DOI: 10.5114/kitp.2013.36129

Application of four complexity stratification tools (Aristotle Basic Score, RACHS-1, STAT Mortality Score, and STAT Mortality Categories) to evaluate early congenital heart surgery outcomes over 16 years at a single institution

Andrzej Kansy[1], Bohdan Maruszewski[1], Jeff Jacobs[2], Przemysław Maruszewski[3]

[1]Department of Cardiac Surgery, Children's Memorial Health Institute, Warsaw
[2]Johns Hopkins Children's Heart Surgery, All Children's Hospital and Florida Hospital for Children, Saint Petersburg, Tampa, and Orlando, Florida, United States of America
[3]Department of Surgery, Children's Memorial Health Institute, Warsaw

Kardiochirurgia i Torakochirurgia Polska 2013; 10 (2): 115–119

Abstract

Introduction: Meaningful evaluation of the quality of care must account for the variations in the population of patients receiving treatment ("case-mix"). In order to analyze mortality after congenital heart surgery over 16 years, we used four complexity stratification tools: Aristotle Basic Complexity Score (ABC Score), Risk Adjustment for Congenital Heart Surgery-1 Categories (RACHS-1 Categories), The Society of Thoracic Surgeons - European Association for Cardio-Thoracic Surgery Congenital Heart Surgery Mortality Score (STAT Mortality Score), and STAT Mortality Categories. Our goal was not only to analyze our institutional results, but also to evaluate the ability of each tool to predict mortality.

Material and methods: Complete and verified data on hospital mortalities that occurred after 8404 operations over 16 years in our institution were included in the study. For evaluating the statistical predictability of each tool, we included only those procedures that were scored by that tool.

Results: Mean hospital mortality was 4.38%, ranging from 0% to 33%. The STAT Mortality Score had the highest discrimination for predicting mortality (C-index = 0.768). The Pearson correlation coefficient between a procedure's STAT Mortality Score and its actual mortality rate was *r* = 0.84. In the subset of procedures which could be classified by all four complexity stratification tools (33 procedures), discrimination was highest for the STAT Mortality Score (C-index = 0.776).

Conclusions: In this single-institution analysis, the STAT Mortality Score had the strongest association with actual mortality. This analysis demonstrates a strategy for the application of complexity stratification tools, based on multi-institutional data, to single-institution results.

Streszczenie

Wstęp: W rzetelnej ocenie wyników leczenia należy brać pod uwagę różnorodność leczonej populacji pacjentów. Bez uwzględnienia tych danych analizowanie, a tym bardziej porównywanie, wyników jest obarczone dużym, nieakceptowalnym błędem. Do analizy śmiertelności po leczeniu wad wrodzonych serca w okresie 16 lat zastosowaliśmy cztery narzędzia do stratyfikacji ich złożoności: *Aristotle Basic Complexity Score* (ABC Score), *Risk Adjustment for Congenital Heart Surgery-1 Categories* (RACHS-1 Categories), *The Society of Thoracic Surgeons - European Association for Cardio-Thoracic Surgery Congenital Heart Surgery Mortality Score* (STAT *Mortality Score*) i STAT *Mortality Categories*. Celem pracy była nie tylko analiza jednoośrodkowych wyników, lecz także określenie zdolności każdego z zastosowanych narzędzi do przewidywania pooperacyjnej śmiertelności.

Materiał i metody: W analizie uwzględniono kompletne i zweryfikowane dane dotyczące śmiertelności szpitalnej po 8404 operacjach wykonanych w okresie 16 lat. W ocenie statystycznej użyteczności każdego z testowanych narzędzi wzięto pod uwagę jedynie procedury, które to narzędzie oceniało.

Wyniki: Średnia śmiertelność szpitalna wynosiła 4,38% (0–33%). STAT *Mortality Score* miał najwyższą zdolność prognozowania w określaniu śmiertelności (*C-index* = 0,768). Współczynnik korelacji Pearsona pomiędzy STAT *Mortality Score* określonej procedury a rzeczywistą śmiertelnością w prezentowanym w niniejszej pracy materiale wyniósł *r* = 0,84. W grupie 33 procedur, które były klasyfikowane przez wszystkie cztery narzędzia, zdolność prognozowania była również najwyższa dla STAT *Mortality Score* (*C-index* = 0,776).

