.77 (95% CI 0.73–0.80), P < 0.0001; H–L test: χ² = 1.39, P = 0.24.

The Statistical Stressing Test confirmed the good predictive accuracy of the PACK2 score, being c-statistic >0.8 in 36%, >0.7 in 84% and non-significant in 8% of the random sets of 1000 patients selected.

Clinical application of PACK2 score

It is noted that 59.66% of patients were categorized as score 0 (n = 15719), 25.54% (n = 6728) as 1 point, 9.99% (n = 2631) as 2 points, 3.61% (n = 950) as 3 points, 1.03% (n = 272) as 4 points and 0.18% (n = 47) as 5 points. The stroke rates were: 0.48, 1.20, 3.57, 6.53, 15.44 and 21.28%, respectively, P < 0.0001 (Fig. 1).

Decision tree CHAID analysis (Fig. 2) was performed by applying the PACK2 score to stratify the perioperative risk of developing a perioperative stroke, in order to assess the benefit of OPCAB vs CPBCAB surgery. Zero and one points were analysed independently, showing no statistical difference in the perioperative stroke rate between off-pump and on-pump surgery (0 points: 0.6 vs 0.4%, P = 0.311; 1 point 1.3 vs 1.1%, P = 0.505, respectively). However, off-pump significantly reduced the perioperative stroke rate for patients with a PACK2 score ≥2, (4.1 vs 6.4%, P = 0.002).

DISCUSSION

This multicentre study developed a clinical grading scale, the PACK2 score, to predict perioperative stroke risk in isolated CABG surgery. Our results suggest the PACK2 score 1) has good accuracy and discriminatory power, with a c-statistic = 0.77 (95% CI 0.74–0.79), P < 0.0001, comparable with other specific scales such as the NNECDSG score; and 2) is simple enough to be used with only four preoperative variables.

Several prognostic perioperative stroke models had previously been developed and validated [2], and as these models have been shown to be highly accurate in predicting outcomes, we opted to use their clinical predictive variables as enrolment criteria for this study. However, several of them use complex scales, or complex algebraic equations in outcome prediction, and none has been simplified into a standard clinical grading scale analogous to the PACK2 score. This scale contains only four preoperative variables and therefore balances a compromise between simplicity and the accuracy of its outcome predictions. To strike an appropriate balance between these two factors, the main purpose of the grading scale must be considered. The
PACK2 score was simply developed to provide a standard assessment tool that can be easily and rapidly applied during preoperative clinical assessment of CABG patients.

How might the PACK2 score be used? Perioperative stroke risk in CABG patients is often a fundamental question, and reported scales are often used to provide initial information regarding this issue [7, 8]. While prognostication is undoubtedly important to assess surgical benefits and risks, any attempt to predict a precise outcome could lead to mistakes. A scale such as the PACK2 score should therefore be incorporated as part of risk stratification for perioperative CABG stroke studies and patient decision management, but not as a precise risk predictor of outcomes. Our aim was, therefore, to show its utility as a clinical tool in assessing stroke risk in CPBCAB surgery. In this study, cases with a score of ≥2 had a reduced incidence of perioperative stroke when undergoing OPCAB vs CPBCAB (P = 0.002).

Several limitations involved in this study should be noted. First, it is limited by its dependence on a retrospective observational study, and the conclusions derived from it are necessarily limited in application. However, we addressed this limitation by enrolling patients from multiple institutions, thereby minimizing any distortions due to the effects of different surgical, anaesthetic, perfusion-related and medical specific practices at the different centres. Secondly, OPCAB performance was not assessed in this study under ‘touch’ or ‘non-touch’ aorta schemes. As ‘touch’ schemes have been reported to limit the benefits of OPCAB over CPBCAB in terms of perioperative stroke incidence, the statistical differences found in high-risk patients may support the thesis of the potential reduction of neurological events with OPCAB surgery. Thirdly, the neurological findings were assessed at each site of origin and not by a single neurologist performing all the preoperative and postoperative examinations, so there could be significant variations in clinical practice, and thus diagnosis, among the 21 centres. Finally, as perioperative diffuse brain encephalopathy was excluded, this may potentially have underestimated the real prevalence of small perioperative strokes; however, other studies have reported diffuse brain encephalopathy, mainly related to intraoperative factors such as cardiopulmonary bypass and general anaesthesia times [6].

