Agreement between three perioperative risk scores

Gilson Soares Feitosa-Filho[1]*, Bruna Melo Coelho Loureiro[2], Jedson dos Santos Nascimento[3]

[1]PhD in Cardiology from Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (InCor-HCFMUSP). Adjunct Professor, Escola Bahiana de Medicina e Saúde Pública. Coordinator of Research Records at Hospital Santa Izabel – Santa Casa de Misericórdia da Bahia, Salvador, BA, Brazil
[2]Medical Student, Escola Bahiana de Medicina e Saúde Pública, Salvador, BA, Brazil
[3]PhD in Anesthesiology – Responsible for the Medical Residency in Anesthesiology at Santa Casa de Misericórdia da Bahia, Salvador, BA, Brazil

Study conducted at Hospital Santa Izabel/Santa Casa de Misericórdia da Bahia, Salvador, BA, Brazil

Article received: 7/26/2015
Accepted for publication: 8/18/2015

***Correspondence:**
Address: Praça Cons. Almeida Couto, 500, Nazaré
Salvador, BA – Brazil
Postal code: 40050-410
Phone: +55 71 8748-8115
gilsonfeitosafilho@yahoo.com.br

http://dx.doi.org/10.1590/1806-9282.62.03.276

Summary

Objective: To evaluate the agreement between the three scores proposed by the II Guideline for Perioperative Evaluation of the Brazilian Society of Cardiology (SBC): the American College of Physicians algorithm (ACP), the Multicenter Study of Perioperative Evaluation (EMAPO) and Lee's Revised Cardiac Risk Index (RCRI).
Method: Patients evaluated preoperatively for non-cardiac surgery by the anesthesiology service were classified as low, moderate or high-risk according to the 3 algorithms suggested by the II Guideline. To calculate the strength of agreement between the scores, the kappa agreement index was used.
Results: Four hundred and one patients were included in the sample. Cohen's kappa inter-rater agreement between scores was 0.270 (CI: 0.222 to 0.318), corresponding to a weak agreement. Analyzing in pairs, the best correlation was between EMAPO and ACP, with kappa = 0.327. Lee's score was the one that classified more patients as low-risk: 98.3%, while EMAPO and ACP classified as low risk 91.3% and 92.5%, respectively.
Conclusion: There is poor correlation among the risk scores proposed by the II Perioperative Evaluation Guideline of the SBC.

Keywords: perioperative care, cardiovascular diseases, postoperative complications.

Introduction

The Perioperative Assessment Guidelines of the Brazilian Society of Cardiology (SBC)[1,2] are excellent guides for medical conduct throughout the perioperative period and have been frequently used in various services in the country. Although different scales assess different outcomes, the II Guideline suggests using one of the following three scales in the preoperative assessment of cardiovascular risk: the algorithm of the American College of Physicians (ACP), the Multicenter Perioperative Assessment Study (EMAPO) and Lee's Revised Cardiac Risk Index (IRCR). These scales have advantages and disadvantages that should be considered during use.

The RCRI, also called the Lee score, is composed of only 6 risk variables, and is therefore easy to apply. In it, patients are divided into four classes of risk: I, II, III or IV based on the sum of the variables presented by them. This index seeks to estimate the probability of heart complications (myocardial infarction, pulmonary edema, primary cardiac arrest or ventricular fibrillation, and complete heart block) up to the 5th day after surgery. This score was derived and validated from a prospective investigation of 4,315 patients, which showed that this is a simple index with good ability to predict events, especially if compared with other indexes.[3,4]

The ACP assigns scores according to clinical and laboratory variables. It puts great value on the presence of clinical symptoms and electrocardiograph changes, with greater accuracy suggested in patients with reduced functional capacity.[5,6] In Brazil, the ACP has been validated in a study conducted at FMUSP's Hospital das Clínicas, in which the likelihood of cardiac events was 61.1, 11.6 and 2.2% for class III (high risk), class II (intermediate risk) and class I (low risk), respectively.[7]

The EMAPO is a scale developed in Brazil. It proposes a classification that includes a large number of vari-

ables, stratifying perioperative risk in 5 levels (very low, low, moderate, high and very high), according to the expected rates of complications and death that occur in the postoperative period prior to discharge from hospital.[8]

This study aimed to assess the correlation between the three scores proposed by the II Guideline for Perioperative Evaluation of the Brazilian Society of Cardiology: Lee, ACP and EMAPO.

