The effect of duration of deep hypothermic circulatory arrest in infant heart surgery on late neurodevelopment: The Boston Circulatory Arrest Trial

David Wypij, PhD^{a,e,f,i}

Jane W. Newburger, MD, MPH^{a,f}

Leonard A. Rappaport, MD^{b,f}

Adre J. duPlessis, MBChB^{c,g}

Richard A. Jonas, MD^{d,h}

Gil Wernovsky, MD^{a,f}

Ming Lin, MS^{a,i}

David C. Bellinger, PhD, MSc^{c,g}

See related article on page 1385.

From the Departments of Cardiology,^a Medicine,^b Neurology,^c and Cardiovascular Surgery,^d and the Clinical Research Program, Children's Hospital^e; the Departments of Pediatrics,^f Neurology,^g and Surgery,^h Harvard Medical School; and the Department of Biostatistics,ⁱ Harvard School of Public Health, Boston, Mass.

Supported by grants HL41786, RR02172, and P30-HD18655 from the National Institutes of Health.

Received for publication June 12, 2002; revisions requested Dec 10, 2002; revisions received May 28, 2003; accepted for publication June 3, 2003.

Address for reprints: David C. Bellinger, PhD, MSc, Neuroepidemiology Unit, Farley Basement 127, Children's Hospital, 300 Longwood Ave, Boston, MA 02115 (E-mail: david.bellinger@tch.harvard.edu).

J Thorac Cardiovasc Surg 2003;126:1397-403

Copyright © 2003 by The American Association for Thoracic Surgery

0022-5223/2003 \$30.00 + 0

doi:10.1016/S0022-5223(03)00940-1

Objectives: Despite the technical advantages of total circulatory arrest for vital organ support during infant heart surgery, many centers have moved away from its use because of the demonstrated effects of circulatory arrest of long duration on neurodevelopmental outcomes. Our goal was to determine the functional form of the association between duration of circulatory arrest and risk of neurodevelopmental dysfunction.

Methods: From 1988 to 1992, in a single-center trial, infants with D-transposition of the great arteries underwent the arterial switch operation after random assignment to circulatory arrest or low-flow bypass. The alpha-stat method was used, and hematocrit on bypass was maintained at 20%. Developmental, neurologic, and speech outcomes were assessed at 8 years of age in 155 of 160 eligible children (97%). Outcomes selected for analysis were Full-Scale, Verbal, and Performance IQ, Reading and Mathematics Composite, time to complete the Grooved Pegboard (dominant hand), and the Mayo Test for Apraxia.

Results: Nonparametric regression and piecewise linear models indicated that neurodevelopmental outcomes were generally not adversely affected unless the duration of circulatory arrest exceeded a threshold of 41 minutes (95% 1-sided lower confidence limit of 32 minutes).

Conclusions: We found that the effect of duration of total circulatory arrest on later neurodevelopmental outcomes is nonlinear, with little influence at shorter durations and with steadily worsening outcomes after longer durations of circulatory arrest. Because the effects of duration of circulatory arrest may vary according to diagnosis, age at surgery, and other bypass and perioperative variables, this study cannot ascertain a universally "safe" duration of total circulatory arrest.

Considerable research has been conducted to determine the duration of total circulatory arrest (CA) that infants undergoing cardiac surgery can tolerate without suffering increased risk of irreversible structural or functional brain damage. The weight of evidence suggests that the maximum "safe" duration is 39 to 65 minutes,^{1,2} although this estimate is based on studies that are heterogeneous in many respects, including the patient population (ie, heart defect), sample size, length of follow-up, end points assessed, and key aspects of surgical management. Because CA has technical advantages from a surgical standpoint and is necessary to repair some complex congenital heart lesions, refinement of this estimate might help to maximize patients' late central nervous system outcomes. Among the challenges in

doing so is the number of other factors that affect neurodevelopmental outcomes, including regional brain temperature during CA, rate of cooling, cerebral blood flow and distribution during cooling, arterial blood pressure during cooling and reperfusion, pH and hematocrit management strategies, and postoperative care.² To render these aspects of surgical management comparable between treatment groups, we began a clinical trial in which children with D-transposition of the great arteries (D-TGA) undergoing the arterial switch operation were randomized to a vital organ support method consisting of deep hypothermia with either predominantly CA or predominantly low-flow (LF) cardiopulmonary bypass. The results of assessments conducted in the perioperative period and at ages 1 and 4 years consistently indicated poorer outcomes among children assigned to CA than to LF.³⁻⁵

