

RESEARCH ARTICLE

Is the EuroSCORE II reliable to estimate operative mortality among octogenarians?

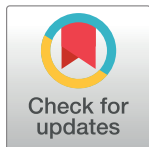
Sophie Provenchère^{1,2}*, Arnaud Chevalier¹‡, Walid Ghodbane³‡, Claire Bouleti⁴‡, Philippe Montravers^{1,5}*, Dan Longrois^{1,5,6}*, Bernard Jung^{4,5,6}

1 Département d'Anesthésie-Réanimation, APHP, Hôpital Bichat-Claude Bernard, Paris, France, **2** INSERM Centre d'Investigation Clinique 1425, APHP, Hôpital Bichat-Claude Bernard, Paris, France, **3** Département de Chirurgie Cardiaque, APHP, Hôpital Bichat-Claude Bernard, Paris, France, **4** Département de Cardiologie, APHP, Hôpital Bichat-Claude Bernard, Paris, France, **5** Université Paris 7-Diderot, Paris, France, **6** INSERM 1148, Paris, France

* These authors contributed equally to this work.

‡ These authors also contributed equally to this work.

* sophie.provenchere@aphp.fr



Abstract

Objectives

Concerns have been raised about the predictive performance (PP) of the EuroSCORE I (ES I) to estimate operative mortality (OM) of patients aged ≥ 80 . The EuroSCORE II (ES II) has been described to have better PP of OM but external validations are scarce. Furthermore, the PP of ES II has not been investigated among the octogenarians. The goal of the study was to compare the PP of ES II and ES I among the overall population and patients ≥ 80 .

Methods

The ES I and ES II were computed for 7161 consecutive patients who underwent major cardiac surgery in a 7-year period. Discrimination was assessed by using the c-index and calibration with the Hosmer-Lemeshow (HL) and calibration plot by comparing predicted and observed mortality.

Results

From the global cohort of 7161 patients, 832 (12%) were ≥ 80 . The mean values of ES I and ES II were 7.4 ± 9.4 and 5.2 ± 9.1 respectively for the whole cohort, 6.3 ± 8.6 and 4.7 ± 8.5 for the patients < 80 , 15.1 ± 11.8 and 8.5 ± 11.0 for the patients ≥ 80 . The mortality was 9.38% (≥ 80) versus 5.18% (< 80). The discriminatory power was good for the two algorithms among the whole population and the < 80 but less satisfying among the ≥ 80 (AUC 0.64 [0.58–0.71] for ES I and 0.67 [0.60–0.73] for the ES II without significant differences ($p = 0.35$) between the two scores. For the octogenarians, the ES II had a fair calibration until 10%-predicted values and over-predicted beyond.

Conclusions

The ES II has a better PP than the ES I among patients < 80 . Its discrimination and calibration are less satisfying in patients ≥ 80 , showing an overestimation in the elderly at very

OPEN ACCESS

Citation: Provenchère S, Chevalier A, Ghodbane W, Bouleti C, Montravers P, Longrois D, et al. (2017) Is the EuroSCORE II reliable to estimate operative mortality among octogenarians? PLoS ONE 12 (11): e0187056. <https://doi.org/10.1371/journal.pone.0187056>

Editor: Giuseppe Sergi, University of Padova, ITALY

Received: March 8, 2017

Accepted: October 12, 2017

Published: November 16, 2017

Copyright: © 2017 Provenchère et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

high-surgical risk. Nevertheless, it shows an acceptable calibration until 10%- predicted mortality.

Introduction

The remarkable gain in life expectancy in most developed countries is challenging for the medical community and for public health decision makers. For both men and women, life expectancy is steadily increasing and the population aged 80 years and over is expected to more than triple in 2060 [1]. Despite disparities among countries, the increased percentage of elderly individuals is largely due to a decrease in cardiovascular disease mortality in high income countries [2]. Consequently, the number of interventions performed in older patients is increasing. Moreover, in this specific population the presence of multiple comorbidities and chronic multiple drugs therapy make preoperative risk assessment crucial, though complex [1, 3].

In cardiac surgery, the number of octogenarians has increased in the past 20 years and now represents more than 10 per cent of the cardiac surgical patients in developed countries. In this specific population at higher risk of major morbidity or mortality [4], risk stratification is even more important since credible alternatives to conventional surgery such as transcatheter aortic valve implantation (TAVI) or percutaneous mitral edge-to-edge repair (MitraClip[®]) exist.

Patient selection is nowadays based on both individual clinical judgment and the Heart Team, a standard of care [5], which integrates current scoring systems, the original EuroSCORE [6] (ES I) and Society of Thoracic Surgeons Predicted Risk of Mortality (STS PROM). It has been demonstrated that the ES I [6] tended to overestimate mortality among high-risk cardiac surgery patients [7]. Consequently, a new scoring system, the EuroSCORE II (ES II), has been developed in order to achieve a better calibration [8]. Numerous studies of external validation have demonstrated that ES II achieved a good discrimination and an improved calibration except among patients with a predicted OM of more than 30 per cent [9, 10]. However, few studies have focused specifically on the predictive performance (PP) of the ES II among high risk patients and in the elderly.

The objective of this study was therefore to evaluate the discrimination and the calibration of the ES II in a contemporary cardiac surgical population and to compare its performance with the ES I among the overall population and the elderly, defined as ≥ 80 years old.

Material and methods

Study population and study design

The study population included all consecutive adult patients who underwent cardiac surgery with cardiopulmonary bypass (CPB) over a 7-year period between January 2006 and December 2012 (7161 patients enrolled) in our institution. Since 2006, patients, surgery characteristics as well as in-hospital outcomes were prospectively entered into an institutional database by trained research assistants (S1 Table).

