Effect of 2 years of high-dose growth hormone therapy on cognitive and psychosocial development in short children born small for gestational age

K Lagrou, J Vanderfaeillie, C Froidecoeur, M Thomas, G Massa, S Tenoutasse, M Craen, M C Lebrethon, D Beckers, I Francois, R Rooman, G Thiry-Counson, C de Beaufort, and J De Schepper

Belgian Study Group for Pediatric Endocrinology, Belgium, 1Department of Pediatrics, University of Antwerp, Antwerp, Belgium, 2Department of Pediatrics, University of Brussels, Brussels, Belgium, 3Department of Pediatrics, University of Brussels, Brussels, Belgium, 4Department of Pediatrics, University of Gent, Gent, Belgium, 5Department of Pediatrics, University of Louvain, Mont-Godinne, Belgium, 6Department of Pediatrics, University of Liege, Liege, Belgium, 7Department of Pediatrics, Virga Jesseziekenhuis, Hasselt, Belgium, 8Department of Pediatrics, Clinique de l’Espérance, Liege, Belgium and 9Clinique de Luxembourg, Luxembourg, Luxembourg

(Correspondence should be addressed to J de Schepper, Department of Pediatrics, Academic Hospital Free University of Brussels, Laarbeeklaan 101, 1090 Brussels, Belgium; Email: jean.deschepper@az.vub.ac.be)

Abstract

Objective and design: Children born small for gestational age (SGA) are not only at risk for short stature, but also for neurodevelopmental and behavioral problems. In this study, we analyzed the effects of high-dose GH therapy on cognitive development and psychosocial functioning in 34 prepubertal (3–8 years) short SGA children, equally randomized into a GH-treated group (TRG) and an untreated group (UTRG).

Methods: At start and after 2 years, children underwent standardized tests measuring the intellectual abilities (Wechsler Preschool and Primary Scale of Intelligence-Revised, or Wechsler Intelligence Scale for Children-Revised); their parents completed a standardized questionnaire evaluating psychosocial functioning (Child Behavior Checklist; CBCL).

Results: At start, total IQ scores were significantly (P < 0.05) lower in the SGA group than in the general population: 32% of the SGA patients had scores below 85. After 2 years, IQ scores remained unchanged in the TRG, but increased significantly (P < 0.05) in the UTRG. After exclusion of children with developmental problems, however, no significant changes in IQ scores occurred in the UTRG as well as the TRG. At baseline, 24% (8/34) children had problematic CBCL total problems scores, equally distributed among the two groups; no significant changes in the different subscale scores occurred after 2 years.

Conclusion: No beneficial effect of 2 years of GH therapy on cognitive and behavioral profile could be observed in a cohort of rather young short SGA children presenting a variable degree of developmental delay and behavioral problems. Subsequent follow-up could reveal potential long-term effects of GH therapy on development and behavior.
Subjects and methods

Patients

After an inclusion period of 1.5 years, 55 prepubertal SGA children consulting for short stature in the pediatric centers were approached to participate in a 2-year trial of high-dose GH, exploring the effects of GH on body growth and body composition, psychosocial status and cognitive functioning. The inclusion criteria were birth weight and/or length below $-2$ S.D. for gestational age (13), chronological age between 3 and 8 years, current height below $-2.5$ S.D., and a height velocity SDS below $+1.0$ S.D. during the last 6–18 months (14). The exclusion criteria were gestational age below 34 weeks, endocrine disease, including GH deficiency, bone disease, severe chronic disease, Turner, Noonan or Down syndrome, or other genetically confirmed syndromes, chromosomal abnormalities, current or previous irradiation therapy, current or previous (up to 18 months before inclusion) treatment with glucocorticoids and severe mental retardation (IQ $\leq 50$). In total 40 children, responding to exclusion and inclusion criteria, were allocated after informed consent into a treated group (TRG) and an untreated group (UTRG) using a minimization randomization, taking into account following four parameters: gender, chronological age, weight SDS, and study center. For ethical reasons, the UTRG did not receive daily placebo injections but underwent the same investigations. The treated children received biosynthetic GH (Genotonorm, Pfizer, New York, NY, USA) for 2 years in a dosage of 0.066 mg/kg per day. Written informed consent was obtained from the parents for enrollment in this study and treatment with growth hormone. The Ethical Review Boards of each participating center approved the study protocol.