Method

This is a single-centered study at a large teaching hospital with a high volume of surgeries. We consecutively reviewed all of the electronic medical records of patients assessed preoperatively for surgery by the Anesthesiology Department at Hospital Santa Izabel, Santa Casa de Misericórdia of Bahia, Brazil, in the months of March, April and May 2012. Before the start of data collection, it was established that the preoperative assessment forms of this hospital were structured as standardized, multiple choice fields covering all the information necessary for the application of the three scores. The forms were completed by the anesthesiologists, after initial training. Although several surgical teams were assisted, there was only one anesthesiology team assessing patients in the hospital. The collection of data from the structured records was performed by a single investigator and organized in SPSS (Statistical Package for Social Science). Based on the information retrieved, surgical risks were calculated for the 3 scores proposed by SBC's II Guideline for Perioperative Evaluation (ACP, EMAPO and Lee). Patients with preoperative assessment for cardiac surgery were excluded.

In order to properly compare the three scores containing risk categories with similar outcomes, they had to be reclassified as proposed by the II Guideline. Risk scores were reclassified into three categories: low risk (up to 3.0% of complications – myocardial infarction, cardiac arrhythmias, pulmonary edema – and/or death due to heart problems up to postoperative hospital discharge); intermediate risk (from 3.0 to 15.0%) and high risk (>15.0%). Low risk included classes I and II from Lee, up to 5 points in the EMAPO and low risk from the ACP. The intermediate risk was represented by classes III and IV from Lee (with heart failure or angina, maximum functional class II), 6 to 10 points in the EMAPO and intermediate risk from the ACP. Lastly, the high risk category was composed by: class III and IV from Lee (with heart failure or angina functional class III or IV), > or equal to 11 points in the EMAPO and a high risk classification from the ACP.[2]

To assess the degree of consistency between the classifications assigned by the 3 risk scores, Cohen's kappa index was calculated, considering as poor agreement a kappa index with values between 0 and 0.19; weak agreement between 0.20 and 0.39; moderate agreement between 0.40 and 0.59; good agreement between 0.60 and 0.79, and excellent agreement between 0.8 and 1.0.[9] Weighted kappa index was calculated to analyze the general agreement between all of the scores. The confidence interval (CI) established was 95%. The statistical procedures were performed using the SPSS software, version 12.0.

The research project was approved by the Prof. Dr. Celso Figueirôa Research Ethics Committee –Hospital Santa Izabel, under opinion CEP No. 68/2012.

Results

We evaluated the medical records of all 421 patients assessed in the anesthesiology clinic. 20 patients with a preoperative assessment for cardiac surgery were excluded. Minor procedures prevailed (295 surgeries – 73.6%), followed by medium (106 surgeries – 26.4%) and no major surgeries. The assessments included a large diversity of procedures, namely: hysteroscopy with biopsy or myomectomy (52 patients), colonoscopy with possible biopsy (47 patients), adenotonsillectomy (31 patients), varicectomy of the limbs (23 patients), hysterectomy (20 patients), prostatectomy (16 patients), nasal septoplasty (14 patients), and cholecystectomy (13 patients).

256 of the total sample in this study (63.8%) were women. The median age of the patients was 46 years (IQR = 30-62 years). The most common comorbidities were hypertension (132 patients - 32.9%), *diabetes mellitus* (53 patients – 13.2%), coronary artery disease (29 patients – 7.2%) and heart failure (21 patients – 5.2%).

320 of the 401 patients stratified by their scores (79.8%) underwent surgery at the hospital in which the study took place, with postoperative records also analyzed. The median amount of days the patients were hospitalized was 1. Only 15 patients (4.7%) were admitted to the intensive care unit (ICU) and 3 (0.9%) died.

The distribution of low-risk patients after reclassification was: 366 (91.3%) patients by the EMAPO; 371 (92.5%) by the ACP; and 394 (98.3%) according to Lee. For intermediate risk, the distribution was: 20 (4.9%) patients by the EMAPO; 30 (7.5%) by the ACP; and 2 (0.5%) according to Lee. High-risk was allocated to patients as follows: 15 (3.7%) in the EMAPO; none in the ACP; and 2 (0.5%) in Lee (Table 1).

