Developmental and Neurological Status of Children at 4 Years of Age After Heart Surgery With Hypothermic Circulatory Arrest or Low-Flow Cardiopulmonary Bypass

David C. Bellinger, PhD, MSc; David Wypij, PhD; Karl C. K. Kuban, MD, MSc; Leonard A. Rappaport, MD; Paul R. Hickey, MD; Gil Wernovsky, MD; Richard A. Jonas, MD; Jane W. Newburger, MD, MPH

Background—It is not known whether developmental and neurological outcomes in the preschool period differ depending on whether the predominant vital organ support strategy used in infant heart surgery was total circulatory arrest (CA) or low-flow cardiopulmonary bypass.

Methods and Results—Infants with D-transposition of the great arteries who underwent an arterial-switch operation were randomly assigned to a support method consisting predominantly of CA or low-flow cardiopulmonary bypass. Developmental and neurological status were evaluated blindly at 4 years of age in 158 of 163 eligible children (97%). Neither IQ scores nor overall neurological status were significantly associated with either treatment group or duration of CA. The CA group scored lower on tests of motor function (gross motor, \( P = 0.01 \); fine motor, \( P = 0.03 \)) and had more severe speech abnormalities (oromotor apraxia, \( P = 0.007 \)). Seizures in the perioperative period, detected either clinically or by continuous electroencephalographic monitoring, were associated with lower mean IQ scores (12.6 and 7.7 points, respectively) and increased risk of neurological abnormalities (odds ratios, 8.4 and 5.6, respectively). The performance of the full cohort was below expectations in several domains, including IQ, expressive language, visual-motor integration, motor function, and oromotor control.

Conclusions—Use of CA to support vital organs during open heart surgery in infancy is associated, at the age of 4 years, with worse motor coordination and planning but not with lower IQ or with worse overall neurological status. 

(Circulation. 1999;100:526-532.)

Key Words: heart defects, congenital ■ thoracic surgery ■ child development ■ brain

Although the surgical morbidity of infants who must undergo cardiac surgery has declined, follow-up studies have identified developmental and neurological abnormalities in as many as 25% of survivors. An important source of morbidity may be operative events, particularly the support techniques used to protect vital organs during cardiac repair. The 2 major techniques of vital organ support are deep hypothermia with either total circulatory arrest (CA) or continuous low-flow cardiopulmonary bypass (LFB). Each is associated with a risk of central nervous system injury. With CA, the risks are primarily related to cerebral hypoxic-ischemic/reperfusion injury, and with LFB, to embolic complications associated with the increased time of extracorporeal circulation.

Previously, we reported early results of the first randomized clinical trial comparing the incidence of brain injury after corrective infant heart surgery using deep hypothermia predominantly with CA or LFB. Children assigned to CA had longer electroencephalographic (EEG) recovery times, greater creatine kinase BB isoenzyme release, and higher prevalences of clinical seizures and ictal activity on continuous EEG monitoring in the early postoperative period. They also had a higher prevalence of neurological abnormalities and poorer motor function at 1 year of age and poorer expressive language and motor development, by parental report, at 2.5 years of age.

Limited inferences can be drawn from these findings about the long-term relative safety of the 2 support methods because of the low predictive validity of infant neurological examinations and developmental tests. Therefore, additional evaluations were conducted to determine whether developmental and neurological differences between treatment groups were still detectable when the children reached preschool age and, if so, to characterize their nature and severity more precisely than was possible using the earlier assessments.
Methods

Patients were enrolled in a prospective randomized single-center trial between April of 1988 and February of 1992. Eligibility criteria included a diagnosis of D-transposition of the great arteries (D-TGA) with intact ventricular septum (IVS) or ventricular septal defect (VSD), scheduled repair by 3 months of age, and coronary-artery anatomy suitable for the arterial-switch operation. Exclusion criteria included a birthweight <2.5 kg, a recognizable syndrome of congenital anomalies, an extracardiac anomaly of greater than minor severity, previous cardiac surgery, or cardiovascular anomalies requiring aortic arch reconstruction or additional open procedures. Children were randomized to receive the arterial-switch operation using deep hypothermia and either total CA or continuous LFB as the predominant method of vital organ support. Perfusion methods applied to children in both treatment groups included surface cooling with ice packs to the head, core cooling to a rectal temperature of 18°C, use of the α-stat strategy of pH management, treatment with methylprednisolone (30 mg/kg) at the beginning of cardiopulmonary bypass and thiopental (10 mg/kg) at 10 minutes after initiation of bypass, and hemodilution to a hematocrit of 20. We did not use ultrafiltration or planned reperfusion during the period of CA. Additional information about operative management is available elsewhere.6,8,9

