Summary
Introduction. The aim of the study was to investigate the prognostic value, sensitivity and specificity of both the logistic and additive European System for Cardiac Operative Risk Evaluation (as well as the European System for Cardiac Operative Risk Evaluation II) and to assess the necessity for developing a local outcome prediction model in cardiac surgery. Material and Methods. The research included 406 consecutive patients who had undergone cardiac surgical procedures at Institute of Cardiovascular Diseases of Vojvodina from January 2012 to July 2012. The authors compared the predicted mortality according to the additive and logistic European Systems for Cardiac Operative Risk Evaluation, the new European System for Cardiac Operative Risk Evaluation II and the observed mortality (30 days after surgery). Results. The difference between the predicted and observed mortality regarding the whole group of 406 operated cardiac patients was not statistically significant for the additive European System for Cardiac Operative Risk Evaluation (p=0.081) and the European System for Cardiac Operative Risk Evaluation II (p=0.164), but it was statistically significant for the logistic European System for Cardiac Operative Risk Evaluation (p=0.031). The areas under the receiver operating characteristic curves are statistically different from 0.5 for both models (additive and logistic European System for Cardiac Operative Risk Evaluation), as well as for the European System for Cardiac Operative Risk Evaluation II. However, the proper classification of the patients has not been observed since their sensitivity and specificity are not satisfactory. Conclusion. The additive and logistic European Systems for Cardiac Operative Risk Evaluation overestimate while the European System for Cardiac Operative Risk Evaluation II underestimates the risk in cardiac surgery. We believe that a locally derived model would be of great use in the everyday clinical practice since it would faithfully illustrate the actual state of patient population of the region where it was developed. At the same time it would provide the accurate prediction of surgical outcome.

Key words: Risk Assessment; Cardiac Surgical Procedures + adverse effects; Cardiac Surgical Procedures + mortality; Predictive Value of Tests

Sažetak

Corresponding author: Prof. dr Bogoljub Mihajlović, Institut za kardiovaskularne bolesti Vojvodine, 21204 Sremska Kamenica, Put doktora Goldmana 4, E-mail: bmihajlovic@ikvbv.ns.ac.rs
Introduction

The need to know the outcome of certain important medical intervention, such as cardiac surgery, has its roots in the human understanding and fear of death [1]. The outcome of a disease or surgery, in terms of survival, is obviously of great importance not only for the patient and his family but for his doctor as well. Mortality is only one of the determinants of the success of an intervention. Other indicators include: complications (morbidity), functional outcome (how fast and to what extent the patient has returned to his every-day activities), long-term survival, length of period before re-intervention, etc [2]. Since the patient population can significantly differ between institutions and geographical areas, neither the comparison of absolute numbers nor the results among institutions seem to suit its purpose [3]. Therefore, various risk stratification models have been developed trying to adjust the differences between the observed groups thus enabling the "real" comparison as well as the prediction of the surgical outcome. These models are means of determining the surgical outcome in relation to the preoperative patient condition [4]. Relatively speaking, a coefficient is assigned to the specific risk factor according to its influence on the outcome in order to provide a more accurate evaluation. Finally, the values of different risk factors are added to calculate the actual risk in terms of outcome (mortality, morbidity, price, etc.) for each patient. According to this value, the patients are classified into groups of low, mean and high risk level. In this way, a more objective comparison of surgical results is made possible and this approach is called risk stratification. Up to date, numerous risk stratification models have been developed – the most commonly used being the European System for Cardiac Operative Risk Evaluation (EuroSCORE), Society of Thoracic Surgeons (STS), Parsonnet, and Cleveland Clinic Score. Each of these models inspects closely a different number of factors, some of which overlap. The very fact that there are numerous models shows that none of them gives an absolutely precise prediction in terms of mortality.

The risk evaluation models in cardiac surgery are more developed than in any other medical field. The experience derived from their extensive use during the last two decades has led to their wide international acceptance and routine use in the outcome prediction. Everyone involved in the health system benefits from using the risk stratification models [5,6]. The new EuroSCORE II model was introduced last year [7].

Health authorities receive data about the number and severity of surgical procedures and they can plan their resources accordingly. The hospital management gains the tool enabling them to follow the success of an institution, the success of individuals and the possibility to evaluate each surgical procedure according to its risk. The doctors are given the opportunity to compare their results and to individualize their approach to each patient according to the severity of the disease. Finally and perhaps most importantly, the patients and their families are given objective information about the severity of disease and the risk which that specific surgical intervention carries.

