Pre- and postoperative developmental attainment in sagittal synostosis

M Bellew, P Chumas, R Mueller, M Liddington, J Russell

Aims: To investigate whether sagittal synostosis (SS) has consequences for children’s mental and psychomotor development, and whether surgery has any impact on this.

Methods: The study involved 28 children with SS who underwent corrective surgery at a mean age of 8.0 (SD 7.16) months, and 28 normal controls. All the children with SS were assessed pre- and postoperatively using the Griffiths Mental Development Scales. The controls were assessed on one occasion, at an age matched with individuals in the patient group at the time of the preoperative assessment. A further control group consisted of 13 children with SS, who had received developmental assessment on two or more occasions without surgical intervention.

Results: The data indicated that children with SS have significantly poorer gross locomotor function than the normal controls. Following surgical intervention the deficit was shown to have resolved; consistent with this a lesser improvement in eye-hand coordination and performance skills was shown. Overall developmental attainment also improved postoperatively. The children with SS who did not receive surgery did not show any improvement in development.

Conclusions: The study shows improved developmental attainment following surgical correction of SS, which may therefore be more than a cosmetic procedure.

The literature regarding development in children with craniosynostosis is controversial, complex, and divided, with widely differing prevalence rates of developmental delay reported. Some studies support the view that children with non-syndromic craniosynostosis are at risk of developmental problems and learning disorders, and that early surgery is required to prevent this, while others have argued that children with non-syndromic craniosynostosis have development within the normal range and that surgery is only indicated for cosmetic reasons. However, the interpretation of some of these data has been controversial and discussion papers on the subject have interpreted the same data differently. Furthermore, even when normal early development that is unaffected by surgery has been shown, mental retardation and specific learning problems at later follow up have been reported. Clearly further research is needed to clarify the situation.

At the Leeds Craniofacial Unit, all children with craniosynostosis undergo routine pre- and postoperative developmental assessment and longer term follow up, enabling the development of individual children to be monitored over time. The current study presents developmental data on 28 children with sagittal synostosis (SS) and examines whether this group of children differs in terms of development from control groups, and whether surgery has any impact on this. The study benefits from a prospective longitudinal design, the assessment of specific areas of development, and the use of control groups rather than relying on comparison with normative data.

METHODS

Patients (surgical SS group)
The study involved 28 children with SS (5 girls and 23 boys) who underwent corrective surgery between March 1996 and September 2002 at a mean age of 8.0 (SD 7.16) months (range 2.8–39.7 months). All children underwent routine pre- and postoperative developmental assessment (see table 1). Eleven patients had a second postoperative assessment, four had a third, and one had a fourth.

Control groups

Normal control group
This consisted of 28 children who had not undergone any surgery and who underwent developmental assessment at an age approximately matched with individuals in the surgical SS group at time of preoperative developmental assessment (comparison with t test showed that this was not significantly different) and were also matched for gender (see table 1). The groups as a whole were also approximately matched for prematurity, geographical and socioeconomic backgrounds, and race (being predominantly Caucasian, with one Asian child in the patient group and two in the control group). One of the control subjects was the twin of one of the patients.

Non-surgical SS control group
As the policy of the Leeds Craniofacial Unit is that all patients with craniosynostosis are referred for developmental assessments, irrespective of age or plans for surgery, a further control group was available that consisted of all children seen during the study period who had SS and who had received developmental assessment on two or more occasions without surgical intervention (13 children; 2 girls and 11 boys). Five of these children did not receive surgery because their conditions were cosmetically “mild”, and two did not receive surgery for both these reasons. A further child was admitted for surgery, but the operation was abandoned because of malignant hyperpyrexia. The remaining child did receive surgery after two preoperative assessments had been undertaken, due to moderately raised intracranial pressure. Table 1 shows the mean ages at the times of the assessments. There was a similar interval between assessments (9.1 months) to that between the surgical group’s pre- and postoperative assessments (8.3 months).
Developmental assessment
All developmental assessments were performed using the Griffiths Mental Development Scales,17 18 and were undertaken by the first author. The assessment consists of five scales, which assess a range of aspects of development: locomotor, personal-social, hearing and speech, eye-hand coordination, and performance. Subquotients are calculated for each of the individual scales, and averaging these yields the general quotient (GQ), which indicates global developmental attainment.

