

Original Article



External Validation of 3 Risk Scores in Adults with Congenital Heart Disease

Bunty K. Ramchandani , MD¹, Luz Polo , MD, PhD², Raúl Sánchez , MD, PhD², Juvenal Rey, MD², Álvaro González , MD, PhD², Jesús Díez, MD, PhD³, and Ángel Aroca, MD, PhD²

¹Section of Cardiac Surgery, Hospital Universitario Clínico San Carlos, Madrid, Spain

²Hospital Universitario La Paz, Madrid, Spain

³Instituto de Investigación, Hospital Universitario La Paz, Madrid, Spain



Received: Jan 8, 2019
Revised: Feb 19, 2019
Accepted: Apr 3, 2019

Correspondence to

Bunty K. Ramchandani, MD
Section of Cardiac Surgery, Hospital Universitario Clínico San Carlos, Calle del Prof Martín Lagos, S/N, Madrid 28040, Spain.
E-mail: bunty.r@gmail.com

Copyright © 2019. The Korean Society of Cardiology

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ORCID iDs

Bunty K. Ramchandani  <https://orcid.org/0000-0002-3304-2317>
Luz Polo  <https://orcid.org/0000-0002-2615-5858>
Raúl Sánchez  <https://orcid.org/0000-0002-0145-9097>
Álvaro González  <https://orcid.org/0000-0001-9257-1174>

Conflict of Interest

The authors have no financial conflicts of interest.

ABSTRACT

Background and Objectives: Adults with congenital heart disease (CHD) are an increasing group of patients thanks to the survival of over 85% of children with CHD. 20% of these patients shall warrant a surgical procedure during their life span. However, currently there is no one risk score that assess correctly the mortality of these procedures. Thus, we analyse the risk scores used at our institution.

Methods: From May 1991 till June 2017, 608 procedures in adults with CHD were performed. The 3 risk scores (risk adjustment for congenital heart surgery [RACHS-1], Aristotle, and Euroscore I) of each procedure were analysed. We used area under the receiver operating characteristic curve (c-index) to measure model discrimination, and Hosmer-Lemeshow (H-L) statistic along with calibration plots to measure calibration.

Results: There was no statistical difference between the area under the curve for the 3 scores ($\chi^2=0.58$ with 2 df, $p=0.750$). There was no evidence of lack of fit for RACHS-1 (H-L, $\chi^2=2.61$; $p=0.271$) and Aristotle score (H-L, $\chi^2=5.69$; $p=0.459$). However, there was evidence in lack of calibration in the Euroscore I scoring system (H-L, $\chi^2=33.69$; $p<0.001$). The calibration slope for RACHS-1 was 0.912, for Aristotle (stratified in risk groups) was -0.14 and for Euroscore I (stratified in risk groups) was 0.46.

Conclusions: RACHS-1 seems to be best risk scoring system for calculating mortality applied to surgery in adults with CHD.

Keywords: Congenital heart defect; Validation studies

INTRODUCTION

In the current era of congenital heart disease (CHD) around 85–90% of paediatric patients survive into adulthood.^{1,2} It is estimated that there are 2.3 million adults with CHD in Europe, in the United States of America this figure lies over a million.² In the western world, this population outnumbers children with CHD.² The greatest survival trend has occurred in patients with more severe diseases.³ Thus, many of these patients shall warrant further surgery in the near future.³

At present, there are few tools to address the case-mix adjustment for congenital cardiac surgery in adults.⁴ In recent years there is a body of literature, which shows the efforts of

Author Contributions

Conceptualization: Ramchandani BK; Data curation: Ramchandani BK, Polo L; Formal analysis: Ramchandani BK, Polo L, Díez J; Methodology: Díez J; Project administration: Ramchandani BK; Supervision: Polo L, Sánchez R, Rey J, González Á, Díez J, Aroca Á; Validation: Ramchandani BK, Díez J; Writing - original draft: Ramchandani BK; Writing - review & editing: Ramchandani BK, Polo L, Sánchez R, Rey J, González Á, Díez J, Aroca Á.

creating a scoring system to address this growing field in cardiac surgery.^{5,6} In a day to day basis, congenital surgeons and cardiologists have to rely on a mixture of risk adjustment for congenital heart surgery (RACHS-1) scoring system,⁷ Aristotle Comprehensive Complexity (ACC) scoring system,⁸ and Euroscore I,⁹ as well as, their personal experience in order to estimate the risk of a certain procedure. The 2 former scores were designed to compare the complexity of different congenital cardiac surgeries. However, the latter is designed to evaluate in-hospital mortality for adults with acquired cardiac pathology.

