JACC: CARDIOVASCULAR INTERVENTIONS VOL. 4, NO. 3, 2011

ISSN 1936-8798/$36.00
PUBLISHED BY ELSEVIER INC. DOI: 10.1016/j.jcin.2010.10.013

Global Risk Classification and Clinical SYNTAX (Synergy between Percutaneous Coronary Intervention with TAXUS and Cardiac Surgery) Score in Patients Undergoing Percutaneous or Surgical Left Main Revascularization

Davide Capodanno, MD,*† Anna Caggegi, MD,* Marco Miano, MD,*
Glauco Cincotta, MD,* Fabio Dipasqua, MD,* Giuseppe Giacchi, MD,*
Piera Capranzano, MD,*† Gianpaolo Ussia, MD,* Maria Elena Di Salvo, MD,*
Alessio La Manna, MD,* Corrado Tamburino, MD, PHD*†

Catania, Italy

Objectives The aim of this study was to investigate the ability to predict cardiac mortality of the Global Risk Classification (GRC) and the Clinical SYNTAX (Synergy between Percutaneous Coronary Intervention with TAXUS and Cardiac Surgery) score (CSS) in left main (LM) patients undergoing percutaneous coronary intervention (PCI) or coronary artery bypass graft (CABG).

Background There is a renewed interest in combining clinical and angiographic information to define the risk of patients undergoing LM revascularization.

Methods The GRC and CSS were assessed in patients undergoing LM PCI (n = 400) or CABG (n = 549). Stand-alone clinical (ACEF [age, creatinine, ejection fraction]), EuroSCORE (European System for Cardiac Operative Risk Evaluation) and angiographic (SYNTAX score) risk scores were also investigated.

Results The GRC (Hosmer-Lemeshow statistic 0.357, p = 0.550; area under the curve 0.743) and the ACEF (Hosmer-Lemeshow 0.426, p = 0.514; area under the curve 0.741) showed the most balanced predictive characteristics in the PCI and CABG cohorts, respectively. In PCI patients, the CSS used fewer data to achieve similar discrimination but poorer calibration than the GRC. Propensity-adjusted outcomes were comparable between PCI and CABG patients with low, intermediate, or high EuroSCORE, ACEF, GRC, and CSS and those with low or intermediate SYNTAX score. Conversely, in the group with the highest SYNTAX score, the risk of cardiac mortality was significantly higher in PCI patients (hazard ratio: 2.323, 95% confidence interval: 1.091 to 4.945, p = 0.029).

Conclusions In LM patients undergoing PCI, combined scores improve the discrimination accuracy of clinical or angiographic stand-alone tools. In LM patients undergoing CABG, the ACEF score has the best prognostic accuracy compared with other stand-alone or combined scores. The good predictive ability for PCI along with the poor predictive ability for CABG make the SYNTAX score the preferable decision-making tool in LM disease. (J Am Coll Cardiol Intv 2011;4:287–97) © 2011 by the American College of Cardiology Foundation

From the *Cardiovascular Department, Ferrarotto Hospital, Catania, Italy; and the †ETNA Foundation, Catania, Italy.
Dr. Ussia is a physical proctor for CoreValve-Medtronic. Dr. Tamburino is a consultant for Abbott Vascular, CeloNova, and Medtronic. All other authors have reported that they have no relationships to disclose.

Manuscript received October 19, 2010, accepted October 29, 2010.

Scoring systems are valuable prognostic tools to predict outcomes, devise tailored therapies, and help patients and their families to get a better understanding of issues relevant to treatment strategies and subsequent risks. In the setting of coronary artery disease (CAD), clinical variables well-correlate with clinical end points, such as death or myocardial infarction (1). The introduction of the SYNTAX (Synergy between Percutaneous Coronary Intervention with TAXUS and Cardiac Surgery) score—which entails 11 angiographic variables to prospectively characterize the coronary vasculature—has recently shifted the attention toward the presumptive advantage of stratifying the individual risk according to lesion complexity, extension, and distribution (2,3). This approach promises to be useful in decision-making of patients with complex CAD, such as those with left main (LM) stenosis (4–6). However, some argue that clinical and angiographic information are both important in defining the risk of patients undergoing LM revascularization (7,8).

Abbreviations and Acronyms

ACEF = age, creatinine, ejection fraction

CABG = coronary artery bypass graft surgery

CAD = coronary artery disease

CI = confidence interval

CSS = Clinical SYNTAX (Synergy between Percutaneous Coronary Intervention with TAXUS and Cardiac Surgery) score

GRC = Global Risk Classification

HR = hazard ratio

IoS = Index of Separation

LM = left main

NRI = net reclassification improvement

PCI = percutaneous coronary intervention

See page 298

Two approaches are on track to incorporate clinical variables into a SYNTAX score-based model: the Global Risk Classification (GRC) and the Clinical SYNTAX score (CSS) (7,9). The GRC is a combination of EuroSCORE (European System for Cardiac Operative Risk Evaluation) and SYNTAX score strata, which identifies 3 categories of risk: high (EuroSCORE >6 and SYNTAX score >26), intermediate (EuroSCORE >6 or SYNTAX score >26), and low (EuroSCORE <6 and SYNTAX score <26) (7). The CSS is the combination of the SYNTAX score and the ACEF (age, creatinine, ejection fraction) score, the latter being a risk model limited to 3 independent pre-operative variables, including age, ejection fraction, and serum creatinine (9,10). Both these combined models have not yet been externally validated in patients with LM disease undergoing percutaneous or surgical revascularization.

To shed more light on the correlation between combined risk models and long-term cardiac mortality of patients with unprotected LM CAD, we applied the GRC and the CSS on a large contemporary cohort of patients treated with LM revascularization by either percutaneous coronary intervention (PCI) or coronary artery bypass graft surgery (CABG). Specific aims were: 1) to validate and compare the performance of the GRC and the CSS in LM patients undergoing PCI or CABG; 2) to evaluate whether the combined risk models would improve the individual ability of the ACEF score, the EuroSCORE, or the SYNTAX score to predict adverse outcomes in LM patients undergoing PCI or CABG; and 3) to investigate the potential for all the investigated scores (combined or component scores) to aid patient selection for CABG.

Methods

Study population. The CUSTOMIZE (Appraise a CUSTOMIZEd strategy for left main revascularization) registry is an ongoing registry holding data from 2 participating centers that performed PCI or CABG for revascularization treatment of consecutive patients with unprotected LM disease (defined as the presence of lesions with stenosis of at least 50% of vessel diameter) between March 2002 and June 2009. Clinical outcomes of a proportion of patients from the CUSTOMIZE registry have been reported elsewhere (5). Patients who had undergone previous CABG were excluded from the present analysis. The local ethics committee at each center approved the use of clinical data for this study, and all patients provided written informed consent. The authors wrote the manuscript and are responsible for the completeness and accuracy of data gathering and analysis.

