Developmental Progress After Cardiac Surgery in Infancy Using Hypothermia and Circulatory Arrest

PATRICIA M. CLARKSON, M.B., BARTON A. MACARTHUR, PH.D., BRIAN G. BARRATT-BOYES, M.B., CH.M., RALPH M. WHITLOCK, M.B., AND JOHN M. NEUTZE, M.D.

SUMMARY Seventy-two of 76 long-term survivors who had surgical repair of congenital heart lesions at 11 days to 26 months of age using profound hypothermia and circulatory arrest underwent psychometric testing (Stanford-Binet) late postoperatively.

The mean IQ of the 72 patients was 92.9 ± 16.5 (SD). Stanford-Binet scores bore no relationship to the duration of circulatory arrest or other aspects of surgical technique. Scores were significantly lower in those who had a low birth weight for gestational age, important neurologic problems preoperatively or were in the lower socioeconomic classes.

An “ideal” control group of 69 children randomly selected from patients satisfying certain criteria based on birth and neonatal characteristics had a mean IQ of 106.2 ± 11.6. Twenty-five patients who had surgical treatment of congenital heart disease met the criteria for the control group except for their heart lesions. Their mean IQ was 101.4 ± 15.0 (NS).

We could not demonstrate any significant deleterious effect that could be attributed to the surgical methods. Rather, the postoperative IQ scores reflected characteristics related to individual patients.

SURGICAL TECHNIQUES using profound hypothermia and circulatory arrest are being used more often in the management of infants with symptomatic congenital heart disease. This study was undertaken to assess the effect of profound hypothermia and circulatory arrest on intellectual development.

Material and Methods

Patients

Seventy-two of 76 consecutive long-term survivors who had surgical repair of congenital heart lesions at 11 days to 26 months of age using profound hypothermia and circulatory arrest from July 1969 until December 1971 were evaluated late postoperatively. Of the four patients not available for psychometric testing, one was assessed clinically to be developmentally normal at 4 years of age, but three are known to be significantly retarded. Developmental progress was suspect in one preoperatively and another was of low birth weight for gestational age; both had an episode of postoperative cardiopulmonary arrest thought to contribute to poor postoperative progress. The remaining patient, reported elsewhere, was shown postoperatively to have a chromosomal abnormality associated with mental retardation.

Surgical Technique

The surface cooling circulatory arrest technique used has been reported in detail elsewhere. In 67 of the 72 patients, hypothermia was provided chiefly or wholly by surface cooling. In five infants the surface cooling technique was replaced almost entirely by core cooling. In this latter group the duration of cooling bypass was much longer (table I) and the rectal temperature higher (average 22.3°C vs 20.5°C) than in the surface-cooled patients. They did not differ significantly in other respects.

Testing Methods

Pre- and postoperative medical history, examination findings and operative records were reviewed. Measurements of head circumference made usually in the week before operation and postoperatively at the time of psychometric testing were used for the purposes of the study (table I), although additional measurements were examined. Head circumference was considered to be abnormally small if it was more than 2 standard deviations below the mean for age, using the data of Vickers and Stuart. Patients were considered to be of low birth weight for gestational age if birth weight was below the tenth percentile.

Seven patients who were of low birth weight for gestational age were considered to have a preoperative problem, as were three other patients with important neurologic lesions. Two of the three had cerebrovascular accidents, and the other, previously treated hydrocephalus. Less well defined events, such as perinatal hypoxia or seizures, were not categorized as neurologic abnormalities, although they may have had an adverse effect on development. Infants who were hypotonic and had poor motor development associated with heart failure or hypoxia were not regarded as having neurologic abnormalities.