CONCLUSIONS

On the basis of data from 21 Spanish medical centres with 26 347 patients enrolled, we conclude that the PACK2 score is an easy-to-use classification scheme that stratifies the risk of perioperative stroke in CABG patients. Further validation from other centres will be necessary before this scale can be considered as a useful tool for planning prognostic and therapeutic strategies to prevent perioperative stroke.

ACKNOWLEDGEMENTS

Investigators and institutions participating in this study representing the Working Group on Arrhythmia Surgery and Cardiac Pacing of the Spanish Society for Cardiovascular and Thoracic Surgery (SECTCV): Hospital Vall d’Hebrón of Barcelona—Rodríguez R; Hospital Clinic i Provincial of Barcelona—Castellà M; Hospital Clínico Universitario Virgen de la Victoria of Málaga—Porras C; Hospital Germans Trias i Pujol of Badalona—Romero B; Hospital Clínico San Carlos of Madrid—Maroto L; Hospital 12 de Octubre of Madrid—Pérez de la Sota E; Hospital Clínico Universitario of Valladolid—Echevarría Ma; Hospital Universitario Marqués de Valdecilla of Santander—Diez L; Complejo Hospitalario de Toledo—Buendía J; Hospital Son Espases of Palma de Mallorca—Enríquez F; Complejo Asistencial Hospital de León—Castaño M; Hospital de La Princesa of Madrid—Reyes G; Hospital de la Santa Creu i Sant Pau of Barcelona—Ginel A; Hospital Universitari i Politécnic La Fe of Valencia—Pérez M; Hospital Universitario Virgen Macarena of Sevilla—García R; Barquero J; Hospital Fundación Jiménez Díaz of...
Madrid—Hereder A, Jiménez A; Hospital Universitario Puerta de Hierro of Madrid—Castedo E; Hospital Xeral of Vigo—Lugo J, Pradas G; Hospital Universitario Puerta del Mar of Cádiz—Gómez M; Hospital General Universitario of Valencia—Hornero F, Martin E, Mené AV and statistical analysis—Rietas JJ.

Conflict of interest: none declared

REFERENCES


APPENDIX. CONFERENCE DISCUSSION

Dr A. Badreddin (Koblenz, Germany): First of all, can you tell us what the difference is between your PACK2 score and the STS score, which tells us in a logistic calculation and percentage, the incidence or the possibility of incidence of stroke postoperatively?

Secondly, regarding your variables, the four variables are already included in the STS scoring system together with other variables; accordingly, the STS score should have more accuracy. Thirdly, we see in the c-statistics of the STS score that were validated externally in independent patient subsets, that the score should have more accuracy. Thirdly, we see in the c-statistics of the STS score together with other variables; accordingly, the STS score should have more accuracy. Thirdly, we see in the c-statistics of the STS score that were validated externally in independent patient subsets, that the score should have more accuracy. Thirdly, we see in the c-statistics of the STS score together with other variables; accordingly, the STS score.

Dr Martin: We based our score on an excellent predictive model which is the Northern New England scale. And the four variables are also included in the EuroSCORE, the STS and many other predictive scores. However, the main point of our score is that it shows an excellent predictive accuracy which you can calculate from memory at the moment when you have a knife in the hand and you have not done it before. So at that moment, when you are in a hurry or have not calculated it before, it allows a very simple system of calculation, so you can decide at a glance if you are going to cannulate the patient or if you are going to ask for the stabilizer to perform the coronary surgery.

Dr Badreddin: Actually, your answer was my second comment. I think the main advantage of your score is the simplicity.

Conflict of interest: none declared

REFERENCES