The study was approved by Bichat hospital Institutional Review Board (“Comité d’Evaluation des Projets de Recherche Biomédicale Paris Nord”; IRB 00006477). It waived the need for patient’s informed consent because the data were analyzed anonymously. In accordance with the French Law, the registry was approved by the Advisory Committee for Information Processing in Health Research (Comité Consultatif sur le Traitement de l’Information en matière

de Recherche dans le domaine de la Santé, Paris, France) and the French National Commission on Computing and Liberty (Commission Nationale Informatique et Liberté, Paris, France). Reporting of the study complies with the Strengthening the Reporting of Observational studies in Epidemiology recommendations statement for reporting. Although data were prospectively acquired for a research purpose, the study should be considered as a retrospective study.

Analyses were performed for all types of surgical procedures including emergency. The types of surgery were defined as follows: isolated coronary artery bypass (CABG), isolated valvular surgery or valvular surgery combined with CABG and other procedures.

The only exclusion criteria were patients less than 18 years old and cardiac transplantations.

The primary endpoint was in-hospital mortality. In-hospital mortality was defined as death occurring at any time in the hospital where the surgery was performed before discharge from hospital. The 30 days mortality was also recorded.

We analyzed separately the two subgroups of patients aged <80 and ≥ 80 years.

Medical records of non survivors among the octogenarians were retrospectively assessed by two senior cardiac anesthesiologists to determine the causes of death. Causes of death were categorized in 3 groups according to timing, i.e. Group 1: Early complications (D0-D2), Group 2: Acute complications (D3-D15) and Group 3: Delayed complications ($\geq D16$). Definitions are given in Table 1.

The ES I and the ES II values were calculated for each patient according to patient characteristics using published equations [6, 8].

Statistical analysis

Results are expressed as mean \pm SD, number (percentage). Comparisons of continuous variables used unpaired Wilcoxon tests and comparisons of discrete variables Chi-2 or Fisher exact tests. The performance of the ES models was analyzed focusing on discrimination and calibration.

Table 1. Causes of death and time to death.

Causes of death
Group 1: Early complications (D0-D2)
<ul style="list-style-type: none"> • Deceased in the operating room • Severe and uncontrolled bleeding with massive transfusion
Group 2: Acute complications (D3-D15)
<ul style="list-style-type: none"> • Cardiogenic shock (e.g., multiple high dose inotropes, mechanical cardio-circulatory support) • Hemorrhagic shock • Sepsis or septic shock • Massive stroke or mesenteric ischemia • Tamponnade • Sudden death
Group 3: Delayed complications ($\geq D16$)
<ul style="list-style-type: none"> • Mediastinitis • Endocarditis • MOF • Prolonged mechanical ventilation • Withdrawal of care

<https://doi.org/10.1371/journal.pone.0187056.t001>

The discrimination performance indicates the extent to which the model distinguishes between patients who will die or survive. It was evaluated by constructing receiver operating characteristic curves for each model and calculating the area under the curve (AUC) with 95% confidence intervals (CI). Numerically, an area of 1.0 indicates the perfect discrimination power, whereas an area of 0.5 indicates no discrimination of the binary outcome.

Calibration refers to the agreement between observed and predicted in-hospital mortality. Overall model calibration was assessed by comparing observed and predicted mortality in 10 equally sized subgroups in increasing order of patient risk, according to the HL test for goodness of fit. A HL p values >0.05 indicates a well-calibrated model for the study population. Model calibration was also assessed using the observed/expected (O/E) mortality ratio. The 95% CI of O/E ratio was calculated using the Byar's method. An O/E ratio >1.0 indicates that the score underpredicts mortality and an O/E ratio <1.0 indicates that the score overpredicts mortality. The O/E ratio indicates good calibration if its 95% CI includes the value 1.0. Finally, model calibration was assessed visually according to the risk level by comparing predicted and observed mortality rates with 95% CI using calibration plots. Whereas the perfect calibrated prediction stays on the 45-degree line, a curve below or above the diagonal reflects overestimation and underestimation respectively. The P-value for statistical significance was 0.05. The statistical analyses were performed with SAS version 9.2 and SPSS version 19.0.

Results

General demographics

A total of 7161 consecutive patients were included in the study.

Table 2 summarizes patient profile with demographics, co morbidities and procedures performed as defined by EuroSCORE algorithms among the entire cohort and for both study groups.

Body mass index (BMI) was calculated using the formula:

$$\text{BMI} = \text{weight}(\text{kg}) / \text{height}^2(\text{m}^2).$$

Creatinine Clearance as an estimate of Glomerular Filtration Rate (eGFR) was calculated using the Cockcroft–Gault formula:

$$\begin{aligned} \text{eGFR}(\text{ml}/\text{min}) \\ = (140 - \text{age}(\text{years})) \times \text{weight}(\text{kg}) \times 0.85(\text{if female}) / 72 \times \text{serum creatinine}(\text{mg}/\text{dL}) \end{aligned}$$

eGFR was standardized for 1.73 m² of body surface area.

All the variables associated with mortality in the study of Nashef et al. [8] [age, sex, extracardiac arteriopathy, chronic lung disease, poor mobility, previous cardiac surgery, creatinine clearance, active endocarditis, critical preoperative state, left ventricular function, systolic pulmonary artery pressure, urgency and type of procedure] are included in Table 2.

Results are given with the item “creatinine clearance” as proposed by the on-line EuroSCORE interactive calculator.

The elderly group (age ≥ 80) embodying 832 patients had a higher prevalence of arterial hypertension, worse NYHA functional class and more frequent atrial fibrillation than those aged < 80. They also had more previous chronic conditions such as malignancy, extra-cardiac arteriopathy, chronic pulmonary disease and previous stroke. Moreover, the only biological variable needed for ES II calculation, the creatinine clearance strongly differed between the two groups emphasizing the high prevalence of chronic kidney disease among the elderly.

Table 2. Characteristics for the whole population and for both age groups.