Wnioski: W tej jednoośrodkowej analizie STAT *Mortality Score* wykazał najsilniejszy statystycznie związek z przewidywaniem

Address for correspondence: Andrzej Kansy, Department of Cardiac Surgery, Children's Memorial Health Institute, Al. Dzieci Polskich 20, 04-730 Warszawa, tel. +48 502 626 907, e-mail: ankansy@wp.pl

Key words: congenital heart surgery, risk stratification.

śmiertelności. Przeprowadzona analiza potwierdza użyteczność narzędzi oceny złożoności procedur, powstałych na podstawie danych wieloośrodkowych, w zastosowaniu do danych pojedynczych instytucji.

Słowa kluczowe: chirurgia wad wrodzonych serca, stratyfikacja ryzyka.

Introduction

The most universally used measures of outcome in congenital heart surgery are operative, 30-day, and hospital mortalities [1, 2]. Two large international multicenter registries for congenital heart surgery, the Society of Thoracic Surgery (STS) Congenital Heart Surgery Database and the European Association for Cardio-Thoracic Surgery (EACTS) Congenital Database, facilitate the collection, analysis, and comparison of the early outcomes of congenital cardiac surgery [3, 4]. Both databases use the same nomenclature for diagnosis, procedures, and complications, as well as the same datasets. After collecting a huge amount of procedural data and outcomes, it became obvious that the next step was to develop validated methods of risk adjustment to facilitate the analysis and reporting of outcomes. Reporting raw unadjusted mortality data can be risky, especially for the surgeons and centers that undertake the challenge of treating the most complex cases [5]. At the beginning of the last decade, two groups of researchers independently developed two tools based primarily on the subjective opinion of experts. The Aristotle Basic Complexity Score (ABC Score) consists in the sum of three components (potential for mortality, potential for morbidity, and technical difficulty) and is able to score 98% of coded operations [6-11]. The Risk Adjustment for Congenital Heart Surgery-1 Categories (RACHS-1 Categories) groups procedures into 6 levels of increasing mortality risk. RACHS-1 scores fewer procedures (92%), but appears to be a slightly stronger predictor of early death [12-15]. In 2009, O'Brien and colleagues derived a new mortality score and mortality categories using the combined data of the STS and EACTS databases. In the Society of Thoracic Surgeons - European Association for Cardio-Thoracic Surgery Congenital Heart Surgery Mortality Score (STAT Mortality Score), each procedure was assigned a numeric score by shifting and rescaling the estimated procedure-specific mortality rates to fit the range from 0.1 to 5.0, and then rounding the values to one decimal place. In the STAT Mortality Categories, procedures are sorted by increasing estimated risk of mortality and divided into 5 relatively homogeneous categories. With this data-driven approach, procedures in the same category are similar with respect to their estimated risk of mortality. The STAT Mortality Categories consist of 5 groups from 1 to 5, with higher numbers indicating higher risk of mortality. The STAT Mortality Score and STAT Mortality Categories appear to be the strongest predictors of early death after congenital heart surgery. O'Brien and colleagues showed strong positive associations between the STAT Mortality Score and STAT Mortality Categories and the observed mortality in the validation sample (the C-Index for the STAT Mortality Score was 0.816, while for the STAT Mortality Categories it was 0.812) [5]. Since 2010, the STAT Mortality Score and STAT Mortality Categories have been implemented in both the STS and EACTS Congenital Heart Surgery Databases as new tools for case-mix adjustment [3, 4].

The aim of our retrospective study was not only to analyze the results achieved by our institution, but also to evaluate the ability of each of the four complexity stratification tools to predict mortality when applied to our single-institution data.

Material and methods

The study was approved by the ethical committee of The Children's Memorial Health Institute. Since the individual patients were not identified, the need for parental consent was waived.

For our retrospective analysis, we used complete and verified data on hospital mortalities that occurred after 8404 operations for congenital heart disease performed in our institution between 1995 and 2010. Patients who were 18 years old and younger were included in the study. For all operations involving combinations of procedures that were not included in the STS-EACTS procedure list, the operation was classified according to the most technically complex procedure, as determined by the difficulty component of the ABC score. Procedures with less than 25 occurrences were excluded.

The ABC score covered the largest number of procedures – 40. The STAT Mortality Score and STAT Mortality Categories covered 37. The number of specific procedures for which the RACHS-1 Categories were defined was 33.

STAT estimated mortality was compared with the mortality observed in our institution. Each of the four tools was examined with regard to the correlation between the mortality observed in our patients and the mortality predicted by each of the scores.