Procedural and post-intervention practices. The decision to perform PCI instead of CABG was considered in the presence of suitable anatomy for stenting and preference by patient and referring physician for the percutaneous approach or in the presence of suitable anatomy and relative contraindications to surgery defined as a EuroSCORE ≥6. The interventional strategy as well as the choice of the various devices and the administration of therapies during the procedure were left to the discretion of the operator and standard practice. Stenting of LM bifurcation was allowed at the discretion of the operator. When a 2-stent strategy was adopted, T-stenting or mini-crush stenting were performed in almost all cases. After the procedure, patients treated with drug-eluting stents were prescribed clopidogrel for at least 6 months. Aspirin was prescribed indefinitely for all patients, irrespective of treatment with PCI or CABG. Surgical revascularization was performed with standard bypass techniques. Whenever possible, the internal thoracic artery was used preferentially for revascularization of the left anterior descending artery. In patients >70 years of age, arterial revascularization was strongly recommended. Patients could be operated either with or without extracorporeal circulation; in on-pump surgeries the type of cardioplegia was left to surgical judgment. The post-procedure medication regimen was chosen according to local clinical practice.

Scoring systems calculation and definition. All the variables required for calculating the different scores were obtained from the clinical-procedural database, assessed for quality and completeness, and introduced in a dedicated database.

Full details on SYNTAX score calculation are reported elsewhere (2). Briefly, all angiographic variables pertinent to SYNTAX score calculation were separately computed by 2 of 3 experienced cardiologists who were blinded to procedural data and clinical outcome by retrospectively analyzing the angiograms obtained before the procedure. In case of disagreement, the opinion of the third observer was obtained, and the final decision was made by consensus. The additive EuroSCORE was calculated on the basis of the original methodology (11). The ACEF score was calculated on the basis of the modified formula proposed by Ranucci et al. (10) (ACEF = [age/left ventricular ejection fraction] + 1 if serum creatinine >2 mg/dl]. The GRC and the CSS were derived as previously described (7,9). Three CSS categories were identified by tertiles (CSS_{LOW} <30, CSS_{MID} 31 to 49, CSS_{HIGH} >50). Three classes of risk were also grouped by tertiles for the ACEF score ($ACEF_{LOW}$ <1.2, $ACEF_{MID}$ 1.2 to 1.4, $ACEF_{HIGH}$ >1.4) and as previously reported for the EuroSCORE (12,13) ($EuroSCORE_{LOW}$ 0 to 2, $EuroSCORE_{MID}$ 2 to 6, $EuroSCORE_{HIGH}$ >6) and the SYNTAX score (4) ($SYNTAX_{LOW}$ 0 to 22, $SYNTAX_{MID}$ 23 to 32, $SYNTAX_{HIGH}$ >32) to compare their performance on the PCI and CABG study populations with those of the GRC and the CSS.

Study end point and data collection. The primary end point was the 2-year cumulative incidence of cardiac mortality. Cardiac mortality was defined as sudden death, fatal myocardial infarction, or death secondary to heart failure. Secondary end points were major adverse cardiac events (MACE) and death from all causes. Major adverse cardiac events were defined as the composite of death from all causes, nonfatal myocardial infarction, or target vessel revascularization. Clinical follow-up data related to medications and clinical status were prospectively collected through scheduled outpatient clinic evaluations. Referring cardiologists, general practitioners, and patients were contacted whenever necessary for further information. All repeated coronary intervention (surgical and percutaneous) and repeat-hospital-stay data were prospectively collected during follow-up with the centralized system of the participating institutions or directly contacting the hospitals where the patients were admitted or referred. Angiographic follow-up was suggested at 6 and 9 months after the index procedure in all consenting patients treated with PCI. It was performed at an earlier time if clinically indicated. However, patients who were at high risk for procedural complications of angiography and had no symptoms or signs of ischemia as well as patients who declined the recommendation did not undergo routine follow-up angiography. For patients who underwent CABG, angiographic follow-up was recommended only if there were ischemic symptoms or signs during follow-up. All outcomes of interest were confirmed by source documentation collected at each center and were centrally adjudicated by an independent, blinded end points committee.

Statistical analysis. All data were processed with the Statistical Package for Social Sciences (version 15, SPSS, Chicago, Illinois). Patient characteristics pertaining to the index procedure (PCI or CABG) were compared across GRC and CSS strata with an analysis of variance for continuous variables (expressed as mean ± SD) and the chi-square test for categorical variables (expressed as percentages). The Spearman's test was used to assess the correlation between the GRC and the CSS. Two-year cumulative rates of cardiac mortality were estimated by the Kaplan-Meier method, and the log-rank test was used to evaluate differences between groups. Although follow-up extended beyond 2 years in a proportion of patients at the time of data analysis, we restricted the follow-up to 2 years in all patients to account for bias introduced by incomplete follow-up. Patients lost to follow-up were considered at risk until the date of last contact, at which point they were censored.

The scores performances were evaluated in terms of calibration and discrimination. From the perspective of goodness-of-fit, calibration evaluates the degree of correspondence between the estimated probabilities produced by a model and the actual observation. For each score, it was measured by the Hosmer-Lemeshow test (the lower the statistic and higher the p values, the more calibrated is the score). A scoring system is expected, when stratifying the study population into 3 groups, to generate an intermediate stratum in which the observed risk ($P_{mid\text{-}observed}$) ideally matches the predicted risk ($P_{mid\text{-}expected}$) defined as (P_{worst} + P_{best})/2. The more the observed risk is close to this predicted value and their difference ($\Delta_{mid} = P_{mid\text{-}observed} - P_{mid\text{-}expected}$) is close to zero, the more the intermediate stratum will be calibrated and the risk stratification will be well-balanced across the groups. In addition, the highest, intermediate, and lowest probabilities must be in the anticipated order, for a score to be considered valid (14).

Discrimination is the probability that the score will assign higher values of risk to patients who will go on to have events compared with those who will not. It was measured with areas under the receiver-operator characteristic curves (AUCs), which range from 0.50 (no discrimination) to 1.0 (perfect discrimination). Discrimination was further assessed by means of the Index of Separation (IoS), defined as $P_{worst} - P_{best}$, assuming P_{worst} as the predicted p value of an event for a patient in the group with the highest score and P_{best} as the predicted p value of the same event for a patient in the group with the lowest score (14). The effects of reclassification with GRC or CSS were assessed with methods that estimate the net reclassification improvement (NRI), as previously described (7).

