Surgery for Congenital Heart Disease

Developmental Outcome After Surgical Versus Interventional Closure of Secundum Atrial Septal Defect in Children

Karen J. Visconti, MA; David P. Bichell, MD; Richard A. Jonas, MD; Jane W. Newburger, MD, MPH; David C. Bellinger, PhD, MSc

Background—The assessment of the impact of cardiopulmonary bypass (CPB) on developmental outcomes in children who undergo open heart surgery is hampered by the absence of a suitable comparison group. The development of interventional catheterization techniques for the repair of certain types of congenital heart lesions provides the opportunity to study children who have not been exposed to CPB.

Methods and Results—We performed standardized neuropsychological testing on children after closure of a secundum atrial septal defect through the use of surgery (n=26) or a transcatheter device (n=19). Device patients, compared with surgical patients, were similar in age at defect closure (mean, 6 years) but older at follow-up testing (12.3 versus 10.6 years). The mean weight percentile at closure was greater and the defect size was smaller in the device patients. Families of device patients tended to have a higher parent IQ, higher level of maternal education, and higher level of maternal occupation. In general, however, children’s IQ and achievement scores were in the normal range for both groups. In regression analyses with adjustment for age at testing and parent IQ, surgical repair was associated with a 9.5-point deficit in Full-Scale IQ (P=0.03) and a 9.7-point deficit in Performance IQ (P=0.05). Block Design was the IQ subtest on which treatment groups differed the most (P=0.01). Surgical patients achieved significantly better scores on errors of commission (P=0.05) and attentiveness index (P=0.03) on a continuous performance test of attention. Scores on tests of achievement and other neuropsychological domains did not differ significantly between the groups. Regression analyses within the surgical group failed to identify significant CPB-related risk factors.

Conclusions—A prospective randomized trial or a study that includes prerepair and postrepair assessments is necessary to establish whether the observed advantages of device closure in neuropsychological outcome represent deleterious effects of CPB or a methodological artifact. (Circulation. 1999;100[suppl II]:II-145–II-150.)

Key Words: heart diseases ■ heart defects, congenital ■ cardiopulmonary bypass ■ pediatrics

Recent studies from our group indicate that neurological development may be adversely affected in children with transposition of the great arteries who undergo relatively long periods of deep hypothermic cardiopulmonary bypass (CPB) with or without total circulatory arrest.1-2 A “safe” duration of CPB has not been established. One strategy for addressing this issue would be to evaluate the late development of a group of children with a congenital heart defect whose repair requires only a relatively brief period of mildly hypothermic bypass. One such group is children with secundum atrial septal defect (ASD2). In contrast to children with transposition of the great arteries, these children are rarely symptomatic or hypoxic.

Efforts to determine the developmental impact of CPB on the late development of children have been hampered by several methodological challenges. One of the most critical challenges involves the identification of an appropriate control group of children who are comparable to a patient sample in all important respects except that they did not experience a period of CPB. Various approaches have been used, including the use of healthy children from the community or siblings, children who undergo surgery for noncardiac disease, and children with cardiac disease who undergo a closed surgical procedure.

Siblings or other healthy controls may not be suitable because any deficits noted among the children who experienced CPB may be due to their underlying disease or to the emotional and physical trauma of surgery and hospitalization rather than specifically to the fact of having experienced CPB. Although the study of noncardiovascular operations might control for the possible developmental impact of hospitalization and surgery, children with noncardiac surgical diagnoses do not have central nervous system risk factors that may be specific to children with certain types of congenital heart lesions. Children with congenital heart disease repaired...
by means of closed surgical procedures would be useful as controls only if the disease could be repaired with equal facility with a closed procedure or an open procedure without the use of CPB.

Recent advances in interventional cardiology have made available a new patient cohort who may afford better control for these factors and thus improve the assessment of the developmental effects of CPB. This cohort consists of children whose congenital cardiac malformation is repaired by means of transcatheter closure, without the need for CPB. 3, 4 We report the results of a retrospective study in which we compared children whose ASD2 was closed with the use of either open heart surgery or a transcatheter device. We hypothesized that because of the short duration of CPB required for surgical closure of an ASD2, the surgical and device groups would not differ in terms of developmental outcome.