TABLE 1 Distribution of patients after reclassification of surgical risk, according to each score.

Classification	Scores		
	EMAPO	ACP	Lee
Low risk	366	371	394
Intermediate risk	20	30	5
High risk	15	0	2

EMAPO: Multicenter Study of Perioperative Evaluation; ACP: American College of Physicians.

The overall kappa correspondence between all the scores (Table 2) was 0.270 (CI: 0.222 – 0.318), corresponding to a weak correspondence. Analyzing the correspondence between the scores in pairs (Table 3), the agreement levels found were: EMAPO and ACP (kappa = 0.327), ACP and Lee (kappa = 0.280) and EMAPO with Lee (kappa = 0.196).

TABLE 2 Overall agreement of all risk scores.

Overall kappa	0.270
Overall p-value	< 0.001
95% confidence interval for kappa	0.222 – 0.318

TABLE 3 kappa values between the combinations of scores.

	ACP	Lee
EMAPO	0.327	0.196
Lee	0.280	(*)

(*) not applicable. EMAPO: Multicenter Study of Perioperative Evaluation; ACP: American College of Physicians.

Discussion

We were able to compare the ACP, EMAPO and Lee indexes, although they presented some differences, on account of the reclassification proposed by II Guideline for Perioperative Evaluation of the Brazilian Society of Cardiology.[3] While the goal of this standardization was to facilitate and enable the choice of any one of these scores, the results showed a low agreement among them. This finding leads us to question the applicability of the reclassification used or even the actual reliability and possible limitations in the use of these indexes to estimate perioperative risk separately. The Lee score rank the most patients as low-risk, suggesting an underestimation when compared to the others. The EMAPO, in turn, may have overestimated the risk after surgery, given that it was the index that attributed the highest risk to the population assessed.

Frequency of death was too low to try to make any comparison between the accuracy of the different scores. Assessing the frequency of perioperative myocardial infarction, arrhythmias or other cardiovascular events was not possible.

Previously published works which compared different risk scores have sought to correlate the accuracy between them but failed to find any significant differences. In Canada, a study including 2,035 patients compared four other indexes of perioperative risk: the American Society of Anesthesiologists index, the Goldman index, the Modified Detsky index and the Canadian Cardiovascular Society index. Different variables were analyzed to compare these indexes; however, there were no significant differences between them.[10]

A study from the state of Santa Catarina was published in the Brazilian Cardiology Archives, assessing 119 patients at a University Hospital using four cardiac risk indexes: the Goldman index, the Detsky index, the Larsen index and Physical Status Classification of the American Society of Anesthesiologists (ASA). It concluded that none of the indexes analyzed were significantly superior to the others and that they did not show better accuracy than that which would be obtained by chance, according to the results found: areas under the ROC curve 0.48 (± 0.03) for the Goldman, ASA and Larsen scale and 0.38 (± 0.03) for the Detsky scale.[11]

Another prospective study compared the EMAPO to the ACP method to determine heart risk in non-cardiac surgeries and to find new variables involved in determining this risk. The results revealed that there was no difference between the two methods, and that the EMAPO was as effective as the American College of Physicians for determining the risk of cardiovascular complications.[8]

The bad performances and low accuracies found in the studies in the literature confirm that the hypothesis of low agreement found in this work, evidenced by divergence in estimating the risks of patients, has been caused by the low capacity of the scores to correctly predict the likelihood of cardiac event or perioperative death in non-cardiac surgeries. Furthermore, the disagreement found in this study may simply have occurred because these 3 scores were not initially proposed to estimate the risk of the same set of events.

This study has some limitations. The main limitation is the fact that, by assessing a vast majority of low-risk patients, the low rates of events did not enable the accuracy of each method to be assessed. Another limitation is the fact that only minor surgeries were predominant. The low frequency of outcomes did not enable an analysis of the accuracy of each score, but merely found that the three scores indicated by the II Guideline do not display good agreement.

