EuroSCORE Performance in Valve Surgery: A Meta-Analysis

Alessandro Parolari, MD, PhD, Lorenzo L. Pesce, PhD, Matteo Trezzi, MD, Laura Cavallotti, MD, Samer Kassem, MD, Claudia Loardi, MD, Davide Pacini, MD, Elena Tremoli, PhD, and Francesco Alamanni, MD

Department of Cardiac Surgery, Unit for Clinical Research in Atherothrombosis, Centro Cardiologico Monzino IRCCS, University of Milan, Milan, Italy; Department of Radiology, The University of Chicago, Chicago, Illinois; and Department of Cardiac Surgery, S. Orsola-Malpighi Hospital, University of Bologna, Bologna, Italy

Background. **The European System for Cardiac Operative Risk Evaluation (EuroSCORE) was developed to predict immediate outcomes after adult cardiac operations, but less than 30% of the cases used to develop this score were valve procedures. We studied EuroSCORE performance in valve procedures.**

Methods. **We performed a meta-analysis of published studies reporting the assessment of discriminatory power of the EuroSCORE by receiver operating characteristics (ROC) curve analysis in adult valve operations. A comparison of observed and predicted mortality rates was also performed.**

Results. **A literature search identified 37 potentially eligible studies, and 12 were selected for meta-analysis comprising 26,621 patients with 1250 events (mortality rate, 4.7%). Meta-analysis of these studies provided an average area under the curve (AUC) value of 0.730 (95% confidence interval [CI], 0.717 to 0.743). The same results were obtained when meta-analyses were performed separately in studies categorized on reliability of uncertainty estimation: in the seven studies reporting reliable uncertainty estimation (8175 patients with 358 events; mortality rate, 4.4%), the ROC curve provided an average AUC value of 0.724 (95% CI, 0.699 to 0.749). The five studies not reporting reliable uncertainty estimation (18,446 patients with 892 events; mortality rate, 4.8%) had an average AUC of 0.732 (95% CI, 0.717 to 0.747). We documented a constant trend to overpredict mortality by EuroSCORE, both in the additive and especially in the logistic form.**

Conclusions. **The EuroSCORE has low discrimination ability for valve surgery, and it sensibly overpredicts risk. Alternative risk scoring algorithms should be seriously considered.**

(Ann Thorac Surg 2010;89:787–93)

Risk stratification in cardiac surgery has become increasingly important in current clinical practice because it can provide benchmarks for hospital performances and can be useful to provide an estimation of the risk related to the operation. The additive and logistic European System for Cardiac Operative Risk Evaluation (EuroSCORE) models are widely used as risk prediction tools in adult cardiac surgery, especially in Europe [1]. These models were based on more than 19,000 consecutive patients who underwent operations at 128 European surgical centers ending in 1995 [2]. At that time, however, coronary artery bypass grafting (CABG) was dominating among the adult cardiac interventions, and less than 30% of operations involved valves [1].

Nevertheless, the EuroSCORE has been widely used to predict risk in valve surgery [3, 4], and it has been recently used to select high-risk patients for transcatheter aortic valve procedures and to compare survival after standard or transcatheter aortic valve replacement [5, 6]. In this study, we proposed a systematic review to answer the question whether it can predict early mortality after valve operations with sufficient accuracy: we assessed EuroSCORE discriminatory power using receiver operating characteristics (ROC) curves analysis and compared observed and EuroSCORE-predicted mortality rates.

Material and Methods

To study EuroSCORE performance in valve operations, a meta-analysis was done following Meta-analysis Of Observational Studies in Epidemiology (MOOSE) [7] and Standards for Reporting of Diagnostic Accuracy (STARD) [8] guidelines. On March 11, 2009, two reviewers searched Medline (1950 through March 7, 2009), and PubMed (up to March 11, 2009), including electronic links to related articles. The text string used was the single word “EuroSCORE.”

To further reduce the probability of losing any major related study, an electronic search was performed of four major cardiothoracic surgery journals in the electronic format, *Interactive CardioVascular and Thoracic Surgery*, *The Annals of Thoracic Surgery*, *The European Journal of Cardiothoracic Surgery*, and *The Journal of Thoracic and Cardiovascular Surgery* (available at http://ats.ctsnetjournals.org/

Accepted for publication Nov 11, 2009.

Address correspondence to Dr Parolari, Department of Cardiac Surgery, University of Milan Centro Cardiologico, Fondazione Monzino IRCCS, Via Parea, 4, Milan, 20138, Italy; e-mail: alessandro.parolari@cardiologicomonzino.it.

Published by Elsevier Inc

0003-4975/10/$36.00
doi:10.1016/j.athoracsur.2009.11.032

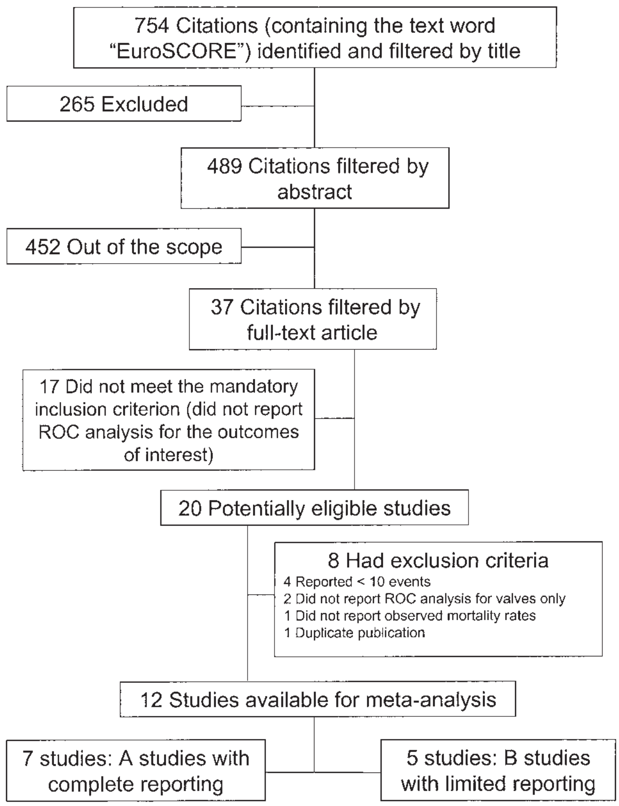

Fig 1. Flow chart of the meta-analysis. (EuroSCORE = European System for Cardiac Operative Risk Evaluation; ROC = receiving operating characteristics.)