In a companion article,⁶ we presented the results of intention-to-treat analyses comparing the neurodevelopmental outcomes of the CA and LF treatment groups at 8 years of age. In this article, we report additional analyses in which we attempted to identify, in more quantitative terms, the functional form of the association between CA duration and key neurodevelopmental outcomes.

Methods

The protocol for the Boston Circulatory Arrest Study has been described previously.^{3,6} Perfusion methods will be reviewed here because they are of particular importance to this article. When rectal temperature reached 18°C or lower and the necessary preliminary surgical dissection had been completed, LF bypass or CA was begun. For patients assigned to the CA strategy, CA was used until continuity was established from the left ventricle to the aorta, the coronary arteries were reimplanted, and the atrial (ASD) or ventricular septal defect (VSD), if any, was repaired. In the LF group, the perfusion rate was reduced to $50 \text{ mL} \cdot \text{min}^{-1} \cdot \text{kg}^{-1}$ (approximately $0.7 \text{ L} \cdot \text{min}^{-1} \cdot \text{m}^{-2}$) for the duration of the aortic and coronary repairs. Infants in both groups underwent repair of ASDs and VSDs with total CA. Perfusion methods were otherwise identical in both groups. We used the alpha-stat pH strategy and crystalloid hemodilution to a hematocrit value of approximately 20% in all patients. The companion article⁶ describes the follow-up rate and the elements of the neurodevelopmental evaluations conducted when the children were age 8 years. Neurodevelopmental outcomes were assessed at 8 years of age in 155 of 160 eligible children (97%).

For analyses looking at the associations with CA duration, continuous outcomes considered were Full-Scale, Verbal, and Performance IQ from the Wechsler Intelligence Scale for Children—Third Edition,⁷ Reading and Mathematics Composite score from the Wechsler Individual Achievement Test,⁸ time to complete the Grooved Pegboard⁹ (dominant hand), and the Mayo Test for Apraxia of Speech and Oral Apraxia—Children's Battery (selected items).¹⁰ The Mayo Test measures the motor planning portion of speech production. These end points were selected from among the large number assessed in the full neurodevelopmental evaluation. The selection was based either on our finding, in intention-to-treat

analyses, that treatment group mean scores differed significantly (ie, Grooved Pegboard, Mayo Test for Apraxia) or based on the importance of the outcome (ie, IQ, academic achievement). We sought to determine the relationship between CA duration and neurodevelopmental test scores measured on a continuous scale. In addition, 2 important binary outcome variables were also considered (presence of apraxia diagnosis and presence of possibly or definitely abnormal overall neurologic examination findings).

Statistical Analysis

Treatment group differences were evaluated by means of intention-to-treat analysis. Analyses assessing the effects of duration of CA as a continuous variable on major continuous outcomes used linear regression and generalized additive models.^{11,12} Generalized additive models are methods of nonparametric regression smoothing that allow for estimation of *P* values for any effect of duration of CA on outcome, as well as *P* values for the departure from linearity or piecewise linearity of effect. Piecewise linear (cut-point, or broken stick) regression models were used to estimate potential thresholds in the effect of duration of CA. Confidence intervals for cut-points were based on inversion of likelihood ratio tests.¹³ A likelihood ratio test showed no statistical evidence for heterogeneity in the cut-points across the multiple outcomes. To estimate a combined threshold across multiple outcomes, we performed a multivariate analysis using maximum likelihood with an arbitrary covariance matrix. All likelihood maximizations and likelihood ratio tests searched over a grid of duration of CA in units of 0.1 minute. Given the precision of our CA durations (to the nearest minute), we report the estimated cut-points and confidence limits to the nearest integer minute. Higher-order polynomial (eg, quadratic) regression models were also considered but did not fit the data better than either generalized additive models or piecewise linear models.

