

Does EuroSCORE work in individual European countries?[☆]

F. Roques*, S.A.M. Nashef, P. Michel, P. Pinna Pintor, M. David, E. Baudet,
The EuroSCORE Study Group

Service de Chirurgie Cardio-vasculaire, CHU de Fort de France, 97200 Fort de France, Martinique, France

Received 8 September 1999; received in revised form 21 February 2000; accepted 29 February 2000

Abstract

Objective: Because of national epidemiological differences in adult heart surgery in Europe, the effectiveness and desirability of a pan-European score for the assessment of quality of surgical care remains controversial. We assessed the predictive value of EuroSCORE in national subsets of the EuroSCORE database. **Methods:** The EuroSCORE development data set was divided into national subsets of which those with 500 or more patients were selected for analysis. The Hosmer–Lemeshow goodness-of-fit test was applied to assess the calibration of the EuroSCORE model on individual national samples and the areas under the receiver operating characteristic (ROC) curve were measured to analyse the EuroSCORE discriminative power on individual death prediction. **Results:** There were 18 676 patients in the six largest national samples: Germany, United Kingdom, Spain, Finland, France and Italy (mean: 3113 patients; range: Finland 1266 to France 4507). Major differences were observed in national distribution of procedures: coronary artery bypass grafting accounted for 77.7% of procedures in Finland but only 46.2% in Spain. The EuroSCORE model goodness-of-fit was satisfactory in all countries (*P*-value overall: 0.4; UK: 0.34; Finland: 0.87; no values less than 0.05). Areas under ROC curves were 0.81 in Germany, 0.79 in the UK, 0.74 in Spain, 0.87 in Finland, 0.82 in France and 0.82 in Italy. **Conclusion:** Despite epidemiological differences between European countries, the discriminative power of EuroSCORE was good in Spain and excellent in all other countries. The system, developed from a merged European database, can therefore be used to assess improvement in quality of care achieved by surgeons and institutions as well as for international European comparison in adult heart surgery. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Hospital mortality; Epidemiology; Open heart surgery; Europe; Risk assessment

1. Introduction

As a result of the need for cost containment in health care, providers of care are increasingly expected to produce reliable information on the cost-effectiveness of their procedures. These trends will affect cardiac surgery throughout Europe. It is likely that, before long, only those institutions which regularly monitor their results by case mix will be accredited by regional and national health agencies. Cardiac surgeons, institutions and specialist professional groups will be required to develop quality-of-care analysis according to available and validated methods. Locally developed but mutually incompatible data collection systems would render national or Europe-wide comparisons of epidemiological and statistical information difficult. As EuroSCORE [1,2] is derived from a cross-section of current European cardiac surgery, it might be proposed as a standard Europe-wide

method allowing institutional quality control as well as pan-European epidemiological analysis. However, legitimate questions can be raised about the effectiveness and accuracy of a pan-European model in every individual European country. We have therefore compared general epidemiological information and the predictive value of the EuroSCORE in national subsets of the EuroSCORE database.

2. Methods

The general protocol of the study (project set-up, data collection and entry) has previously been described in detail [1] and is summarized in Appendix A. The database has already been used to identify individual operative risk factors in the European Cardiac Surgery [1] and to develop the European System for Cardiac Operative Risk Evaluation [2]. For this study, the Overall EuroSCORE database (including the development and the validation subset of the previously described EuroSCORE trial) was divided into national subsets of which those with 500 or more

[☆] Presented at the 13th Annual Meeting of the European Association for Cardio-thoracic Surgery, Glasgow, Scotland, UK, September 5–8, 1999.

* Corresponding author. Tel.: +596 552271; fax: +596-758438.

E-mail address: f.r.fwi@wanadoo.fr (F. Roques).