Formal psychometric testing was undertaken at 29–84 months (mean 47.9 months) of age. The Stanford-Binet scale was selected as a measurement of in-
TABLE 1. Clinical Data

<table>
<thead>
<tr>
<th></th>
<th>Mean ± SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at operation (months)</td>
<td>8.3 ± 6.8</td>
<td>0.2-26</td>
</tr>
<tr>
<td>Prearrest CP bypass (min)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface cooled (n = 55)</td>
<td>7.9 ± 4.2</td>
<td>2-18</td>
</tr>
<tr>
<td>Core cooled (n = 5)</td>
<td>27.2 ± 5.9</td>
<td>18-32</td>
</tr>
<tr>
<td>Postarrest CP bypass (min)</td>
<td></td>
<td>11-46</td>
</tr>
<tr>
<td>Total CP bypass (min)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface cooled (n = 67)</td>
<td>28.7 ± 9.0</td>
<td>11-56</td>
</tr>
<tr>
<td>Core cooled (n = 5)</td>
<td>59.0 ± 7.1</td>
<td>47-64</td>
</tr>
<tr>
<td>Temperature circulatory arrest (nasopharyngeal, °C)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surface cooled</td>
<td>22.0 ± 1.4</td>
<td>18.5-24.6</td>
</tr>
<tr>
<td>Core cooled</td>
<td>21.8 ± 0.5</td>
<td>21.4-22.5</td>
</tr>
<tr>
<td>Duration circulatory arrest (min)</td>
<td>48.4 ± 12.4</td>
<td>28-89</td>
</tr>
<tr>
<td>Age (months) postoperative follow-up (months)</td>
<td>47.9 ± 12.3</td>
<td>29-84</td>
</tr>
<tr>
<td>IQ (Stanford-Binet)</td>
<td>92.9 ± 16.5</td>
<td>52-128</td>
</tr>
<tr>
<td>Preoperative head circumference (mean from mean for age and sex)</td>
<td>-1.9 ± 1.3</td>
<td>-5.6 to +1.4</td>
</tr>
<tr>
<td>Postoperative head circumference</td>
<td>-0.8 ± 1.2</td>
<td>-4.0 to +2.1</td>
</tr>
</tbody>
</table>

Intelligence because of its wide use and its particular application to the age range of these children. Psychometric testing of 70 of the 72 children was performed by one experienced person under relatively uniform conditions. Two patients were assessed by other psychologists. All children performed satisfactorily in the test situation. Results of the Stanford-Binet test were also analyzed according to Sattler's classification, and speech and language were assessed by the psychologist at the time of testing.

Socioeconomic class 1-6 was based on a local classification, with the paternal occupation used most often. Class 1 represents the highest socioeconomic group.

Linear regressions were constructed by the least-squares method in an attempt to determine whether variables related to preoperative status and perioperative events were associated with postoperative performance as judged by IQ. Results were considered to show significant correlation at the 5% level. Comparisons between groups were made using the unpaired 1 test. Differences were considered significant at 2p < 0.05.

When this study was undertaken another group of 69 children also had Stanford-Binet tests administered by the same investigator. This group was randomly selected from a sequential group of children born at one hospital who met certain characteristics at birth and in the neonatal period, with a view to providing an "ideal" group for comparison with other infants who had pre- or perinatal problems. It included Caucasian singletons, 38-42 weeks gestation, who weighed 3000-4000 g and who had no evidence of congenital malformation, neonatal illnesses or other adverse features. They were tested as preschoolers at an average age of 57 months.

Results

Relation to Operative Events

The mean IQ of the 72 patients tested was 92.9 ± 16.5 (table 1). Stanford-Binet scores bore no relationship either to the duration of circulatory arrest (fig. 1) or to other aspects of surgical technique, namely the duration of cardiopulmonary bypass, blood pressure during cooling or minimum nasopharyngeal or rectal temperature (fig. 2). The mean score of those who had ventilatory or circulatory support in the postoperative period was similar to the score of those who did not require this assistance. There was no demonstrable relationship between intelligence test results and age at operation, age at assessment, race (Caucasian vs Polynesian) or sex.