Conclusion

In conclusion, the ACP, EMAPO and Lee scores presented significantly different agreements, showing that the choice of score to be used may lead to differences in the estimation of a patient's risk. This finding suggests that these scores should not be regrouped into 3 risk groups of similar outcomes, and that this unification must be reassessed in the coming guidelines.

Resumo

Concordância entre três escores de risco perioperatório

Objetivo: avaliar a concordância entre os três escores propostos pela II Diretriz de Avaliação Perioperatória da Sociedade Brasileira de Cardiologia (SBC): algoritmo do American College of Physicians (ACP), Estudo Multicêntrico de Avaliação Perioperatória (Emapo) e Índice de Risco Cardíaco Revisado de Lee (IRCR).
Método: pacientes avaliados no pré-operatório para cirurgia não cardíaca em serviço de anestesiologia foram classificados em baixo, moderado ou alto risco pelas três escalas sugeridas pela II Diretriz. Para avaliar o grau de concordância entre as classificações, calculou-se o índice de concordância kappa.
Resultados: quatrocentos e um pacientes foram incluídos. O índice kappa de Cohen de concordância entre os três escores foi de 0,270 (IC: 0,222-0,318), correspondendo a uma concordância fraca. Analisando aos pares, a melhor correlação foi entre Emapo e ACP, com kappa de 0,327. O escore de Lee foi o que classificou mais pacientes como baixo risco: 98,3%, ao passo que Emapo e ACP classificaram como baixo risco 91,3 e 92,5%, respectivamente.

Conclusão: há uma baixa concordância entre os escores de risco propostos pela II Diretriz de Avaliação Perioperatória da SBC.

Palavras-chave: assistência perioperatória, doenças cardiovasculares, complicações pós-operatórias.

References

1. Caramelli B, Pinho C, Calderaro D, Gualandro DM, Yu PC. I Diretriz de Avaliação Perioperatória da Sociedade Brasileira de Cardiologia. Arq Bras Cardiol. 2007; 88(5):e139-e178.
2. Gualandro DM, Yu PC, Calderaro D, Caramelli B. II Diretriz de Avaliação Perioperatória da Sociedade Brasileira de Cardiologia. Arq Bras Cardiol. 2011; 96(3 supl.1):1-68.
3. Lee TH, Marcantonio ER, Mangione CM, Thomas EJ, Polanczyk CA, Cook EF, et al. Derivation and prospective validation of a simple index for prediction of cardiac risk of major noncardiac surgery. Circulation. 1999; 100(10):1043-9.
4. Goldman L, Caldera DL, Nussbaun SR, Southwick FS, Krogstad D, Murray B, et al. Multifactorial index of cardiac risk in noncardiac surgical procedures. N Engl J Med. 1977; 297(16):845-50.
5. Palda AV, Detsky AS. Guidelines for assessing and managing the perioperative risk from coronary artery disease associated with major noncardiac surgery. American College of Physicians. Ann Intern Med. 1997; 127(4):309-12.
6. Detsky AS, Abrams HB, McLaughlin JR, Drucker DJ, Sasson Z, Johnston N, et al. Predicting cardiac complications in patients undergoing non-cardiac surgery. J Gen Intern Med. 1986; 1(4):211-9.
7. Machado FS. Determinantes clínicos das complicações cardíacas pós-operatórias e de mortalidade geral em até 30 dias após cirurgia não cardíaca. [Tese de Doutorado]. Faculdade de Medicina da Universidade de São Paulo. USP/FM/SBD-054/2001.
8. Pinho C, Grandini PC, Gualandro DM, Calderaro D, Monachini M, Caramelli B. Multicenter study of perioperative evaluation for noncardiac surgeries in Brazil (EMAPO). Clinics (Sao Paulo). 2007; 62(1):17-22.
9. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33(1):159-74.
10. Gilbert K, Larocque BJ, Patrick LT. Prospective evaluation of cardiac risk indices for patients undergoing noncardiac surgery. Ann Intern Med. 2000; 133(5):356-9.
11. Heinisch RH, Barbieri CF, Nunes Filho JR, Oliveira GL, Heinisch LMM. Prospective assessment of different indices of cardiac risk for patients undergoing noncardiac surgeries. Arq Bras Cardiol. 2002; 79(4):327-38.