Patients characteristics	Overall population N(%)	Patients <80 N(%)	Patients ≥ 80 N(%)	p-value
Age (year)	63±14 [18–94]	61±13 [18–79]	83±3 [80–94]	<0.0001
Male gender	4869 (68)	4447 (70)	422 (51)	<0.0001
Body mass index (Kg/m²)	26.5±4.6 [12–50.2]	26.5±4.6 [12–50]	26.1±4.25 [14.4–47.3]	0.0157
Medical and surgical history				
Smoker				
Current	1116 (15.6)	1099(17.3)	17 (2)	<0.001
Past	2534 (37.4)	2267 (35.8)	267(32)	
Unknown	30 (0.4)	24 (0.38)	6 (0,7)	
No	3481(48.6)	2939 (46.4)	542 (65)	
Arterial hypertension	4025 (56.2)	3416 (53)	609 (73)	<0.0001
Diabetes status				
Oral therapy only	1247(17.4)	1125 (18)	122 (14)	0.001
Insulin	565 (7.9)	519 (8)	46 (6)	
No	5349 (74.7)	4685 (74)	664 (80)	
Dyslipemia	3591 (50)	3145 (50)	445 (53)	0.039
History of coronary artery disease				
Stable angina	1279 (17.9)	1148 (18)	131 (16)	0.0002
STEMI	722 (10)	666 (10)	56 (7)	
Non STEMI	735 (10.3)	658 (10)	56 (7)	
No	4425 (61.8)	3857 (61)	568 (68)	
ACS <90 days	994 (13.8)	901 (14)	93 (11)	0.016
Previous cardiac surgery				
Others	53 (0.7)	49 (0,8)	4 (0,5)	<0.0001
CABG	68 (0.9)	59 (1)	9 (1)	
Combined surgery	30 (0.4)	27 (0.4)	3 (0.3)	
Valvular surgery	574 (8)	550 (9)	24 (3)	
No	6436 (89)	5644 (89)	792 (95)	
Extracardiac arteriopathy				
Lower limbs	337 (4.7)	297 (5)	40 (5)	<0.0001
Previous vascular surgery	264 (3.7)	228 (4)	36 (4)	
Carotid stenosis >50%	382 (5.3)	308 (5)	74 (9)	
No	6178 (86.3)	5496 (86)	682 (82)	
Chronic pulmonary disease				
Untreated	282 (3.9)	240 (4)	42 (5)	<0.0001
Treated	413 (5.8)	337 (5)	76 (9)	
No	6466 (90.3)	5752 (91)	714 (86)	
On dialysis	71 (1)	70 (1)	1 (0,1)	0.0025
Cancer				
<5years	209 (2.9)	169 (3)	40 (5)	<0.0001
>5 years	260 (3.6)	210 (3)	50 (6)	
Metastasis	20 (0.3)	16 (0.2)	4 (0.5)	
No Metastasis	76 (0.1)	60 (0.11)	16 (0.2)	
Atrial fibrillation	850 (11.9)	718 (11)	132 (16)	0.0002
Neurological or musculoskeletal dysfunction^a	212 (3)	193 (3)	19 (2)	0.270
Previous stroke				

(Continued)

Table 2. (Continued)

Patients characteristics	Overall population N(%)	Patients <80 N(%)	Patients ≥ 80 N(%)	p-value
TIA	199 (2.8)	167 (2)	32 (4)	0.1318
Hemorrhagic stroke	62 (0.9)	54 (1)	8 (1)	
Ischemic stroke	341 (4.8)	309 (5)	32 (4)	
No	6559 (91.6)	5799 (92)	760 (91)	
Cirrhosis				
Uncomplicated	30 (0.4)	30 (0.4)	0	0.044
PHT	45 (0.6)	43 (0.68)	2 (0.2)	
No	7086 (99)	6256 (99)	830 (99)	
Active endocarditis	321 (4.5)	299 (4.7)	22 (2.6)	0.006
NYHA class				
I	1538 (21.5)	1449 (23)	89(10)	<0.0001
II	672 (9.4)	639 (10)	33 (4)	
III	2789 (38.9)	2453 (39)	336 (40)	
IV	2162 (30.2)	1788 (28)	374 (45)	
SPP >55 mmHg	333 (4.7)	288 (5)	45 (5)	0.269
SPP (mmHg)	42± 14 [10–125]	41±14 [10–125]	44±12 [20–90]	0.269
Critical preoperative state^b	218 (3)	198 (3)	20 (2)	0.253
Type of procedure				
CABG	2648 (37)	2485 (39)	163 (20)	<0.0001
Valvular surgery	4134 (57.7)	3490 (55)	645 (78)	
Other	378 (5.3)	353 (6)	24 (2)	
Urgency^c	409 (5.7)	353 (6)	56 (7)	0.1778
Creatinine clearance (ml/min/m²SC)^d	77.2±32 [5.5–262]	81.1±31 [5.5–262]	47±16 [6–108]	<0.0001
LVEF (%)	57.8±12 [10–89]	57.9±12 [10–89]	57.6±12 [15–84]	0.5313

Abbreviations: MI: Myocardial Infarction; TIA: Transient Ischemic Attack; STEMI: ST elevation Myocardial Infarction; ACS: Acute Coronary Syndrome; CABG: Coronary Artery Bypass Grafting; PHT: Portal HyperTension; NYHA: New York Heart Association; SPP: Systolic Pulmonary Pressure; BMI: Body Mass Index; LVEF: Left Ventricular Ejection Fraction
^{a,b,c,d}: as defined by EuroSCORE algorithms

Continuous variables are presented as means ±SD and [minimum-maximum], categorial data are presented as number (percentage) of patients. P-values are from Wilcoxon tests and Chi-2 or Fisher exact tests.

<https://doi.org/10.1371/journal.pone.0187056.t002>

Finally, they differed from the younger patients by undergoing more frequently valvular surgery.

ES and mortality

Our in-hospital mortality rate was 5.7%, 5.2% and 9.4% in the overall population, in patients aged <80 and ≥ 80 respectively.