All comparisons were adjusted for diagnosis (intact ventricular septum [IVS] versus VSD) and family social class (Hollingshead Four-Factor Index of Social Class, unpublished manuscript). We also assessed potential interactions between diagnosis and duration of CA. For the generalized additive model and cut-point regression analyses, we used the average Reading Composite and Mathematics Composite scores, because the models for these outcomes considered individually were so similar. For the time to complete the Grooved Pegboard, proportional hazards regression was used to compare treatment groups. For the generalized additive model and cut-point regression analyses, the natural logarithm of the time to complete the Grooved Pegboard was used because this transformation yielded scores that were approximately normally distributed.

Analyses assessing the effects of duration of CA as a continuous variable on binary outcomes used logistic regression adjusting for diagnosis. Fisher exact tests were used to compare binomial proportions. All *P* values are 2-sided. Because we wished to determine the minimum duration of CA at which adverse effects could be detected under the conditions particular to the era of our study, we report 1-sided 95% lower confidence limits for cut-points. These were calculated as the lower limits of 2-sided 90% confidence intervals.

TABLE 1. Neurodevelopment scores according to ventricular septal status and treatment group

Variable	Intact ventricular septum		Ventricular septal defect		P value*
	Circulatory arrest (n = 59)	Low-flow bypass (n = 58)	Circulatory arrest (n = 20)	Low-flow bypass (n = 18)	
Wechsler Intelligence Scale for Children, 3rd edition					
Full-Scale IQ	98.0 ± 16.3	99.3 ± 13.7	88.5 ± 14.6	96.9 ± 15.2	.46
Verbal IQ	101.9 ± 18.4	101.4 ± 14.1	89.1 ± 15.1	99.5 ± 15.9	.33†
Performance IQ	94.4 ± 14.5	97.3 ± 13.3	90.0 ± 15.5	94.6 ± 15.1	.03
Wechsler Individual Achievement Test, summary scores					.31
Reading Composite	95.6 ± 17.7	96.8 ± 12.4	89.5 ± 12.9	95.6 ± 15.5	.66
Mathematics Composite	97.2 ± 19.2	98.9 ± 14.4	90.4 ± 14.4	94.1 ± 20.8	.79
Fine Motor					
Grooved Pegboard, time to complete: dominant hand (s)	104 ± 35	94 ± 26	111 ± 28	90 ± 12	.06‡
Mayo Test for Apraxia	190 ± 13	195 ± 7	188 ± 13	197 ± 8	.002

Values are mean ± SD.

*P values, determined by linear regression, are for differences between treatment groups, with adjustment for diagnosis and social class.

†Upper P value is for patients with an intact ventricular septum, and lower P value is for patients with a ventricular septal defect, due to a significant treatment group by diagnostic group interaction. Based on linear regression, with adjustment for social class.

‡P value, determined by proportional hazards regression, is for differences between treatment groups, with adjustment for diagnosis and social class.

Results

General Intelligence

Treatment groups did not differ significantly in terms of Full-Scale or Performance IQ (Table 1). Among children with D-TGA/VSD only, assignment to CA was associated with lower Verbal IQ ($P = .03$). In nonparametric smoothing regression analyses of the entire cohort using duration of CA as a continuous variable, we found that longer CA was significantly associated with lower Full-Scale IQ ($P = .04$, Figure 1, A) and Performance IQ ($P = .02$, graph not shown). In addition, these generalized additive models showed marginal evidence for departures from linearity for both Full-Scale IQ and Performance IQ ($P = .08$ for each).