search.dtl). The journals were searched from January 1999 to March 2009 for the single word "EuroSCORE" in the full text of all articles. The title of every article was considered first, then selected abstracts were searched to identify reports about risk prediction by the additive or logistic EuroSCORE, or heart valve operations. The full texts of these articles were retrieved and searched for in-hospital or 30-day mortality data and for mortality prediction by the EuroSCORE. Figure 1 reports the flow chart of paper selection.

Once papers were identified, the selection criterion for meta-analysis inclusion for each study was the presence of an assessment of the discriminatory power of the EuroSCORE (additive or logistic model, or both) by ROC analysis in valve operations, with or without concomitant procedures such as CABG. In addition, exclusion criteria were (1) studies reporting ROC analysis in valve operations with fewer than 10 events occurring in the study period, (2) evidence of duplicate publication, (3) confounded patient population (ie, when the ROC analysis included other patient categories such as isolated CABG and they could not be differentiated from valves), and (4) failure to report raw mortality data.

Studies selected for meta-analysis were then classified as (1) studies with reliable uncertainty estimation (group A), when the figure of merit defined as the area under the curve (AUC), was reported with a reliable dispersion parameter (standard error or 95% confidence interval [CI], or both); and (2) articles without reliable uncertainty estimation (group B), if only the AUC was reported and its dispersion parameter was missing or unreliable [8]. The dispersion variables of these last studies were estimated using an approximate formula, derived from the asymptotic form of the standard error for the estimate of a population proportion [9]. Meta-analysis models were developed for all articles, and dividing articles based on the presence of uncertainty estimation to verify whether our estimate of the AUC did not seem to influence the results of the meta-analysis. Because the purpose of this meta-analysis was to gain insight in the general discriminatory characteristics of the EuroSCORE models (ie, whether they can discriminate) we included only the AUC instead of building a complete summary ROC curve. Our analysis follows the work of McClish [10] and Zhou [11]. We assessed heterogeneity using a one-sided χ^2 measure and estimated the overall performance using inverse variance weights, as described in Parolari and colleagues [12].

The 37 potentially eligible studies (Fig 1) were also searched for reporting the observed and predicted (by either additive or logistic model) mortality rates to assess the ratio between observed and expected mortality rates; in this case, subanalyses considering different categories of patients (eg, only valves, valves plus CABG, high-risk patients) were considered only when five or more studies reported data for different categories of patients. Observed and predicted mortality rates were compared with the χ^2 test with continuity correction. Multiple comparisons calculations were done following the Holm method [13].

All calculations were done with University of Chicago ROC software (http://xray.bsd.uchicago.edu/krl/), Excel 2003 (Microsoft Inc, Redmond, WA) or Mathematica 6 software (Wolfram Research Inc, Champaign, IL).

Results

Literature Search and Article Selection

A detailed description of patient selection is reported in the Appendix*. From these selection criteria, we could identify 12 studies; of these, 7 were classified as group A studies, including 8175 patients with 358 events, for a mortality rate of 4.4% [3, 14–19]; and 5 as group B, including 18446 patients with 892 events, for a mortality rate of 4.8% [20–24] (Table 1).

Meta-Analysis

Meta-analysis of all the 12 studies (Fig 2) provided an average AUC value of 0.730 (95% CI, 0.717 to 0.743) and the Q statistic (21.9, $p = 0.056$) showed a nonsignificant amount of heterogeneity. The same results were obtained when meta-analyses were performed separately on the studies categorized on reliability of uncertainty estimation. Group A studies provided an average AUC of 0.724 (95% CI, 0.699 to 0.749) and the Q statistic (7.71, $p = 0.36$) showed very little evidence of heterogeneity. Likewise,

*See note at end of article regarding e-only Appendix.

Table 1. Main Features of the Studies Assessing EuroSCORE Performance in Valve Surgery