Data analysis
Comparison 1: Surgical SS group and normal control group
The preoperative data from the surgical SS group were compared with the data from the normal control group. Firstly, the overall GQs were compared using a $t$ test. The individual subquotients from which the GQ is derived were then compared using a mixed two way ANOVA, to determine any differences in scores within the individual scales of the Griffiths or in interaction between the two groups.

Comparison 2: Pre- to postoperative data from the surgical SS group
(2a) Between pre- and first postoperative
The pre- and initial postoperative GQs of the surgical SS group were compared using a related $t$ test, and the individual subquotients analysed using a two way repeated measures ANOVA, to determine any differences in scores obtained within the individual scales of the Griffiths, and for any main or interaction effect pre- to postoperatively.

(2b) Between pre- and latest postoperative
The analysis was repeated, but this time using the latest postoperative developmental data to give an indication of any longer term effects on development. This meant that for 11 patients their second postoperative assessment was used, for four their third postoperative assessment, and for one their fourth. This gave a mean age of 33.2 months (SD 22.93) and a mean interval of 25.3 months (SD 24.03; range 5.8–66.5 months) after surgery for this analysis.

(2c) Between pre- and second postoperative
The analysis was repeated for just the 11 patients who had undergone a second postoperative assessment.

Comparison 3: Between surgical and non-surgical SS groups
The preoperative developmental data from the surgical SS group was compared with the developmental data from the non-surgical SS control group at the first occasion of testing. The overall GQs were analysed using a $t$ test and the subquotients were analysed using a two way mixed ANOVA. This was to establish how similar the initial data sets were in order to determine whether comparisons between how the groups fared over time would be valid.

Comparison 4: Between occasions 1 and 2 for the non-surgical SS group
(4a) Whole group ($n = 13$)
The data from the two occasions of testing from the non-surgical SS control group were compared. The overall GQs were analysed using a $t$ test and the subquotients were analysed using a repeated measures two way ANOVA.

(4b) Excluding “mild” cases ($n = 6$)
The analyses were then repeated but with the “mild” cases, and a child who had had learning difficulties since birth unrelated to her synostosis, excluded. This gave a group of six children, comprising four children who would have been surgical candidates if they had presented early enough, the child who had had anaesthesia problems, and the child who subsequently underwent surgery. As such they were more similar to the surgical SS group in terms of diagnosis and therefore more directly comparable.

### Table 1 Mean ages and times of developmental assessments for the three groups

<table>
<thead>
<tr>
<th></th>
<th>Surgical SS group</th>
<th>Normal control group</th>
<th>Non-surgical SS control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preoperative developmental assessment/occasion 1</td>
<td>$n = 28$ (5 girls, 23 boys) Mean age: 6.9 months (SD 5.71; range 2.4–30.2)</td>
<td>$n = 28$ (5 girls, 23 boys) Mean age: 7.2 months (SD 6.14; range 1.7–33.7)</td>
<td>$n = 13$ (2 girls, 11 boys) Mean age: 25.7 months (SD 19.98; range 5.3–76.6)</td>
</tr>
<tr>
<td>Mean interval: 1.1 months before surgery (SD 1.77; range 0.03–9.5)</td>
<td>Mean interval: 7.2 months after surgery (SD 2.70; range 2.9–15.6)</td>
<td>Mean interval: 41.0 months after surgery (SD 21.84; range 8.1–66.5)</td>
<td></td>
</tr>
<tr>
<td>1st postoperative developmental assessment/occasion 2</td>
<td>$n = 28$ Mean age: 15.2 months (SD 7.93; range 5.7–45.9)</td>
<td>$n = 13$ Mean age: 34.8 months (SD 19.31; range 15.8–84.5)</td>
<td></td>
</tr>
<tr>
<td>Mean interval: 7.2 months after surgery (SD 2.70; range 2.9–15.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd postoperative developmental assessment</td>
<td>$n = 11$ Mean age: 46.5 months (SD 22.06; range 10.9–69.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean interval: 41.0 months after surgery (SD 21.84; range 8.1–66.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3rd postoperative developmental assessment</td>
<td>$n = 4$ Mean age: 49.9 months (SD 19.28; range 23.3–66.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean interval: 45.3 months after surgery (SD 18.36; range 20.5–62.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4th postoperative developmental assessment</td>
<td>$n = 1$ Age: 60.1 months Interval: 54.5 months after surgery</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
RESULTS
Comparison 1: Surgical SS group and normal control group
Comparison of the preoperative developmental data from the surgical SS group with the data from the normal controls (table 2), did not reveal a significant difference in terms of overall GQ, but analysis of the subquotients from the five individual scales that the GQ is derived from, revealed a significant interaction effect on two way mixed ANOVA (\( p = 0.033 \)). To identify the source of the significance, the data were further investigated using the Bonferroni method for unplanned pairwise multiple comparisons.\(^*\) This revealed that the significant effect was due to the poorer performance by the surgical SS group on the locomotor scale, and this reached a Bonferroni level of significance (\( p = 0.009 \)). No differences were found with the other scales. The delay on the locomotor scale was severe in some cases. Twelve children (43%) obtained locomotor scores within the learning difficulties (LD) or borderline learning difficulties (BLD) range of ability (36% below a GQ of 76). Locomotor delay was, however, seen in only four children (14%) in the control group.