The purpose of this study is to validate these risk scoring systems in a Spanish cohort of adult patients with CHD.

METHODS

We present here a retrospective cohort study of 608 consecutive cardiac operations on adult patients (age > 18 years), with previous history of CHD, who underwent cardiac surgery at a single tertiary, academic centre between May 1991 and June 2017. The Institutional Review Board and Ethics Committee approved of the study. Individual patient consent was not necessary.

Scoring systems

The RACHS-1, the ACC, and Euroscore I scoring systems were revised. The RACHS-1 score ranges from 1–6 (mortality rates were 0.4% in category 1, 3.8% in 2, 8.5% in 3, 19.4% in 4, and 47.7% in 6).¹⁰ The ACC score ranges from 1.5–15 and is divided into 4 complexity levels¹¹ (level 1, 1.5–5.6; level 2, 6–7.8; level 3, 8–9.5; and level 4, 10–15). The logistic Euroscore I ranges from 0–100% and is broken down into 3 risk categories¹² (low, 0–2 points with a mortality of 1.3%; medium, 3–5 points with a mortality of 3%; and high, >6 points with a mortality of 10%).

Outcomes

The primary outcome was in-hospital mortality defined as death occurring during the surgical procedure or the first 30 days or postoperative hospitalization stay.

Statistical analysis

Normal distribution of quantitative variables were analysed with the Kolmogorov-Smirnov test. Continuous variables with normal distribution were expressed with mean and standard deviation. Variables that did not have a normal distribution were expressed with median and interquartile range. Categorical variables were expressed by absolute and relative frequency (%). Statistical significance was considered when the probability for a type I error (alpha) was lesser than 0.05.

The performance of the 3 risk scores was compared using discrimination and calibration. Discrimination refers to the ability of the model to distinguish patients with the endpoint (dead) from patients without (alive). This is quantified with the concordance (c) statistic. For a binary endpoint, c-statistic is equal to the area under the curve (AUC) of a receiver operating characteristic curve, which represent sensitivity (true positive rate) against 1-specificity (false positive rate) for consecutive cut-offs for the predicted risk,^{13,14} these results were expressed with a 95% confidence interval (CI).

Calibration on the other hand refers to agreement between observed endpoints and predictions, in other words, it measures the accuracy of a prediction. Traditionally, Hosmer-

Lemeshow (H-L) goodness-of-fit tests were performed to assess calibration. However, this test has several disadvantages; it doesn't measure the magnitude of miscalibration, deviations in calibration are sensitive to sample size and the test is dependant to arbitrary groupings.¹⁵⁾

In our case the standard H-L test was complemented with calibration plots of observed events plotted against expected events.¹⁶⁾

These tests were repeated with the first half and the second half validation period (13 years) of the cohort to detect Hawthorn effect.¹³⁾ In the case of Aristotle score and logistic Euroscore 1, the analysis was done with and without stratification in risk categories.¹⁷⁾

Missing data for calculation of the different risk score was considered as missing completely at random as these were due to inability to access patient history. Multiple imputation was used to handle missing data.¹⁸⁾

RESULTS

Demographics

Table 1 shows our demographic data. Six hundred and eight procedures were performed from May 1991 till June 2017. The overall in-hospital mortality was of 20 deaths (3.29%), these cases are summarized in **Table 2**. Nearly half (46.4%) of the patients were males, the median age was of 36 years which shows a very young adult population. About half (43.26%) of our cohort had previous operations and a fifth (19.07%) had two or more sternotomies (**Table 3**).