*2p is the two-tailed probability of a chance difference between two groups. This is appropriate when there is no a priori reason to expect that any difference would be in a particular direction.
Effect of Pre- and Postoperative Events

The 10 patients who were of low birth weight for gestational age or had important neurologic problems preoperatively had a mean IQ of 76.5 ± 20.0, which was significantly lower (2p < 0.001) than the mean IQ of those in whom these problems were absent (95.5 ± 14.4).

Transient seizures occurred early postoperatively in three patients and later postoperatively in association with an E. coli enteritis in another. No permanent adverse effects were noted in these four patients.

One patient who had a total circulatory arrest time of 89 minutes, interrupted by a period of car-diopulmonary bypass, had a cardiac arrest postoperatively and showed important neurologic abnormalities on discharge. These largely resolved, and IQ at review was 79. Two patients had choreoathetotic movements postoperatively. One had a family history of mental retardation, was of low birth weight for gestational age and had a disproportionately small head circumference before operation. Arrest time was 71 minutes. At the time of postoperative review, head circumference remained abnormally small; there was mild ataxia and the IQ was 65. The other had prolonged seizures in the neonatal period, and developmental retardation was recognized before operation. Arrest time was 67 minutes. At follow-up the patient had mild neurologic abnormalities thought to be due, at least in part, to neonatal events. The IQ was 82.

General

No major differences were found in mean intelligence scores among patients in socioeconomic classes 1–4 (table 2). Patients in socioeconomic classes 5 or 6 had a mean IQ of 85.6 ± 15.3, compared with 97.6 ± 15.4 in classes 1–4 (2p < 0.005). The difference between these groups remained significant (2p < 0.01) when patients with preoperative problems were excluded. The proportion of patients in socioeconomic class 6 (26%) was greater than that of the general population (12%).

There was a significant positive correlation between pre- and postoperative head circumference and intelligence test results. No patient showed a relative decrease in head circumference (expressed as the number of standard deviations from the mean for age and sex) postoperatively compared with preoperatively. Mean head circumference (table 1) increased significantly after operation (2p < 0.001). In 14 patients the increase was 2 standard deviations or more. Head circumference was abnormally small in 32 patients before operation, but in 21 of these it was in the normal range after operation. Intelligence test scores in patients with a small preoperative but normal postoperative head circumference were compared with those in patients whose head circumference was normal on both measurements. Differences were not significant when allowance was made for birth characteristics, the presence of preoperative problems and socioeconomic class. Although IQ was assessed at over 100 in two of the 11 patients whose head circumference remained abnormally small postoperatively, the mean score in this subgroup was abnormally low even allowing for adverse preoperative factors.

Seventy-five percent of the patients had either ventricular septal defect, tetralogy of Fallot (including tetralogy-like lesions) or transposition of the great arteries (table 3). These groups showed no significant difference in mean IQ. Although patients with tetralogy of Fallot tended to have a lower mean IQ, differences between them and patients with transposition of the great arteries or ventricular septal defect were not significant whether or not allowance was made.
made for the presence of preoperative problems.

The results of the Stanford-Binet test were also analyzed according to Sattler’s classification using seven categories (language, memory, conceptual thinking, reasoning, numerical reasoning, visual motor and social intelligence), and profiles constructed for the different diagnostic groups showed no evidence of substantial fall away in performance in any of the categories. As a group, patients score less well than expected when language was examined in more detail and in terms of quality of language and quantity and intelligibility of speech.

Comparison with a Control Group

Twenty-five of the patients who had congenital heart disease met all the criteria of the “ideal” control group except for their congenital heart disease. Psychometric testing was undertaken in this subgroup of patients at an earlier mean age (43 months), but this was not considered to affect the comparison between them and the ideal group. The ideal control group of 69 children had a mean IQ of 106.2 ± 11.6, while the 25 children comparable to them except for their congenital heart disease had a mean IQ of 101.4 ± 15.0. This difference is not statistically significant.