Control of potential confounders when comparing PCI and CABG outcomes after score stratification was attempted by developing a propensity score with logistic regression. The propensity score was the conditional probability of receiving either PCI or CABG, given a set of measured covariates. In our context, it was computed for each of the patients with a logistic regression model including diabetes mellitus, the ACEF score, the EuroSCORE, and the SYNTAX score. The selection of the variables, which formed a "minimum relevant" information set according to standards of propensity score application in health-care outcome, was based on a close relation with both treatment effect and the choice of treatment as assessed by univariate analysis. The population was then divided into quintiles according to the propensity score. Within each quintile, the mean propensity scores of PCI and CABG groups were compared, as were their clinical and procedural characteristics. Covariate interactions and higher-order terms for the continuous variables proved unnecessary for the balance of baseline characteristics across quintiles. The model was well-calibrated (Hosmer-Lemeshow test = 0.72) and showed a good discrimination (c-statistic = 0.71). The resulting propensity score was then included in the Cox proportional hazard models for 2-year cardiac mortality as a linear term, with the treatment group (PCI or CABG) as a covariate. No relevant changes were noted after forcing time-dependent variables (i.e., tertile of study period) to account for potential changes in practice, including use of longer durations of dual antiplatelet therapy as well as use of drug-eluting stents. According to this procedure, final results were presented as adjusted hazard ratio (HR) and 95% confidence interval (CI).

Results

Baseline characteristics of 949 consecutive patients with LM disease stratified by GRC or CSS categories and treatment type (PCI, n = 400; CABG, n = 549) are summarized in Tables 1 and 2. Among patients treated with PCI, the mean SYNTAX score, ACEF score, and additive EuroSCORE were 26.3 ± 11.7, 1.6 ± 0.7, and 5.6 ± 3.4, respectively. The same figures for patients treated with CABG were 33.2 ± 12.6 ($p < 0.001$), 1.3 ± 0.4 ($p < 0.001$), and 4.6 ± 2.7 ($p < 0.001$), respectively.

Significance of different risk models for prognostic stratification in PCI. The 2-year cumulative incidences of cardiac mortality across EuroSCORE, ACEF, and SYNTAX score risk strata of patients included in the PCI cohort are shown in Figure 1. The ACEF score displayed the best calibration characteristics (Hosmer-Lemeshow statistic 0.216, p = 0.642), whereas each score showed an acceptable discrimination ability, with AUCs ranging from 0.687 to 0.729 and IoS ranging from 0.140 to 0.167 (Table 3).

When assessed in the PCI cohort, the GRC and the CSS showed a strong correlation ($R = 0.751$, $p < 0.001$). Patients at low, intermediate, and high risk were 43.3%, 38.5%, and 18.3% on the basis of the GRC and 42.5%, 27.0%, and 30.5% on the basis of the CSS score, respectively (Table 1). A total of 130 patients (32.5%) were differently categorized on the basis of the 2 combined scores. The cumulative incidences of cardiac mortality were 0.6%, 8.7%, and 24.2% in the GRC_{LOW}, GRC_{MID}, and GRC_{HIGH} strata, respectively, and 2.6%, 1.0%, and 25.6% in the CSS_{LOW}, CSS_{MID}, and CSS_{HIGH} strata, respectively (Fig. 2). In 22 patients who experienced cardiac mortality at 2 years, compared with GRC, CSS improved classification in 7 and worsened it in 1, with a net gain in reclassification of 27.3%. In the 378 patients who did not die of cardiac mortality at 2 years, CSS reclassified 39 downward and 83 upward, with a net gain in reclassification of −11.6%. Therefore, the NRI for CSS over GRC was estimated to be 15.6%.

The 2 scores showed a similar index of separation (GRC IoS 0.236, CSS IoS 0.230)—higher than those described in the preceding text for the ACEF, the EuroSCORE, and the SYNTAX score—thus reflecting a better discrimination ability than the component scores. This was also in line with the finding of an AUC of 0.743 (95% CI: 0.645 to 0.842) for GRC and 0.762 (95% CI: 0.659 to 0.866) for CSS (Table 3). In terms of calibration, the observed event rate of the intermediate stratum was closest to the anticipated rate when the GRC was used (GRC Δ_{mid} −3.7%, CSS Δ_{mid} −13.1%), reflecting a better calibration ability of the GRC compared with the CSS. Consistently, Hosmer-Lemeshow statistics were 0.357 for GRC (p = 0.550) and 3.833 (p = 0.05) for CSS.

Online Tables 1 and 2 show the discrimination and calibration abilities of the 5 risk models when used to predict the risk of MACE or death from all causes. In patients undergoing PCI, all models displayed a poorer discrimination for predicting MACE than that observed for cardiac mortality. The SYNTAX score, the GRC, and the CSS also showed a poorer discrimination for predicting death from all causes compared with cardiac death; conversely, the EuroSCORE showed a similar discrimination and the ACEF score showed a higher discrimination for death from all causes compared with cardiac death. Calibration was generally unsatisfactory for both secondary end points, with the possible exception of the ACEF score and the GRC for MACE and the ACEF score for death from all causes.

Significance of different risk models for prognostic stratification in CABG. When applied on the CABG cohort, the GRC and CSS showed a moderate correlation ($R = 0.693$, $p < 0.001$). Patients at low, intermediate, and high risk were 28.4%, 53.9%, and 17.7% on the basis of the GRC and 26.6%, 38.1%, and 35.3% on the basis of the CSS score, respectively. A total of 194 patients (35.3%) were differently categorized on the basis of the 2 scores. At 2 years, the cumulative incidences of cardiac mortality were 2.7%, 5.3%,

Table 1. Baseline Characteristics in Patients With Left Main Disease Treated With PCI