The distribution of in-hospital mortality of the elderly according to the 3 groups of timing of deaths as defined in Table 1 was 25.6% (N = 20) in group 1 (mean ES II: 17.6±22.3), 38.4% (N = 30) in group 2 (mean ES II: 11.3±13.9) and 35.8% (N = 28) in group 3 (mean ES II: 24.1 ±28.2). Twelve patients (60%) in group 1 died in the operating room and 4 (15%) in group 3 died after withdrawal of care.

Table 3 reports observed mortality and ES values of our patient cohort and according to age subgroups:

Table 3. Observed and predicted mortality of the study population.

	Overall population	<80	≥80
N =	7161	6329	832
In-hospital mortality	406 (5.67)	328 (5.18)	78 (9.38)
30 days mortality	327 (4.57)	265 (4.18)	62 (7.45)
EuroSCORE I	7.36±9.43 [0.88–88.48]	6.35±8.57 [0.88–88.48]	15.06±11.85 [3.44–85.88]
EuroSCORE II	5.17±9.07 [0.49–94.42]	4.73±8.51 [0.49–94.42]	8.51±10.98 [1.19–89.31]

Continuous variables are presented as means ±SD and [95% CI]; categorial data are presented as percentage of patients.

<https://doi.org/10.1371/journal.pone.0187056.t003>

For the whole cohort and the two subgroups (<80 and ≥80 respectively), there was a difference of nearly 1% between 30 day mortality and in-hospital mortality.

The mean values of ES I and ES II were 7.4±9.4 and 5.2±9.1 respectively for the whole cohort, 6.3±8.6 and 4.7±8.5 for the patients aged <80, 15.1±11.8 and 8.5±11.0 for the patients aged ≥80.

The ES I tended to overestimate in-hospital mortality among the whole cohort and for both study groups. There was in particular a strong overestimation of in-hospital mortality among the elderly.

Conversely, the predicted mortality of the ES II was close to the observed in-hospital mortality, even in the elderly.

Performance of EuroSCORE

Discrimination. The ROC curves for ES I and ES II are plotted in Fig 1.

The AUC was high in the overall population and among patients aged <80, being 0.79 [0.77–0.81] and 0.80 [0.78–0.83] for the ES I, respectively. When considering the ES II, the AUC was 0.80 [0.78–0.82] and 0.81 [0.79–0.84], respectively. The comparison among the two ES highlighted a trend to a better discrimination for ES II and the previous original version for the overall population (p = 0.07) but there was no significant differences for the patients aged <80 (p = 0.21).

Conversely, the discrimination was less satisfying for patients aged ≥80, being 0.64 [0.58–0.71] for ES I and 0.67 [0.60–0.73] for the ES II without significant differences (p = 0.35) between the two scores (Table 4).

Calibration. PP of the ES I and II is detailed in Table 5 for the overall population and according to age groups.

Values of O/E ratio and their 95% CI show a significant over-prediction of in-hospital mortality with the ES I in the overall population and in both age groups. This is consistent with the highly significant p values of the HL test. Calibration was better with the ES II, as shown by the values of O/E ratio and the fact that 95% CI included 1.0, although HL test was significant in the overall population and in patients aged over 80.

The calibration plots are represented for the whole population and for both age study groups in Fig 2.

The ES I shows a constant trend to over-prediction of in-hospital mortality, which is particularly marked in the high-risk group (predicted mortality >30%).

Conversely, the ES II had a better calibration than ES I. It had a tendency to under-estimate the in-hospital mortality but was otherwise close to the ideal diagonal except for the high-risk group.