For Verbal IQ, the diagnosis by duration of CA interaction term fell just short of significance ($P = .052$), with the magnitude of the effect of duration of CA being slightly more negative for the VSD group compared with the IVS group. The association between duration of CA and Verbal IQ in the combined diagnosis groups was not significant ($P = .19$). In addition, we found no significant departures from linearity for Verbal IQ ($P = .53$ for the IVS group, $P = .99$ for the VSD group).

Academic Achievement

Treatment groups did not differ significantly on the Reading and Mathematics Composite scores. The average of the Reading and Mathematics Composite scores was not significantly related to duration of CA ($P = .26$, Figure 1, B) and showed no evidence for departure from linearity ($P = .99$).

Fine Motor Function

Children assigned to CA tended to take longer than the children assigned to LF to complete the Grooved Pegboard for the dominant hand ($P = .06$, Table 1). Longer duration of CA was associated with longer time to complete ($P = .01$, Figure 1, C), and completion times showed a significant departure from linearity ($P = .02$).

Speech Production

Children assigned to CA did significantly worse than children in the LF group on the Mayo Test for Apraxia of Speech and Oral Apraxia ($P = .002$, Table 1). Longer duration of CA was significantly associated with worse scores ($P < .001$, Figure 1, D), and the Mayo Test for Apraxia showed a significant departure from linearity ($P = .002$).

Determination of Circulatory Arrest Thresholds

Among the 6 outcome variables, significance or tendency toward departures from linearity was observed for Full-Scale IQ ($P = .08$), Performance IQ ($P = .08$), Grooved Pegboard ($P = .02$), and the Mayo Test for Apraxia of Speech and Oral Apraxia ($P = .002$). We explored the potential threshold effects of CA duration on these outcomes using piecewise linear regression models to determine the best single cut-point beyond which scores declined. For example, for Full-Scale IQ scores a cut-point of 42 minutes was estimated (1-sided 95% lower confidence limit of 26 minutes, Table 2). The linear effect of CA duration on Full-Scale IQ was quite flat (not significantly

Figure 1. Neurodevelopmental outcome as a function of duration of CA. The *solid lines* were derived from generalized additive models as a smooth function of duration of CA, with adjustment for diagnosis and family social class. The *dashed lines* present 95% confidence intervals, with the scatter plots denoting the raw data values. The *P* values assess whether there is any effect of duration of CA on outcomes, with adjustment for diagnosis and family social class. The panels represent Full-Scale IQ (A), Average Achievement (of Wechsler Reading and Mathematics Composite scores) (B), Grooved Pegboard (Dominant Hand) (C), and Mayo Test of Apraxia of Speech (D). The natural logarithm of the time to complete the Grooved Pegboard was used because this transformation yielded scores that were approximately normally distributed. The graph then exponentiates the results to return to the raw scale.

different from zero) for duration of CA less than 42 minutes but significantly negative beyond 42 minutes.

We found remarkable consistency in the cut-points beyond which scores worsened across these outcome variables (Table 2), with estimated cut-points ranging between 35 and 43 minutes. The 1-sided 95% lower confidence limits vary among the outcome variables, depending on the strength of the cut-point model (relative to a simple linear regression model) and the inherent variability of the outcome variables. To sharpen inferences, we performed a multivariate

analysis using all 6 outcome variables. We found no evidence for heterogeneity in the 6 cut-points, and so fit a model with a common cut-point. In this analysis, the point estimate of the common cut-point was 41 minutes, with a 1-sided 95% lower confidence limit of 32 minutes (Table 2). As expected, the common cut-point estimate is in the middle of the individual cut-point estimates. The lower confidence limit is tighter (closer to the cut-point estimate) due to additional information gained across the 6 outcome variables.

To explore these models further, we fit separate models for these outcome variables using a common cut-point of 41 minutes, and we compared the slopes of duration of CA before and after this cut-point (Table 3). As this common cut-point was estimated from our data, we view these analyses as exploratory in nature. For each of the 6 outcome variables, we found no significant association between duration of CA and outcome for durations of less than 41 minutes. However, we found significant associations between duration of CA and outcome for durations of more than 41 minutes for Full-Scale IQ ($P = .006$), Verbal IQ ($P = .03$), Performance IQ ($P = .009$), Grooved Pegboard ($P = .04$), and Mayo Test of Apraxia ($P < .001$). In no case did nonparametric regression smoothing or higher-order polynomial regression models find significant departures from the piecewise linear model. In these cut-point models using duration of CA and adjusting for diagnosis and family social class, we found no statistically significant effects of total support time on outcomes.