	GRC				CSS			
	Low (n = 173)	Middle (n = 154)	High (n = 73)	p Value	Low (n = 170)	Middle (n = 108)	High (n = 122)	p Value
Age, yrs	61.5 ± 9.5	69.9 ± 9.4	76.8 ± 6.6	<0.001	61.8 ± 10.5	69.1 ± 8.6	74.1 ± 8.4	<0.001
Male	81.5	76.6	67.1	0.05	78.2	74.1	77.9	0.69
Risk factors								
Systemic hypertension	64.7	68.8	80.8	0.04	65.9	71.3	72.1	0.45
Hypercholesterolemia	60.7	55.8	43.8	0.05	59.4	53.7	52.5	0.44
Smoking habitus	51.4	35.7	35.6	0.007	48.8	40.7	35.2	0.06
Diabetes mellitus	27.7	33.1	35.6	0.39	27.6	30.6	36.9	0.24
Creatinine >2 mg/dl	1.7	1.9	11.0	0.001	0.6	1.9	9.0	<0.001
Medical history								
Previous MI	35.3	31.8	41.1	0.39	32.9	33.3	39.3	0.48
Peripheral artery disease	12.7	18.2	30.1	0.005	17.6	15.7	20.5	0.64
Previous PCI	37.0	19.5	16.4	<0.001	38.8	19.4	15.6	<0.001
Clinical presentation								
Stable angina	47.4	36.4	19.2	<0.001	46.5	33.3	30.3	0.01
UA/NSTEMI	47.4	54.5	76.7	<0.001	50.6	54.6	63.1	0.10
Acute MI	5.2	9.1	4.1	0.23	2.9	12.0	6.6	0.01
LVEF	52.6 ± 7.9	46.1 ± 11.3	41.3 ± 12.7	<0.001	53.4 ± 7	49.6 ± 9.3	39.3 ± 11.9	<0.001
EuroSCORE								
<3	35.8	14.9	0	<0.001	37.1	13.0	6.6	<0.001
3–6	64.2	40.3	0	<0.001	48.8	53.7	26.2	<0.001
>6	0	44.8	100	<0.001	14.1	33.3	67.2	<0.001
Lesion location								
Ostium	32.9	34.4	35.6	0.91	36.5	26.9	36.9	0.18
Shaft	13.3	13.6	13.7	0.99	13.5	13.0	13.9	0.98
Distal	53.8	51.9	50.7	0.89	50	60.2	49.2	0.17
Extent of CAD								
Isolated LMCA disease	13.0	5.5	1.8	0.01	16.0	3.2	2.0	<0.001
LMCA plus 1-vessel disease	50.0	22.3	14.3	<0.001	53.3	24.4	13.9	<0.001
LMCA plus 2-vessel disease	28.8	36.6	28.6	0.36	23.0	45.3	31.6	0.002
LMCA plus 3-vessel disease	7.5	35.7	54.8	<0.001	6.7	26.7	53.2	<0.001
SYNTAX score	18 ± 5.2	29.5 ± 11	39.3 ± 8.3	<0.001	16.9 ± 5.2	27.3 ± 6.8	38.4 ± 9.8	<0.001

Values are mean ± SD or %.

CAD = coronary artery disease; CSS = clinical SYNTAX (Synergy between Percutaneous Coronary Intervention with TAXUS and Cardiac Surgery) score; EuroSCORE = European System for Cardiac Operative Risk Evaluation; GRC = Global Risk Classification; LMCA = left main coronary artery; LVEF = left ventricular ejection fraction; MI = myocardial infarction; NSTEMI = non–ST-segment elevation myocardial infarction; PCI = percutaneous coronary intervention; UA = unstable angina.

and 9.6% in the GRC_{LOW}, GRC_{MID}, and GRC_{HIGH} strata, respectively, and 3.9%, 2.7%, and 9.1% in the CSS_{LOW}, CSS_{MID}, and CSS_{HIGH} strata, respectively (Fig. 2). In 22 patients who experienced cardiac mortality at 2 years, compared with GRC, CSS improved classification in 7 and worsened it in 1, with a net gain in reclassification of 27.3%. In the 527 patients who did not die of cardiac mortality at 2 years, CSS reclassified 44 downward and 142 upward, with a net gain in reclassification of −18.6%. Therefore, the NRI for CSS over GRC was estimated to be 8.7%.

Both scores showed a mild-to-moderate discrimination ability (GRC AUC: 0.616, 95% CI: 0.499 to 0.734; CSS AUC: 0.637, 95% CI: 0.513 to 0.761), with narrow IoS. Hosmer-Lemeshow statistics were 0.019 for GRC (p = 0.891) and 2.534 (p = 0.111) for CSS, reflecting a better calibration ability of GRC.

Interestingly, the best discrimination in CABG patients was provided by the ACEF score (AUC 0.741, 95% CI: 0.650 to 0.832), which also showed the larger IoS (0.111) compared with all the other models, including the GRC and the CSS (Table 3). The SYNTAX score showed the poorest discrimination ability, with large superimposition and no statistically significant separation of the Kaplan Meier curves for cardiac mortality at 2 years (p = 0.589) (Fig. 1).

When used to predict MACE in subjects undergoing CABG, all 5 scores showed poor discrimination and calibration abilities. When used to predict death from all causes, the only score with satisfactory discrimination was the ACEF score (Online Tables 1 and 2).

JACC: CARDIOVASCULAR INTERVENTIONS, VOL. 4, NO. 3, 2011
MARCH 2011:287–97

Table 2. Baseline Characteristics in Patients With Left Main Disease Treated With CABG

	GRC				CSS			
	Low (n = 156)	Middle (n = 296)	High (n = 97)	p Value	Low (n = 146)	Middle (n = 209)	High (n = 194)	p Value
Age, yrs	60.9 ± 10.1	65.7 ± 9.2	73.6 ± 7.2	<0.001	59 ± 10.4	66 ± 8.7	70.3 ± 8.1	<0.001
Male	80.8	81.4	69.1	0.03	78.1	77.5	81.4	0.59
Risk factors								
Systemic hypertension	67.3	73.6	82.5	0.03	66.4	76.6	75.3	0.08
Hypercholesterolemia	52.6	57.4	49.5	0.33	53.4	57.4	52.6	0.59
Smoking habitus	57.1	47.0	35.1	0.003	60.3	39.7	46.9	0.001
Diabetes mellitus	30.8	44.6	44.3	0.12	28.1	42.1	48.5	0.001
Creatinine >2 mg/dl	1.3	1.0	2.1	0.73	0.0	0.5	3.1	0.02
Medical history								
Previous MI	26.9	28.7	45.4	0.004	26.7	23.4	42.8	<0.001
Peripheral artery disease	5.8	14.2	27.8	<0.001	6.8	16.3	17.5	0.01
Previous PCI	19.9	9.8	9.3	0.005	23.3	10.5	6.7	<0.001
Clinical presentation								
Stable angina	49.4	46.6	17.5	<0.001	49.3	42.6	36.6	0.06
UA/NSTEMI	50.6	53.4	79.4	<0.001	50.7	56.9	62.4	0.10
Acute MI	0	0	3.1	0.001	0	0.5	1.0	0.44
LVEF	53.4 ± 7.4	52.8 ± 7.6	45.8 ± 10.1	<0.001	55.5 ± 4.6	53.7 ± 6.9	46.9 ± 10	<0.001
EuroSCORE								
<3	37.8	25.0	0	<0.001	47.9	21.1	9.8	<0.001
3–6	62.2	61.8	0	<0.001	43.8	59.8	46.9	0.005
>6	0	13.2	100	<0.001	8.2	19.1	43.3	<0.001
Lesion location								
Ostium	20.7	13.4	11.1	0.07	20.9	13.5	12.5	0.08
Shaft	5.3	4.0	12.2	0.01	5.8	5.2	6.5	0.86
Distal	74	82.6	76.7	0.09	73.4	81.3	81.0	0.16
Extent of CAD								
Isolated LMCA disease	13.1	1.8	0	<0.001	14.8	1.5	0.5	<0.001
LMCA plus 1-vessel disease	24.8	9.6	6.8	<0.001	26.8	12.0	4.7	<0.001
LMCA plus 2-vessel disease	32.7	31.9	26.9	0.57	34.5	36.0	23.6	0.02
LMCA plus 3-vessel disease	29.4	56.7	66.3	<0.001	23.9	50.0	70.7	<0.001
SYNTAX score	19.7 ± 4.9	37.5 ± 10.6	41.8 ± 9.6	<0.001	19.6 ± 5.6	32 ± 7.1	44.9 ± 9.7	<0.001