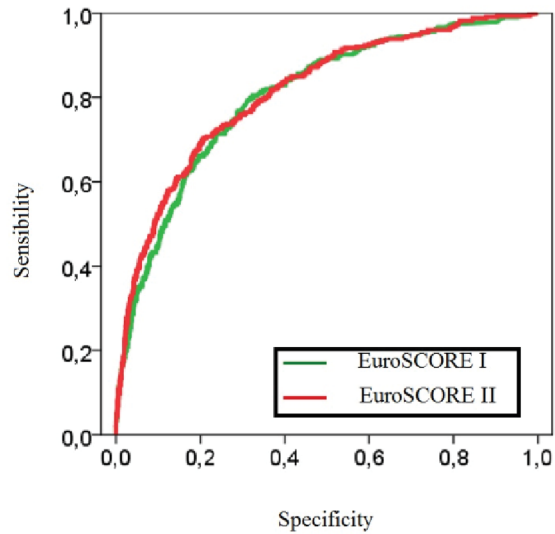


Fig 1A: ROC curve for the entire population

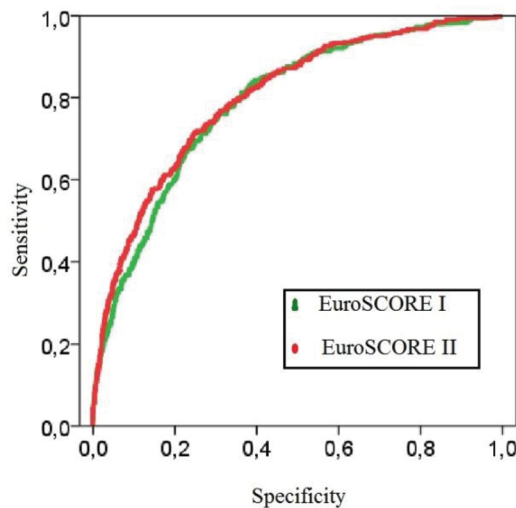


Fig 1B: ROC curve for the ≥ 80

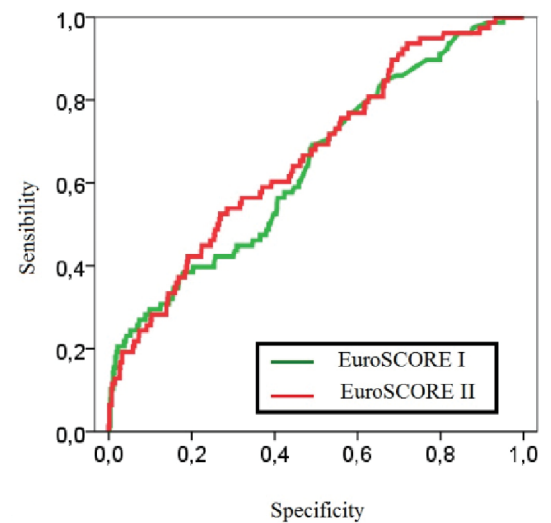


Fig 1C: ROC curve for the < 80

Fig 1. Receiver operating characteristics curves for the two models, for the entire population (Fig 1A) and for the two studied groups (Fig 1B: ≥ 80 and Fig 1C: < 80).

<https://doi.org/10.1371/journal.pone.0187056.g001>

The best predictive performance of ES II was observed in patients aged < 80 with a more satisfying calibration among all risk deciles, showing a good agreement between observed and predicted in-hospital mortality, which is consistent with O/E values and the non-significant HL test.

Table 4. Discrimination for in-hospital mortality among the overall population and according to age.

	Overall population	<80	≥80
N =	7161	6329	832
c-index ES I	0.79 [0.77–0.81]	0.80 [0.78–0.83]	0.64 [0.58–0.71]
c-index ES II	0.80 [0.78–0.82]	0.81 [0.79–0.84]	0.67 [0.60–0.73]
p ES I/ES II	0.07	0.21	0.35

Continuous variables are presented as means ±SD and [95% CI]; P-values are from Wilcoxon tests and Chi-2 or Fisher exact tests

<https://doi.org/10.1371/journal.pone.0187056.t004>

In patients aged over 80, calibration was better with ES II than ES I as assessed by O/E ratios and 95% CI, although HL test was significant with both ES I and ES II. There was better agreement between observed and predicted mortality, as illustrated by CI of calibration plots (Fig 2) when considering patients with a predicted operative mortality of < 10% that represented 653 patients (78.4%) of those ≥ 80.

Discussion

Cardiac surgery care providers are confronted with the constant increase of octogenarians in their daily practice, since such elderly patients represent almost 10% of their surgical procedures. Despite the increase in number of high-risk patients, the mortality rate has decreased, indicating an improvement in surgical, anesthetic and perioperative care.

Preoperative risk stratification of elderly patients is crucial because of medical, socio-economic and ethical considerations. Clearly, if the elderly are at higher risk of in-hospital complications and mortality [11], they must no longer be denied for surgery on the isolated age criteria, even in case of complex or emergent surgery [12–14]. If they leave the hospital alive, they could expect a longer life expectancy, a high discharge home rate after appropriate cardiac rehabilitation, and a good quality of life [15, 16].

Accurate risk prediction tools are essential because the elderly are more and more candidates for invasive procedures, except those with severe neurological or mental disorders or limited life expectancy. Credible therapeutic alternatives to conventional surgery such as percutaneous treatment of valvular heart disease are currently available. In the light of this recent evolution, accurate and objective information is mandatory for the patients and their families,

Table 5. Calibration of ES I and ES II for in-hospital mortality among the overall population and according to age.

	Overall population	<80	≥80
N =	7161	6329	832
EuroSCORE I			
O/E [°]	0.77 [0.77–0.85]	0.82 [0.73–0.91]	0.62 [0.49–0.78]
Hosmer-Lemeshow*			
p calibration ES I	<0.0001	0.001	<0.0001
EuroSCORE II			
O/E [°]	1.10 [0.99–1.21]	1.10 [0.98–1.22]	1.10 [0.87–1.37]
Hosmer-Lemeshow*			
p calibration ES II	0.006	0.13	<0.0001

[°]: Observed and Expected mortality ratio and [95% CI]