The different types of congenital anomalies needing surgical procedures are represented in **Table 4**. The most frequent procedure were septal defect repairs (30.8%) with a very low mortality (0.005%), being atrial septal defect the most frequent procedure (23.36%), followed by pulmonary valve replacement (PVR) with a bioprosthesis (12.83%). Most of our deaths were in the left sided lesions surgery (10 cases), in this group the most frequent procedure was aortic valve replacement with mechanical valves (10% and 20%). However, the highest risk scores were in the thoracic arteries or veins, single ventricle and miscellaneous groups. A large proportion of our cohort had concomitant procedures, which increased the complexity of the cases and this is directly related to the surgical risk.

Missing data

There were no missing data regarding mortality. However, 7 cases did not have the RACHS-1 score calculation (1.15%), 3 cases did not have the Aristotle score (0.49%) and 35 cases did not have the logistic Euroscore I calculated (5.76%)

Table 1. Characteristics of the adult CHD cohort

Characteristic	Mean±standard deviation	Min-max	Frequency (%)
Age (years)	38.193±13.98	16.08-77.41	-
Extracorporeal time (mins)	113.5±69.66	20-545	-
Cross clamp time (mins)	64.35±49.29	0-352	-
RACHS-1	2.146±0.867	1-4	-
Aristotle	6.256±1.98	3-12.5	-
Euroscore I	5.037±5.95	0-74.3	-
Sex (male)	-	-	46.4
Previous operation	-	-	43.3

CHD = congenital heart disease; RACHS-1 = risk adjustment for congenital heart surgery.

Table 2. Description of in-hospital mortality cases

Sex	Age (years)	Diagnosis	Redo	Description	RACHs	Aristotle	Euroscore 1	Cause of death
F	23.9	TGA	3	Systemic-pulmonary shunt	3	6.8	2.08	Stroke
M	32	Right side lesion	2	Redo conduit+AVR+VSD	3	8	3.87	Incoercible bleeding
F	23.3	Thoracic vasculature	3	Aortoplasty	3	8.8	8.10	Intraoperative death
M	49	Left side lesion	1	AVR+PFO	3	7	37.31	Sepsis
F	70	Right side lesion	0	TVR	3	7.5	4.24	ARDS
M	59	Right side lesion	1	Fallot repair	2	8	4	TRALI
F	51	Left side lesion	0	MVR+extra-anatomic conduit	4	7.8	4.47	Intraoperative death
F	45	Left side lesion	0	AVR+MVR+membrane	3	7.5	6.35	Tamponade in ward
F	63.9	Left side lesion	0	Sub-aortic stenosis+miectomy	2	6.3	2.08	Incoercible bleeding
M	35	Left side lesion	0	AVR+ascending aorta	3	8.8	6.9	Incoercible bleeding
M	32	Right side lesion	1	PVR	2	6.5	8.87	Myocardial infarction
F	32.8	Left side lesion	3	Bentall+VSD	3	8.8	9.71	Stroke
M	29.2	Single ventricle	1	Single ventricular physiology	3	9	4.59	PE
M	25.8	Single ventricle	0	APVD repair+Glenn	2	9	6.58	Intraoperative death
M	57.08	Left side lesion	1	MVR+AVR+TR	3	7.5	4.34	MOF
M	32.2	Single ventricle	1	Fontan conversion+MAZE	3	9	11.1	Low cardiac output
F	26.9	Septal defect	0	Ostium primum	2	4	2.08	Ventricular tachycardia
F	26	Thoracic vasc.	1	Aortoplasty	3	8.8	3.15	Massive hemoptysis
F	18.8	Left side lesion	2	MVR	3	7.5	6.25	Sepsis
F	42.3	Left side lesion	4	MVR+MVR+de Vega	3	7.5	12.54	Sepsis
F	77.4	Left side lesion	0	AVR	3	7	6.28	Data not available

APVD = anomalous pulmonary vein drainage; ARDS = acute respiratory distress syndrome; AVR = aortic valve replacement; MOF = multi-organ failure; MVR = mitral valve replacement; PE = pulmonary embolism; PFO = patent foramen ovale; PVR = pulmonary valve replacement; Redo = reoperation; TGA = transposition of great arteries; TR = tricuspid ring; TRALI = transfusion related acute lung injury; TVR = tricuspid valve replacement; VSD = ventricular septal defects.