The arterial switch operation was performed in 171 infants, of whom 167 were alive at 4 years of age. Four patients lived outside the United States and were not contacted. Of the 163 eligible patients, 158 (97%) participated. Five families refused (3%). Parents were invited to return to Boston to have their children evaluated at the age of 4 years by a psychologist, pediatric neurologist, speech pathologist, and audiologist. The psychologist traveled to 5 children whose families were unable to return to Boston. For these children, only evaluations of general intelligence, motor function, and language function were completed. All examiners were blinded to treatment assignment and clinical course. This study was approved by the Institutional Review Board and conducted in accordance with institutional guidelines. Parents of all children provided informed consent.

Developmental Testing

General intelligence was assessed using the Wechsler Preschool and Primary Scale of Intelligence–Revised.10 Motor function was assessed using the Peabody Developmental Motor Scales11 and the Grooved Pegboard.12

Language was assessed using the Test for Auditory Comprehension of Language,13 the Receptive One-Word Picture Vocabulary Test,14 the Expressive One-Word Picture Vocabulary Test,15 and the Grammatic Closure subtest of the Illinois Test of Psycholinguistic Abilities.16

Neurological Examination

Findings on the neurological examination17,18 were classified as possible, mild (no functional impairment), moderate (functional impairment requiring intervention/therapy), or severe (dependent on assistance). Abnormalities were subclassified as disorders of head shape and growth, neurocognitive abilities, special senses, cranial nerves, motor system, and gait. Children could be classified as having more than 1 type of abnormality.

Speech Evaluation

Speech was assessed using the Oral and Speech Motor Control test.19 Mayo Tests for Apraxia of Speech and Oral Apraxia-Children’s Battery20 (selected items), and the Goldman-Fristoe Test of Articulation.21 The speech pathologist made a summary judgment regarding volitional oral movement abnormalities and apraxia of speech. If either was present, it was classified as mild, moderate, or severe.

Audiological Evaluation

Hearing acuity was assessed by conditioned play audiometry or sound field audiometry. Tympanic membrane compliance was also evaluated. Abnormal hearing was defined as a bilateral increase in threshold of ≥16 dB for frequencies of 1 to 4 kHz.

Statistical Analysis

The primary outcomes were full-scale IQ and status on neurological examination. Other outcomes were considered secondary. Treatment group differences were evaluated by means of intention-to-treat analyses. Secondary analyses evaluated the effect of duration of CA on the outcomes. A child diagnosed as autistic (assigned to LFB) was included in analyses of neurological outcomes but not developmental and speech outcomes, which could not be completed. All comparisons were adjusted for diagnosis (IVS versus VSD). Comparisons of IQ, language, motor, and continuous speech variables were also adjusted for family social class.22

Continuous outcomes were analyzed using linear regression. Paired t tests were used for intrasubject comparisons of scores. Because standard scores on the Peabody Developmental Motor Scales were skewed, analyses were based on raw scores adjusted for age at testing. Time to complete the Grooved Pegboard was analyzed using the Cox proportional hazards model. This task was stopped after 180 s if a child had not finished. Fisher’s exact tests and logistic regression were used to analyze binary variables, and exact trend tests were used for ordered categorical variables.

We had expected to follow-up approximately 148 patients, providing 86% power to detect a difference of half a standard deviation in full-scale IQ and 88% power to detect a difference of 25% in the prevalence of possible or definite neurological abnormalities.

Results

Within each diagnostic group, patients randomized to CA and LFB were comparable with respect to preoperative and sociodemographic characteristics (Table 1). Families were predominantly middle-class, with parental IQ scores23 in the average range. Children’s mean height and weight were within normal limits and comparable in the treatment groups. No child was reported to have significant activity limitations. Two children had undergone additional heart surgery since the arterial-switch operation. One child was currently on a cardiac-related medication (digoxin).

Developmental Evaluation

General Intelligence

In the full cohort, full-scale, verbal, and performance IQ scores (mean±SD) were 92.6±14.7, 95.1±15.0, and 91.6±14.5, respectively. All 3 scores were significantly lower than the population mean of 100 (P<0.001). Performance IQ was significantly lower than verbal IQ (P<0.001). The subtests on which scores tended to be lowest were those that assessed visual-spatial and visual-motor integration skills. On all subtests but 2 (Comprehension, Sentences), mean scores were significantly lower than those in the population.