*: Hosmer-Lemeshow goodness-of-fit test

<https://doi.org/10.1371/journal.pone.0187056.t005>

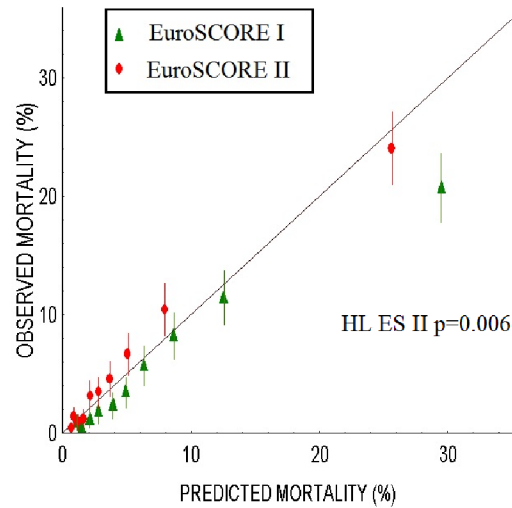


Fig 2A: Calibration plots for the entire population

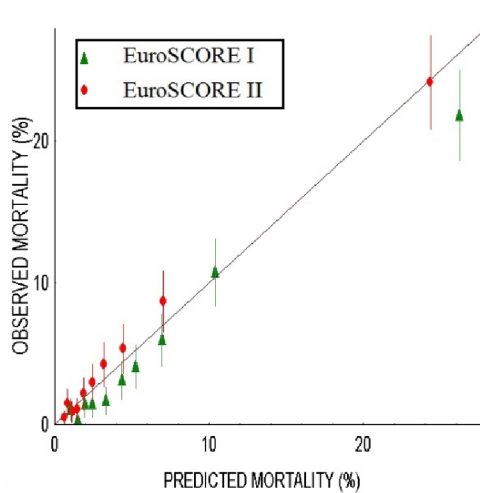


Fig 2B: Calibration plots for ≥ 80

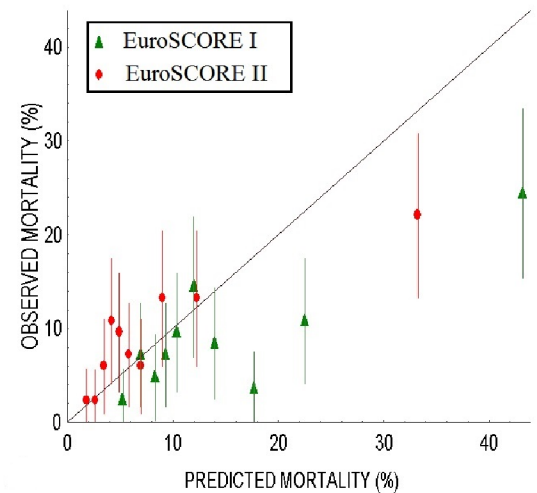


Fig 2C: Calibration plots for < 80

Fig 2. Calibration plots for the two models, for the entire population (Fig 2A) and the two studied groups (Fig 2B: ≥ 80 and Fig 2C: < 80).

<https://doi.org/10.1371/journal.pone.0187056.g002>

guiding clinical decision making and, probably offering several choices to the intermediate-risk patients in a near future.

Obviously, a risk score should only be one of the components of the final decision. Europeans' recent guidelines privilege the clinical judgment of multidisciplinary heart team for the evaluation of high risk patients [17]. Nevertheless, multifactorial risk models developed through multiple logistic regression analyses have been proposed to improve risk prediction,

mainly the STS Risk Score and the ES II and they are widely used to provide an individual assessment of OM and therefore contribute to decision making for surgery.

The term of reference for the ideal risk score should be the simplicity of use (restricted numbers of parameters) and a robust PP.

The Society of Thoracic Surgeons (STS) has developed risk models to predict mortality and morbidity for CABG, valve, and combined valve and CABG operations from a large multicenter dataset of patients. The STS risk is regularly updated and used an on-line risk calculator including a large number of covariates. Even though it is rather time consuming to calculate it for each patient, the STS score seems to be a more reliable tool to assess the OM of the high risk patient than the ES I [18].

The ES II scoring system [8] was derived from a multicenter database of more than 20000 consecutive patients. It replaced the previous version elaborated 15 years earlier, for which poor calibration was demonstrated when it was applied to contemporary data sets. As recently reported [9, 19, 20], the ES II achieves good discrimination for OM but a lower calibration for high-risk surgery. Interestingly, Barili et al. [9] reported a high discriminatory power (0.82, 95% CI: 0.80–0.85) similar to the previous original version, and an optimal calibration up to 30%-predicted mortality.

Ours results are close to the discrimination values reported by Barili et al. with an AUC of 0.79 for the ES I and 0.80 for the ES II and 0.80 for the ES I and 0.81 for the ES II for the whole population and the patients aged <80 respectively.

At the age of 80 or more, we also report moderate discrimination of ES II with an AUC below 0.70 and a fair calibration (assessed with calibration plot) without improvement between ES I and ES II.

In a recent paper Poullis et al. [21] underlined an unacceptably low c-statistic index (<0.70) in patients aged ≥ 70 for both ES I and ES II but a good calibration regardless of age. For the first time, the authors choose not to sub-analyze by operation type but specifically by age, as we did.

However, we thought it would be better to focus on the elderly patients with a higher cutoff value for age (≥ 80 in our study as compared to ≥ 70 in their analysis). Although there are commonly used definitions of old age, there is no general consensus regarding the age at which a person becomes old.

Clearly, the first conclusion that could be drawn from our results and those of the literature is that the ES II has a poor PP among the octogenarians, as the others high-risk subjects [9, 20, 22].

What reasons could be put forward to explain such a failure of risk scoring among the elderly?

Experts explain this lack of performance by the fact that octogenarians are underrepresented in contemporary dataset with patients' characteristics and operative techniques moving over time [17]. Moreover, specific risks factors (extensive ascending aortic or valvular calcifications, tissue fragility) among the elderly have relevant impact on OM, but are not applicable to the general population. We could speculate that the high rate of patients (more than 25%) who died in the operating theater or very soon thereafter from direct surgical causes illustrate these unpredictable risks.

Finally, assessment of cognitive and functional capacity and indices of frailty are not evaluated by current risk scores. Recent works [23] have shown the negative impact on mortality of what is currently named the "frailty phenotype". Preoperative evaluation with validated indices is mandated and is one of the components of the Heart Team specifications' [5].

Does that mean that all octogenarians (or more) need such a complex and, on a practical point of view, time consuming evaluation?

First of all, in order to assess all the calibration properties of the ES II system, we performed a graphical representation of O/E mortality by deciles of predicted probability (HL goodness-of-fit test, Fig 2) for the whole population and for the two age groups (<80 and ≥ 80). Using the same graphical representation, Barili et al. evidenced an optimal calibration until 30%-predicted mortality. Our results slightly differ from those of Barili in our specific elderly population: miscalibration of ES II is mainly observed when considering patients with a predicted OM of more than 10 percent (that is 20% of our population). In octogenarians, ES II provides a reliable estimation of OM, provided it is $< 10\%$, which represents the vast majority of candidates to surgery.

To summarize clinical judgment and risk score calculation would be sufficient for the low-intermediary risk patients. Finally, further studies are needed to investigate if biological variables (such as hemoglobin level [24], serum albumin and high sensitivity C-reactive protein [25]) used as surrogate markers of frailty could add to the PP of the ES II.

Limitations

The main limitations of our study are its retrospective and single-center nature. Furthermore, our results may not represent general practice or outcomes: we have to admit the dynamic nature of the outcome and its dependency on perioperative factors, which are modifiable in experts' hands. Our institutional in-hospital mortality rate was 5.7%, very close to the predicted mortality rate expected by the ES II (5.17%), higher than those of previous studies [8, 9, 22] but similar to that of a recent Spanish cohort [26], reflecting a much higher risk profile population.