Studies, Year	Period	Hospitals	Patient Categories	CABG? (%)	High-Risk?	Model	Pts.	Events	(%)	95% CI	Pred.	O/E	AUC	95% CI (SE)
Group A studies														
Au, 2007 [14]	11/99–7/05	Single center	Valve pts.	N/A	No	ADDI	1406	67	4.77	(3.74–6.05)	5.2	0.917	0.77	0.70–0.83 (N/A)
Florath, 2006 [15]	3/96–12/03	Single center	AVR ± CABG	Yes (41)	No	LOGI	2198	84	3.82	(3.08–4.47)	N/A		0.7	0.66–0.75 (NA)
Heikkinen, 2007 [3]	1/93–12/00	Single center	MVR (regurg) ± CABG	Yes (37)	No	ADDI	180	18	10	(6.2–15.6)	5	2.000	0.804	0.689–0.919 (0.059)
Langanay, 2006 [16]	1/00–12/04	Single center	AVR ± CABG	Yes (19)	Yes (>80 y)	ADDI	442	33	7.47	(5.27–10.4)	N/A		0.75	0.65–0.86 (N/A)
Osswald, 2009 [17]	1/94–3/06	Single center	AVR only	No	No	ADDI	1545	34	2.2	(1.55–3.10)	6.1	0.361	0.677	0.606–0.748 (N/A)
Toumpoulis, 2005 [18]	1/92–3/02	Single center	Valve ± CABG	Yes (45)	No	ADDI	1105	70	6.33	(5.00–7.98)	8	0.791	0.72	0.66–0.78 (N/A)
Van Gameren, 2008 [19]	1/03–1/07	Single center	Valves only	No	No	ADDI	904	25	2.77	(1.84–4.12)	5.3	0.523	0.77	0.67–0.87 (N/A)
Van Gameren, 2008 [19]	1/03–1/07	Single center	Valve + CABG	Yes (all)	No	ADDI	395	27	6.84	(4.64–9.91)	6.4	1.069	0.71	0.62–0.80 (N/A)
						Total	8175	358	4.4					
Group B studies														
Bhatti, 2006 [20]	4/02–3/04	Multicenter	Valves only	No	No	LOGI	1523	53	3.48	(2.64–4.56)	7.9	0.441	0.79	N/A
Bhatti, 2006 [20]	4/02–3/04	Multicenter	Valves + CABG	Yes (all)	No	LOGI	984	71	7.22	(5.71–9.06)	9.6	0.752	0.73	N/A
Gummert, 2009 [21]	1/06–12/07	Multicenter	AVR only	No	No	LOGI	6305	245	3.89	(3.43–4.40)	7.03	0.48125	0.693	N/A
Karthik, 2004 [22]	4/97–3/02	Multicenter	AVR + CABG	Yes (all)	No	ADDI	1769	154	8.71	(7.45–10.1)	6.7	1.300	0.73	N/A
Roques, 2001 [23]	9/95–12/95	Multicenter	Valve ± CABG	Yes (21)	No	ADDI	5672	344	6.06	(5.46–6.73)	N/A		0.75	N/A
Xu, 2007 [24]	1/04–1/06	Single center	Valve ± CABG	Yes (8)	No	ADDI	2193	25	1.14	(0.75–1.70)	N/A		0.69	N/A
						Total	18,446	892	4.8					
All studies included in meta-analysis						Total	26,621	1250	4.7					

ADDI = additive EuroSCORE; AUC = area under the curve; AVR only = isolated aortic valve replacement; AVR = aortic valve replacement; AVR + CABG = aortic valve replacement with concomitant CABG; AVR ± CABG = aortic valve replacement with or without concomitant CABG; CABG = coronary artery bypass grafting; CI = confidence interval; EuroSCORE = European System for Cardiac Operative Risk Evaluation; LOGI = logistic EuroSCORE; MVR = mitral valve replacement or repair (non specified); N/A = not available; O/E = observed/expected ratio; Pred. = predicted (%) in-hospital mortality by EuroSCORE; Pts. = number of patients; SE = standard error of the mean; Valve pts. = valve procedures not otherwise specified; Valve + CABG = single, double, triple valve procedure with concomitant CABG; Valve ± CABG = single, double, triple valve procedure with or without concomitant CABG; Valves only = single, double, triple valve procedures without concomitant CABG.

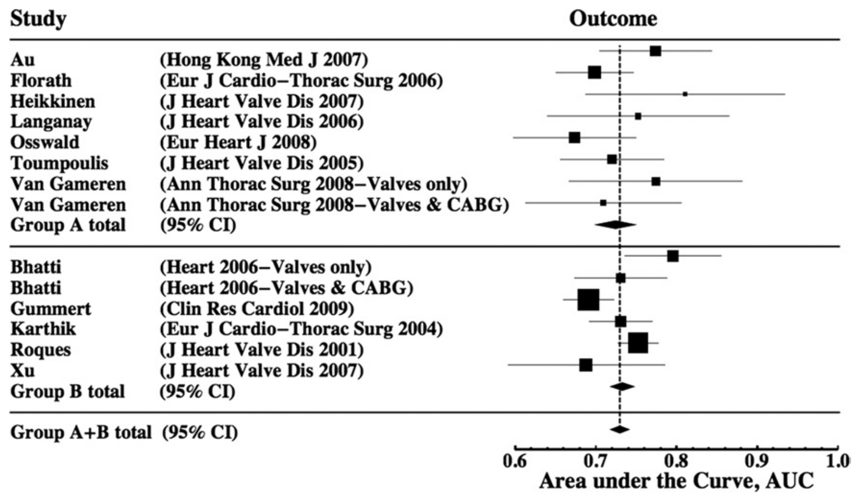

Fig 2. Meta-analysis of all studies, and of studies classified based on reliability of uncertainty estimation (please see Methods for further details) denoted by first author, journal acronym, and publication year, assessing European System for Cardiac Operative Risk Evaluation (EuroSCORE) performance in valve surgery with ROC analysis. The area under the curve (AUC) and 95% confidence intervals (95% CI) for each study are displayed. Squares indicating individual trial differences are scaled according to weighting in the meta-analysis. The width of the diamond for pooled data denotes the lower and upper 95% CI. Please note that, although 12 studies were included in this meta-analysis, in this figure there are 14 different AUC estimates because two studies that were selected (Bhatti and Van Gameren) reported the AUCs separately for patients undergoing valve operations without CABG and for patients undergoing valve operations concomitantly with CABG.

group B studies had an average AUC of 0.732 (95% CI, 0.717 to 0.747).