For 34 children (15 boys and 19 girls), all IQ scores (total, verbal, and performance IQ) were available at start of the study (T0) and after 2 years (T2); IQ could not be reliably assessed in four patients at start of the study because of a serious language delay and the IQ scores were missing for two other patients after 2 years (one drop-out and one incomplete IQ test). The auxological and the clinical characteristics of these 34 children with complete data are presented in Table 1. Twenty-one patients were preschoolers (3–5 years; 8 in the TRG and 13 in the UTRG) and 13 were school-age children (6 years or more; nine in the TRG and four in the UTRG). Six children (18%; six girls versus zero boy, $P < 0.05$) presented minor (mainly facial) dysmorphic features. Some dysmorphisms were indicative of Silver Russell syndrome in three children (two in the treated group and one in the untreated group), for 3M syndrome in one treated and Renpenning syndrome in one untreated subject. At start of the study, parents reported chronic health problems in eight children (24%), mainly recurrent respiratory, digestive and orthopedic problems. Fourteen SGA children (45%) had developmental problems, defined as a delay in learning to walk, to talk and in toilet-training. One-third of the children (32%) needed specialized help (physical, speech, or psychomotor therapy). At start, no significant differences were observed for any of the baseline clinical and auxological parameters between treated and untreated patients, between boys and girls and between children with and without dysmorphic features except for birth weight: subjects with dysmorphic features having a significantly lower birth weight SDS than children without minor dysmorphic features ($-3.44$ vs $-2.44$, $P < 0.05$). The socioeconomic and educational status of the mothers, each subdivided into three levels (low, medium, and high), were also equally distributed within the TRG and UTRG, the boys and the girls and the children with and without dysmorphic features (data not shown).

Methods

The standard auxological assessment consisted of height, weight, and head circumference measurements every 6 months by the same clinical monitor. Height SDS were calculated using the British references of Freeman et al. (14).

The cognitive and psychological evaluation was performed at start of the study, before the assignment of treatment modality, and after 2 years. Intellectual abilities were assessed by the age-appropriate Wechsler scales: Wechsler Preschool and Primary Scale of Intelligence-Revised (WPPSI-R) for children between 3 and 6 years of age and the Wechsler Intelligence Scale for Children-Revised (WISC-R), when subjects were older than 6 years (15, 16). French- and Dutch-translated and adapted versions were used (WISC-R 1981 and 1986, WPPSI-R 1995). In 20 patients (13 in the UTRG and 7 in the TRG), a change in IQ test instrument was made because of an age older than 6 years at retesting after 2 years. Behavioral and emotional problems are measured using the Child Behavior Checklist (CBCL): a standardized measure of social competence, behavioral, and emotional problems in children aged 4–18 years developed by Achenbach (17). French- and Dutch-translated and adapted versions of the CBCL were completed by the parents at start and after 2 years (17, 18). The CBCL consists of a competence and a problem behavior section. In this study, we report the results of the behavioral scale scores: total problem score, externalizing score (regrouping aggressive behavior and delinquent behavior) and internalizing score (regrouping somatic complaints, withdrawn, anxious depressed) as well as the separate subscale scores: somatic complaints, social problems, withdrawn, anxious depressed, attention problems, thought problems, aggressive behavior, and
delinquent behavior. Higher scores at the CBCL indicate more problems. In order to differentiate between ‘normal’ and ‘problematic’ functioning cut-off points are used. According to the Achenbach’s manual, for the syndrome scales T-scores < 67 are considered in the normal range, T-scores ranging from 67 to 70 are considered to be borderline clinical, and T-scores above 70 are in the clinical range, whereas for the broad-band scales T-scores < 60 are considered in the normal range, 60–63 represent borderline scores, and scores greater than 63 are in the clinical range. The difference in cut-off points for the syndrome subscales and the broad-band scales is due to the truncation of T-scores for the syndrome subscales, whereas T-scores are not truncated in the broad-band scales.