Comparison 2: Pre- to postoperative data from the surgical SS group
(2a) Between pre- and first postoperative
Comparison of developmental data from the surgical SS group pre- to postoperatively (table 2) shows that in terms of overall development (GQ) there was a significant improvement after surgery (\( p = 0.001 \)). Analysis of the subquotients from the five individual scales by two way ANOVA showed a significant effect for time (\( p = 0.006 \)) and for the interaction effect between time and the individual scales of the Griffiths (\( p = 0.002 \)). Bonferroni \( t \) tests revealed that the postoperative improvement in development was largely due to the preoperative delay on locomotor skills resolving postoperatively (\( p = 0.0001 \)). In fact 23 of the 28 surgical SS patients showed a postoperative improvement on this scale. This included all 12 of the children who had obtained preoperative locomotor scores within the LD/BLD range of ability. By the time of the first postoperative assessment, these deficits had all resolved to within the normal range. Furthermore, for three children who had an overall GQ in the LD/BLD range of ability preoperatively, this also improved to fall within the normal range. A lesser improvement was shown on the eye-hand coordination and performance scales, though this did not quite reach Bonferroni significance (both \( p = 0.05 \)).

(2b) Between pre- and latest postoperative
When the analysis was repeated incorporating the longer term follow up data (table 2), the pre- to postoperative improvement in overall GQ was more significant (\( p = 0.0001 \)) as the developmental attainment of the children had continued to improve with time. Analysis of the subquotients from the five individual scales by two way ANOVA showed a more significant effect for time (\( p = 0.00001 \)) as well as a more significant interaction effect between time and the individual scales of the Griffiths (\( p = 0.0001 \)). Further analysis using the Bonferroni method for unplanned multiple comparisons again showed that the postoperative improvement in development was largely due to improvement in locomotor skills (\( p = 0.0001 \)), with 24 of the 28 SS patients now showing a postoperative improvement on this scale. There was again a smaller pre- to postoperative improvement for the eye-hand coordination scale (\( p = 0.03 \)) and also the hearing and speech

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Mean developmental data for the normal controls and surgical SS group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GQ</td>
</tr>
<tr>
<td>Surgical SS group (n = 28)</td>
<td></td>
</tr>
<tr>
<td>Preop</td>
<td>101.8 (SD 15.61)</td>
</tr>
<tr>
<td></td>
<td>99.1 (SD 21.92)</td>
</tr>
<tr>
<td></td>
<td>99.5 (SD 22.42)</td>
</tr>
<tr>
<td></td>
<td>105.2 (SD 20.16)</td>
</tr>
<tr>
<td></td>
<td>103.5 (SD 16.59)</td>
</tr>
<tr>
<td></td>
<td>103.9 (SD 18.58)</td>
</tr>
<tr>
<td></td>
<td>102.9 (SD 17.21)</td>
</tr>
<tr>
<td></td>
<td>100.0 (SD 16.94)</td>
</tr>
</tbody>
</table>

**p = 0.0001, comparison of pre- and postoperative GQ in surgical SS patients.**
*Cp = 0.0001, comparison of pre- and postoperative locomotor score in surgical SS patients.
*Dp = 0.002, comparison of pre- and postoperative performance score in surgical SS patients.
*This uses ordinary \( t \) tests but with a stricter criterion for significance, calculated by dividing 0.05 by the number of tests that will be made. Five tests were made (to compare each of the five subquotients pre- to postoperatively) and so the significance level was set at 0.01.
scale (p = 0.05), and an improvement on the performance scale which reached Bonferroni significance (p = 0.002).