We investigated whether we could propose an operation-type analysis, but unfortunately, the studies differed in too many ways to allow us to reach any reliable conclusion. Given that studies involving different types of operation (valves only without CABG, high-risk patients, aortic valves, mitral valves) produced homogeneous outcomes relative to the estimation uncertainties, reliable statements about potential differences in the performance of the EuroSCORE among them could not be made. This does not mean the EuroSCORE performs identically, but rather, that much larger experiments are necessary to achieve discrimination. For this reason we did not present a subcategory analysis of the data. However, there is no reason to expect a much better performance for any of the procedures involved, and therefore, such a study might be of little relevance from the clinical perspective.

Observed vs Expected Mortality Rates

Of the 37 potentially eligible studies, 24 reported the observed and predicted mortality rates for the additive or logistic model, or for both, in valve operations (Table 2

Table 2. Observed vs Expected Mortality Rates for the Additive and Logistic EuroSCORE in Different Subsets of Patients

EuroSCORE Obs vs Exp	Patient Type	Studies No.	Patients No.	Events	%	95% CI	Predicted	%	95% CI	O/E Ratio	*p* Value
Additive	Valve pts.	20	10,387	619	5.96	5.52–6.44	698.8	6.73	6.26–7.23	0.886	0.024[a]
Logistic	Valve pts.	22	18,010	925	5.14	4.82–5.47	1551.0	8.61	8.21–9.03	0.596	<0.0001
Additive	Valve pts.-HR	8	1369	124	9.06	7.62–10.73	130.1	9.50	8.03–11.21	0.953	0.74
Logistic	Valve pts.-HR	8	1527	128	8.38	7.06–9.91	234.8	15.4	13.6–17.3	0.545	<0.0001
Additive	Valves only	7	3956	147	3.72	3.16–4.37	273.9	6.92	6.16–7.77	0.537	<0.0001
Logistic	Valves only	8	11,708	442	3.78	3.44–4.14	988.8	8.44	7.95–8.97	0.447	<0.0001
Additive	AVR ± CABG	9	5457	345	6.32	5.70–7.00	391.8	7.18	6.51–7.90	0.881	0.08
Logistic	AVR ± CABG	11	12,628	639	5.06	4.69–5.47	1133.5	8.98	8.49–9.49	0.564	<0.0001
Additive	AVR only	5	3042	121	3.98	3.32–4.75	225	7.40	6.50–8.40	0.538	<0.0001
Logistic	AVR only	5	9258	362	3.91	3.53–4.33	809.4	8.74	8.18–9.34	0.447	<0.0001

[a] Nonsignificant after multiple comparison correction.

AUC = area under the curve; AVR only = isolated aortic valve replacement; AVR ± CABG = aortic valve replacement with or without concomitant CABG; CABG = coronary artery bypass; CI = confidence interval; EuroSCORE = European System for Cardiac Operative Risk Evaluation; Exp. = expected; HR = high-risk; O/E = observed/expected ratio; Obs. = observed; Pts. = number of patients; Valve pts. = every valve procedure with or without concomitant CABG; Valves only = single, double, triple valve procedures without concomitant CABG.

and Appendix*). The following patient categories were analyzed because at least five studies were available that reported data about the observed and the expected mortality rates: (1) all patients undergoing valve operations with or without concomitant CABG; (2) high-risk patients undergoing valve operations with or without concomitant CABG; (3) patients undergoing valve surgery without concomitant CABG; (4) patients undergoing aortic valve replacement with or without concomitant CABG; and (5) patients undergoing isolated aortic valve replacement.

Overall, data show a constant and statistically significant propensity of the logistic EuroSCORE to overpredict mortality in all the five patient categories (even after multiple-comparisons correction), whereas the additive EuroSCORE overpredicts in three of five subcategories: all patients undergoing valve surgery with or without concomitant CABG (the only one nonsignificant after multiple comparisons correction), patients undergoing valve surgery without concomitant CABG, and patients undergoing isolated aortic valve replacement. In other words, the additive EuroSCORE does not seem to overestimate risk for cases of higher complexity (high-risk patients and valves plus CABG); on the other hand, the logistic EuroSCORE performs poorly in all subcategories of patients, constantly overpredicting mortality.

Comment

Scoring systems that allow perioperative risk estimation are important tools in medical quality control and in performance evaluation of hospitals and surgeons. Several risk scoring systems are currently used in adult cardiac surgery, and the EuroSCORE, both in its additive and logistic version, is the one most frequently used in Europe. This method, which requires a relatively simple data collection (17 items), has been shown to have relatively good discriminatory performance in adult cardiac operations in European [25] and North American patients [26]. It is still unknown, however, whether the best discriminatory performance by the EuroSCORE is achieved in case of CABG procedures [21, 27] or in valves [20, 28], and whether the differences in performance depend on the type of procedure. Given that at the time of its development in 1995 isolated CABG was by far the most common indication for adult cardiac operations—almost two-thirds of the procedures—whereas valve procedures accounted for little less than 30% of the procedures [2], some authors have raised questions about the reliability of this scoring system in valve procedures [19], especially in aortic interventions and in higher-risk patients [29, 30].

The discriminatory power and precision in risk prediction of the EuroSCORE in valve surgery has recently become increasingly important for two reasons. The first is that in the most centers, valve procedures—either isolated or combined—actually represent more than 50% of the total caseload; therefore, accurate risk estimation in this patient population—mainly elderly and very elderly people—has become much more important. The second reason is strictly related to the recent evolution in technical options in aortic valve operations that has led to a steady increase in the adoption of transcatheter aortic valve procedures in patients at the highest risk or in very elderly people [31]. A correct risk prediction is essential to select the patients who might benefit the most from this new and still experimental technique while at the same time managing effectively health program expenditures, because these new devices are quite expensive. Although some authors have already suggested that the EuroSCORE might be an effective tool for the selection of these patients [5, 6], the question of appropriateness of the use of these models must be addressed before such a decision process can be implemented [32].