Table 1
Distribution of patients among the six largest EuroSCORE national cohorts

Country	No. of patients
Germany	4277
France	4507
United Kingdom	3567
Italy	2637
Spain	2422
Finland	1266
Total	18676

patients were selected for analysis. National demographic and epidemiological characteristics were compared using usual bivariate methods (chi-square for categorical covariates and Kruskal–Wallis test for continuous covariates). To analyse the risk profile of the national cohorts, National mean predicted mortality (mean EuroSCORE values) were calculated. They were compared with observed mortality in the national subsets. To assess the calibration of the EuroSCORE model on individual national samples, the Hosmer–Lemeshow chi-square test was applied (this test compared the number of observed and predicted death in deciles of risk in the entire range of probability of death in every national subset). To analyse the discriminative power of the EuroSCORE on individual death prediction, areas under the receiver operating characteristic (ROC) curves were measured on national samples. To refine the analysis and identify potential differences in the performance of the model related to variation in national distribution of valvular and coronary surgery, an identical analysis was carried out on national cohorts of patients who had isolated coronary artery bypass grafting (CABG).

3. Results

Among the 19 030 patients in the database, 18 676 were from the six largest national cohorts (Germany, France, United Kingdom, Italy, Spain, Finland), ranging from 4507 patients in France to 1266 in Finland (Table 1). As a result of the methodology previously described [1], infor-

mation for all the mandatory fields of the EuroSCORE was available in 14 724 patients. Epidemiological comparison of national samples is detailed in Table 2. There were significant differences in a number of areas. Mean age ranged from 61.4 years in Italy to 64.5 years in Germany. Female gender accounted for only 25.8% of patients in Finland but 32.7% in Spain. Valve surgery still accounts for between a third and a half of the procedures in the countries in southern Europe (35.5, 41.7 and 51.5% in Italy, France and Spain, respectively) but less than a quarter in the north (21.7% in Germany, 24.4% in the UK, 18.6% in Finland). The risk profile of patients was also different: mean EuroSCORE was relatively low in Finland (3.4), and higher in France (4.6) and Spain (4.7). These differences in case mix were reflected in the differences in observed operative mortality ranging from 3.2% in Finland to 8.3% in Spain. Calibration was good in all countries: the EuroSCORE logistic regression model fitted to each national sample and gave the following *P*-values in the Hosmer–Lemeshow chi-square test: Germany 0.72, United Kingdom 0.34, Spain 0.38, France 0.88, Finland 0.84 and Italy 0.43 (*P* never less than 0.05).

Table 3 details the area under the ROC curve observed in the national samples. In the overall population, the discriminative value of the EuroSCORE was excellent in Germany, Finland, Italy and France (area under ROC curves > 0.8), very good in the UK (0.79) and good in Spain (0.74). In the CABG subgroups the predictive value of the EuroSCORE was very good in the UK and Spain and excellent in all other countries.

4. Discussion

Within the European Union, there are major national differences in the quality of the information available on cardiac surgical practice. Some countries have already implemented central national databases which collect annual information on operations and case mix. Others only collect procedural registers and some have no central data collection. A European register of procedures was set up and widely published in 1996 [3] but this information is

Table 2
Comparisons of demographic and epidemiological data among the six largest EuroSCORE national cohorts

Variable	Germany	United Kingdom	Spain	Finland	France	Italy	<i>P</i> -value
Age (years)	63.5	61.9	61.2	61.9	63.4	61.4	0.0001
Female gender (%)	26.7	26.9	32.7	25.8	29.1	28.6	0.0001
Impaired LV function (LVEF < 0.3) (%) ^a	3.4	11.7	4.6	5.4	4.8	4.8	0.0001
Isolated coronary surgery (%)	73.4	71.7	46.2	77.7	53.6	57.6	0.0001
Mitral surgery (%)	5.7	9	23	6	14.9	16.1	0.0001
Aortic surgery (%)	16	15.4	28.5	12.6	26.8	19.4	0.0001
EuroSCORE (mean ± SD)	3.7 ± 2.6	4.1 ± 3	4.7 ± 2.9	3.4 ± 2.6	4.6 ± 2.9	4.3 ± 2.9	0.0001
Observed mortality (%)	3.6	5.2	8.3	3.2	5.1	3.5	0.0001

^a LV, Left ventricle; EF, ejection fraction.