Table 3. Frequency of re-sternotomies

Sternotomies	Frequency
0	345 (56.74)
1	147 (24.18)
2	80 (13.16)
3	23 (3.78)
4	10 (1.64)
5	3 (0.49)

Values are presented as number (%).

Discrimination

In **Figure 1**, we compare the AUC of the 3 risk scores. The AUC for RACHS-1 was 0.712 (95% CI, 0.64–0.78), for European Association for Cardiothoracic Surgery the AUC was 0.735 (95% CI, 0.64–0.83) and in Euroscore the AUC was 0.696 (95% CI, 0.60–0.79), There was no statistical difference between the AUC for the 3 scores ($\chi^2=0.58$ with 2 df, $p=0.750$).

Table 4. Main lesions needing surgical treatment with average risk scores

Lesions	Count/mortality	Frequency	Euroscore I	RACHS-1	Aristotle	Concomitant surgery
Septal defects	187/1	30.76/0.005	2.54	1.38	4.09	62 (33.16)
Right side lesions	173/5	28.45/0.03	5.21	2.43	7.04	132 (76.3)
Left side lesions	143/10	23.52/0.07	7.43	2.89	7.61	90 (62.94)
Thoracic arteries and veins	42/2	6.92/0.05	8.39	2.24	7.62	8 (19.05)
Pulmonary vein anomalies	35/0	5.76/0	2.93	1.40	5.50	23 (65.7)
Single ventricle physiology	19/2	3.12/0.11	7.9	2.67	7.99	10 (52.63)
TGA	4/1	0.66/0.25	5.13	3	7.58	1 (25)
Double outlet right ventricle	2/0	0.33/0	1.80	3	10.3	2 (100)
Miscellaneous	2/0	0.33/0	31.14	0	6	0 (0)
Systemic vein anomalies	1/0	0.16/0	1.51	0	6	1 (100)

Values are presented as number (%).

RACHS-1 = risk adjustment for congenital heart surgery; TGA = transposition of great arteries.

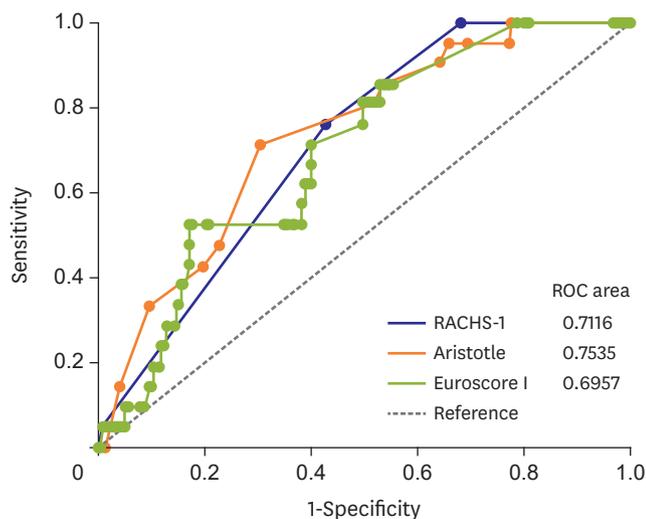


Figure 1. Discrimination of the 3 risk score scales: ROC curves of the RACHS-1, Aristotle, and Euroscore I. All 3 risk scores showed a fair discrimination without any statistically significant differences. ROC = receiver operating characteristic; RACHS-1 = risk adjustment for congenital heart surgery.

Calibration

There was no evidence of lack of fit for RACHS-1 (H-L, $\chi^2=2.61$; $p=0.271$) and Aristotle score (H-L, $\chi^2=5.69$; $p=0.459$). However, there was evidence in lack of calibration in the Euroscore I scoring system (H-L, $\chi^2=33.69$; $p<0.001$). **Figure 2** presents the calibration plots for the 3 scores. A perfect calibration has a calibration slope of 1, when the risk is overestimated the value is greater than 1 and vice versa in cases of underestimation. In **Figure 3**, the Aristotle and Euroscore I are analysed stratified in risk categories, rather than the point system.