This study has been designed mainly to define the role of the EuroSCORE in risk prediction for the current clinical cardiac surgical practice that progressively involves valve patients who are older. It was designed to attempt providing useful insights in the use of additive and logistic EuroSCORE models in defining which patients among the candidates to aortic valve replacement are at highest risk and therefore might be the best candidates for alternative and innovative transcatheter procedures.

Unfortunately, our study strongly suggests that the EuroSCORE might not be the appropriate tool for risk prediction in isolated valve operations or those combined with other cardiac procedures. The AUC derived from the meta-analysis provided estimates of 0.72 to 0.74, which are in a range of a performance considered less than satisfactory for a risk stratification algorithm [33]. EuroSCORE discrimination is also substantially lower with respect to the performance of the Society of Thoracic Surgery (STS) algorithm, which is about 0.8 for isolated valve operations [34] and about 0.75 for valves plus CABG [35]. The explanation for this is that the STS score is updated almost annually, and, for this reason, it may better follow the changes occurring in valve patient population with relative ease, whereas the EuroSCORE is now undergoing its first revision since its introduction.

That the heterogeneity test was almost significant is consistent with differences actually existing among the different procedures studies here, which should be expected, but given the large number of alternatives, it would require a large and detailed study to investigate it. Perhaps these resources would be better used to design a new more specific risk prediction scheme. Interestingly, a preliminary analysis of our data in terms of ROC analysis showed similar discriminatory performance for additive and logistic EuroSCORE models, suggesting a near equivalence for these two models from this perspective. In addition, the analysis of the ratio between observed and expected deaths by the additive and logistic models suggest overall a tendency for both models to significantly overpredict the mortality risk, which appears to be more marked for the logistic model.

In detail, the logistic EuroSCORE constantly and significantly overpredicted mortality in all five subcatego-

*See note at end of article regarding e-only Appendix.

ries of patients used in our analysis. The older method, the additive model, showed significant overprediction in three of five subcategories, the ones with somehow smaller complexity of the case/lower risk, whereas the additive model performed at least fairly in the identification of the global risk of patient population in the most complex subcategories.

This appears to contrast with previous findings suggesting that the more recent logistic model is more appropriate for risk prediction, especially in more complex cases. In fact, it appears that we are observing a EuroSCORE paradox: the older and less sophisticated algorithm—the additive EuroSCORE—outperforms the more recent and complex method implemented just to estimate with more precision the risk in complicated cases. Our study adds further evidence that in current cardiac surgical practice, there is a compelling need for perhaps even a complete reengineering of EuroSCORE. Almost 15 years have elapsed since its development, and the clinical profile of cardiac surgical patients, especially for valve procedures, has changed substantially and so should the methods used to assess their risk.