This monocentric study insures the quality and the exhaustiveness of data collection particularly concerning the mortality rate. Our results differ from those previously reported in the ES II study: for the three study groups, we report a gap of 1% between 30 day mortality and in-hospital mortality for the whole population and, of 2% among the elderly, with in-hospital mortality being superior to the 30 days mortality. This raises question about the way in which the mortality was recorded for the ES II study, (the death in the hospital where the operation took place being the outcome measure) and the robustness of their observed mortality rate. A recent editorial underlines the crucial issue of a precise definition of such named OM among studies, for benchmarking between different cardiac surgical centers [27].

Conclusion

The ES II has a better PP than the ES I in patients aged < 80 . Its discrimination and calibration are less satisfying in patients aged ≥ 80 , showing an overestimation in the elderly at very high-surgical risk. Nevertheless, it shows an acceptable calibration until 10%- predicted mortality and could be used for the low-intermediate risk octogenarians, who represent the majority of patients in this age group. The poor discrimination and calibration of the ES II in very high-risk patients ≥ 80 should be an incentive to use the ES values with extreme caution when taking medical decision (e.g. TAVI versus conventional surgery) or informing the patients/families.

Supporting information

S1 Table. Dataset. Patient's data.
(XLS)

Acknowledgments

The authors thank INSERM CIC 1425, for monitoring and data management of the study.

Author Contributions

Conceptualization: Sophie Provenchère, Bernard Iung.

Data curation: Sophie Provenchère, Arnaud Chevalier, Claire Bouleti, Bernard Iung.

Formal analysis: Sophie Provenchère, Claire Bouleti, Bernard Iung.

Investigation: Sophie Provenchère, Arnaud Chevalier, Walid Ghodbane, Claire Bouleti, Bernard Iung.

Methodology: Sophie Provenchère, Bernard Iung.

Project administration: Sophie Provenchère, Bernard Iung.

Resources: Sophie Provenchère, Bernard Iung.

Software: Sophie Provenchère, Bernard Iung.

Supervision: Sophie Provenchère, Bernard Iung.

Validation: Sophie Provenchère, Arnaud Chevalier, Walid Ghodbane, Claire Bouleti, Philippe Montravers, Dan Longrois, Bernard Iung.

Visualization: Sophie Provenchère, Arnaud Chevalier, Walid Ghodbane, Claire Bouleti, Philippe Montravers, Dan Longrois, Bernard Iung.

Writing – original draft: Sophie Provenchère.

Writing – review & editing: Sophie Provenchère, Philippe Montravers, Dan Longrois, Bernard Iung.

References

1. Rechel B, Grundy E, Robine JM, Cylus J, Mackenbach JP, Knai C, et al. Ageing in the European Union. *Lancet*. 2013; 381(9874):1312–22. [https://doi.org/10.1016/S0140-6736\(12\)62087-X](https://doi.org/10.1016/S0140-6736(12)62087-X) PMID: 23541057.
2. Mathers CD, Stevens GA, Boerma T, White RA, Tobias MI. Causes of international increases in older age life expectancy. *Lancet*. 2015; 385(9967):540–8. [https://doi.org/10.1016/S0140-6736\(14\)60569-9](https://doi.org/10.1016/S0140-6736(14)60569-9) PMID: 25468166.
3. Qato DM, Wilder J, Schumm LP, Gillet V, Alexander GC. Changes in Prescription and Over-the-Counter Medication and Dietary Supplement Use Among Older Adults in the United States, 2005 vs 2011. *JAMA Intern Med*. 2016. <https://doi.org/10.1001/jamainternmed.2015.8581> PMID: 26998708.
4. Curiel-Balsera E, Mora-Ordóñez JM, Castillo-Lorente E, Benitez-Parejo J, Herruzo-Aviles A, Ravina-Sanz JJ, et al. Mortality and complications in elderly patients undergoing cardiac surgery. *J Crit Care*. 2013; 28(4):397–404. <https://doi.org/10.1016/j.jcrc.2012.12.011> PMID: 23428711.
5. Holmes DR Jr., Rich JB, Zoghbi WA, Mack MJ. The heart team of cardiovascular care. *J Am Coll Cardiol*. 2013; 61(9):903–7. <https://doi.org/10.1016/j.jacc.2012.08.1034> PMID: 23449424.
6. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). *Eur J Cardiothorac Surg*. 1999; 16(1):9–13. Epub 1999/08/24. PMID: 10456395.
7. Chhor V, Merceron S, Ricome S, Baron G, Daoud O, Dilly MP, et al. Poor performances of EuroSCORE and CARE score for prediction of perioperative mortality in octogenarians undergoing aortic valve replacement for aortic stenosis. *Eur J Anaesthesiol*. 2010; 27(8):702–7. <https://doi.org/10.1097/EJA.0b013e32833a45de> PMID: 20520558.
8. Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, et al. EuroSCORE II. *Eur J Cardiothorac Surg*. 2012; 41(4):734–44; discussion 44–5. Epub 2012/03/02. <https://doi.org/10.1093/ejcts/ezs043> PMID: 22378855.