Statistical analyses

Auxological results and test scores are expressed as mean ± s.d., clinical and interview data as percentages. Differences of continuous variables between subgroups were evaluated by the Student’s unpaired t-test or by the Mann–Whitney U test as appropriate and differences of categorical variables by the Pearson χ² test or by the Fisher’s exact test as appropriate. Changes in test scores were evaluated by the Student’s paired t-test or Wilcoxon rank test as appropriate. According to the manual, differences between and within subjects are examined using the raw CBCL subscale scores (18). Correlations were examined by the Pearson correlation coefficient. The level of significance of difference was set at P < 0.05. Statistical analyses were carried out using the SPSS software (version 13.0, SPSS, Inc., Chicago, IL, USA).

Results

Intellectual abilities and school functioning

Table 2 shows the results for the intellectual abilities. At baseline evaluation, the children had significantly different IQ scores than the general population. Of the total IQ scores, 65% were below 100 and 32% were below 85; 59% of the verbal IQ scores were below 100 and 24% below 85; 65% of the performance IQ scores were below 100 and 29% below 85. The children with mental retardation (total IQ ≥ 54 and < 70) were equally distributed over the treated (1/17) and the untreated group (3/17). At baseline, the total, verbal, and the performance IQ scores of the treated children were different from those of the untreated group. In addition, no significant differences were observed between boys and girls and between children with and without minor dysmorphic features.

As shown in Table 1, mean head circumference SDS of the whole group was −2.5 ± 1.4 with 85% (29/34) of the children having a head circumference below −1 s.d. Children with developmental problems (14/31) had a significantly (P < 0.01) lower head circumference SDS (−3.4 ± 1.3 vs −1.9 ± 1.2 respectively). At baseline, the head circumference SDS were

Table 1 Comparison of the baseline auxological and clinical data between the treated group (TRG; n = 17) and the untreated group (UTRG; n = 17).

<table>
<thead>
<tr>
<th></th>
<th>All children (n = 34)</th>
<th>TRG (n = 17)</th>
<th>UTRG (n = 17)</th>
<th>P TRG–UTRG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birth weight (SDS)a</td>
<td>−2.6 (0.8)</td>
<td>−2.7 (0.9)</td>
<td>−2.5 (0.7)</td>
<td>NS</td>
</tr>
<tr>
<td>Gestational age (weeks)a</td>
<td>37.7 (1.9)</td>
<td>37.2 (2.1)</td>
<td>38.3 (1.5)</td>
<td>NS</td>
</tr>
<tr>
<td>Chronological age (years)a</td>
<td>5.5 (1.4)</td>
<td>5.7 (1.6)</td>
<td>5.3 (1.3)</td>
<td>NS</td>
</tr>
<tr>
<td>Height (SDS)a</td>
<td>−3.2 (0.8)</td>
<td>−3.4 (0.6)</td>
<td>−3.1 (0.9)</td>
<td>NS</td>
</tr>
<tr>
<td>Height velocity (SDS)a</td>
<td>−1.2 (1.1)</td>
<td>−1.4 (1.3)</td>
<td>−0.9 (0.8)</td>
<td>NS</td>
</tr>
<tr>
<td>Head circumference (SDS)a</td>
<td>−2.5 (1.4)</td>
<td>−2.6 (1.3)</td>
<td>−2.4 (1.4)</td>
<td>NS</td>
</tr>
<tr>
<td>Male/female</td>
<td>15/19</td>
<td>9/8</td>
<td>6/11</td>
<td>NS</td>
</tr>
<tr>
<td>Chronic health problemsb</td>
<td>8/34 (24%)</td>
<td>4/17 (24%)</td>
<td>4/17 (24%)</td>
<td>NS</td>
</tr>
<tr>
<td>Developmental problemsb</td>
<td>14/31 (45%)</td>
<td>7/15 (47%)</td>
<td>7/16 (44%)</td>
<td>NS</td>
</tr>
<tr>
<td>Specialized help neededb</td>
<td>10/31 (30%)</td>
<td>6/15 (40%)</td>
<td>4/16 (25%)</td>
<td>NS</td>
</tr>
<tr>
<td>Dysmorphic featuresb</td>
<td>6/34 (18%)</td>
<td>3/17 (18%)</td>
<td>3/17 (18%)</td>
<td>NS</td>
</tr>
</tbody>
</table>