References

1. Roques F, Nashef SA, Michel P, et al. Risk factors and outcome in European cardiac surgery: analysis of the EuroSCORE multinational database of 19030 patients. Eur J Cardiothorac Surg 1999;15:816–22; discussion 22–3.
2. 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: 9–13.
3. Heikkinen J, Biancari F, Satta J, et al. Predicting immediate and late outcome after surgery for mitral valve regurgitation with EuroSCORE. J Heart Valve Dis 2007;16:116–21.
4. Pedrazzini GB, Masson S, Latini R, et al. Comparison of brain natriuretic peptide plasma levels versus logistic EuroSCORE in predicting in-hospital and late postoperative mortality in patients undergoing aortic valve replacement for symptomatic aortic stenosis. Am J Cardiol 2008;102:749–54.
5. Grube E, Schuler G, Buellesfeld L, et al. Percutaneous aortic valve replacement for severe aortic stenosis in high-risk patients using the second- and current third-generation self-expanding CoreValve prosthesis: device success and 30-day clinical outcome. J Am Coll Cardiol 2007;50:69–76.
6. Leontyev S, Walther T, Borger MA, et al. Aortic valve replacement in octogenarians: utility of risk stratification with EuroSCORE. Ann Thorac Surg 2009;87:1440–5.
7. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000;283:2008–12.
8. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. Standards for Reporting of Diagnostic Accuracy. Clin Chem 2003;49:1–6.
9. Wagner RF, Metz CE, Campbell G. Assessment of medical imaging systems and computer aids: a tutorial review. Acad Radiol 2007;14:723–48.
10. McClish DK. Combining and comparing area estimates across studies or strata. Med Decis Making 1992;12:274–9.
11. Zhou XH. Empirical Bayes combination of estimated areas under ROC curves using estimating equations. Med Decis Making 1996;16:24–8.
12. Parolari A, Pesce LL, Trezzi M, et al. Performance of EuroSCORE in CABG and off-pump coronary artery bypass grafting: single institution experience and meta-analysis. Eur Heart J 2009;30:297–304.
13. Glantz SA. Primer of biostatistics. New York: McGraw-Hill; 2002.
14. Au WK, Sun MP, Lam KT, Cheng LC, Chiu SW, Das SR. Mortality prediction in adult cardiac surgery patients: comparison of two risk stratification models. Hong Kong Med J 2007;13:293–7.
15. Florath I, Albert A, Hassanein W, et al. Current determinants of 30-day and 3-month mortality in over 2000 aortic valve replacements: Impact of routine laboratory parameters. Eur J Cardiothorac Surg 2006;30:716–21.
16. Langanay T, Verhoye JP, Ocampo G, et al. Current hospital mortality of aortic valve replacement in octogenarians. J Heart Valve Dis 2006;15:630–7; discussion 7.
17. Osswald BR, Gegouskov V, Badowski-Zyla D, et al. Overestimation of aortic valve replacement risk by EuroSCORE: implications for percutaneous valve replacement. Eur Heart J Eur Heart J 2009;30:74–80.
18. Toumpoulis IK, Anagnostopoulos CE. Does EuroSCORE predict length of stay and specific postoperative complications after heart valve surgery? J Heart Valve Dis 2005;14: 243–50.
19. van Gameren M, Kappetein AP, Steyerberg EW, et al. Do we need separate risk stratification models for hospital mortality after heart valve surgery? Ann Thorac Surg 2008;85: 921–30.
20. Bhatti F, Grayson AD, Grotte G, et al. The logistic EuroSCORE in cardiac surgery: how well does it predict operative risk? Heart 2006;92:1817–20.
21. Gummert JF, Funkat A, Osswald B, et al. EuroSCORE overestimates the risk of cardiac surgery: results from the national registry of the German Society of Thoracic and Cardiovascular Surgery. Clin Res Cardiol 2009;98:363–9.
22. Karthik S, Srinivasan AK, Grayson AD, et al. Limitations of additive EuroSCORE for measuring risk stratified mortality in combined coronary and valve surgery. Eur J Cardiothorac Surg 2004;26:318–22.
23. Roques F, Nashef SA, Michel P. Risk factors for early mortality after valve surgery in Europe in the 1990s: lessons from the EuroSCORE pilot program. J Heart Valve Dis 2001;10:572–7; discussion 577–8.
24. Xu J, Ge Y, Hu S, Song Y, Sun H, Liu P. A simple predictive model of prolonged intensive care unit stay after surgery for acquired heart valve disease. J Heart Valve Dis 2007;16: 109–15.
25. Roques F, Nashef SA, Michel P, et al. Does EuroSCORE work in individual European countries? Eur J Cardiothorac Surg 2000;18:27–30.
26. Nashef SA, Roques F, Hammill BG, et al. Validation of European System for Cardiac Operative Risk Evaluation (EuroSCORE) in North American cardiac surgery. Eur J Cardiothorac Surg 2002;22:101–5.
27. Berman M, Stamler A, Sahar G, et al. Validation of the 2000 Bernstein-Parsonnet score versus the EuroSCORE as a prognostic tool in cardiac surgery. Ann Thorac Surg 2006;81: 537–40.
28. Kawachi Y, Nakashima A, Toshima Y, Arinaga K, Kawano H. Risk stratification analysis of operative mortality in heart and thoracic aorta surgery: comparison between Parsonnet and EuroSCORE additive model. Eur J Cardiothorac Surg 2001;20:961–6.
29. Dewey TM, Brown D, Ryan WH, Herbert MA, Prince SL, Mack MJ. Reliability of risk algorithms in predicting early and late operative outcomes in high-risk patients undergoing aortic valve replacement. J Thorac Cardiovasc Surg 2008;135:180–7.
30. Grossi EA, Schwartz CF, Yu PJ, et al. High-risk aortic valve replacement: are the outcomes as bad as predicted? Ann Thorac Surg 2008;85:102–6; discussion 7.
31. Fusari M, Alamanni F, Bona V, et al. Transcatheter aortic valve implantation in the operating room: early experience. J Cardiovasc Med (Hagerstown) 2009;10:383–93.

32. Osswald BR, Gegouskov V, Badowski-Zyla D, et al. Overestimation of aortic valve replacement risk by EuroSCORE: implications for percutaneous valve replacement. Eur Heart J 2009;30:74–80.
33. Pepe M. The statistical evaluation of medical tests for classification and prediction. Oxford, NY: Oxford University Press; 2003.
34. O'Brien SM, Shahian DM, Filardo G, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 2—isolated valve surgery. Ann Thorac Surg 2009;88:S23–42.
35. Shahian DM, O'Brien SM, Filardo G, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 3—valve plus coronary artery bypass grafting surgery. Ann Thorac Surg 2009;88:S43–62.

*The Appendix is available only online. To access it, please visit http://ats.ctsnetjournals.org and search for the article by Parolari, Vol. 89, pages 787–93.

Ann Thorac Surg
2010;89:793.e1–2

Appendix

Literature Search and Article Selection

The literature search identified 37 potentially eligible studies [1–37]; of these, 17 did not meet the mandatory inclusion criterion [4, 5, 7, 8, 10–13, 16, 22, 23, 25, 26, 31, 32, 34, 36], and 8 were excluded because of the presence of exclusion criteria (Fig 1) [2, 6, 18–20, 24, 28, 30].

Twelve studies were then selected for meta-analysis purposes [1, 3, 9, 14, 15, 17, 21, 27, 29, 33, 35, 37]; two studies [3, 35] reported the area under the curve (AUC) for patients undergoing valve operations and valve operations plus coronary artery bypass grafting (CABG) separately, and data concerning both these European System for Cardiac Operative Risk Evaluation (EuroSCORE) performances have been included in the meta-analysis. In addition, it should also be noted that among the articles selected for meta-analysis, five [15, 17, 27, 35, 37] reported the performance of both additive and logistic EuroSCORE on the same patients subsets (Appendix Table).

Interestingly, all the AUC values computed using the additive or the logistic models for the same cases produced essentially identical results [15, 17, 27, 35, 37], even when the analysis was done only for subcategories [35]. This result, consistent with our previous findings [38] suggests that for the purpose of assessing discriminating ability using the AUC, we can safely use the two types of measurement interchangeably. For this reason every study was used once, favoring the additive model (because there are larger number of these studies), with the addition of only one study that reported logistic, but not additive, EuroSCORE performance [9]. Finally, one article [37] provided confidence intervals of the AUC for the logistic and additive EuroSCORE that were unrealistically small. This was probably because they used a method for computing variances that is known to be biased low when samples are very unbalanced (very different number of positive and negative cases). We therefore considered the method as providing no estimate of measurement error and classified the study as low quality.