Statistical analysis

Statistical analysis was performed using Statistica for Windows and MedCalc Software. Descriptive statistics were presented as mean values and 95% confidence intervals (95% CI). Each score was modeled as a continuous variable, while hospital mortality was modeled as a binary variable. Univariate and multivariate logistic regression was used to assess the risk factors for hospital mortality in our dataset. The discrimination of the model was assessed by calculating the c-statistic. The ability of the scores to predict the risk of individual procedures was quantified by calculating the Pearson correlation coefficient between the score and the actual calculated procedure-specific mortality rate in our patients. Because of sampling variations, procedures with fewer than 40 occurrences were excluded when calculating the Pearson correlation coefficient.

Results

In our experience between 1995 and 2010, the overall mean hospital mortality was 4.38%, 95% CI (3.9–4.8). After exclusion of procedures with incidence lower than 25, 7493 operations remained for further analysis (89%).

The ABC Score was applied to 89% of operations. The STAT Mortality Score and Categories covered 83.2% of operations. The RACHS-1 Categories covered 74.4%.

The frequency of in-hospital mortality for the different procedures ranged from 0 to 33.3%; there were 8 procedures with no deaths. The highest mortality occurred after Norwood operations (33.3%). The comparison of the mortality rates observed in our institution and the mortality rates estimated by the STAT risk model is presented in Table I.

Examples of the regression models of the scores and categories are summarized in Table II.

The STAT Mortality Score had the highest discrimination for predicting mortality (C-index = 0.768). The STAT Mortality Categories, RACHS-1, and ABC Score were weaker predictors of death (C-index: 0.750, 0.765, and 0.746 respectively).

To assess which of the tools discriminate mortality better, each of them was evaluated in the sample using the subset of procedures to which all four tools were applicable. This sample contained 6534 operations. As summarized in Table III, discrimination was highest for the STAT Mortality Score (C-index = 0.777), followed by the STAT Mortality Categories

Tab. I. Mortality observed in our material and estimated by the STS-EACTS risk model

Procedure name	All operations	Observed mortality risk % (95% CI)	STS EACTS estimated mortality risk model based % (95% CI)
Coarctation repair, patch aortoplasty	34	0 (0-10.8)	4.3 (2.6-6.5)
Vascular ring repair	79	0 (0-4.67)	0.9 (0.4-1.6)
Pacemaker implantation, permanent	154	0 (0-2.4)	2.2 (1.4-3.1)
Diaphragm plication	32	0 (0-11.53)	
Delayed sternal closure	143	0 (0-2.58)	
AVR, mechanical	61	0 (0-6.05)	1.7 (0.7-3.2)
AVC (AVSD) repair, intermediate (transitional)	26	0 (0-14.19)	1.6 (0.7-3.0)
Valvuloplasty, aortic	92	0 (0-4.01)	1.9 (1.1-2.9)
ASD repair, patch	616	0.16 (0.01-0.9)	0.3 (0.1-0.5)
VSD repair, primary closure	302	0.33 (0.01-1.84)	1.2 (0.6-2.1)
ASD repair, primary closure	570	0.36 (0.01-1.27)	0.9 (0.5-1.3)
Coarctation repair, end to end	247	1.2 (0.25-3.55)	1.0 (0.6-1.5)
Pacemaker procedure	510	1.37 (0.55-2.82)	1.4 (0.9-2.1)
AVC (AVSD) repair, partial (incomplete) (PAVSD)	218	1.38 (0.3-4.02)	0.5 (0.2-0.9)
TOF repair, ventriculotomy, non-transannular patch	134	1.49 (0.18-5.39)	1.5 (0.8-2.4)
Pericardial drainage procedure	326	1.53 (0.50-3.58)	
VSD repair, patch	1037	1.93 (1.18-2.98)	0.9 (0.7-1.1)
PDA closure, surgical	417	2.40 (1.15-4.41)	1.9 (1.3-2.5)
TOF repair, ventriculotomy, transannular patch	340	2.94 (1.41-5.41)	2.7 (2.1-3.4)
Aortic stenosis, subvalvular, repair	66	3.03 (0.37-10.94)	0.6 (0.3-1.0)
Valvuloplasty, mitral	29	3.45 (0.09-19.21)	1.9 (1.3-2.6)
Fontan, TCPC, external conduit, fenestrated	27	3.7 (0.09-20.63)	3.0 (2.1-4.0)
Coarctation repair, end to end, extended	143	4.19 (1.54-9.13)	2.5 (1.9-3.3)
BDCPA (bidirectional Glenn)	137	5.11 (2.05-10.53)	2.7 (2.1-3.4)
VSD, multiple, repair	28	7.14 (0.86-25.80)	4.0 (2.2-6.3)
Shunt, systemic to pulmonary, MBTS	639	7.82 (5.81-10.32)	8.9 (7.9-10.0)
ASO	212	8.50 (5.03-13.42)	4.8 (3.9-5.7)
PAB	149	9.39 (5.14-15.76)	8.0 (3.7-13.7)
AVC (AVSD) repair, complete (CAVSD)	339	10.03 (6.95-14.02)	4.6 (3.9-5.4)
RVOT procedure	38	10.53 (2.87-26.95)	2.6 (1.9-3.5)