Methods

Patient Selection

The sampling frame consisted of children who underwent transcather closure of an ASD2 at Children's Hospital (Boston, Mass) between January 1989 and December 1991 and who (1) were born after 1980, (2) resided in New England or eastern New York State, and (3) had no associated anomalies. Thirty-three children met these criteria. Each child was matched by age at repair, length of time since repair, and to a child whose ASD2 was repaired surgically. When initial attempts to locate and recruit these 66 families yielded only 11 and 17 families in the device and surgical groups, respectively, we modified our residence criterion and sought to recruit families who had relocated from the New England/New York State area. In addition, we considered for inclusion patients whose operations were conducted between 1989 and 1996. The final sample consisted of 19 children whose ASD2s were repaired by means of a catheter-delivered device and 26 children whose ASD2s were repaired surgically, representing participation rates of 46% and 43%, respectively (Table 1). All 26 surgical patients, compared with 11 (58%) of the device patients, resided in New England or eastern New York State.

This study was approved by the Children's Hospital Committee on Clinical Investigation and was conducted in accordance with institutional guidelines. Informed consent was obtained from the parents of all the children.

Medical Record Review

Information was obtained from a child's medical record about the size of the ASD2 as determined with preoperative echocardiography; height, weight, and body surface area at the time of repair; duration of the repair; duration of anesthesia; time on mechanical ventilation; and days of hospitalization after the procedure. For surgical patients, the following information about the period on CPB was abstracted from perfusion records: minimum temperature, minimum pump flow rate, mean arterial pressure, minimum hematocrit, minimum arterial P O 2 , and minimum and maximum pH.

Developmental Assessment

Patients were administered a comprehensive 4-hour battery of tests. The primary outcomes were general intelligence, assessed with use of the Wechsler Intelligence Scale for Children–III (WISC-III), 5 and academic achievement, assessed with use of the Wechsler Individual Achievement Test–Screener. 6 Secondary outcomes were tests of specific neuropsychological domains, including attention (Connors' Continuous Performance Test), 7 memory (Wide Range Assessment of Memory and Learning–Screener), 8 visual-motor integration (Developmental Test of Visual-Motor Integration), 9 visual-spatial skills (Test of Visual-Perceptual Skills), 10 executive function (Trail-Making Test, 11 Wisconsin Card Sorting Test 12 ), and language (Clinical Evaluation of Language Fundamentals–3–Screener). 13

At the time a child was evaluated, a parent completed a questionnaire pertaining to family demographics and child developmental history. Parent IQ was assessed with use of the Kaufman Brief Intelligence Test. 14 Parents completed the Child Behavior Checklist (CBCL), 15 an assessment of child behavior and social competence, and teachers completed the Teacher's Report Form 16 of the CBCL.

For 32 children (71%), the evaluation was conducted at Children’s Hospital. For 13 children (device, n = 9; surgical, n = 4) who were unable to travel to Boston, the evaluation was conducted at an institution near the child's home. All evaluations were conducted by the same psychologist (K.J.V.).

Statistical Analysis

The associations between continuous outcome variables and treatment group were estimated with multiple linear regression analysis, with adjustment for age at testing and parent IQ. Treatment groups were compared with respect to medical variables, family demographic characteristics, and child developmental history with the use of t tests for continuous variables and χ 2 tests for categorical variables. Following the recommendations provided in the manuals, a T score of 64 (90th percentile) was used as a cutpoint for the identification of children with extreme scores on the 3 summary scales of the CBCL and Teacher’s Report Form (ie, total problem behaviors, internalizing behaviors, externalizing behaviors) and a T score of 67 (95th percentile) for the identification of children with extreme scores on the syndrome (ie, narrow-band) scales. All P values are 2-tailed.