Observed vs Expected Mortality Rates

The studies that reported the observed and predicted mortality rates for the additive or logistic model, or for both, in valve operations were references 1, 3–5, 7, 8, 10–17, 20, 22, 23, 25–27, and 32–35.

Appendix Table. Studies Included in the Meta-Analysis Reporting EuroSCORE Performance for the Additive and Logistic Models

Study, Year	Model	Pts.	Events	(%)	95% CI	Pred.	O/E	AUC	95% CI (SE)
Karthik, 2004 [17]	ADDI	1769	154	8.71	(7.45–10.1)	6.7	1.300	0.73	N/A
Karthik, 2004 [17]	LOGI	1769	154	8.71	(7.45–10.1)	9.4	0.927	0.73	N/A
Heikkinen, 2007 [15]	ADDI	180	18	10	(6.2–15.6)	5	2.000	0.804	0.689–0.919 (0.059)
Heikkinen, 2007 [15]	LOGI	180	18	10	(6.2–15.6)	3.7	2.703	0.806	0.695–0.918 (0.057)
Osswald, 2009 [27]	ADDI	1545	34	2.2	(1.55–3.10)	6.1	0.361	0.677	0.606–0.748 (N/A)
Osswald, 2009 [27]	LOGI	1545	34	2.2	(1.55–3.10)	9.3	0.237	0.666	0.593–0.740 (N/A)
Van Gameren, 2008 [35]	ADDI	904	25	2.77	(1.84–4.12)	5.3	0.523	0.77	0.67–0.87 (N/A)
Van Gameren, 2008 [35]	LOGI	904	25	2.77	(1.84–4.12)	6.1	0.454	0.76	0.66–0.86 (N/A)
Van Gameren, 2008 35]	ADDI	395	27	6.84	(4.64–9.91)	6.4	1.069	0.71	0.62–0.80 (N/A)
Van Gameren, 2008 [35]	LOGI	395	27	6.84	(4.64–9.91)	7.8	0.877	0.72	0.63–0.81 (N/A)
Xu, 2007 [37]	ADDI	2193	25	1.14	(0.75–1.70)	N/A	N/A	0.69	0.67–0.71
Xu, 2007 [37]	LOGI	2193	25	1.14	(0.75–1.70)	N/A	N/A	0.71	0.69–0.73

ADDI = additive EuroSCORE; AUC = area under the curve; CI = confidence interval; EuroSCORE = European System for Cardiac Operative Risk Evaluation; LOGI = logistic EuroSCORE; N/A = not available; O/E = observed/expected ratio; Pred. = predicted (%) in-hospital mortality by EuroSCORE; Pts. = number of patients; SE = standard error of the mean.

References

1. Au WK, Sun MP, Lam KT, Cheng LC, Chiu SW, Das SR. Mortality prediction in adult cardiac surgery patients: comparison of two risk stratification models. Hong Kong Med J 2007;13:293–7.
2. Berman M, Stamler A, Sahar G, et al. Validation of the 2000 Bernstein-Parsonnet score versus the EuroSCORE as a prognostic tool in cardiac surgery. Ann Thorac Surg 2006;81:537–40.
3. Bhatti F, Grayson AD, Grotte G, et al. The logistic EuroSCORE in cardiac surgery: how well does it predict operative risk? Heart 2006;92:1817–20.
4. Bose AK, Aitchison JD, Dark JH. Aortic valve replacement in octogenarians. J Cardiothorac Surg 2007;2:33.
5. Cladellas M, Bruguera J, Comin J, et al. Is pre-operative anaemia a risk marker for in-hospital mortality and morbidity after valve replacement? Eur Heart J 2006;27:1093–9.
6. Collart F, Feier H, Kerbaul F, et al. Primary valvular surgery in octogenarians: perioperative outcome. J Heart Valve Dis 2005;14:238–42; discussion 42.
7. Collart F, Feier H, Kerbaul F, et al. Valvular surgery in octogenarians: operative risks factors, evaluation of Euroscore and long term results. Eur J Cardiothorac Surg 2005;27:276–80.
8. Dewey TM, Brown D, Ryan WH, Herbert MA, Prince SL, Mack MJ. Reliability of risk algorithms in predicting early and late operative outcomes in high-risk patients undergoing aortic valve replacement. J Thorac Cardiovasc Surg 2008;135:180–7.
9. Florath I, Albert A, Hassanein W, et al. Current determinants of 30-day and 3-month mortality in over 2000 aortic valve replacements: Impact of routine laboratory parameters. Eur J Cardiothorac Surg 2006;30:716–21.