The aim of the study was to investigate the prognostic value, sensitivity and specificity of both additive and logistic EuroSCORE as well as EuroSCORE II and to assess the necessity for developing a local outcome prediction model in cardiac surgery given all the specifics of the local population as well as customized healthcare system.

Material and Methods

Out of 406 consecutive patients from the study sample, 266 (65.5%), 78 (19.2%), and 62 (15.3%) patients had undergone coronary, valvular and combined surgery, respectively at the Institute of Cardiovascular Diseases of Vojvodina from January 2012 to July 2012. The authors analyzed the predicted mortality according to the EuroSCORE (additive and logistic), the EuroSCORE II and the observed mortality. Since the 30th postoperative day is of importance for the evaluation of results (operative risk evaluation) after cardiac surgery, all patients were contacted by phone in order to evaluate their status. Data were collected prospectively and analyzed retrospectively.

Statistical Analysis

All results for continuous variables are expressed as median (the 25-th percentile–75-th percentile). The need to know the outcome of certain important medical intervention, such as cardiac surgery, has its roots in the human understanding and fear of death [1]. The outcome of a disease or surgery, in terms of survival, is obviously of great importance not only for the patient and his family but for his doctor as well. Mortality is only one of the determinants of the success of an intervention. Other indicators include: complications (morbidity), functional outcome (how fast and to what extent the patient has returned to his every-day activities), long-term survival, length of period before re-intervention, etc [2]. Since the patient population can significantly differ between institutions and geographical areas, neither the comparison of absolute numbers nor the results among institutions seem to suit its purpose [3]. Therefore, various risk stratification models have been developed trying to adjust the differences between the observed groups thus enabling the "real" comparison as well as the prediction of the surgical outcome. These models are means of determining the surgical outcome in relation to the preoperative patient condition [4]. Relatively speaking, a coefficient is assigned to the specific risk factor according to its influence on the outcome in order to provide a more accurate evaluation. Finally, the values of different risk factors are added to calculate the actual risk in terms of outcome (mortality, morbidity, price, etc.) for each patient. According to this value, the patients are classified into groups of low, mean and high risk level. In this way, a more objective comparison of surgical results is made possible and this approach is called risk stratification. Up to date, numerous risk stratification models have been developed – the most commonly used being the European System for Cardiac Operative Risk Evaluation (EuroSCORE), Society of Thoracic Surgeons (STS), Parsonnet, and Cleveland Clinic Score. Each of these models inspects closely a different number of factors, some of which overlap. The very fact that there are numerous models shows that none of them gives an absolutely precise prediction in terms of mortality.

The risk evaluation models in cardiac surgery are more developed than in any other medical field. The experience derived from their extensive use during the last two decades has led to their wide international acceptance and routine use in the outcome prediction. Everyone involved in the health system benefits from using the risk stratification models [5,6]. The new EuroSCORE II model was introduced last year [7].

Health authorities receive data about the number and severity of surgical procedures and they can plan their resources accordingly. The hospital management gains the tool enabling them to follow the success of an institution, the success of individuals and the possibility to evaluate each surgical procedure according to its risk. The doctors are given the opportunity to compare their results and to individualize their approach to each patient according to the severity of the disease. Finally and perhaps most importantly, the patients and their families are given objective information about the severity of disease and the risk which that specific surgical intervention carries.

The aim of the study was to investigate the prognostic value, sensitivity and specificity of both additive and logistic EuroSCORE as well as EuroSCORE II and to assess the necessity for developing a local outcome prediction model in cardiac surgery given all the specifics of the local population as well as customized healthcare system.

Material and Methods

Out of 406 consecutive patients from the study sample, 266 (65.5%), 78 (19.2%), and 62 (15.3%) patients had undergone coronary, valvular and combined surgery, respectively at the Institute of Cardiovascular Diseases of Vojvodina from January 2012 to July 2012. The authors analyzed the predicted mortality according to the EuroSCORE (additive and logistic), the EuroSCORE II and the observed mortality. Since the 30th postoperative day is of importance for the evaluation of results (operative risk evaluation) after cardiac surgery, all patients were contacted by phone in order to evaluate their status. Data were collected prospectively and analyzed retrospectively.