aData are expressed as mean ± S.D. 
Data are expressed as number of children with mentioned problem.
correlated with total IQ scores ($r=0.37$, $P<0.05$), verbal IQ scores ($r=0.34$, $P=0.05$), and performance IQ scores ($r=0.35$, $P<0.05$).

After 2 years, the mean total, verbal, and the performance IQ scores remained unchanged in the treated group, but increased significantly ($P<0.05$) in the untreated group. However, after the exclusion of children with developmental problems, no significant changes in the total, verbal, and performance IQ scores were observed in none of the two groups (TRG ($n=8$): TIQ = 99.8, VIQ = 99.0, PIQ = 101.0 at T0 versus TIQ = 101.3, VIQ = 101.6, PIQ = 101.1 at T2; UTRG ($n=9$): TIQ = 99.0, VIQ = 99.7, PIQ = 98.7 at T0 versus TIQ = 102.3, VIQ = 104.1, PIQ = 100.3 at T2). No significant changes in IQ scores were observed between children with and without minor dysmorphic features and between boys and girls (data not shown). In those children with a change from WPPSI-R to WISC-R, the change in performance IQ score was higher compared with those without a change of test instrument (7.1 vs $-2.4$, $P<0.01$).

The head circumference SDS increased significantly in the TRG (from $-2.6 \pm 1.3$ to $-1.9 \pm 1.3$, $P<0.001$) and remained unchanged in the UTRG (from $-2.4 \pm 1.4$ to $-2.4 \pm 1.5$, $P=ns$). Height SDS increased significantly in the TRG ($P<0.001$) as well as in the UTRG ($P<0.05$), but the change was much greater in the TRG (TRG: from $-3.4 \pm 0.6$ to $-1.8 \pm 0.7$ versus UTRG: from $-3.1 \pm 0.9$ to $-2.9 \pm 0.8$). In the TRG as well as in the UTRG, changes in IQ scores after 2 years were neither correlated with changes in head circumference SDS nor in height SDS. After 2 years, 6/17 (35%) of the treated and 4/17 (23%) of the untreated children still had developmental problems and needed specialized help. Table 3 shows the number of children with school delay after 2 years of observation: three (17%) treated and seven (41%) of the untreated children had to repeat the first or the second grade of primary school. Eight (47%) of the treated and seven (41%) of the untreated children had already repeated a grade at school at the end of the study.

### Psychosocial functioning

Baseline CBCL raw scores at the behavioral subscales were not different between the treated and the untreated groups, except for the somatic complaints, attention problem, and the aggressive behavior subscale scores, which were significantly higher in the untreated group (data not shown).

Table 3 shows the number of children with school delay at the final evaluation in the treated group (TRG; $n=17$) and in the untreated group (UTRG; $n=17$).

#### Table 3 Number of children with school delay at the final evaluation in the treated group (TRG; $n=17$) and in the untreated group (UTRG; $n=17$).