10. Ghosh S, Jutley RS, Wraighte P, Shajar M, Naik SK. Beating-heart mitral valve surgery in patients with poor left ventricular function. J Heart Valve Dis 2004;13:622–7; discussion 7–9.
11. Gibson PH, Croal BL, Cuthbertson BH, et al. The relationship between renal function and outcome from heart valve surgery. Am Heart J 2008;156:893–9.
12. Gonzalez-Vilchez F, Vazquez de Prada JA, Nistal F, et al. [Current surgical treatment of calcified aortic stenosis. Rev Esp Cardiol 2008;61:84–7.
13. Grossi EA, Schwartz CF, Yu PJ, et al. High-risk aortic valve replacement: are the outcomes as bad as predicted? Ann Thorac Surg 2008;85:102–6; discussion 107.
14. Gummert JF, Funkat A, Osswald B, et al. EuroSCORE overestimates the risk of cardiac surgery: results from the national registry of the German Society of Thoracic and Cardiovascular Surgery. Clin Res Cardiol 2009;98:363–9.
15. Heikkinen J, Biancari F, Satta J, et al. Predicting immediate and late outcome after surgery for mitral valve regurgitation with EuroSCORE. J Heart Valve Dis 2007;16:116–21.
16. Kaartama T, Heikkinen L, Vento A. An evaluation of mitral valve procedures using the European system for cardiac operative risk evaluation. Scand J Surg 2008;97:254–8.
17. Karthik S, Srinivasan AK, Grayson AD, et al. Limitations of additive EuroSCORE for measuring risk stratified mortality in combined coronary and valve surgery. Eur J Cardiothorac Surg 2004;26:318–22.
18. Kasimir MT, Bialy J, Moidl R, et al. EuroSCORE predicts mid-term outcome after combined valve and coronary bypass surgery. J Heart Valve Dis 2004;13:439–43.
19. Kawachi Y, Nakashima A, Toshima Y, Arinaga K, Kawano H. Risk stratification analysis of operative mortality in heart and thoracic aorta surgery: comparison between Parsonnet and EuroSCORE additive model. Eur J Cardiothorac Surg 2001;20:961–6.
20. Khaladj N, Shrestha M, Peterss S, et al. Isolated surgical aortic valve replacement after previous coronary artery bypass grafting with patent grafts: is this old-fashioned technique obsolete? Eur J Cardiothorac Surg 2009;35:260–4; discussion 264.
21. Langanay T, Verhoye JP, Ocampo G, et al. Current hospital mortality of aortic valve replacement in octogenarians. J Heart Valve Dis 2006;15:630–7; discussion 637.
22. Levy F, Laurent M, Monin JL, et al. Aortic valve replacement for low-flow/low-gradient aortic stenosis operative risk stratification and long-term outcome: a European multicenter study. J Am Coll Cardiol 2008;51:1466–72.
23. Mascherbauer J, Rosenhek R, Fuchs C, et al. Moderate patient-prosthesis mismatch after valve replacement for severe aortic stenosis has no impact on short-term and long-term mortality. Heart 2008;94:1639–45.
24. Mestres CA, Castro MA, Bernabeu E, et al. Preoperative risk stratification in infective endocarditis. Does the EuroSCORE model work? Preliminary results. Eur J Cardiothorac Surg 2007;32:281–5.
25. Monin JL, Monchi M, Kirsch ME, et al. Low-gradient aortic stenosis: impact of prosthesis-patient mismatch on survival. Eur Heart J 2007;28:2620–6.
26. Ngaage DL, Cowen ME, Griffin S, Guvendik L, Cale AR. Are initial valve operations in octogenarians still high-risk in the current era? J Heart Valve Dis 2008;17:227–32.
27. Osswald BR, Gegouskov V, Badowski-Zyla D, Tochtermann U, Thomas G, Hagl S, Blackstone EH. Overestimation of aortic valve replacement risk by EuroSCORE: implications for percutaneous valve replacement. Eur Heart J Eur Heart J 2009;30:74–80.
28. Pedrazzini GB, Masson S, Latini R, et al. Comparison of brain natriuretic peptide plasma levels versus logistic EuroSCORE in predicting in-hospital and late postoperative mortality in patients undergoing aortic valve replacement for symptomatic aortic stenosis. Am J Cardiol 2008;102: 749–54.
29. Roques F, Nashef SA, Michel P. Risk factors for early mortality after valve surgery in Europe in the 1990s: lessons from the EuroSCORE pilot program. J Heart Valve Dis 2001;10:572–7; discussion 577–8.
30. Suojaranta-Ylinen RT, Kuitunen AH, Kukkonen SI, Vento AE, Salminen US. Risk evaluation of cardiac surgery in octogenarians. J Cardiothorac Vasc Anesth 2006;20:526–30.
31. Toumpoulis IK, Chamogeorgakis TP, Angouras DC, Swistel DG, Anagnostopoulos CE, Rokkas CK. Independent predictors for early and long-term mortality after heart valve surgery. J Heart Valve Dis 2008;17:548–56.
32. Toumpoulis IK, Anagnostopoulos CE, Toumpoulis SK, DeRose JJ Jr, Swistel DG. EuroSCORE predicts long-term mortality after heart valve surgery. Ann Thorac Surg 2005; 79:1902–8.
33. Toumpoulis IK, Anagnostopoulos CE. Does EuroSCORE predict length of stay and specific postoperative complications after heart valve surgery? J Heart Valve Dis 2005;14: 243–50.
34. Urso S, Sadaba R, Greco E, et al. One-hundred aortic valve replacements in octogenarians: outcomes and risk factors for early mortality. J Heart Valve Dis 2007;16:139–44.
35. van Gameren M, Kappetein AP, Steyerberg EW, et al. Do we need separate risk stratification models for hospital mortality after heart valve surgery? Ann Thorac Surg 2008;85: 921–30.
36. Xu J, Ge Y, Pan S, Liu F, Shi Y. A preoperative and intraoperative predictive model of prolonged intensive care unit stay for valvular surgery. J Heart Valve Dis 2006;15: 219–24.
37. Xu J, Ge Y, Hu S, Song Y, Sun H, Liu P. A simple predictive model of prolonged intensive care unit stay after surgery for acquired heart valve disease. J Heart Valve Dis 2007; 16:109–15.
38. Parolari A, Pesce LL, Trezzi M, et al. Performance of EuroSCORE in CABG and off-pump coronary artery bypass grafting: single institution experience and meta-analysis. Eur Heart J 2009;30:297–304.