Treatment group differences were not significant for full-scale, verbal, or performance IQ (Table 2) or for any subtest. Results were similar when duration of CA replaced treatment group in the regression model. Social class accounted for more of the variation in IQ (24%) than did treatment assignment and diagnosis (3%).

Motor Function

In the full cohort, mean raw gross and fine motor scores on the Peabody Developmental Motor Scales corresponded to the 9th and 4th percentiles for age, respectively. Assignment to CA was associated with significantly lower gross motor
(P=0.01) and fine motor (P=0.03) scores (Table 2). Duration of CA was inversely associated with gross motor (P=0.06) but not fine motor score (P=0.23). The CA group scored lower on 3 subtests, Balance (P=0.05), Nonlocomotor Ability (P=0.008), and Manual Dexterity (P=0.05), and took longer to complete the Grooved Pegboard (P=0.006).

### Language Function
On all tests but Grammatic Closure, scores in the full cohort were significantly below population means. Receptive One-Word Picture Vocabulary Test scores (97.1±15.6) were significantly higher than Expressive One-Word Picture Vocabulary Test scores (92.5±15.7) (P<0.001). Treatment group differences were not significant for any language test (Table 2).

### Neurological Evaluation
Forty-two children (28%) had possible neurological abnormalities and 45 (30%) had definite abnormalities (Table 3). Most definite abnormalities (87%) were considered mild. Abnormalities were more common in the CA group, but the difference did not reach statistical significance (P=0.19).

Most abnormalities involved neurocognitive functions (eg, language, attention) or motor functions (eg, balance, hopping). Cranial nerve abnormalities were noted more often in the CA group (12% versus 1%; P=0.009). Among the 10 children with such abnormalities, 5 had abnormal phonation (articulation), 7 had asymmetric facial movements produced either spontaneously or in response to a specific command, and 2 had dysconjugate eye movements due to strabismus.

### Speech Evaluation
Assignment to CA was associated with reduced ability to imitate oral movements and speech sounds (Total Functional Score; P<0.001) (Table 4). Similar but nonsignificant treatment group differences were noted on the Mayo Test (P=0.10), which also assessed the ability to perform specific oral movements. The severity of abnormalities of volitional oral movements (eg, responses to commands such as “Stick out your tongue”) was greater among children assigned to CA (P=0.02). This group also made more articulation errors (Goldman-Fristoe Test of Articulation; P=0.002) and performed less well on polysyllabic repetitions in terms of rate...
and duration (P = 0.03). Treatment groups did not differ on monosyllabic repetitions (P = 0.28). Apraxia of speech was both more prevalent among children assigned to CA than to LFB (33% versus 18%, respectively; P = 0.03) and more severe (P = 0.007). The risk of apraxia increased with the duration of CA (odds ratio, 1.8 for an increase of 30 minutes; 95% CI, 1.01 to 3.2; P = 0.045).

The poorer speech outcomes in the CA group were not attributable to a higher prevalence of abnormalities in the structures used in sound and speech production (Total Structural Score; P = 0.38). In all groups, the mean score was close to the maximum possible (24). The prevalence of abnormal hearing was also similar in the 2 groups (12% in CA versus 8% in LFB; P = 0.43). All cases of bilateral hearing loss were conductive. One child (assigned to CA) had a profound unilateral sensorineural hearing loss.

### Other Predictors
Presence of a VSD was an independent risk factor for lower IQ scores (full-scale IQ: mean difference, 5.4 points; P = 0.03; verbal IQ: mean difference, 4.6 points, P = 0.07; performance IQ: mean difference, 5.5 points, P = 0.03). VSD was also associated with apraxia of speech (odds ratio, 2.8; 95% CI, 1.2 to 6.7; P = 0.02). Clinical seizures postoperatively were associated with lower IQ scores and with increased risk of possible or definite neurological abnormalities (Table 5). EEG seizures postoperatively also increased a child’s risks of these outcomes. Abnormal hearing was associated with deficits of approximately 8 points on each IQ scale (P < 0.05). Although some preoperative variables (eg, Apgar scores, acidosis) were significantly associated with 1 or more outcomes, they were not associated with treatment group assignment and did not confound treatment group effects.