Tab. I. Continue

Procedure name	All operations	Observed mortality risk % (95% CI)	STS EACTS estimated mortality risk model based % (95% CI)
Shunt, systemic to pulmonary, central (from aorta to main pulmonary artery)	26	11.54 (2.38-33.72)	12.1 (9.7-14.6)
BBDCPA (bilateral, bidirectional Glenn)	50	12.0 (4.40-26.12)	2.4 (1.2-3.8)
Fontan, TCPC, external conduit, non-fenestrated	25	12.0 (2.47-35.07)	3.9 (1.3-7.9)
Conduit, placement, RV to PA	27	14.81 (4.04-37.93)	6.7 (5.2-8.4)
ASO and VSD repair	111	15.32 (8.92-24.52)	8.2 (6.6-10.0)
Valve replacement, mitral (MVR)	42	16.67 (6.70-34.34)	7.3 (5.4-9.4)
Anomalous origin of coronary artery from pulmonary artery repair	44	18.18 (7.85-35.83)	2.6 (1.2-4.4)
Interrupted aortic arch repair	36	19.44 (7.82-40.06)	12.2 (9.6-15.1)
TAPVC repair	92	19.57 (11.60-30.92)	11.2 (9.5-12.8)
Aortic arch repair	73	22.22 (13.17-35.12)	7.8 (6.1-9.8)
Truncus arteriosus repair	47	29.79 (16.28-49.98)	14.1 (11.4-16.8)
Norwood procedure	54	33.33 (19.76-52.68)	23.6 (21.9-25.3)

AVR – aortic valve replacement; AVC – aortic valve calcification; AVSD – atrioventricular septal defect; TOF – tetralogy of Fallot; RVOT – right ventricular outflow tract; BBDCPA – bilateral bidirectional cavopulmonary anastomosis; TCPC – total cavopulmonary connection; ASD – atrial septal defect; VSD – ventricular septal defect; ASO – arterial switch operation; BDCPA – bidirectional cavopulmonary anastomosis; MBTS – modified Blalock-Taussig shunt; PAB – pulmonary artery banding; TAPVC – total anomalous pulmonary venous connection

Tab. II. Summary of logistic regression models of the tested tools

Tool	Odds ratio (95% CI)	*p* Value
ABC Score	1.34 (1.28–1.40)	< 0.0001
STS-EACTS Categories	2.23 (2.03–2.46)	< 0.0001
RACHS-1 Categories	2.61 (2.34–2.91)	< 0.0001
STS-EACTS Mortality Score	2.93 (2.58–3.31)	< 0.0001

Tab. III. C-index for the tested tools

Tool	C-index	SE^a	95% CI^b
STS-EACTS Mortality Score	0.777	0.0121	0.766-0.787
STS-EACTS Categories	0.772	0.0118	0.762-0.782
RACHS-1 Categories	0.765	0.0127	0.755–0.776
ABC Score	0.721	0.0137	0.710–0.732

^{a} DeLong et al. 1988; ^{b} Binominal exACT

(C-index = 0.772), RACHS-1 Categories (C-index = 0.765), and ABC Score (C-index = 0.721), all differences – $p < 0.0001$.

Receiver operating characteristic curves for the four tools are displayed in Fig. 1.

The pairwise comparisons of ROC curves between the four tools are summarized in Table IV. Only the difference between the areas of the STAT Mortality Score and the ABC Score was significant, $p < 0.0001$.

With regard to procedures with at least 40 occurrences, the Pearson correlation coefficient between the scores of a procedure and its actual observed mortality rate was highest for the STAT Mortality Score: $r = 0.8379$ (Fig. 2).