<table>
<thead>
<tr>
<th>Number of children who repeated a grade</th>
<th>TRG ($n=17$)</th>
<th>UTRG ($n=17$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nursery school/Kindergarten</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First grade</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Second grade</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Third grade</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Primary school</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First grade</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Second grade</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Third grade</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>8/17</td>
<td>7/17</td>
</tr>
</tbody>
</table>

*Data are expressed as mean $\pm$ S.D.
compared with the treated group (median scores 1 vs 0, 4.5 vs 2.0, and 7.0 vs 3.0 respectively, P<0.05). As shown in Table 4, also the T-scores on the subscales somatic complaints, attention problems, and aggressive behavior were significantly (P<0.05) higher in the UTRG compared with the TRG. A considerable number of children had deviant CBCL T-scores at baseline evaluation: at the total problem scale 24% (8/34) of the children had deviant T-scores (3/17 untreated children were within the clinical and 2/17 within the borderline clinical range; none of the treated children in the clinical and 3/17 treated children were within the borderline clinical range); at the internalizing scale 29% (10/34) children had deviant T-scores (3/17 untreated children were within the clinical and 4/17 within the borderline clinical range; 2/17 treated children were within the clinical range and 1/17 within the borderline clinical range); and at the externalizing scale 21% (7/34) children had deviant scores (3/17 untreated children were within the clinical range and 2/17 within the borderline clinical range; none of the treated children were within the clinical range and 2/17 treated children were within the borderline clinical range).

After 2 years, the absolute scores of the different behavioral subscales did not change significantly in the treated and in the untreated groups, with the exception of the median withdrawal subscale score which decreased significantly (P<0.05) in the UTRG. The median raw attention problem scores and raw aggressive behavior scores in the treated group increased to a level comparable with the untreated group after 2 years (data not shown). No significant differences were observed between boys and girls and between children with and without minor dysmorphic features for any of the CBCL raw subscales scores at baseline as well as after 2 years. After 2 years, 21% (7/33) children had deviant T-scores at the total problem scale (3/17 untreated children within the clinical and 1/17 within the borderline clinical range; 2/17 treated children were within the clinical range and 1/17 within the borderline clinical range); at the internalizing scale 15% (5/33) children had deviant T-scores (1/17 untreated children were within the clinical and 2/17 within the borderline clinical range; 1/16 treated children were within the clinical range and 1/16 within the borderline clinical range); and at the externalizing scale 15% (5/33) children had deviant scores (2/17 untreated children were within the clinical range and 1/17 within the borderline clinical range; none of the treated children were within the clinical range and 2/17 treated children were within the borderline clinical range).

**Discussion**

This is the first study reporting on the effects of high-dose GH therapy on the developmental and behavioral profile of short SGA children including an untreated control group. In half of the SGA children, addressed for short stature at a Belgian endocrine clinic, developmental problems were reported by the parents and one-third had a total IQ score below 85 at initial IQ testing. This percentage might be an underestimation given the use of rather older test instruments in our study, as mean scores on IQ tests increase each decade (known as the Flynn effect). The findings of lower intellectual capacities and recurrent health problems are consistent with those reported in several other publications on the cognitive development of SGA children (3, 10, 19–29). It is somewhat surprising that none of these previous studies referred to the need of specialized help in these children, which was noted to be provided in one-third of the patients in the present study, excluding, however, children with severe mental retardation, but including some children with minor dysmorphisms.

We could not find any beneficial effect of GH therapy neither on intellectual functioning nor on school functioning, despite a significant effect on head growth in comparison with an untreated group. Although it remains unclear to which extent changes in head circumference might reflect changes in brain growth during GH therapy, our data clearly show that an accelerated head growth during GH therapy does not necessarily result in improved cognitive abilities. We suggest that prenatal factors, which cannot be corrected by GH therapy or specialized help, have a persistent postnatal influence on the impaired brain development.