**Discussion**

Cerebral damage may result from several operative events. On the basis of experience with cardiopulmonary bypass techniques, the most common cause would be cerebral air embolism, although particulate emboli can also cause brain damage. Despite meticulous attention to surgical technique, including ligation of the ductus arteriosus when closure had not been demonstrated by preliminary studies and careful control of temperature gradients, the completeness of de-airing of the left ventricle remains conjectural, and air from this source could cause early or late intellectual impairment. Surgical technique is also designed to minimize particulate embolization, but the arterial

---

**Table 2. Socioeconomic Class**

<table>
<thead>
<tr>
<th>Socioeconomic class</th>
<th>All cases</th>
<th>No. preop problems</th>
<th>Ideal with CHD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>IQ</td>
<td>No.</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>90.7 ± 21.4</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>103.5 ± 16.6</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>97.4 ± 11.9</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td>97.8 ± 14.0</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>92.1 ± 7.4</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>19</td>
<td>82.2 ± 17.4</td>
<td>14</td>
</tr>
<tr>
<td>1-4</td>
<td>43</td>
<td>97.6 ± 15.4</td>
<td>38</td>
</tr>
<tr>
<td>5-6</td>
<td>29</td>
<td>85.6 ± 15.3</td>
<td>24</td>
</tr>
<tr>
<td>1-6</td>
<td>72</td>
<td>92.9 ± 16.5</td>
<td>62</td>
</tr>
</tbody>
</table>

*Values are mean ± sd. Abbreviation: CHD = congenital heart disease.*

**Table 3. Type of Cardiac Lesion**

<table>
<thead>
<tr>
<th>All cases</th>
<th>No preop problem</th>
<th>Ideal with CHD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>IQ</td>
</tr>
<tr>
<td>Ventricular septal defect</td>
<td>16</td>
<td>93.5 ± 14.6</td>
</tr>
<tr>
<td>Transposition of great arteries</td>
<td>22</td>
<td>90.6 ± 18.5</td>
</tr>
<tr>
<td>Tetralogy of Fallot</td>
<td>17</td>
<td>88.5 ± 13.9</td>
</tr>
<tr>
<td>Total anomalous PVD</td>
<td>7</td>
<td>101.4 ± 22.1</td>
</tr>
<tr>
<td>Endocardial cushion defect</td>
<td>5</td>
<td>102.4 ± 16.8</td>
</tr>
<tr>
<td>Other</td>
<td>5</td>
<td>95.2 ± 10.3</td>
</tr>
<tr>
<td>Totals</td>
<td>72</td>
<td>92.9 ± 16.5</td>
</tr>
</tbody>
</table>

*Values are mean ± sd. Abbreviations: CHD = congenital heart disease; PVD = pulmonary vascular disease.*
filter used in the study group was less effective than those now available.

Apart from the risk of air or particulate emboli, these infants were vulnerable to cerebral damage from the circulatory arrest per se, as the safe arrest period for any given temperature is uncertain. In the early series of infants reported here, we based arrest time on that recommended by Hikasa et al.\textsuperscript{13} They suggested that for a 60-minute arrest period, a rectal temperature of 20°C was appropriate. In fact, we preferred nasopharyngeal temperature to rectal in the belief that it more accurately reflects brain temperature. We were surprised that the mean nasopharyngeal arrest temperature in this group of patients was 22°C. This was in part caused by the selection of a temperature of 22–24°C in most patients where the surgeon determined that the defect could be repaired within a relatively short arrest time (e.g., ventricular septal or atrial septal defect). The relatively poor correlation between arrest temperature and arrest time (fig. 2) indicates that this practice is not wise and for some time now we have stabilized the prearrest nasopharyngeal temperature at 18°C in almost every infant. This has been made easier by the use of a circulating water bed for surface cooling,\textsuperscript{3} which is far more efficient than the conventional cooling blanket used in the study group. The series reported here is thus not comparable in this respect to current practice.