Comments

The evaluation of the predictability of operative mortality provided by each of the complexity stratification tools was one of the two major aims of our study. The strength of our study is that it was performed using complete and validated data.

This study began in 1995, which was before many of the major improvements of outcomes of the most complex operations for congenital heart disease took place. Nevertheless, in-hospital mortality in our cohort of patients is similar to that reported by O'Brien and colleagues (4.38% vs. 4.3%). In some complex procedures, the mortality in our institution was higher. The cause of the higher mortality in

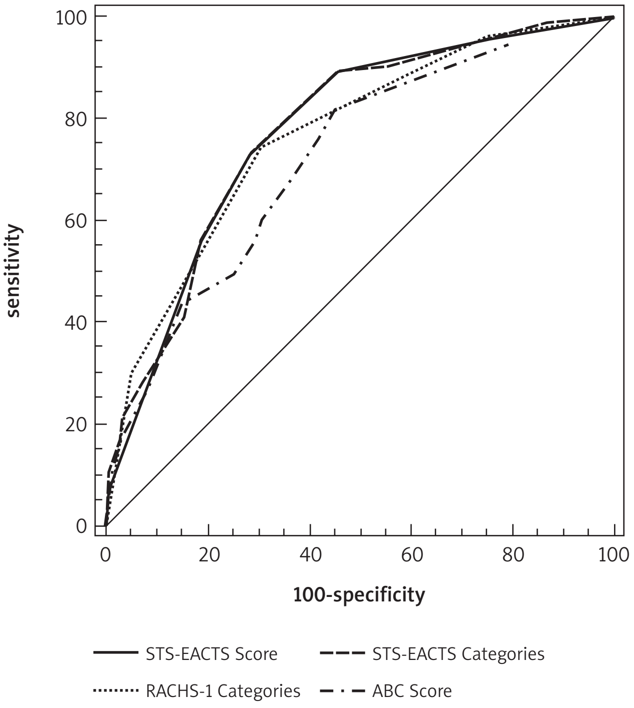

Fig. 1. Receiver operating characteristic curves for the tested tools

Tab. IV. Pairwise comparison of ROC curves between the tested tools

Tool	Difference between areas	SE ^{a}	*P* value
STS-EACTS Score vs. STS_EACTS Categories	0.0045	0.0033	0.1701
STS-EACTS Score vs. RACHS-1 Categories	0.0113	0.0082	0.1687
STS-EACTS Score vs. ABC Score	0.0553	0.0135	< 0.0001

^{a} *DeLong et al. 1988*

patients undergoing these more complex operations could be related to both the early study time frame and the small sample size of several individual procedures.

All four stratification tools have good discrimination for predicting mortality. This discrimination is higher when the sample of procedures is limited to those scored by all tools. The tool with the highest discrimination is the STAT Mortality Score followed by the STAT Mortality Categories, RACHS-1 Categories, and ABC Score. The STAT Mortality Score has the highest correlation with the observed operative mortality in our study. This finding is consistent with the data reported by O'Brien and colleagues [5]. The differences in discrimination in the four stratification tools are small, and only the difference between the STAT Mortality Score and the ABC Score is statistically significant.

Conclusions

In this single-institution analysis, the STAT Mortality Score had the strongest correlation with the observed mortality prior to discharge from the hospital. This study demonstrates a strategy for the application of complexity stratification tools, based on multi-institutional data, to single-institution results.

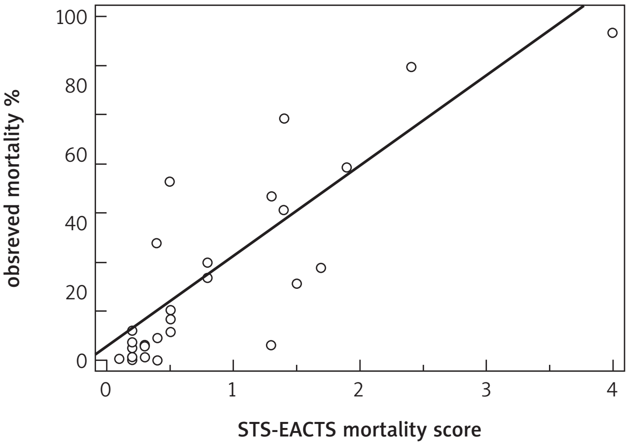

Fig. 2. Association between the STS-EACTS Mortality Score and the observed in-hospital mortality