We do not have a satisfactory explanation for our rather unexpected finding of a significant increase in IQ scores in the untreated group of SGA children. In addition, another parameter evaluating cognitive development, i.e. school achievement, did not reveal differences between treated and untreated children: a similar percentage in both groups had repeated a grade in school at final evaluation. Although the clinical characteristics, including gestational age, birth weight, current head circumference and stature, and the socioeconomic background, defined by the educational and professional status of the mothers, of the two groups were comparable at baseline, potential interaction effects of a series of unknown intervening variables could have played a role in the changes of IQ scores from baseline to after 2 years. We suspect that among others, the seriousness or the kind of developmental delay might interact with the changes in IQ scores, since no significant increase in IQ scores was found in the untreated group anymore after exclusion of children with a developmental delay. Furthermore, the change of test instrument during the study could also have favored the positive evolution of IQ scores in the untreated group since a higher number had a change of test instrument, which was found to be related to a significant increase in PIQ scores. The absence of a beneficial effect of GH therapy on IQ scores in our study is in contradiction with the data of two other studies, reporting a significant increase in performance IQ scores in short SGA children.
after 2 and 8 years of GH therapy respectively (10, 12).

Some caution is, however, needed when comparing our results with these two studies, since they were both uncontrolled and included children of an older age range. Furthermore, only 2 of the 12 originally provided WISC-R subscales were used for evaluating the intellectual abilities in the long-term study.

In our study, one quarter of the children presented behavioral problems as assessed by the Child Behavior Checklist. At start of the study, i.e. before the instauration of GH therapy, children assigned to the untreated group showed slightly more somatic complaints, attention problems, and aggressive behavior than children assigned to the treated group. The higher prevalence of these specific behavioral problems could be due to the slightly higher number of very young children (3–5 years) in this group. Several items of the attention problems and the aggressive behavior subscale describe behavior or symptoms rather typical for young children (difficult to concentrate, difficult to sit still, impulsive, clumsy on the one hand and demands attention, noisy, angry, yells, stubborn, disobedient, on the other), but are disturbing for parents and as such easily perceived as behavioral problems. We found that after 2 years, attention problems and aggressive behavior scores increased in the treated group to a level comparable with the untreated group. It is unclear if and to which extent the constraints of daily injections could have played a role in the increase of aggressive problems in the treated group. Our CBCL data indicating problematic behavior in short SGA children irrespective of GH therapy are in line with the findings of van Pareren et al. (12), revealing higher CBCL scores for externalizing and total problems both at start and after 2 years of GH therapy. A potential explanation of the rather elevated aggressive behavior scores in the 2-year evaluation of SGA children could be a parental perception effect: at the age of 6–8 years children show more self-affirmation, stick more firmly to their own idea and will, which could be perceived as aggressive behavior by parents. Finally, it is worth mentioning that the only longitudinal study in short SGA children published up to now, did show an improvement in behavioral problem scores after 8 years of GH treatment (12).

This study certainly has his limitations. The inclusion of children with dysmorphic signs (although in limited number), the use of an older instrument for IQ testing (chosen because of unavailability of the WISC III in Belgium at the start of the study), the low number of subjects included (despite a prolonged inclusion period of 1.5 years) are among the most important limitations. We cannot exclude that beside the limited power of the study and the relative short time of follow-up for evaluating changes in neuropsychological development, the inclusion of patients with dysmorphic signs and mental retardation and the change of test instrument might have obscured a small positive effect of GH administration on cognitive and psychological functioning.

In conclusion, the findings of our study do not support an improvement of cognitive function and behavioral profile after 2 years of GH therapy in young prepubertal short SGA children. Follow-up examinations might reveal eventual long-term effects of GH therapy on development and behavior in this young cohort of SGA children, presenting a variable degree of developmental delay and behavioral problems at start of GH therapy.

Acknowledgements

This study was supported by a grant of the BSGPE. Growth hormone was provided by Pfizer Belgium. The authors express their thanks to all the members of the BSGPE, who invited patients to participate in this study. This work was supported by a research grant from the Belgian Study Group for Pediatric Endocrinology.

References


20 Harvey D, Prince J, Bunting J, Parkinson C & Campbell S. Abilities of children who were small-for-gestational age babies. Pediatrics 1982 60 296–300.