Table 3
 Predictive ability^a of the EuroSCORE model for mortality prediction in individual European national cohorts

Country	CABG alone		All cardiac surgery	
	No. of patients	Area under ROC curve	No. of patients	Area under ROC curve
Germany	2447	0.83	3257	0.81
France	2200	0.85	3723	0.82
Italy	1324	0.88	2223	0.82
United Kingdom	1799	0.76	2387	0.79
Finland	857	0.85	1109	0.87
Spain	1006	0.78	2025	0.74
Total	9633		14724	

^a Good predictive ability: $0.7 < \text{area under ROC curves} < 0.75$; very good predictive ability: $0.75 < \text{area under ROC curves} < 0.8$; excellent predictive ability: $\text{area under ROC curves} > 0.8$.

not sufficient to demonstrate regional differences in clinical characteristics and case mix in Europe.

The EuroSCORE study was the first large Europe-wide attempt to collect information on the risk profile of patients, procedures and outcome in adult heart surgery. Major differences have been observed in the risk profile of national samples. It can be argued that these differences can be linked to the different prevalence of valve surgery within countries (Table 2), but there may be other factors such as, for example, other epidemiological variation in co-morbidity and variable surgical decision-making on high-risk patients. Further explanations on epidemiological differences in the risk profile of CABG patients included in the EuroSCORE database is analysed elsewhere [4].

Even if controversy still exists about methods of quality-of-care assessment in adult heart surgery, it is widely accepted that monitoring of risk-adjusted mortality is one of the simplest methods of such assessment. Indeed, it can be argued that there is little point in more sophisticated quality measurement tools until basic, robust data about risk-adjusted mortality is in place. Whatever the method of choice (simple additive score, multivariate logistic regression model or Bayesian analysis), a system with proven predictive ability is essential for pertinent conclusions about the quality of care provided. It can therefore be argued that a pan-European method is not sufficiently sensitive to regional demographics and epidemiology to yield valuable quality-of-care assessment. One might then defend the concept of local quality control models, as already proposed [5]. On the other hand, sporadic, unintegrated regional quality monitoring would hardly allow the pan-European analysis essential to the improvement of European cardiac surgery at least from a public health point of view. The EuroSCORE study group took into consideration the major epidemiological differences within European cardiac surgery and yet deliberately proposed a global score which included risk factors for coronary surgery, valve surgery and other forms of adult open cardiac surgery. In particular, the model gives sufficient weight for valve surgery to work well in countries in the south of Europe. As a consequence, the predictive value of the

system, as assessed by the area under ROC curves, is not higher in the north of Europe than in the south (Table 3) nor does it perform better in the CABG subsets than in the overall population. As the size of the national samples in the EuroSCORE database were unequal, differences in national predictive value of the system could have been expected but where not observed. Even if the large size of national samples (from 1109 to 3723 patients) supports the argument that the model is strongly representative on a national basis, we fully accept that the voluntary enrolment in the study may have introduced a bias. The predictive power of the score observed in the six major participating countries should therefore be confirmed by new independent national trials. The model is currently under investigation in the UK: its performance, annually assessed in the national Adult Cardiac Surgical Database report, should soon be published. We propose similar trials in other European countries, whether or not they had participated in the original project. Important differences in observed mortality among national samples of the EuroSCORE database are presented in Table 2. These are matched by concordant variation in risk profile as assessed by EuroSCORE. This might be considered as the first attempt to compare the quality of care provided in European countries, using national risk-adjusted mortality as the basis for such comparison. We believe that, in addition to the methodological cautions previously described [1], reliable conclusions about quality of care provided in individual countries would require total or random data collection with on-site validation and cannot therefore be discussed in the present work. Nevertheless, one can imagine the quality of information that could be delivered by a European institute devoted to the analysis and validation of national cardiac surgical outcome data with the objective of assessing the quality of care in Europe.

5. Conclusion

Major epidemiological differences have been observed in national samples of the EuroSCORE trial. Nevertheless, the predictive value of the EuroSCORE model assessed by

ROC curves on national cohorts was good to excellent in the six major countries participating in the project. We expect that future national studies will confirm the predictive value of this pan-European method of quality-of-care assessment. The evaluation and comparison of the quality of care provided in the different countries of the European Union can be achieved by the establishment of a dedicated and independent European institute.