Results

Patient Characteristics

Device and surgical groups did not differ significantly in age at repair. The interval between repair and follow-up was longer for the device patients, however, so they were significantly older than the surgical patients at the time of the evaluation (Table 2). The distributions of ASD2 sizes differed significantly between groups, with 10 of 26 surgical patients having a defect size of >2 cm compared with none of the device patients (P < 0.05). Weight percentile at the time of repair was significantly greater among the device patients (P < 0.01), although the groups did not differ in height percentile. The duration of mechanical ventilation and the number of days from the procedure until discharge were also significantly greater for the surgical patients.

Table 3 presents data on perfusion variables for children in the surgical group. Membrane oxygenators were used for all patients in this group. None underwent a period of circulatory arrest, nor was intraoperative transesophageal echocardiography performed.

For all children in the device group, ASD2 closure was achieved with a Clamshell I device (range, 17 to 40 mm). Of the 10 patients for whom fluoroscopy times were available, the mean time was 36 minutes (SD, 14 minutes;
range, 19 to 60 minutes). At the last recorded follow-up visit, a residual shunt was classified as “trivial to absent” for 13 device patients, “small” for 4 device patients, and “more than small” for 2 device patients. For 13 children, the device was known to have at least 1 fractured arm. For 2 children, the device status was unknown.

In general, the families of the 2 patient groups did not differ significantly in sociodemographic characteristics, such as parent education, occupation, and IQ (Table 2), although by most indicators, the families of children in the device group tended to be of higher socioeconomic status. Families were predominantly intact, white, and middle class. Both patient groups consisted predominantly of girls (83% and 88% for device and surgical groups, respectively).

### Developmental Assessment

#### General Intelligence

Mean IQ scores for both groups were in the average range. In the cohort as a whole, mean Full-Scale IQ was 103.8 (SD, 14.5; range, 69 to 137). Mean Verbal IQ was 104.0 (SD, 14.3; range, 74 to 138), and mean Performance IQ was 103.2 (SD, 15.3; range, 68 to 141). The device group scored significantly higher than the surgical group on several of the WISC-III composite scores, including Full-Scale IQ ($P=0.03$), Performance IQ ($P=0.05$), and Perceptual Organization ($P=0.04$) (Table 4). The difference in the adjusted mean Full-Scale IQ scores for the 2 groups was 9.5 points, corresponding to $0.6 SD$. The difference in adjusted mean Performance IQ scores was 9.7 points.

Group differences on the verbal subtests of the WISC-III were modest, with only 1, Vocabulary, reaching the 0.10 level of significance. On the performance or nonverbal subtests, the group difference in Block Design was significant ($P=0.01$), with the mean difference in scaled scores (3.3) $>1$ SD (3 points for WISC-III subtests). On another performance subtest, Object Assembly, the group difference approached significance ($P=0.09$). Thus, the treatment group difference in Full-Scale IQ in large part reflected group differences in visual-motor and visual-spatial skills rather than in language-based skills.

To assess whether residual confounding attributable to the slight advantage of the device group in terms of mean