References

1. Jacobs JP, Benavidez OJ, Bacha EA, Walters HL, Jacobs ML. The nomenclature of safety and quality of care for patients with congenital cardiac diseases: a report of Society of Thoracic Surgeons Congenital Database Taskforce Subcommittee on Patient Safety. Cardiol Young 2008; 18 suppl. 2: 81-91.
2. Maruszewski B, Lacour-Gayet FG, Monro JL, Keogh B, Tobota Z, Kansy A. An attempt at data verification in the EACTS Congenital Database. Eur J Cardiothorac Surg 2005; 28: 400-404.
3. STS Congenital Heart Surgery Database. Available at: www.sts.org.
4. EACTS Congenital Database. Available at: www. eactscongenitaldb.org.
5. O'Brien SM, Clarke DR, Jacobs JP, Jacobs ML, Lacour-Gayet FG, Pizarro C, Welke KF, Maruszewski B, Tobota Z, Miller WJ, Hamilton L, Peterson ED, Mavroudis C, Edwards FH. An empirically based tool for analyzing mortality associated with congenital heart surgery. J Thorac Cardiovasc Surg 2009; 138: 1139-1153.
6. Lacour-Gayet F, Clarke D, Jacobs J, Comas J, Daebritz S, Daenen W, Gaynor W, Hamilton L, Jacobs M, Maruszsewski B, Pozzi M, Spray T, Stellin G, Tchervenkov C, Mavroudis And C; Aristotle Committee. The Aristotle Score: a complexity – adjusted method to evaluate surgical results. Eur J Cariothorac Surg 2004; 25: 911-924.
7. Lacour-Gayet F, Clarke D, Jacobs J, Gaynor W, Hamilton L, Jacobs M, Maruszewski B, Pozzi M, Spray T, Tchervenkov C, Mavroudis C; Aristotle Committee. The Aristotle score for congenital heart surgery. Semin Thorac Cardiovasc Surg Pediatr Card Surg Annu 2004; 7: 185-191.
8. Maruszewski B, Tobota Z, Kansy A, Lacour-Gayet F, Jacobs JP, Clarke D, Elliott MJ. The Aristotle Score methogology for evaluation of outcomes in Congenital Heart Surgery. Standardy Med Pediatria 2005; 7 supl. 22: 29-33.
9. Kansy A, Maruszewski B, Burczyński P, Pastuszko A, Birbach M, Lipiński W, Mirkowicz-Małek M, Mozol K, Kozłowski M, Manowska M, Tobota Z. Analiza wczesnych wyników 2542 operacji wykonanych w referencyjnym ośrodku kardiochirurgii dziecięcej w latach 2001–2005. Kardiochir Torakochir Pol 2006; 4: 372-376.
10. Kansy A, Tobota Z, Burczyński P, Birbach M, Maruszewski P, Maruszewski B. Analiza wyników leczenia wad wrodzonych serca u noworodków z niską masą ciała na podstawie danych EACTS. Przegl Chir Dziec 2008; 3: 81-86.
11. Kansy A, Tobota Z, Maruszewski P, Maruszewski B. Analysis of 14843 neonatal congenital heart surgical procedures in the European Association for Cardiothoracic Surgery Congenital Database. Ann Thorac Surg 2010; 89: 1255-1259.
12. Jenkins KJ. Risk adjustment for congenital heart surgery: the RACHS-1 method. Semi Thorac Cardiovasc Surg Pediatr Card Surg Annu 2004; 7: 180-184.
13. Jenkins KJ, Gauvreau K. Center-specific differences in mortality: preliminary analyses using the Risk Adjustment in Congenital Heart Surgery (RACHS-1) method. J Thorac Cardiovasc Surg 2002; 124: 97-104.
14. Al-Radi OO, Harrell FE Jr, Caldarone CA, McCrindle BW, Jacobs JP, Williams MG, Van Arsdell GS, Williams WG. Case complexity scores in congenital heart surgery: a comparative study of the Aristotle Basic Complexity score and the Risk Adjustment in Congenital Heart Surgery (RACHS-1) system. J Thorac Cardiovasc Surg 2007; 133: 865-875.
15. Kang N, Tsang VT, Elliott MJ, de Leval MR, Cole TJ. Dose the Aristotle score predict outcome in congenital heart surgery? Eur J Cardiothorac Surg 2006; 29: 986-988.