The study group included 13 infants in whom the circulatory arrest time was longer than 60 minutes, and in three it exceeded 70 minutes. While the intelligence test results in patients without preoperative problems do not show any statistically significant difference from those in patients arrested for longer than 60 minutes (mean IQ 91.1 ± 16.9 compared with those arrested for a shorter time (mean IQ 96.4, ± 13.9), the data do show a trend toward a lower mean score in the former group (figs. 2 and 3). Sixty minutes should probably be regarded as the maximum arrest time for a nasopharyngeal temperature of 20°C. With a nasopharyngeal temperature of 16–18°C, current results suggest that an arrest time of as long as 70 minutes is safe. Should it appear likely that the arrest time is going to exceed these limits, the operation must be planned to allow a period of reperfusion at a blood temperature of 18°C for at least 15 minutes. The possibility that the circulatory arrest time may exceed the safe limits in more complex repairs is the strongest argument in favor of using intermittent, low-flow perfusion at approximately 20°C, punctuated with short periods of circulatory arrest only when absolutely necessary, rather than total circulatory arrest.\textsuperscript{14} It is not known how much protection such a technique provides the brain, as there may be ineffective capillary perfusion with very low flows at very low temperatures. No data are available concerning the incidence of cerebral damage in infants so managed.

Our data do not provide conclusive information on the incidence or severity of brain damage caused by profound hypothermic circulatory arrest techniques. To obtain this it would be necessary to assess the intelligence in individual patients both before and after surgery. The latter is not usually feasible because of the young age of the subjects before surgery and their often critical illness. The presence of even severe developmental retardation, let alone subtler degrees, may be difficult to determine in tiny, sick infants. The problems are compounded by the significant incidence of preoperative events known to influence later intellectual behavior, such as neurologic abnormalities and low birth weight for gestational age, which 15% (10 of 72) of the patients had. The early-year levels of the Stanford-Binet scale measure such functions as visual motor capacities, reasoning, social intelligence and language ability. These four areas make up 85% of the tests in the 2–5-year age group. Language function over this age range centers around simple descriptive phrases concerning common objects in the child's environment. The stimuli used in the tests are perceptual, and one-phrase definitions suffice. However, at 6 years of age and in all succeeding years, more complex
verbal descriptions are required for success. As all patients in both the control and congenital heart disease ideal groups were tested before 6 years of age, we thought that the mean age difference between the two groups would not seriously affect comparison. The four patients not included in the study would not have affected comparisons with the control group, as none would have been included in the ideal group. They would, however, lower the mean IQ of the total group by 1 or 2 points if one assumes an IQ of 40–60 for the three retarded children and an IQ equal to the mean for similar children tested for the clinically normal patient.

The data do show that the mean intelligence score of the study group as a whole is different from normal. The difference between patients and controls with similar birth and neonatal characteristics is, however, negligible. In particular, there was no correlation between the IQ and duration of circulatory arrest. Postoperative choreoathetosis, a syndrome known to be associated with deep hypothermia techniques, was noted in only two patients (2.7%), and both were neurologically and developmentally suspect preoperatively. These preoperative abnormalities, combined with longer-than-ideal circulatory arrest times (67 and 71 minutes), may have been responsible. Choreoathetosis was not recorded in infants with a circulatory arrest time less than 65 minutes. This incidence is less than that recorded by Brumberg et al., who attributed this complication to the hypothermic technique. Choreoathetosis was encountered in the early experience with cardiopulmonary bypass alone.

Pathologic studies in children who die after operations during which profound hypothermia and circulatory arrest were used have not shown any specific cerebral lesions that could be related to the operative technique; rather, the type and location of the lesions appear to depend on shock-like states that sometimes occur before operation.

As expected, the Stanford-Binet intelligence test results were lower in patients who had a smaller head circumference and in patients with significant preoperative problems. However, an abnormally small head circumference preoperatively may be a manifestation of growth failure, become normal postoperatively and may be associated with a normal IQ. Preoperative head circumference would not seem to be a good predictor of subsequent developmental progress unless it is both abnormally small and considerably further below the mean for age than weight. Scores were also lower in patients and controls in the lower socioeconomic classes, with the proportion of patients in the lowest class being higher than for the general population. They were unrelated to the type of congenital heart defect.