### Discussion
We found that, at 4 years of age, children assigned to CA and LFB groups did not differ significantly with respect to the primary outcomes, IQ and overall status on neurological examination. The CA group did, however, perform significantly worse on assessments of gross motor, fine motor, and speech functions. Specifically, these children demonstrated higher frequencies of oromotor and facial movement abnormalities, developmental immaturities of hand use and gait,
and abnormalities of speech production. Cranial abnormalities, the subtype of neurological abnormalities that did differ in frequency between treatment groups, usually involved oromotor function (phonation, use of facial muscles). Use of CA thus resulted in children with diminished coordination and motor planning abilities, but cognitive skills were similar to those of children whose surgery was performed using LFB. These findings are consistent with and extend the findings of our previous evaluations of this study cohort,4–6 providing more detailed characterizations of the treatment group differences in outcome. The motor and speech deficits of the CA group will likely predict later academic and behavior problems.24–27 We speculate that higher-order neuronal and associative white matter pathways were affected by perioperative events, as suggested by the fact that cranial magnetic resonance imaging studies in these children at the age of 1 year

### TABLE 4. Speech Production Outcomes According to Ventricular Septal Status and Treatment Group

<table>
<thead>
<tr>
<th>Variable</th>
<th>Intact Ventricular Septum</th>
<th>Ventricular Septal Defect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CA (n=56)</td>
<td>LFB (n=59)</td>
</tr>
<tr>
<td>Oral and speech motor control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total functional score</td>
<td>100.8±10.1</td>
<td>104.9±5.7</td>
</tr>
<tr>
<td>Total structural score</td>
<td>23.6±0.8</td>
<td>23.6±0.7</td>
</tr>
<tr>
<td>Mayo test for apraxia</td>
<td>158.4±21.2</td>
<td>164.7±18.4</td>
</tr>
<tr>
<td>Goldman-Fristoe test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. errors</td>
<td>19.2±13.0</td>
<td>14.6±9.9</td>
</tr>
<tr>
<td>Rate and duration of repetitions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monosyllabic</td>
<td>9.9±3.7</td>
<td>10.6±3.6</td>
</tr>
<tr>
<td>Polysyllabic</td>
<td>2.9±1.1</td>
<td>3.4±1.3</td>
</tr>
<tr>
<td>Abnormalities of volitional oral movements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>15/49 (31)</td>
<td>19/56 (34)</td>
</tr>
<tr>
<td>Moderate</td>
<td>6/49 (12)</td>
<td>3/56 (5)</td>
</tr>
<tr>
<td>Severe</td>
<td>3/49 (6)</td>
<td>0/56</td>
</tr>
<tr>
<td>Apraxia of speech</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>8/54 (15)</td>
<td>6/59 (10)</td>
</tr>
<tr>
<td>Moderate</td>
<td>4/54 (7)</td>
<td>1/59 (2)</td>
</tr>
<tr>
<td>Severe</td>
<td>3/54 (6)</td>
<td>0/59</td>
</tr>
</tbody>
</table>

*Mean deficits, 95% confidence intervals, and P values are for effects of treatment group, with adjustment for diagnosis and social class.
†Exact P value for trend is for difference between treatment groups, with adjustment for diagnosis.

Values are mean ±SD or No. with abnormalities (%).

### TABLE 5. Associations Between Seizure Status and Developmental, Neurological, and Speech Outcomes

<table>
<thead>
<tr>
<th>Continuous Outcomes</th>
<th>Clinical Seizures</th>
<th>EEG Seizures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Deficit*</td>
<td>Confidence Interval P</td>
</tr>
<tr>
<td>Full-Scale IQ</td>
<td>12.6 (3.8,21.4)</td>
<td>0.005</td>
</tr>
<tr>
<td>Verbal IQ</td>
<td>13.4 (4.1,22.6)</td>
<td>0.005</td>
</tr>
<tr>
<td>Performance IQ</td>
<td>7.8 (−0.9,16.5)</td>
<td>0.08</td>
</tr>
<tr>
<td>Gross motor scale</td>
<td>8.6 (−5.4,22.7)</td>
<td>0.22</td>
</tr>
<tr>
<td>Fine motor scale</td>
<td>2.1 (−6.1,10.3)</td>
<td>0.62</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dichotomous Outcomes</th>
<th>Odds Ratio†</th>
<th>Confidence Interval P</th>
<th>Odds Ratio</th>
<th>Confidence Interval P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Possible or definite neurological abnormalities</td>
<td>8.4 (1.0,71.5)</td>
<td>0.05</td>
<td>5.6 (1.7,18.8)</td>
<td>0.005</td>
</tr>
<tr>
<td>Apraxia of speech</td>
<td>5.0 (0.7,35.0)</td>
<td>0.10</td>
<td>3.3 (0.9,12.2)</td>
<td>0.07</td>
</tr>
</tbody>
</table>