Values are mean ± SD or %.
CABG = coronary artery bypass graft surgery; other abbreviations as in Table 1.

Significance of different risk models for treatment selection. Baseline differences in clinical and angiographic characteristics of patients treated with PCI or CABG are displayed in Online Table 3. Patients undergoing PCI were older, had lower left ventricular ejection fraction, and more frequently presented with acute myocardial infarction and EuroSCORE >6. The unadjusted rates of 2-year cardiac mortality were 8.0% for PCI and 5.3% for CABG ($p = 0.305$). Patients undergoing PCI experienced higher rates of MACE (23.1% vs. 9.7%, $p < 0.001$) and death from all causes (12.6% vs. 6.5%, $p = 0.016$).

Differences between PCI and CABG were investigated by baseline SYNTAX score, EuroSCORE, ACEF score, GRC, and CSS strata for 2-year cardiac mortality. The propensity score adjusted outcomes were comparable between PCI and CABG with either low, intermediate, or high EuroSCORE and ACEF score (Fig. 3). Cardiac mortality was also similar between PCI and CABG in the $SYNTAX_{LOW}$ and $SYNTAX_{MID}$ strata; but in the $SYNTAX_{HIGH}$ group, the risk of cardiac mortality was significantly higher in patients treated with PCI (HR: 2.323, 95% CI: 1.091 to 4.945, $p = 0.029$). Although a gradient in the cardiac mortality risk was seen across risk categories, neither GRC (HR: 1.760, 95% CI: 0.648 to 4.779, $p = 0.267$) nor CSS (HR: 1.711, 95% CI: 0.810 to 3.615, $p = 0.159$) were successful in displaying statistically significant differences between PCI and CABG in the high-risk subgroup (Fig. 3).

Discussion

This study adds to the evidence on prognostic scores in LM CAD with the following observations. First, in patients

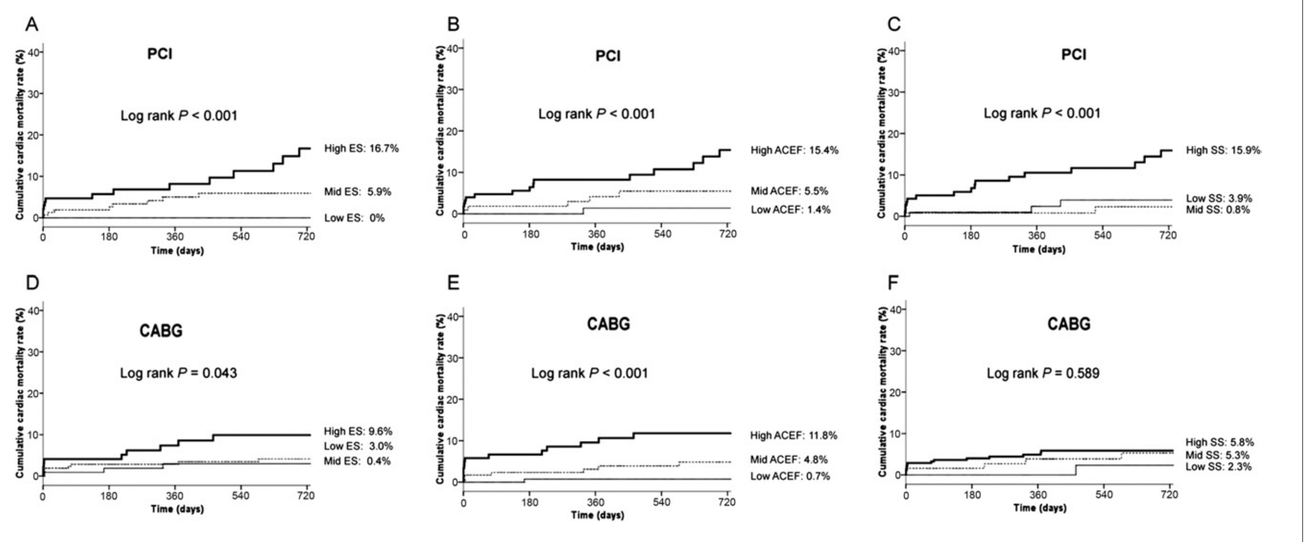


Figure 1. Risk Prediction of Stand-Alone Models

Kaplan-Meier estimates of 2-year cardiac mortality by baseline EuroSCORE (ES); ACEF (age, creatinine, ejection fraction) score; and SYNTAX (Synergy between Percutaneous Coronary Intervention with TAXUS and Cardiac Surgery) score (SS) risk categories in patients treated by percutaneous coronary intervention (PCI) **(A to C)** or coronary artery bypass grafting (CABG) **(D to F)**. The p values from log-rank test.

treated with PCI, the use of combined clinical and angiographic scores such as the GRC and the CSS yields a greater discrimination than that obtained by stand-alone tools such as the SYNTAX score, the EuroSCORE, and the ACEF score for assessing the risk of cardiac mortality. Of note, CSS uses fewer data to achieve similar discrimination and poorer calibration than GRC. Second, in patients treated with CABG, the ACEF and the SYNTAX scores yield the best and the worst prediction abilities for cardiac mortality, respectively, whereas the GRC and the CSS are somewhat discriminative but to a lesser extent than that observed in the PCI cohort. Third, all prognostic models generally yield superior results when used to assess cardiac mortality and mortality from all causes; conversely, their performance worsens when used to predict the risk of MACE. Finally, patient selection for CABG is aided by the SYNTAX score but not by the GRC, the CSS, the ACEF, and the EuroSCORE.