References

- [1] Roques F, Nashef SAM, Michel P, Gauducheau E, de Vincentiis C, Baudet E, Cortina J, David M, Faichney A, Gabrielle F, Gams E, Harjula A, Jones MT, Pinna Pintor P, Salamon R, Thulin L. Risk factors and outcome in European cardiac surgery: analysis of the EuroSCORE multinational database of 19 030 patients. *Eur J Cardio-thorac Surg* 1999;15:816–823.
- [2] Nashef SAM, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R, The EuroSCORE Study Group. European System for Cardiac Operative Risk Evaluation (EuroSCORE). *Eur J Cardio-thorac Surg* 1999;16:9–13.
- [3] Unger F. Open heart surgery in Europe 1993. *Eur J Cardio-thorac Surg* 1996;10:120–128.
- [4] Nashef SAM, Roques F, Michel P, Cortina J, Faichney A, Gams E, Harjula A, Jones MT. Coronary surgery in Europe: comparison of the national subsets of the EuroSCORE database. *Eur J Cardio-thorac Surg* 2000;17:396–399.
- [5] Pons JMV, Granados A, Espinas JA, Borrás JM, Martín I, Moreno V. Assessing open heart surgery mortality in Catalonia (Spain) through a predictive risk model. *Eur J Cardio-thorac Surg* 1997;11:415–423.

Appendix A.

Information on risk factors (68 preoperative and 29 operative) and outcome, displayed on a simple A4 data collection form, was collected in 128 voluntary participating centres from eight European countries during the period from September to December 1995. After double entry and submission to quality checks and completion control, the database was used to identify risk factors for operative mortality in Europe [1]. A logistic regression model was then applied to select 17 variables used to develop an additive scoring system: the European System for Cardiac Operative Risk Evaluation, or EuroSCORE [2]. This system

is devoted to quality of care analysis in European Cardiac surgery.

Appendix B. Conference discussion

Mr K. Dhital (London, UK): Is it at all possible for you to comment from the data you have whether the outcome is in any way related to different systems of health care provision in the various countries? And I say that with particular reference to the UK where, for example, the long waiting list may reflect in the poorer ejection fractions in that subgroup of patients. The question really is whether or not the outcome and your patient groups in any way reflect the different health care provisions in the various countries?

Dr Roques: We are not allowed to say that differences in outcome observed during our study reflect differences in health care provision among European nations because we do not know if our samples really represent the countries. To do that, on-site control during the study would have been required, and the distribution of private, public, teaching hospitals in the samples and the countries would have to be compared. The information that we can produce is that the EuroSCORE worked very well in any of the national samples of the study.

Dr M. Turina (Zurich, Switzerland): You use EuroSCORE as a fixed set of values in the rapidly changing world of cardiac surgery. I have been observing these changes in the last 5–6 years where our waiting list went from 6 months to 3 days, and there is the major change in the mortality which is inversely related to the length of the waiting list. Does your data base consider the length of the waiting list, because this has an impact on mortality. Are you implementing the measures to change some of your parameters?

Dr Roques: To the first question, the system is not designed to answer to all the questions regarding quality of care assessment. Its purpose is to be simple, objective and related to the patient. Therefore we did not analyze the appropriateness of the choice of the procedure for example, nor did we analyze the consequences of the waiting list on the results. Our feeling is that ‘Waiting list’ is not a sufficiently objective criterion to be included in a risk stratification system. One can game the system with such a criterion. On the other hand ‘Waiting list’ is indirectly considered in the EuroSCORE with 2 patients related risk-factors: recent myocardial infarction and surgery for unstable angina.

To the second question, we would like to implement the measures, of course. There is a very interesting analysis of the EuroSCORE made by the British Society of Cardiothoracic Surgeons, and we invite other national societies to try the system on their databases. Connecting information could help us to implement our system and to update it in the future.

Dr Turina: But you will have to change the parameters of the EuroSCORE as time goes by, because they are not immovable values. Just look at the Parsonnet score which is becoming outdated.

Dr Roques: Yes we intend to do it. Your support and the help of the European Association will be necessary to merge and update ongoing experiences of the EuroSCORE.