### TABLE 2. Comparison of Medical and Sociodemographic Variables for Device and Surgical Groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>Device Group (n=19)</th>
<th>Surgical Group (n=26)</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at repair, mo</td>
<td>65.6±22.4</td>
<td>69.7±21.9</td>
<td>0.54</td>
</tr>
<tr>
<td>Age at follow-up, mo</td>
<td>147.5±27.8</td>
<td>127.1±25.5</td>
<td>0.01</td>
</tr>
<tr>
<td>ASD size, % &gt;2 cm</td>
<td>0</td>
<td>38.5</td>
<td>0.01</td>
</tr>
<tr>
<td>Weight at repair, %</td>
<td>60.1±28.0</td>
<td>43.1±28.4</td>
<td>0.05</td>
</tr>
<tr>
<td>Height at repair, %</td>
<td>48.2±33.8</td>
<td>43.8±33.0</td>
<td>0.67</td>
</tr>
<tr>
<td>Body surface area at repair</td>
<td>0.79±0.14</td>
<td>0.78±0.16</td>
<td>0.88</td>
</tr>
<tr>
<td>Duration of repair procedure, min</td>
<td>143.1±49.6</td>
<td>127.1±29.0</td>
<td>0.24</td>
</tr>
<tr>
<td>Duration of anesthesia, min</td>
<td>163.6±54.8</td>
<td>176.2±25.8</td>
<td>0.38</td>
</tr>
<tr>
<td>Duration of mechanical ventilation, min</td>
<td>50.5±52.7</td>
<td>356.4±216.8</td>
<td>0.0001</td>
</tr>
<tr>
<td>Days after repair to discharge</td>
<td>1.1±0.2</td>
<td>3.5±0.7</td>
<td>0.0001</td>
</tr>
<tr>
<td>Sociodemographic characteristic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marital status, % married</td>
<td>89</td>
<td>85</td>
<td>0.46</td>
</tr>
<tr>
<td>Father occupation*</td>
<td>5.8±2.9</td>
<td>5.7±2.2</td>
<td>0.85</td>
</tr>
<tr>
<td>Father education, y</td>
<td>15.2±4.3</td>
<td>14.2±4.4</td>
<td>0.44</td>
</tr>
<tr>
<td>Mother employed, % yes</td>
<td>74</td>
<td>73</td>
<td>0.96</td>
</tr>
<tr>
<td>Mother occupation*</td>
<td>6.4±2.0</td>
<td>5.2±1.6</td>
<td>0.06</td>
</tr>
<tr>
<td>Mother education, y</td>
<td>15.3±2.5</td>
<td>13.8±2.7</td>
<td>0.07</td>
</tr>
<tr>
<td>Ethnicity, % white</td>
<td>89</td>
<td>88</td>
<td>0.84</td>
</tr>
<tr>
<td>Birth order, % first born</td>
<td>68</td>
<td>54</td>
<td>0.41</td>
</tr>
<tr>
<td>Current grade in school</td>
<td>5.6±2.5</td>
<td>4.2±2.0</td>
<td>0.04</td>
</tr>
<tr>
<td>Parent IQ†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite</td>
<td>107.1±8.1</td>
<td>103.5±10.5</td>
<td>0.27</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>106.2±10.3</td>
<td>105.3±10.6</td>
<td>0.80</td>
</tr>
<tr>
<td>Matrices</td>
<td>106.8±7.8</td>
<td>101.6±10.9</td>
<td>0.12</td>
</tr>
</tbody>
</table>

*Score on the Hollingshead Four Factor Index of Social Class, with higher score indicating higher social class.

†Score on the Kaufman-Brief Intelligence Test. Values are mean ± SD.
parental IQ was responsible for the group differences in child IQ, additional multiple regressions were carried out after efforts were made to render the parent IQ distributions more similar in the device and surgical groups. Two approaches were used. In the first approach, each child in the device group was matched to a child in the surgical group on the basis of parental IQ (for 6 children, parent IQ was identical; for 5, there was a ±1-point difference; for 1, there was a ±2-point difference). For 2 device children, 2 matches were available in the surgical group. Separate analyses were conducted to evaluate whether the result differed depending on which of the possible matches was included (match 1 and match 2). In both analyses (n=12 matched pairs), the mean parent IQ scores in the device and surgical groups were 107.4 (SD, 6.1) and 107.2 (SD, 5.7), respectively. In the second approach, the parent IQ distribution for the surgical group was truncated at the lower end to make its mean value comparable to the mean value in the device group (Truncation). In this analysis (n=35), the mean parent IQ scores in the device and surgical groups were 107.1 (SD, 8.1; range, 88 to 119) and 107.0 (SD, 7.6; range, 94 to 127), respectively. In these additional analyses, in which treatment groups were more closely matched in terms of parent IQ, group differences in the children’s Full-Scale IQ scores, with adjustment for age at testing and parent IQ, were generally as large as and, in many instances, larger than those evident in analyses of the full cohort, although the probability value in the subset analyses tended to be less extreme due to the smaller numbers of subjects included (Table 5).

To explore whether the lower scores among the surgical patients were attributable to patients with larger defect sizes, we repeated the main analyses and excluded patients with larger defect sizes of ≥2 cm. The results were unchanged. Within the group of surgical patients, Full-Scale IQ score was not significantly associated with any of the CPB-related variables measured, including duration of CPB, aortic cross-clamping time, minimum pump flow rate, minimum tympanic temperature, minimum or maximum pH, minimum or maximum PCO₂, minimum arterial Po₂, or minimum hematocrit (Table 6). Minimum hematocrit was the variable that bore the strongest relation to Full-Scale IQ, with higher values associated with higher scores (P=0.14).