No difference in calibration was detected in the first and second half (13 years) of the validation period. However, there was an increase in the average risk scores during these 2 time spans. In the first half (n=248), the average risk scores were: RACHS-1, 1.93 ± 0.9 ; Aristotle, 5.71 ± 2.19 ; and Euroscore I, 3.39 ± 3.23 with a total mortality of 5.24%. The second half (n=360), the average risk scores were: RACHS-1, 2.3 ± 0.81 ; Aristotle, 6.63 ± 1.73 ; and

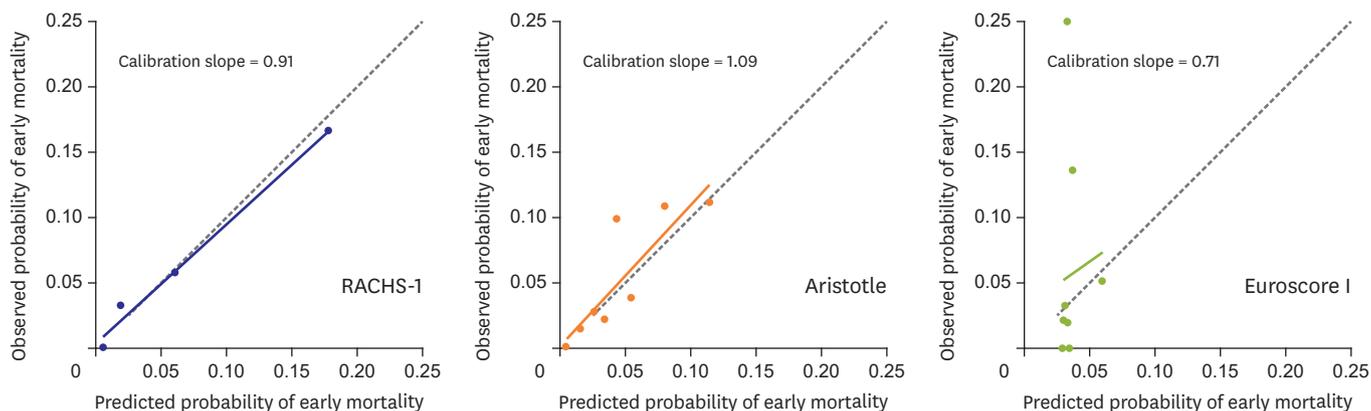


Figure 2. Calibration plots of RACHS-1, Aristotle, and Euroscore I risk score scales. RACHS-1 seems to slightly underestimate the in-hospital mortality, while Aristotle slightly overestimates it. Euroscore-I shows an important miscalibration. RACHS-1 = risk adjustment for congenital heart surgery.

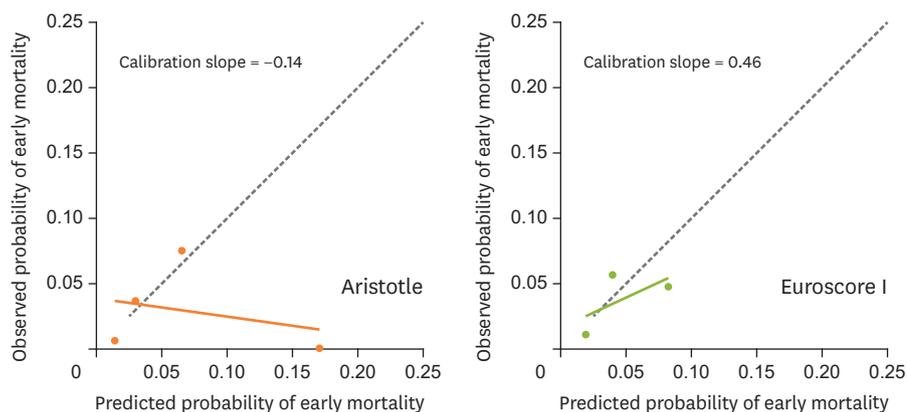


Figure 3. Calibration plots of Aristotle (1–5) and Euroscore I (low, medium, and high) stratified by risk categories. When the Aristotle and Euroscore-I are grouped in risk categories both show an important miscalibration.