Additional analyses focused on comparing the IQ and achievement scores for those subjects with less than 41 minutes of CA to population norms. The mean scores of these 76 subjects were 98.9 ± 14.1 (mean $\pm$ SD) for Full-Scale IQ, 100.8 ± 14.9 for Verbal IQ, 97.1 ± 13.6 for Performance IQ, 96.8 ± 13.4 for Reading Composite score, and 97.6 ± 16.4 for Mathematics Composite score, compared with population norms of 100 ± 15 . This subgroup performed similarly to the population norms, except that they had lower Reading Composite scores ($P = .04$) and a tendency for lower Performance IQ ($P = .07$). Subjects with greater than or equal to 41 minutes of CA, compared with those with less than 41 minutes, were slightly more likely to have received remedial services (44% versus 30%) and to have repeated a grade (12% versus 8%), but these differences were not statistically significant.

Binary Outcomes

Adjusting for diagnosis, duration of CA was significantly associated with presence of apraxia diagnosis ($P = .02$, Figure 2, A) but not with possibly or definitely abnormal overall neurologic examination findings ($P = .40$, Figure 2, B). As seen in Figure 2, A, the risk of apraxia diagnosis begins to increase at approximately 30 to 40 minutes of CA, consistent with the cut-point analyses for the continuous outcomes.

Discussion

We found that important late neurodevelopmental outcomes, such as IQ, academic achievement, fine motor function, and speech production, were generally unaffected by use of shorter duration of CA. In our study cohort who underwent surgery from 1988 to 1992 using perfusion methods particular to that era, nonparametric regression methods

TABLE 2. Estimation of threshold effects of duration of circulatory arrest based on piecewise linear models*

Variable	Cut-point estimate (min)	95% lower confidence limit (min)
Full-Scale IQ	42	27
Verbal IQ	41	23
Performance IQ	47	31
Average Achievement	43	4
Grooved Pegboard	35	13
Mayo Test for Apraxia	40	29
Combined analysis (over all 6 outcomes)	41	32

*All models were based on piecewise linear regression, with adjustment for diagnosis and social class.

indicated worse performance in children whose CA duration exceeded 41 minutes, after which performance declined monotonically. The 1-sided 95% lower confidence limit for this estimate was 32 minutes. Our study findings, the first obtained within an experimental setting, thus indicate that the neurodevelopmental sequelae of total CA are related to the duration of its use in a nonlinear fashion. During the era in which our patients underwent surgery, the “safe” duration of CA appeared to lie at the lower end of the ranges that have been suggested previously.^{1,2}

The correlation/linear regression methods used in previous studies,¹⁴⁻¹⁸ which force a linear fit to the relationship between CA and neurodevelopment, are problematic if a threshold exists below which duration of CA is not related to outcome. The overall slope or correlation is a weighted average of the flat slope at CA durations below the threshold and the slope that describes the relationship at CA durations that exceed the threshold. The latter slope is thus necessarily equal to or greater than the slope calculated for the entire range of CA duration under the assumption of linearity. Our use of nonparametric regression methods thus provided an opportunity to identify a threshold or cut-point in this association.

On the basis of previous reports,¹⁻⁵ some groups undertaking repair of congenital heart anomalies have attempted to avoid any use of CA. The present study suggests that it may not be necessary to take special measures to avoid deep hypothermic CA periods of shorter duration, and cardiovascular surgeons should balance any technical advantages of the various perfusion techniques with their potential risks on an individual basis. Innovative ancillary methods for cerebral support have been introduced by many groups as alternatives to deep hypothermic CA. These methods, such as retrograde cerebral perfusion, are hypothesized to provide improved and homogeneous cerebral protection relative to deep hypothermic CA. The efficacies of these techniques should be tested in randomized, controlled trials, however.