9. Barili F, Pacini D, Capo A, Rasovic O, Grossi C, Alamanni F, et al. Does EuroSCORE II perform better than its original versions? A multicentre validation study. *Eur Heart J*. 2013; 34(1):22–9. <https://doi.org/10.1093/eurheartj/ehs342> PMID: 23028171.
10. lung B, Vahanian A. Towards improved risk scores: the quest for the grail continues. *Eur Heart J*. 2013; 34(1):10–2. Epub 2012/10/03. <https://doi.org/10.1093/eurheartj/ehs343> PMID: 23028170.
11. Krane M, Voss B, Hiebinger A, Deutsch MA, Wottke M, Hapfelmeier A, et al. Twenty years of cardiac surgery in patients aged 80 years and older: risks and benefits. *Ann Thorac Surg*. 2011; 91(2):506–13. <https://doi.org/10.1016/j.athoracsur.2010.10.041> PMID: 21256302.
12. Onorati F, Biancari F, De Feo M, Mariscalco G, Messina A, Santarpino G, et al. Outcome of redo surgical aortic valve replacement in patients 80 years and older: results from the Multicenter RECORD Initiative. *Ann Thorac Surg*. 2014; 97(2):537–43. <https://doi.org/10.1016/j.athoracsur.2013.09.007> PMID: 24036070.
13. Pagni S, Ganzel BL, Singh R, Austin EH, Mascio C, Williams ML, et al. Clinical outcome after triple-valve operations in the modern era: are elderly patients at increased surgical risk? *Ann Thorac Surg*. 2014; 97(2):569–76. <https://doi.org/10.1016/j.athoracsur.2013.07.083> PMID: 24140216.
14. Ghanta RK, Shekar PS, McGurk S, Rosborough DM, Aranki SF. Nonelective cardiac surgery in the elderly: is it justified? *J Thorac Cardiovasc Surg*. 2010; 140(1):103–9, 9 e1. <https://doi.org/10.1016/j.jtcvs.2009.10.001> PMID: 20005528.
15. Pasquali SK, Alexander KP, Peterson ED. Cardiac rehabilitation in the elderly. *Am Heart J*. 2001; 142(5):748–55. <https://doi.org/10.1067/mhj.2001.119134> PMID: 11685158.
16. Vasques F, Lucenteforte E, Paone R, Mugelli A, Biancari F. Outcome of patients aged \geq 80 years undergoing combined aortic valve replacement and coronary artery bypass grafting: a systematic review and meta-analysis of 40 studies. *Am Heart J*. 2012; 164(3):410–8 e1. <https://doi.org/10.1016/j.ahj.2012.06.019> PMID: 22980309.
17. Rosenhek R, lung B, Tornos P, Antunes MJ, Prendergast BD, Otto CM, et al. ESC Working Group on Valvular Heart Disease Position Paper: assessing the risk of interventions in patients with valvular heart disease. *Eur Heart J*. 2012; 33(7):822–8, 8a, 8b. <https://doi.org/10.1093/eurheartj/ehr061> PMID: 21406443.
18. Wendt D, Osswald BR, Kayser K, Thielmann M, Tossios P, Massoudy P, et al. Society of Thoracic Surgeons score is superior to the EuroSCORE determining mortality in high risk patients undergoing isolated aortic valve replacement. *Ann Thorac Surg*. 2009; 88(2):468–74; discussion 74–5. <https://doi.org/10.1016/j.athoracsur.2009.04.059> PMID: 19632395.
19. Guida P, Mastro F, Scrascia G, Whitlock R, Paparella D. Performance of the European System for Cardiac Operative Risk Evaluation II: a meta-analysis of 22 studies involving 145,592 cardiac surgery procedures. *J Thorac Cardiovasc Surg*. 2014; 148(6):3049–57 e1. <https://doi.org/10.1016/j.jtcvs.2014.07.039> PMID: 25161130.
20. Howell NJ, Head SJ, Freemantle N, van der Meulen TA, Senanayake E, Menon A, et al. The new EuroSCORE II does not improve prediction of mortality in high-risk patients undergoing cardiac surgery: a collaborative analysis of two European centres. *Eur J Cardiothorac Surg*. 2013; 44(6):1006–11; discussion 11. <https://doi.org/10.1093/ejcts/ezt174> PMID: 23536616.
21. Poullis M, Pullan M, Chalmers J, Mediratta N. The validity of the original EuroSCORE and EuroSCORE II in patients over the age of seventy. *Interact Cardiovasc Thorac Surg*. 2015; 20(2):172–7. <https://doi.org/10.1093/icvts/ivu345> PMID: 25348730.
22. Di Dedda U, Pelissero G, Agnelli B, De Vincentis C, Castelvechchio S, Ranucci M. Accuracy, calibration and clinical performance of the new EuroSCORE II risk stratification system. *Eur J Cardiothorac Surg*. 2013; 43(1):27–32. <https://doi.org/10.1093/ejcts/ezs196> PMID: 22822108.
23. Sepehri A, Beggs T, Hassan A, Rigatto C, Shaw-Daigle C, Tangri N, et al. The impact of frailty on outcomes after cardiac surgery: a systematic review. *J Thorac Cardiovasc Surg*. 2014; 148(6):3110–7. <https://doi.org/10.1016/j.jtcvs.2014.07.087> PMID: 25199821.
24. Scrascia G, Guida P, Caparrotti SM, Capone G, Contini M, Cassese M, et al. Incremental value of anemia in cardiac surgical risk prediction with the European System for Cardiac Operative Risk Evaluation (EuroSCORE) II model. *Ann Thorac Surg*. 2014; 98(3):869–75. <https://doi.org/10.1016/j.athoracsur.2014.04.124> PMID: 25085558.
25. Mitnitski A, Collerton J, Martin-Ruiz C, Jagger C, von Zglinicki T, Rockwood K, et al. Age-related frailty and its association with biological markers of ageing. *BMC Med*. 2015; 13:161. <https://doi.org/10.1186/s12916-015-0400-x> PMID: 26166298.
26. Garcia-Valentin A, Mestres CA, Bernabeu E, Bahamonde JA, Martin I, Rueda C, et al. Validation and quality measurements for EuroSCORE and EuroSCORE II in the Spanish cardiac surgical population: a prospective, multicentre study. *Eur J Cardiothorac Surg*. 2016; 49(2):399–405. PMID: 25762397.
27. Swinkels BM, Plokker HW. Evaluating operative mortality of cardiac surgery: first define operative mortality. *Neth Heart J*. 2010; 18(7–8):344–5. PMID: 20729999.