(2c) Between pre- and second postoperative
When the analysis was repeated for just the 11 patients who had had a second postoperative assessment, the results were even more significant: improvement in overall GQ from preoperative to second postoperative assessment (p = 0.009), significant effect for time on the two way ANOVA (p = 0.011) and for the interaction (p = 0.001), largely accountable for by postoperative improvement on the locomotor scale (Bonferroni significance of p = 0.002).

Comparison 3: Between surgical and non-surgical SS groups
Comparison of the preoperative data from the surgical patients and occasion 1 of the non-surgical SS control group showed that the initial data sets from these two groups of patients were very similar. No significant differences were found, except that the non-surgical group were older at time of first assessment (p = 0.005).

Comparison 4: Between occasions 1 and 2 for the non-surgical SS group
(4a) Whole group (n = 13)
When the non-surgical patients' first assessment was compared with their second (Table 3), unlike the group who received surgery, there was no significant change in GQ over the two occasions of testing. Analysis of the subquotients from the individual scales by two way ANOVA also showed no significant effects. Further examination of the data revealed two children to have GQs within the LD range of ability on the first occasion. One of these was the child with LD unrelated to her synostosis, and this had not resolved by the second occasion. For the other, there was only slight improvement. Examining the locomotor scale in particular revealed three patients with scores within the LD range of ability. One of these remained the same on the second occasion of testing and the other two had only improved very slightly and were still BLD.

(4b) Excluding ‘‘mild” cases (n = 6)
When the analyses were repeated excluding the “mild” cases and the child with LD, making the surgical and non-surgical groups more comparable, the scores were generally slightly lower than in the full non-surgical SS group and comparison between the first and second occasions of testing still showed no statistically significant improvement over time.

DISCUSSION
This study indicates that children with SS have significantly poorer gross locomotor function than age and gender matched healthy children. This was severely delayed in some cases. Following surgical intervention the deficit was shown to have resolved, which raises the possibility that surgery has an impact on development and offers more than just cosmetic improvement. Certainly it has been striking during the preoperative assessments that these babies have difficulty with head control. Many parents have commented on how locomotor ability has “suddenly” and noticeably improved postoperatively. The longer term follow up showed that this improvement is maintained and indeed continues to improve with time.

Overall GQ also improved for this group postoperatively and a lesser improvement was shown for the eye-hand coordination and performance scales. This is consistent with the data for the locomotor scale, as it would be expected that a child with problems of head control would also have difficulty with eye-hand coordination and that improvement on the former would affect the latter. This is in turn required for the items on the performance scale (practical application of manipulative skills in novel situations) and an improvement on these skills increased as the children got older. This consistency between the scales adds weight to the findings.

The fact that the normal control group did not show the same gross locomotor deficits, shows that these difficulties were specific to the children with SS and were not simply attributable to normative error on the test, error in the assessment procedure by the clinician, or error introduced due to inherent difficulties in assessing intelligent quotients in babies (though these would be expected to have a global effect in any case and would be unlikely to cause delay solely on the locomotor scale). As it has been shown that normative developmental trends have led to higher attainment on the Griffiths than when the test was devised, particularly on the locomotor scale, the fact that the SS group actually show delay on this scale is even more noteworthy.

The fact that the non-surgical SS group did not show any improvement in development over two occasions of testing on the locomotor scale or otherwise despite similar data on their first occasions of testing, suggests that the improvement in development with the surgical SS group is not simply due to the passage of time, maturity, or any practice effects because they had done the test previously. Furthermore, Griffiths has previously shown good retest reliability in young subjects (highly significant test-retest correlation of r = 0.87, n = 60), the fact that there is no overlap of test items in babies (as they develop new skills so rapidly at this age) being an important factor in this. Therefore an improvement in developmental attainment such has been shown in the current study, would not be expected without an additional factor operating. In this case it seems that improvement in development follows surgical intervention.