Estimates of individual and population-based outcomes after invasive procedures are of paramount importance in medicine (15). The SYNTAX LM substudy has recently confirmed the excellent concordance between clinical outcomes and baseline SYNTAX score (4). In fact, PCI mortality was nicely stratified when analyzing the 1-year outcomes, according to the baseline SYNTAX score tertile. This result is consistent with those of other study populations (16,17). In the SYNTAX LM cohort, 1-year

Table 3. Calibration and Discrimination Parameters

	SYNTAX Score	EuroSCORE	ACEF Score	GRC	CSS
PCI					
HL statistic (p value)	2.448 (0.118)	1.607 (0.205)	0.216 (0.642)	0.357 (0.550)	3.833 (0.05)
AUC (95% CI)	0.729 (0.617–0.841)	0.687 (0.591–0.783)	0.688 (0.592–0.785)	0.743 (0.645–0.842)	0.762 (0.659–0.866)
IoS*	0.151	0.167	0.140	0.236	0.230
Δ_{mid}†	−4.5%	−2.5%	−2.9%	−3.7%	−13.1%
CABG					
HL statistic (p value)	0.098 (0.754)	0.321 (0.571)	0.426 (0.514)	0.019 (0.891)	2.534 (0.111)
AUC (95% CI)	0.556 (0.439–0.673)	0.622 (0.500–0.744)	0.741 (0.650–0.832)	0.616 (0.499–0.734)	0.637 (0.513–0.761)
IoS*	0.035	0.095	0.111	0.069	0.052
Δ_{mid}†	1.3%	−2%	−1.5%	−0.9%	−3.8%

*The index of separation (IoS) is defined as $P_{worst} - P_{best}$, assuming P_{worst} as the predicted p value of event for a patient in the group with the highest score and P_{best} as the predicted p value of the same event for a patient in the group with the lowest score. †The Δ_{mid} is defined as $P_{mid\text{-}observed} - P_{mid\text{-}expected}$, assuming $P_{mid\text{-}observed}$ as the observed p value of event for a patients in the group with the intermediate score and $P_{mid\text{-}expected}$ as the anticipated p value of event for a patients in the group with the intermediate score.

ACEF = age, creatinine, ejection fraction; AUC = area under the curve; CI = confidence interval; HL = Hosmer-Lemeshow; other abbreviations as in Tables 1 and 2.

JACC: CARDIOVASCULAR INTERVENTIONS, VOL. 4, NO. 3, 2011
MARCH 2011:287–97

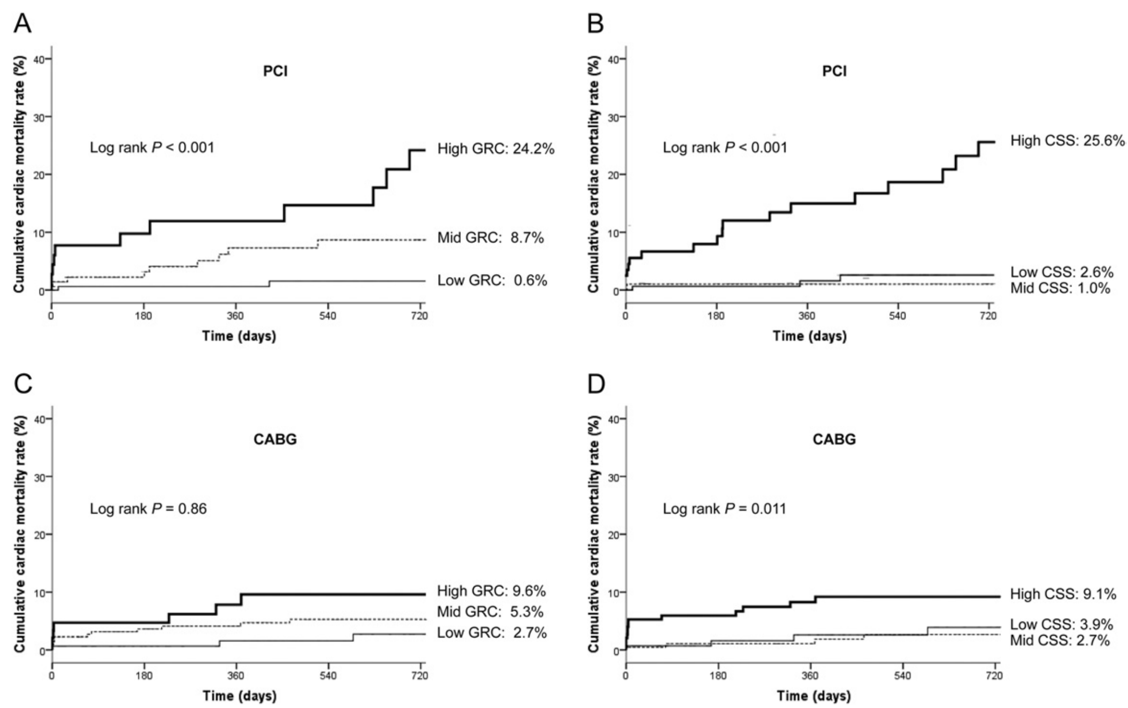


Figure 2. Risk Prediction of Combined Models

Kaplan-Meier estimates of 2-year cardiac mortality by baseline Global Risk Classification (GRC) and Clinical SYNTAX score (CSS) risk categories in patients treated by PCI **(A, B)** or CABG **(D, E)**. The p values from log-rank test. Abbreviations as in Figure 1.

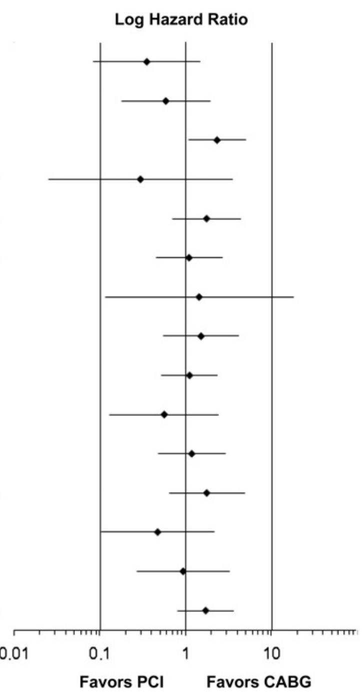
mortality after CABG did not stratify according to the baseline SYNTAX score tertile (4), in contrast to PCI.

A key concern that currently limits the application of the SYNTAX score is that the score algorithm does not entail any clinical variable. Comorbidities are known to impact early outcomes of patients undergoing CABG, but a role in worsening the operative morbidity of patients undergoing PCI has also been suggested (18). We have previously demonstrated in a small cohort of patients undergoing LM PCI that the addition of clinical variables to the SYNTAX score improves its accuracy in terms of both discrimination and calibration (7). An improvement in the ability of the SYNTAX score to predict mortality by incorporating the ACEF score to obtain the CSS has also been demonstrated in patients with multivessel disease (9).