Euroscore I, 6.17 ± 7.04 with a total mortality of 2.22%. During the last 13 years, despite the increase in risk complexity the mortality decreased.

DISCUSSION

Our institution is a tertiary centre and since 1991 it is one of the national referral centres for CHD. We have a consolidated multidisciplinary unit taking care of adults with CHD with over two decades of experience. The volume load of this specific aspect of our speciality has been steadily increasing, from 20–25 cases per year in the first decade to over 50 cases per year in the last few years. The complexity of the cases has also started to increase,¹⁹⁾ in the first decade septal defects were the most frequent cases referred to our centre. Nowadays, right sided lesions in reoperative scenarios like PVR and tricuspid valve repair after a Fallot repair is our daily bread and butter. There are around 100,000 adults with CHD in Spain. If 85% of the paediatric population with CHD survive to adulthood, it is calculated that there will be an increase of 2,500 adult patients with CHD per year. Approximately 20% of these patients will need an intervention in their life span, which in most cases is surgical.²⁰⁾ Thus, adult CHD is a growing field, a by-product of the successes of paediatric heart surgery. As our counterparts in acquired heart disease, it is necessary for us to have a specific risk score in order compare our outcomes with other institutions.

To our knowledge this is the first validation study in a retrospective Spanish cohort of adults with CHD. All three scoring systems fared marginally good for discrimination. RACHS-1 and Aristotle scoring systems were well calibrated, whereas Euroscore I underestimated the mortality risk. When the scores were further stratified into risk categories, both Aristotle and Euroscore I performed poorly in calibration. The Society of Thoracic Surgeons-European Association for Cardio-Thoracic Surgery mortality score was not used in this study because the number of missing values was too large to be addressed by imputation methods.

Approximately half (10 cases) of our hospital mortality were due to left cardiac lesions, with average RACHS-1 score of 3, Aristotle score of 7.57 and Euroscore I of 9%. All of the cases had concomitant procedures, either the aorta, septal defects or right side procedures and half of these cases were re-sternotomies. The second most frequent in-hospital mortality group were right cardiac lesions (4 cases), with an average RACHS 1 score of 2.75, Aristotle score 7.5

and Euroscore 5.25%. Three of these cases (75%) were re-sternotomies, and all of these had concomitant septal defect repair.

The case of a risk scoring system has always been a challenge in congenital heart surgery. Diagnosis and procedures are numerous, with small patient samples. The situation of adult CHD has the added complexity of treating acquired cardiac diseases in patients with CHD of long evolution which has not been repaired or patients with prior surgeries which now also require a further intervention in order to repair residual lesions, complications or sequels derived from the first surgery. Thanks to the herculean efforts of scholars and data collectors do we have the RACHS-1 and Aristotle scoring systems in the paediatric arena. But, variables contemplated in these score such as low weight, or prematurity do not account for the risk in these adult patients with acquired pathologies. On the other, hand Euroscore I fails to appreciate the nature of these patients, many with genetic disorders, which are prone to higher mortality and morbidity.

The limitations of this study are its retrospective nature which spreads over 26 years. Although it is of a single institution and the surgical team has not changed, but management policies have varied and so has the case-mix.

In conclusion, there is a growing necessity of a specific risk score for adults with CHD,¹³⁾ modelled with the unique case-mix of this special cohort of patients in a prospective manner. Meanwhile RACHS-1 seems to be best risk scoring system for calculating mortality in our medium.