Figure 2. Estimated probabilities of apraxia diagnosis (A) and abnormal neurologic examination (B) as a function of the duration of total CA. The *solid lines* were derived from logistic regression curves, and the *dashed lines* present 95% confidence intervals. In addition, point estimates and exact 95% confidence intervals for outcome probabilities are plotted for the mean of each quartile of duration of CA. The *P* values were based on logistic regression for the effect of duration of CA on outcome, with adjustment for diagnosis.

TABLE 3. Exploratory analysis of regression slopes for neurodevelopmental scores as a function of duration of circulatory arrest for a common threshold of 41 minutes*

Variable	Circulatory arrest <41 min			Circulatory arrest >41 min			<i>P</i> value comparing slopes
	Regression slope	95% CI	<i>P</i> value	Regression slope	95% CI	<i>P</i> value	
Full-Scale IQ	.09	-.09, .28	.31	-.38	-.66, -.11	.006	.02
Verbal IQ	.14	-.07, .34	.19	-.34	-.64, -.04	.03	.04
Performance IQ	.04	-.14, .22	.68	-.36	-.62, -.09	.009	.06
Average Achievement	.01	-.17, .20	.90	-.17	-.45, .11	.23	.39
Grooved Pegboard	.00	-.001, .001	.99	.002	.00, .005	.04	.17
Mayo Test for Apraxia	.09	-.06, .23	.24	-.47	-.68, -.25	<.001	.001

*All models were based on piecewise linear regression, with adjustment for diagnosis and social class.

Grooved Pegboard results are based on the natural logarithm transformation, with higher scores representing longer (worse) times to completion. For all other outcomes, lower scores represent worse outcomes.

Our analyses should be interpreted in light of several limitations. Our modeling of the effects of duration of total CA in this trial, including delineation of cut-points beyond which neurodevelopmental scores declined, is specific to the era in which surgery was conducted at our institution,³ when our cardiopulmonary bypass techniques included hemodilution to a hematocrit value of 20%, the alpha-stat strategy of acid-base management,¹⁹ absence of arterial filtration, and hardware that would be considered outmoded today. Postoperative events and length of stay can also affect later neurodevelopmental outcome. In the era of our study, postoperative care used techniques including the routine use of high-dose fentanyl and prolonged neuromuscular blockade, which are used less frequently currently. Thus, as a better understand-

ing is achieved of the contributions of these other factors to neurodevelopmental outcomes, we might learn that durations of CA longer than those suggested in our models can be tolerated. All children in our study had some period of CA; thus, we cannot assess the benefit of a strategy that completely circumvents its use. Furthermore, we studied a very homogeneous group of children with D-TGA from a single institution. Because diagnosis can affect both cognitive outcome and surgical conduct, the results of these analyses may not be generalizable to a population with mixed diagnoses or age at surgery beyond infancy. Indeed, the technique of deep hypothermic CA is now used with much greater frequency than it was 15 years ago by surgeons repairing adults with acquired heart disease such as ascending and arch aortic aneurysms; the effects of CA

should be studied further in older age groups and other diagnostic groups.

In summary, we found that the effect of duration of total CA on later neurodevelopmental outcomes is nonlinear, with little influence at shorter durations and with steadily worsening outcomes after longer durations of CA. This study cannot ascertain a universally "safe" duration of total CA because outcomes may vary according to diagnosis, age at surgery, and other aspects of perfusion on cardiopulmonary bypass and postoperative care. In particular, the relative contributions of genetic polymorphisms and mutations, preoperative health, other operative factors, and postoperative events to late neurodevelopmental status remain to be clarified.