Statistical Analysis

All results for continuous variables are expressed as median (the 25-th percentile–75-th percentile).
The comparisons between the groups were analysed by Mann-Whitney test. The percentages of 30-days mortality were compared by Chi-square test. The receiver operating characteristic (ROC) curve was generated and the area under the curve (AUC) was calculated. This method was used to investigate the prognostic value of additive, logistic EuroSCORE and EuroSCORE II. The sensitivity and specificity for optimal cut-off values were calculated. The differences were considered significant at p<0.05. The statistical analysis was performed using SPSS Version 18.

Results

The difference between the predicted and observed mortality regarding the whole group of 406 operated cardiac patients was not statistically significant for the additive EuroSCORE (p=0.081), and the EuroSCORE II (p=0.164), but it was statistically significant for the logistic EuroSCORE (Table 1).

The additive and logistic EuroSCORE overestimate, while the EuroSCORE II underestimates the risk in the whole group of 406 operated patients (Graph 1).

In coronary surgery, the difference between the predicted and observed mortality according to the additive and logistic EuroSCORE was significant, while it was not significant concerning the EuroSCORE II (Table 2). The additive and logistic EuroSCORE overestimate the risk in coronary surgery.

In valvular surgery, the difference between the predicted and observed mortality according to the additive EuroSCORE, logistic EuroSCORE and the EuroSCORE II was not significant (p=0.979; p=0.927 and p=0.114, respectively – Table 3).

In combined, coronary and valvular surgery the difference between the predicted and observed mortality according to the additive EuroSCORE, logistic EuroSCORE and EuroSCORE II was not significant (p=0.661; p=0.466 and p=0.221, respectively – Table 4).

The areas under the receiver operating characteristics curves are statistically different from 0.5 for both models (the additive and logistic EuroSCORE), as well as for the EuroSCORE II. However, no reliable classification of the patients was

Table 2. Coronary surgery
*Tabela 2. Koronarna hirurgija*

<table>
<thead>
<tr>
<th></th>
<th>Predicted mortality</th>
<th>Observed mortality</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additive EuroSCORE/Aditivni EvroSKOR</td>
<td>3.73%</td>
<td>1.5%</td>
<td>0.055</td>
</tr>
<tr>
<td>Logistic EuroSCORE/Logistički EvroSKOR</td>
<td>4.66%</td>
<td>1.5%</td>
<td>0.029</td>
</tr>
<tr>
<td>EuroSCORE II/EvroSKOR II</td>
<td>1.57%</td>
<td>1.5%</td>
<td>0.828</td>
</tr>
</tbody>
</table>

Graph 1. Scores
*Grafikon 1. Skorovi*

Table 3. Valvular surgery
*Tabela 3. Valvularna hirurgija*

<table>
<thead>
<tr>
<th></th>
<th>Predicted mortality</th>
<th>Observed mortality</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additive EuroSCORE/Aditivni EvroSKOR</td>
<td>3.79%</td>
<td>3.8%</td>
<td>0.979</td>
</tr>
<tr>
<td>Logistic EuroSCORE/Logistički EvroSKOR</td>
<td>4.05%</td>
<td>3.8%</td>
<td>0.927</td>
</tr>
<tr>
<td>EuroSCORE II/EvroSKOR II</td>
<td>1.6%</td>
<td>3.8%</td>
<td>0.114</td>
</tr>
</tbody>
</table>

Table 4. Combined surgery
*Tabela 4. Kombinovana hirurgija*

<table>
<thead>
<tr>
<th></th>
<th>Predicted mortality</th>
<th>Observed mortality</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additive EuroSCORE/Aditivni EvroSKOR</td>
<td>6.18%</td>
<td>4.8%</td>
<td>0.661</td>
</tr>
<tr>
<td>Logistic EuroSCORE/Logistički EvroSKOR</td>
<td>7.25%</td>
<td>4.8%</td>
<td>0.466</td>
</tr>
<tr>
<td>EuroSCORE II/EvroSKOR II</td>
<td>2.44%</td>
<td>4.8%</td>
<td>0.221</td>
</tr>
</tbody>
</table>
observed since the sensitivity and specificity are not satisfactory (Table 5).

Discussion

Choosing the most reliable model among many other models raises a question about how good the model really is in terms of effectiveness in relation to other models. Numerous factors can influence the model’s predictive power: differences in risk factor definitions, the management of incomplete data, surgical procedure selection criteria, geographical differences etc. The prevalence of some risk factors can also change over time.