REFERENCES

1. Verheugt CL, Uiterwaal CS, Vaartjes I, et al. Chance of surgery in adult congenital heart disease. *Eur J Prev Cardiol* 2017;24:1319-27.
[PUBMED](#) | [CROSSREF](#)
2. Zomer AC, Verheugt CL, Vaartjes I, et al. Surgery in adults with congenital heart disease. *Circulation* 2011;124:2195-201.
[PUBMED](#) | [CROSSREF](#)
3. Kogon B, Oster M. Assessing surgical risk for adults with congenital heart disease: are pediatric scoring systems appropriate? *J Thorac Cardiovasc Surg* 2014;147:666-71.
[PUBMED](#) | [CROSSREF](#)
4. Fuller S, Jacobs JP. Congenital heart surgery in adults: the challenge of estimation of risk of mortality. *World J Pediatr Congenit Heart Surg* 2016;7:436-9.
[PUBMED](#) | [CROSSREF](#)
5. Hörer J, Kasnar-Samprec J, Cleuziou J, et al. Mortality following congenital heart surgery in adults can be predicted accurately by combining expert-based and evidence-based pediatric risk scores. *World J Pediatr Congenit Heart Surg* 2016;7:425-35.
[PUBMED](#) | [CROSSREF](#)
6. Fuller SM, He X, Jacobs JP, et al. Estimating mortality risk for adult congenital heart surgery: an analysis of the society of thoracic surgeons congenital heart surgery database. *Ann Thorac Surg* 2015;100:1728-35.
[PUBMED](#) | [CROSSREF](#)
7. Jenkins KJ, Gauvreau K, Newburger JW, Spray TL, Moller JH, Iezzoni LI. Consensus-based method for risk adjustment for surgery for congenital heart disease. *J Thorac Cardiovasc Surg* 2002;123:110-8.
[PUBMED](#) | [CROSSREF](#)
8. Lacour-Gayet F, Clarke D, Jacobs J, et al. The Aristotle score: a complexity-adjusted method to evaluate surgical results. *Eur J Cardiothorac Surg* 2004;25:911-24.
[PUBMED](#) | [CROSSREF](#)

9. 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.
[PUBMED](#)
10. Jacobs ML, Jacobs JP, Jenkins KJ, Gauvreau K, Clarke DR, Lacour-gayet F. Stratification of complexity: the risk adjustment for congenital heart surgery-1 method and the Aristotle Complexity Score--past, present, and future. *Cardiol Young* 2008;18:163-8.
[PUBMED](#) | [CROSSREF](#)
11. Jacobs JP, Jacobs ML, Maruszewski B, et al. Initial application in the EACTS and STS Congenital Heart Surgery Databases of an empirically derived methodology of complexity adjustment to evaluate surgical case mix and results. *Eur J Cardiothorac Surg* 2012;42:775-9.
[PUBMED](#) | [CROSSREF](#)
12. Lafuente S, Trilla A, Bruni L, et al. Validation of the EuroSCORE probabilistic model in patients undergoing coronary bypass grafting. *Rev Esp Cardiol* 2008;61:589-94.
[PUBMED](#)
13. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J* 2014;35:1925-31.
[PUBMED](#) | [CROSSREF](#)
14. Hickey GL, Blackstone EH. External model validation of binary clinical risk prediction models in cardiovascular and thoracic surgery. *J Thorac Cardiovasc Surg* 2016;152:351-5.
[PUBMED](#) | [CROSSREF](#)
15. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. *BMJ* 2009;338:b605.
[PUBMED](#) | [CROSSREF](#)
16. Nashef SA. Risk scores and how to evaluate them. *Eur J Cardiothorac Surg* 2016;50:519.
[PUBMED](#) | [CROSSREF](#)
17. Kieser TM, Rose MS, Head SJ. Comparison of logistic EuroSCORE and EuroSCORE II in predicting operative mortality of 1125 total arterial operations. *Eur J Cardiothorac Surg* 2016;50:509-18.
[PUBMED](#) | [CROSSREF](#)
18. Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res* 1999;8:3-15.
[PUBMED](#) | [CROSSREF](#)
19. Aroca A, Polo L, Pérez-Farinós N, et al. Risk factors for surgery of congenital heart disease in adults: twenty-two years of experience. Who should operate them? *Arch Cardiol Mex* 2014;84:262-72.
[PUBMED](#)
20. Aroca Á, Segovia J, Polo L, et al. Reintervenciones quirúrgicas y trasplante de órganos torácicos. *Rev Esp Cardiol Supl* 2009;9:98-107.