However, although improved postoperative developmental attainment has been shown, a causal link between SS, locomotor delay, and surgery has not been established, and any suggested mechanisms as to why the locomotor attainment improved postoperatively would be pure conjecture. One possible explanation could be that it is a purely mechanical effect of correcting the head shape. Another explanation could be increased locomotor stimulation from the parents subsequent to the developmental assessment. However, as many parents reported a sudden and noticeable improvement in locomotor attainment postoperatively, this better supports the improvement being related to the surgery in some way, especially as most of them had noticed the locomotor delay themselves prior to preoperative assessment, and had been unable to effect any change then. It seems equally unlikely that it is in any way related to raised intracranial pressure (ICP). This has a suggested inverse relation with IQ, but the incidence has been shown to increase with the number of sutures involved, and it is generally accepted that raised ICP is unusual in single suture craniosynostosis. It has been reported to occur in only 7–13% of patients with SS. Furthermore, Arnaud and colleagues found that the risk of having raised ICP was age dependent (20% if child >1 year at presentation compared to 2% if <1 year). The mean age of the surgical SS children in our study was 6.9 months (SD 5.71) at the time of the first assessment, and therefore ICP is unlikely to have been a major confounding factor. In any case, one would have expected a more global effect rather than a specific effect on locomotor function. Clearly further studies are required to establish the aetiology of the developmental improvement.

The conclusions that can be drawn from this control group are undoubtedly limited by the fact that they are not age matched with the surgical group, as the patient assessments are driven by clinical need, not research protocols.
The specific quantifiable early developmental delay identified in this study has not been reported before, even in studies that suggest developmental delay as the children get older. Early studies have referred to children being “definitely retarded” preoperatively, but have not involved formal assessment or defined the developmental delay. Statistically significant resolution of developmental delay postoperatively has also not been shown previously. However, most other studies have only considered mental development and have not specifically examined gross locomotor development. Kapp-Simon states that although the Bayley Scales of Infant Development in which they used provide scores of both mental and psychomotor development, they only used the mental development index. Speltz and colleagues did examine the Kohen-Raz subscales of the Bayley Scales of Infant Development, but although this did include specific areas of psychomotor functioning, it did not include gross locomotor functioning, and again no particular deficits were detected.

Only one other study appears to have examined gross locomotor ability in young children with craniosynostosis, and the results of that study support the current findings. Kaiser reported that 35.5% of 50 children with SS (operated on at an average age of 5.5 months) had clinical signs such as cerebral palsy, psychomotor retardation, and/or neurological signs. Postoperatively many of these problems had resolved, and only 14.5% had clinical signs, all of them minor or unrelated to the craniosynostosis. Kaiser’s percentage of 35.5 is strikingly similar to the figure of 35.7% of surgical SS patients in the current study falling into the LD range of ability on the locomotor scale.

The conclusions from the current study are that there is a risk of delay in locomotor development in children with SS but that this is observed to improve following surgical intervention, and is accompanied by a lesser improvement in eye-hand coordination and performance skills. Overall developmental attainment also improves. The use of control groups indicates that the developmental deficit was specific to the children with SS, and improvement in development over two occasions of testing does not occur if surgery is not undertaken. These findings therefore support the contention that surgery for SS may be more than purely cosmetic.

ACKNOWLEDGEMENTS
This work was undertaken by the Leeds Teaching Hospitals NHS Trust who received funding from the NHS Executive; the views expressed in this publication are those of the authors and not necessarily those of the NHS Executive. We would like to thank Professor A Boyston for suggestions regarding data analysis and preparation of the manuscript.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Mean developmental data, occasions 1 and 2, for the non-surgical SS group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Occasion 1 (n = 13)</td>
<td>Occasion 2 (n = 13)</td>
</tr>
<tr>
<td>GQ</td>
<td>98.6 (SD 18.61)</td>
</tr>
<tr>
<td>Locomotor</td>
<td>99.4 (SD 25.04)</td>
</tr>
<tr>
<td>Personal-social</td>
<td>96.1 (SD 20.49)</td>
</tr>
<tr>
<td>Hearing and speech</td>
<td>97.0 (SD 28.43)</td>
</tr>
<tr>
<td>Eye/hand coordination</td>
<td>98.5 (SD 17.99)</td>
</tr>
<tr>
<td>Performance</td>
<td>99.6 (SD 16.36)</td>
</tr>
</tbody>
</table>

*R: Results of two way ANOVA: time, p = 0.633; subquotients, p = 0.760; interaction, p = 0.415.

R Mueller, Department of Clinical Genetics, St James’s University Hospital, Leeds, UK
J Russell, Department of Maxillofacial Surgery, Leeds General Infirmary, Leeds, UK

Competing interests: none declared

REFERENCES