The present study retrospectively validates for the first time the combined (clinical and angiographic) GRC and CSS in a contemporary dataset of LM patients treated by PCI or CABG. Although both scores confirmed superior discrimination compared with other stand-alone scores in the setting of PCI, the CSS showed an NRI over GRC (15.6% for PCI and 8.7% for CABG) but lower calibration. In particular, although the CSS was associated with an approximately 27% positive reclassification of patients who experienced cardiac mortality, the GRC was associated with an approximately 12% to 19% positive reclassification of patients who survived. This is in line with the observation that the GRC—due to its higher calibration ability—tends to enhance the prognostic utility of each risk category, including the intermediate stratum, whereas the CSS is possibly associated with a "threshold effect" that prompts the identification of a subpopulation of patients at high chance of dying from cardiac causes, thus making the intermediate stratum less important.

We found that the ACEF score emerged as the best predictive tool for long-term cardiac mortality in the setting of CABG. The potential for parsimonious models to offer the same or even better level of accuracy of more complex clinical scores in the CABG scenario is consistent with previous studies (10,19). The notion that the GRC and the CSS suboptimally stratified cardiac mortality in CABG compared with PCI is likely to be the reflection of the less-essential impact of the SYNTAX score in lesions bypassed surgically (4,6).

In aggregate, the comparison of 5 scores in 2 PCI and CABG populations from the CUSTOMIZE registry supports the understanding that the best balance in terms of discrimination and calibration for cardiac mortality might be offered by the GRC for LM patients undergoing PCI and the ACEF score for those undergoing CABG (Fig. 4). Interestingly, the ability of the different prognostic models to predict hard end points, such as cardiac mortality and



	HR	LCL	UCL	p
SYNTAX score 0-22	0.348	0.084	1.445	0.146
SYNTAX score 23-32	0.582	0.177	1.910	0.372
SYNTAX score >32	2.323	1.091	4.945	0.029
EuroSCORE 0-2	0.293	0.025	3.482	0.331
EuroSCORE 3-6	1.753	0.709	4.339	0.224
EuroSCORE >6	1.091	0.452	2.638	0.846
ACEF < 1.2	1.431	0.115	17.804	0.781
ACEF 1.2-1.4	1.511	0.552	4.134	0.421
ACEF > 1.4	1.107	0.521	2.352	0.793
GRC low	0.557	0.129	2.407	0.434
GRC mid	1.174	0.475	2.904	0.728
GRC high	1.760	0.648	4.779	0.267
CSS low	0.467	0.103	2.122	0.324
CSS mid	0.929	0.268	3.218	0.908
CSS high	1.711	0.810	3.615	0.159

Figure 3. Usefulness of Different Risk Models on Decision-Making

Forrest plot of propensity score adjusted risk of 2-year cardiac mortality across risk strata of SYNTAX score, EuroSCORE, ACEF score, GRC, and CSS. HR = hazard ratio; LCL = lower confidence limit; UCL = upper confidence limit; other abbreviations as in Figure 1.

mortality from all causes, was superior to their ability to predict soft end points, such as MACE. This trend has been consistently noted across previous studies (16,18,20,21).

There is another consideration that might be drawn from this study. The ideal features of a score for being helpful in guiding decision-making between 2 different procedures do not necessarily match those features that characterize scores with superior discrimination and calibration ability. In fact, a score is supportive in driving treatment strategies if risk strata exist that display a significant separation of the outcomes between different procedures. This is more likely to occur if the score is predictive for a particular strategy (e.g., PCI) and not predictive for another strategy (e.g., CABG). In our study, this combination occurred only with the SYNTAX score, whereas the other scores showed a certain grade of accuracy, albeit different in magnitude, both in the PCI and the CABG scenario. As a result, it is not surprising that the SYNTAX score was the only tool highlighting the potential to aid decision-making in the LM setting, consistent with previous reports (4,5). Conversely, the addition of clinical variables to the SYNTAX score improves the prognostic accuracy not only in patients undergoing PCI but also in patients treated with CABG and therefore paradoxically exerts a detrimental effect in separating the effects of the 2 procedures as far as the coronary complexity increases.

Study limitations. This study is limited by its post-hoc nature. The CUSTOMIZE registry is an observational study run over a long period, in which the treatment strategies for the PCI cohort might have changed over time. We cannot tell what impact these changes might have had on the results. However, time-dependent variables did not prove to be necessary to improve the accuracy of our model. Information about secondary preventive medications in the PCI and CABG cohorts were not fully available. Statistical adjustment when comparing PCI and CABG across risk strata was attempted by means of a parsimonious propensity-score model with good discrimination and calibration abilities. However, the impact of unidentified confounders on the results of this comparison cannot be entirely ruled out. Finally, a significant reduction of 2-year cardiac mortality was associated with CABG in the highest SYNTAX score stratum, at a $p = 0.029$ level. Because 5 subgroups analyses were performed for the PCI versus CABG comparison, a more conservative alpha value (i.e., $0.05/5 = 0.01$) would have been indicated. Therefore, play of chance for this result cannot be entirely ruled out. However, the finding that LM patients with higher SYNTAX score experience more adverse outcomes with PCI compared with CABG is consistent with previous studies (4,5).