Further insight into the performance of our patients is, we believe, given by the use of a control group of "ideal" children who were tested by the same experienced observer and compared with the 25 patients in the study group with congenital heart disease who fulfilled similar "ideal" criteria. The proportion of children with congenital heart disease who met the characteristics of the ideal group was smaller than expected for our general population (35% vs 45%), a difference largely accounted for because only three of 22 patients with tetralogy of Fallot fulfilled the "ideal" criteria. We cannot explain this finding. Patients with congenital heart disease categorized as ideal did not, of course, meet the strict criteria used for the selection of the control group, as all had congenital malformations. Moreover, they were not excluded from the ideal group when cyanosis or other symptoms caused by congenital heart disease were present in the neonatal period, regardless of whether surgical treatment was required at this time. Thus, some may have been at a disadvantage compared with the ideal controls. These differences may explain the slightly lower mean IQ in the ideal congenital heart disease group (101.4) than in the ideal control group (106.2), a difference that is not statistically significant.

The method of selection of ideal subjects excludes the influence of low birth weight (whether caused by prematurity or inappropriately low birth weight for gestational age) and associated extracardiac anomalies on late developmental progress, but only partially allows for the effect of socioeconomic status. All these features have been noted to have an effect on late developmental progress in children with congenital heart disease.

The present study group is the largest series of children to be psychometrically tested after profound hypothermia–circulatory arrest techniques. The results are similar to those from other studies. Stevenson et al. undertook developmental assessment using a variety of tests in 32 children after heart surgery in infancy, using surface-induced hypothermia and circulatory arrest. They did not find any striking correlation between late mean IQ and the minimum temperature or duration of circulatory arrest, but commented that postoperative problems were more frequent in infants with circulatory arrest times longer than 50 minutes. They concluded that the hypothermic technique did not cause neurologic or intellectual impairment. Other investigators have reached a similar conclusion in smaller groups of patients. Subramanian et al. evaluated 36 children after this type of surgery using the Stanford-Binet or, in those 6 years of age or older, the WISC-R. The IQ in their group of children was found to have a normal distribution, with 50% scoring 90–110 and 77% 80–120. Results in our study group were comparable, with 44% scoring 90–110 and 78% 80–120. Similar results have been reported in a small group of infants who had repair of tetralogy of Fallot using cardiopulmonary bypass alone, and in patients in whom profound hypothermia was induced using core cooling. As a group, our patients scored less well than expected when language was considered in terms of quality of language and quantity and intelligibility of speech. However, when preoperative problems and the influence of social classes 5 and 6 were allowed for these differences largely, although not completely, disappeared. This is chiefly a subjective assessment and, like other workers, we are uncertain that this finding
is of any significance. Wright and associates\(^7\) reported findings in patients who had surgical closure of a ventricular septal defect using either cardiopulmonary bypass alone or profound hypothermia (core cooling) and circulatory arrest. They concluded that physical and intellectual development were more likely to be normal in patients operated upon using cardiopulmonary bypass. The observations presented on late developmental progress were, however, predominantly those of the patients' parents. No data on formal psychometric testing were presented. Attributes considered to indicate "slow" mental progress were nonspecific and may be encountered in children whose developmental progress is normal as well as in those in whom it is delayed. No significant difference was shown in intellectual development between patients whose operation was done using hypothermic arrest and those in whom cardiopulmonary bypass was used.

In conclusion, we could not demonstrate any deleterious effect attributable to the operative technique used in our patients. Rather, the postoperative intelligence test scores appear to reflect characteristics related to individual patients.

References

Developmental progress after cardiac surgery in infancy using hypothermia and circulatory arrest.
P M Clarkson, B A MacArthur, B G Barratt-Boyes, R M Whitlock and J M Neutze

Circulation. 1980;62:855-861
doi: 10.1161/01.CIR.62.4.855

Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 1980 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/62/4/855