We are indebted to Ludmila Kyn for database programming; to the study nurses (Kristin C. Lucius, RN, MS, Amy Z. Walsh, RN, BSN, Jodi Bartlett, RN, and Ellen McGrath, RN) for data collection assistance; to Kathleen M. Alexander for project coordination; to Donna M. Duva and Donna M. Donati for scheduling evaluations, arranging patient travel, and for data management assistance; and to Lisa-Jean Buckley for data management assistance.

References

1. Kirkham FJ. Recognition and prevention of neurological complications in pediatric cardiac surgery. *Pediatr Cardiol*. 1998;19:331-45.
2. Kirklin JW, Barratt-Boyes BG. Cardiac surgery: morphology, diagnostic criteria, natural history, techniques, results, and indications. Vol 1, 2nd ed. New York: Churchill Livingstone; 1993.
3. Newburger JW, Jonas RA, Wernovsky G, et al. A comparison of the perioperative neurologic effects of hypothermic circulatory arrest versus low-flow cardiopulmonary bypass in infant heart surgery. *N Engl J Med*. 1993;329:1057-64.
4. Bellinger DC, Jonas RA, Rappaport LA, et al. Developmental and neurologic status of children after heart surgery with hypothermic circulatory arrest or low-flow cardiopulmonary bypass. *N Engl J Med*. 1995;332:549-55.
5. Bellinger D, Wypij D, Kuban KC, et al. Developmental and neurologic status of children at 4 years of age after heart surgery with hypothermic circulatory arrest or low-flow cardiopulmonary bypass. *Circulation*. 1999;100:526-32.
6. Bellinger D, Wypij D, duPlessis AJ, et al. Neurodevelopmental status at eight years in children with dextro-transposition of the great arteries: The Boston Circulatory Arrest Study. *J Thorac Cardiovasc Surg*. 2003;126:.
7. Wechsler D. Wechsler Intelligence Scale for children. 3rd ed. San Antonio (Tex): The Psychological Corporation; 1991.
8. The Psychological Corporation. Wechsler Individual Achievement Test manual. San Antonio (Tex): Harcourt, Brace & Company; 1992.
9. Reitan RM, Davison LA. Clinical neuropsychology: current status and applications. New York: Winston/Wiley; 1974.
10. Darley FL, Aronson AW, Brown J. Motor speech disorders. Boston: Little Brown; 1975.
11. Hastie TJ, Tibshirani KJ. Generalized additive models. London: Chapman and Hall; 1990.
12. Statistical Sciences. S-PLUS guide to statistical and mathematical analysis, version 3.2. Seattle: StatSci, a division of MathSoft, Inc; 1993.
13. Casella G, Berger RL. Statistical inference. 2nd ed. Belmont (Calif): Duxbury Press; 2002.
14. Oates RK, Simpson JM, Turnbull JAB, Cartmill TB. The relationship between intelligence and duration of circulatory arrest with deep hypothermia. *J Thorac Cardiovasc Surg*. 1995;110:786-92.
15. Kern FH, Hinton VJ, Nereo NE, Hayes CJ, Gersony WM. Early developmental outcome after the Norwood procedure for hypoplastic left heart syndrome. *Pediatrics*. 1998;102:1148-52.
16. Hovels-Gurich HH, Seghaye MC, Dabritz S, Messmer BJ, von Bernuth G. Cognitive and motor development in preschool and school-aged children after neonatal arterial switch operation. *J Thorac Cardiovasc Surg*. 1997;114:578-85.
17. Ellerbeck KA, Smith ML, Holden EW, et al. Neurodevelopmental outcomes in children surviving D-transposition of the great arteries. *J Dev Behav Pediatr*. 1998;19:335-41.
18. Sakamoto T, Hatsuoka S, Stock UA, et al. Prediction of safe duration of hypothermic circulatory arrest by near-infrared spectroscopy. *J Thorac Cardiovasc Surg*. 2001;122:339-50.
19. Sakamoto T, Zurakowski D, Duebener LF, et al. Combination of alpha-stat strategy and hemodilution exacerbates neurologic injury in a survival piglet model with deep hypothermic circulatory arrest. *Ann Thorac Surg*. 2002;73:180-90.