### Academic Achievement

In the cohort as a whole, mean composite score on the Weschler Individual Achievement Test (Screener) was 103.4 (SD, 16.2; range, 67 to 135). Treatment group differences were not significant for the composite score or for any of the 3 subtest scores (Reading, Mathematical Reasoning, Spelling), although the mean values for the surgical group were lower on each of them (Table 4).

### Neuropsychological Outcomes

Few group differences were found on the secondary neuropsychological outcomes. On the test of sustained attention, Connors’ Continuous Performance Test, the device group, compared with the surgery group, committed significantly
TABLE 5.  Additional Analyses of Treatment Group Differences in Child IQ

<table>
<thead>
<tr>
<th>Variable</th>
<th>Regression Coefficient in Subset Analysis</th>
<th>Match 1*</th>
<th>Match 2</th>
<th>Truncation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full-Scale IQ</td>
<td>-13.4 (5.7)</td>
<td>-9.8 (5.4)</td>
<td>-10.1 (4.4)</td>
<td></td>
</tr>
<tr>
<td>Verbal IQ</td>
<td>-13.4 (5.9)</td>
<td>-6.8 (5.5)</td>
<td>-6.7 (4.6)</td>
<td></td>
</tr>
<tr>
<td>Performance IQ</td>
<td>-13.8 (6.0)</td>
<td>-10.4 (5.9)</td>
<td>-11.0 (4.6)</td>
<td></td>
</tr>
<tr>
<td>Verbal Comprehension</td>
<td>-10.4 (5.5)</td>
<td>-8.6 (5.2)</td>
<td>-7.2 (4.3)</td>
<td></td>
</tr>
<tr>
<td>Perceptual Organization</td>
<td>-13.9 (6.1)</td>
<td>-11.1 (6.0)</td>
<td>-11.5 (4.6)</td>
<td></td>
</tr>
<tr>
<td>Freedom from Distractibility</td>
<td>-7.3 (6.6)</td>
<td>-1.1 (6.0)</td>
<td>-3.3 (5.5)</td>
<td></td>
</tr>
<tr>
<td>Processing Speed</td>
<td>-2.8 (6.3)</td>
<td>1.1 (6.7)</td>
<td>0.2 (5.2)</td>
<td></td>
</tr>
</tbody>
</table>

Methods used to define subsets included in Match 1, Match 2, and Truncation analyses are described in text.

*Estimated difference (SE) between mean scores of the device and surgical groups. Negative sign indicates a lower score in surgical group. These regression coefficients should be compared with those obtained when all patients were included in the analyses (see Table 4).

Discussion

Although mean IQ and academic achievement scores were solidly within the normal range, regardless of whether a child’s ASD2 was repaired surgically or by means of a catheter-delivered device, WISC-III Full-Scale IQ and Performance IQ scores were nevertheless significantly lower in the surgical group than in the device group. The group differences were generally >0.5 SD. Group differences in visual-spatial/visual-motor skills were largely responsible for the differences in WISC-III summary scores. The difference in scores on Block Design represents >1 SD; Block Design is generally viewed as the Wechsler IQ subtest that is the best measure of visual-spatial organization. Block Design scores tend to be vulnerable to many types of brain injury. In previous studies of the neuropsychological sequelae of cardiovascular surgery in children, deficits in visual-motor and visual-spatial functions have been among the most consistent findings. The device group performed significantly worse than the surgery group on a test of sustained attention, committing more errors of commission and achieving a lower score on an index of attentiveness. Consistent treatment-group differences were not found, however, on tests of academic achievement, executive functions, language, or memory.