In our previous studies [8-11], we analyzed the predictive value of the EuroSCORE model in coronary surgery, as well as trends of risk factors included in the EuroSCORE model. It was observed that the profile of coronary patients undergoing surgery in one of the cardiac surgery centres is drastically changing primarily due to the significantly advanced percutaneous techniques for myocardial revascularization.

Nilsson et al. compared the characteristics of 19 different risk stratification models in cardiac surgery [12]. They followed both 30-day and 1-year outcome. This study involved 6222 patients who had undergone cardiac surgery in a single Swedish hospital from 1996 until 2001. The ROC curve analysis was used to test the performance and accuracy of different models. The EuroSCORE model was given the preference over other models because it was notably more accurate and reliable, included the acceptable number of involved variables and was widely spread all over the world.

The question of the optimal number of risk factors included in the outcome prediction model was raised. The model must be concise and able to give more accurate prediction with the least possible number of variables included in the outcome prediction model was confirmed for both additive and logistic EuroSCORE model. These values put the EuroSCORE model among the models with a great discriminative power. However, the poor calibration of the discriminative power had values from 0.74 to 0.87 on various samples while its highest value was on the Finnish population [20]. The results of this study showed that the AUC value for the additive EuroSCORE model was 0.813, while it was 0.815 for the logistic EuroSCORE model. These values put the EuroSCORE model among the models with a great discriminating power. However, the poor calibration was confirmed for both additive and logistic models in the last three independent series of patients [21-23]. This could be explained by the fact that in time some advances have been made in surgical techniques, anaesthesiology approach, perioperative medication therapy, adequate patient selection, and perioperative patient care. This technological advance raised the question of the prediction power of the EuroSCORE model [24]. However, proofs from several European national registers for operated patients show that, in the same cases, mortality was reduced to half in spite of the fact that the patient profile has significantly changed when risk factors are concerned [25]. Patients undergoing surgery today have more risk factors on average, which results in higher values of the EuroSCORE [26].

<table>
<thead>
<tr>
<th>Table 5. Sensitivity and specificity</th>
<th>AUROC</th>
<th>p value</th>
<th>Cut-off value</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additive EuroSCORE/Aditivni EvroSKOR</td>
<td>0.700</td>
<td>0.031</td>
<td>4.4</td>
<td>70%</td>
<td>60.5%</td>
</tr>
<tr>
<td>Logistic EuroSCORE/Logistički EvroSKOR</td>
<td>0.731</td>
<td>0.013</td>
<td>4.6</td>
<td>70%</td>
<td>71%</td>
</tr>
<tr>
<td>EuroSCORE II/EvroSKOR II</td>
<td>0.682</td>
<td>0.020</td>
<td>1.46</td>
<td>71.4%</td>
<td>63.7%</td>
</tr>
</tbody>
</table>

AUROC - Area Under the receiver operating characteristic curve
It is possible to correct the EuroSCORE model in relation to the success of the specific hospital using the following formula: expected mortality = (the value of the logistic EuroSCORE model x Average hospital mortality)/average value of the logistic EuroSCORE model [27].

We advocate the development of self-made model for a number of reasons. A self-made model can usually handle input data (specific patient profile, constraints and advantages of healthcare environment) more reliably yielding better risk estimation. A self-made model depicts the "real" status of unique healthcare process.

This study is not intended to deny the validity of the existing EuroSCORE. Clinical benchmarking and comparison of the results with other hospitals around the world is extremely important and only possible through standardized models such as the EuroSCORE. However, certain risk factors, not included in the EuroSCORE, have significant impact on the postoperative outcome.

The results of our study show that the EuroSCORE models (additive and logistic) as well as the EuroSCORE II have good prognostic value, but low sensitivity and specificity. This was the reason why we decided to design our local model for cardiac operative risk evaluation „VOJVODINASKOR“, based on four-year Project, supported by the Provincial Secretariat for Science and Technological Development - Vojvodina.

Limitation of the study: The EuroSCORE II model was created last year and it has been in clinical use since the beginning of 2012 [7]. After six months of experience we compared the additive EuroSCORE, logistic EuroSCORE and the EuroSCORE II, although this group of 406 operated patients is relatively small for statistical computation.

Conclusion

The additive and logistic European System for Cardiac Operative Risk Evaluation overestimate, while the European System for Cardiac Operative Risk Evaluation II underestimates the risk in cardiac surgery. We believe that a locally derived model would be of great use in everyday clinical practice since it faithfully illustrates the actual state of patient population of the region where it was developed. At the same time it would provide an accurate prediction of surgical outcome.

References