*Mean deficits, 95% confidence intervals, and P values are for effects of seizures, with adjustment for treatment group, diagnosis, and social class. Adjustment was also made for child’s age at testing in analysis of Peabody Developmental Motor Scales.
†Odds ratios, 95% confidence intervals, and P values are for effects of seizures, with adjustment for treatment group and diagnosis.
identified few recognizable specific vascular bed infarctions and no association between such infarctions and neurological deficits.5

Previous studies on the neurodevelopmental impact of CA have yielded conflicting findings. Some have reported deficits in intelligence related to duration of CA,28–33 although others have not.34–36 Methodological limitations of published studies include small sample size, diverse cardiac defects and ages at repair, retrospective study design, comparison of techniques used at different time periods, and lack of uniformity in the age at follow-up. In most studies, the sampling frame is unspecified and the extent of ascertainment bias uncharacterized. In our prospective randomized trial, treatment groups were homogeneous in all measured respects and nearly all eligible children were evaluated at 4 years.

Although this study was designed as a randomized trial and did not include an untreated control group such as siblings, the levels of performance of the cohort as a whole were generally significantly below population norms. Weaknesses were noted in several domains, including IQ, expressive language, visual-motor integration, motor planning and organization, and oromotor control. Additional analyses of a subgroup of the cohort revealed impairments in important preliteracy skills.37 Both strategies of vital organ support involve the use of cardiopulmonary bypass and, thus, risk of central nervous system damage from microemboli, macroemboli, and hypoperfusion. Preoperative cyanosis or hemodynamic abnormalities in the preoperative or postoperative period may have contributed to the cohort’s reduced performance. We cannot exclude the possibility that D-TGA is associated with neurodevelopmental abnormalities independent of operative or perioperative events, although the prevalence of known genetic abnormalities with central nervous system involvement, such as chromosome 22q11 microdeletion, is exceedingly low among children with D-TGA.38 Furthermore, children with recognized syndromes of congenital anomalies were excluded from our sample.

The enrollment period for this trial extended from 1988 to 1992. Our findings are thus specific to the intraoperative protocols used at our institution during this era. Nevertheless, most aspects of these protocols continue to be widely used at centers with expertise in infant heart surgery. Perfusion methods in current use, including ultrafiltration, and future novel neuroprotective strategies may improve developmental and neurological outcomes of children who must undergo periods of CA.

The outcomes of patients with an associated diagnosis of VSD were generally worse than those of patients with one of IVS, despite their more optimal preoperative status, including higher oxygen saturation and lower proportion requiring preoperative intubation.4 Among the factors that may account for the poorer outcomes among these children, relative to children without a VSD, are their generally longer total support time, their slightly older age at surgery, and as-yet unidentified genetic differences. The poorer outcomes among these children were already evident in the early postoperative period, when they were significantly more likely to have clinical and EEG seizures.4

In summary, these data suggest that, compared with LFB, a predominant CA strategy or a longer duration of total CA used with deep hypothermia to support vital organs during open heart surgery in infancy is associated with worse motor coordination and planning but not with significantly lower IQ or worse overall neurological status at the age of 4 years. In the cohort as a whole, cognitive, language, and motor performance were significantly reduced relative to the general population. In the future, improved strategies for neuroprotection should be developed for children with critical congenital heart disease who require open heart surgery in infancy.

Acknowledgements
Supported, in part, by NIH grants HL41786, RR02172, and P30-HD18655.

References
22. Hollingshead A. Four Factor Index of Social Status. New Haven, Conn: Department of Sociology, Yale University; 1975.
Developmental and Neurological Status of Children at 4 Years of Age After Heart Surgery With Hypothermic Circulatory Arrest or Low-Flow Cardiopulmonary Bypass

Circulation. 1999;100:526-532
doi: 10.1161/01.CIR.100.5.526

Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 1999 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/100/5/526

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org//subscriptions/