Few data are available on the neuropsychological outcomes of children with ASD2. Children who undergo surgical repair of an ASD do not undergo circulatory arrest or long periods of CPB and thus are presumed to be less subject to the increased risk of neurological injury associated with these support techniques. Nevertheless, the desire to avoid blood transfusion in these children frequently resulted in a severe degree of dilutional anemia (mean minimum hematocrit, 18.2; range, 13 to 22). This may have limited oxygen delivery during the period of CPB, as indicated by recent studies from our laboratory. Indeed, the perfusion variable most strongly related to IQ was minimum hematocrit. For most of the period during which the patients in this cohort underwent surgery (1989 to 1996), the a-stat pH strategy was used, which may have further restricted cerebral oxygen delivery. Although children with cyanotic heart disease tend to have better outcomes than those with cyanotic heart disease, they appear to perform worse than noncardiac, healthy controls. In 1 of the few studies examining the development of children with cyanotic heart disease, Yang et al reported that children with either an ASD or a ventricular septal defect performed below age-matched nonsurgical controls on tests of intelligence, complex problem solving, visual discrimination, sustained attention, and spatial memory. Compared with controls, they also appeared to have a higher prevalence of parent-reported behavior problems, particularly problems of an externalizing nature (eg, hyperactive, impulsive, or aggressive).
The treatment group differences we found are of surprising magnitude given the brief periods of CPB used during the surgical closure of an ASD2. The children recover rapidly. In current practice, they are routinely extubated in the operating room and are generally discharged within 2 to 3 days of the procedure. Thus, if brain injury does occur intraoperatively, it is clinically silent. It is important, therefore, to identify several limitations of our study and possible alternative explanations for our findings. First, the design was observational. It is possible that preexisting differences between the surgical and catheter groups rather than differences in medical management of the ASD2 are responsible for the group differences in developmental outcome. Even though we adjusted our analyses for parent IQ and conducted additional analyses on subsets of the device and surgical groups more closely matched on parent IQ, we cannot exclude the possibility that the group differences in outcome reflect residual confounding by unmeasured characteristics that distinguish families who chose device closure over surgical closure. The components of the WISC-III on which group differences were greatest, Performance IQ and Block Design, test skills very similar to those assessed with the Kaufman Brief Intelligence Test subtest, Matrices, on which the parents of the 2 groups of children differed most. The most effective strategy for the reduction of the possibility of residual confounding would be to conduct a clinical trial involving random assignment of children with ASDs to the surgical and device groups. If this is not feasible, a design that provides for prerepair and postrepair developmental assessments of a child would also be informative. Serial developmental evaluations during the postrepair period would characterize the natural history of any associated deficits. Based on studies of adults who underwent cardiac surgery, many of the neuropsychological complications that are apparent in the early postoperative period resolve within a few months.25 Similar studies have not been carried out in children.

Second, our sample size was small, which both limited the power of our hypothesis tests and reduced our ability to identify associations between perfusion variables and developmental outcome within the surgery group.

Third, our findings may reflect selection bias insofar as we were able to obtain follow-up information on fewer than 50% of the eligible children in each group. We cannot be sure that within each treatment group, the children who participated were comparable to those who did not. We do not have any information on the outcomes of children who did not participate, precluding an evaluation of this possibility.

Fourth, our battery included many tests and many outcomes. Some of the group differences we found could be the result of chance. On the other hand, the major differences between treatment groups, visual-spatial and visual-motor functioning, are similar to those reported in other studies of patient groups undergoing procedures involving the use of CPB.

In summary, our preliminary data suggest that future prospective studies should be conducted to determine whether the observed advantages of transcatheter closure over surgical repair of an ASD2 represent deleterious effects of CPB or a methodological artifact attributable to residual confounding, chance, patient selection bias, or follow-up bias.

Acknowledgment

We would like to thank the study families and children for their time and effort.

References


17. Hollingshead A. Four Factor Index of Social Status. New Haven, Conn: Department of Sociology, Yale University; 1975.


Developmental Outcome After Surgical Versus Interventional Closure of Secundum Atrial Septal Defect in Children
Karen J. Visconti, David P. Bichell, Richard A. Jonas, Jane W. Newburger and David C. Bellinger

_Circulation_. 1999;100:II-145-II-150
doi: 10.1161/01.CIR.100.suppl_2.II-145

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/suppl_2/II-145