JACC: CARDIOVASCULAR INTERVENTIONS, VOL. 4, NO. 3, 2011
MARCH 2011:287–97

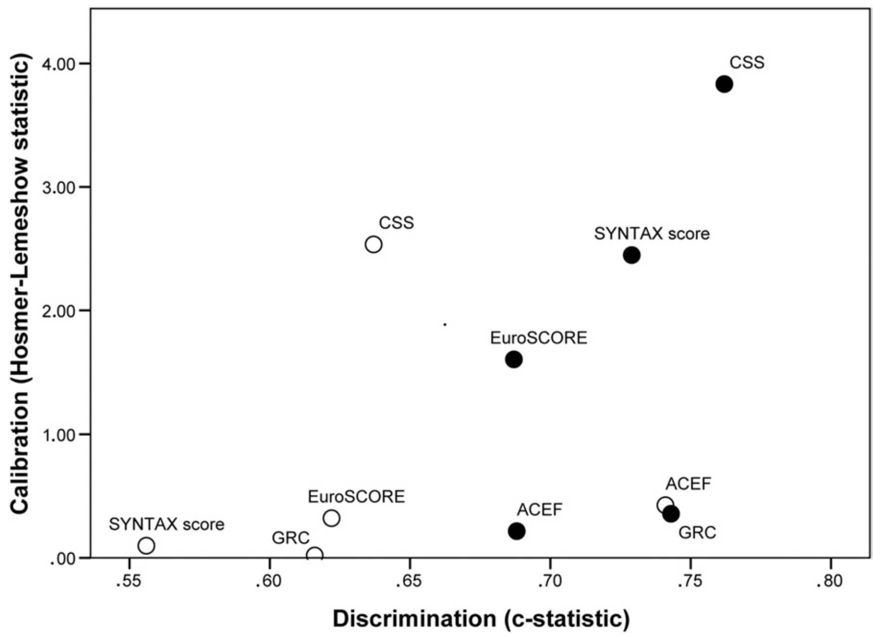


Figure 4. Calibration and Discrimination Abilities of Different Risk Models

Prediction accuracy of SYNTAX score, EuroSCORE, ACEF score, GRC, and CSS in patients treated with PCI **(solid circles)** or CABG **(empty circles)**. The position of the score in the graph depends on the balance between the score calibration (assessed by the Hosmer-Lemeshow statistic) and the discrimination (assessed by the area under the receiver-operator characteristic curve). The prognostic accuracy of each score increases from top to the bottom and from left to right. Abbreviations as in Figure 1.

Conclusions

In patients with LM CAD undergoing PCI, combined clinical and angiographic scoring systems improve the prognostic accuracy of clinical or angiographic stand-alone tools. In this context, the CSS uses fewer data to achieve similar discrimination but poorer calibration than the GRC. In LM patients undergoing CABG, the ACEF score has the best prognostic accuracy compared with other stand-alone or combined scoring systems. The good predictive ability in the PCI scenario along with the poor predictive ability in the CABG scenario make the SYNTAX score the preferable tool to guide decision-making in LM CAD.

Reprint requests and correspondence: Dr. Davide Capodanno, Cardiology Department, Ferrarotto Hospital, University of Catania, Via Citelli 6, 95124 Catania, Italy. E-mail: dcapodanno@gmail.com.

REFERENCES

1. Singh M, Rihal CS, Lennon RJ, Garratt KN, Holmes DR Jr. Comparison of Mayo Clinic risk score and American College of Cardiology/American Heart Association lesion classification in the prediction of adverse cardiovascular outcome following percutaneous coronary interventions. J Am Coll Cardiol 2004;44:357–61.
2. Sianos G, Morel MA, Kappetein AP, et al. The SYNTAX Score: an angiographic tool grading the complexity of coronary artery disease. EuroIntervention 2005;1:219–27.
3. Dawkins KD, Morel MA, Serruys PW. Counting the score: the SYNTAX Score and coronary risk. EuroIntervention 2009;5:33–5.
4. Morice MC, Serruys PW, Kappetein AP, et al. Outcomes in patients with de novo left main disease treated with either percutaneous coronary intervention using paclitaxel-eluting stents or coronary artery bypass graft treatment in the Synergy between Percutaneous Coronary intervention with TAXUS and Cardiac Surgery (SYNTAX) trial. Circulation 2010;121:2645–53.
5. Capodanno D, Capranzano P, Di Salvo ME, et al. Usefulness of SYNTAX score to select patients with left main coronary artery disease to be treated with coronary artery bypass graft. J Am Coll Cardiol Intv 2009;2:731–8.
6. Feldman T. The SYNTAX score in practice: an aid for patient selection for complex PCI. Catheter Cardiovasc Interv 2009;73:618–9.
7. Capodanno D, Miano M, Cincotta G, et al. EuroSCORE refines the predictive ability of SYNTAX score in patients undergoing left main percutaneous coronary intervention. Am Heart J 2010;159:103–9.
8. Lozano I, Avanzas P, Rondan J. SYNTAX score and left main stenting: do we need clinical variables to predict outcomes? Am Heart J 2010;159:e25.
9. Garg S, Sarno G, Garcia-Garcia HM, et al., ARTS-II Investigators. A new tool for the risk stratification of patients with complex coronary artery disease: the Clinical SYNTAX score. Circ Cardiovasc Interv 2010;3:317–26.
10. Ranucci M, Castelvecchio S, Menicanti L, Frigiola A, Pelissero G. Risk of assessing mortality risk in elective cardiac operations: age, creatinine, ejection fraction, and the law of parsimony. Circulation 2009;119:3053–61.
11. Singh M, Rihal CS, Lennon RJ. A critical appraisal of current models of risk stratification for percutaneous coronary interventions. Am Heart J 2005;149:753–60.

12. Calafiore AM, Di Mauro M, Canosa C, Di Giammarco G, Iaco AL, Contini M. Early and late outcome of myocardial revascularization with and without cardiopulmonary bypass in high risk patients (EuroSCORE > or = 6). Eur J Cardiothorac Surg 2003;23:360–7.
13. Pavei A, Oreglia JA, Martin G, et al. Long-term follow-up of percutaneous coronary intervention of unprotected left main lesions with drug eluting stents: predictors of clinical outcome. EuroIntervention 2009;4:457–63.
14. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med 2000;19:453–73.
15. Kereiakes DJ. Cultivating prognosis following percutaneous coronary intervention: the American College of Cardiology/National Cardiovascular Data Registry risk score. J Am Coll Cardiol 2010;55:1933–5.
16. Capodanno D, Di Salvo ME, Cincotta G, Miano M, Tamburino C, Tamburino C. Usefulness of the SYNTAX score for predicting clinical outcome after percutaneous coronary intervention of unprotected left main coronary artery disease. Circ Cardiovasc Interv 2009;2:302–8.
17. Brito J, Teles R, Sousa P, et al. Predictive value of SYNTAX score in risk stratification of patients undergoing left main coronary artery angioplasty (abstr). EuroIntervention 2010;6:H69.
18. Romagnoli E, Burzotta F, Trani C, et al. EuroSCORE as predictor of in-hospital mortality after percutaneous coronary intervention. Heart 2009;95:43–8.
19. Ranucci M, Castelvecchio S, Menicanti L, Frigiola A, Pelissero G. Accuracy, calibration and clinical performance of the EuroSCORE: can we reduce the number of variables? Eur J Cardiothorac Surg 2010;37:724–9.
20. Garg S, Stone GW, Kappetein AP, Sabik JF III, Simonton C, Serruys PW. Clinical and angiographic risk assessment in patients with left main stem lesions. J Am Coll Cardiol Intv 2010;3:891–901.
21. Kim YH, Park DW, Kim WJ, et al. Validation of SYNTAX (Synergy between PCI with Taxus and Cardiac Surgery) score for prediction of outcomes after unprotected left main coronary revascularization. J Am Coll Cardiol Intv 2010;3:612–23.

Key Words: coronary artery bypass graft (CABG) ■ percutaneous coronary intervention (PCI) ■ revascularization ■ risk score.

APPENDIX

For supplementary tables, please see the online version of this article.