

# Individualized treatment to optimize eventual cognitive outcome in congenital hypothyroidism

Jacoba J. Bongers-Schokking<sup>1</sup>, Wilma C.M. Resing<sup>2</sup>, Wilma Oostdijk<sup>3</sup>, Yolanda B. de Rijke<sup>4,5</sup> and Sabine M.P.F. de Muinck Keizer-Schrama<sup>1</sup>

**BACKGROUND:** In congenital hypothyroidism (CH), age-specific reference ranges (asRR) for  $fT_4$  and thyrotropine (TSH) are usually used to signal over/under-treatment. We compared the consequences of individual  $fT_4$  steady-state concentrations (SSC's) and asRR regarding over-treatment signaling and intelligence quotient at 11 y ( $IQ_{11}$ ) and the effect of early over-treatment with high L- $T_4$  dosages on  $IQ_{11}$ .

**METHODS:** Sixty-one patients (27 severe, 34 mild CH) were psychologically tested at 1.8, 6, and 11 y. Development scores were related to over-treatment in the period 0–24 mo, relative to either individual  $fT_4$  SSC's or asRR. Three groups were formed, based on severity of over/under-treatment 0–5 mo (severe, mild, and no over/under-treatment).

**RESULTS:**  $fT_4$  and TSH asRR missed 41–50% of the over-treatment episodes and consequently 22% of the over-treated patients, classified as such by  $fT_4$  SSC's. Severe over-treatment 0–5 mo led to lowered  $IQ_{11}$ 's and to a 5.5-fold higher risk of  $IQ_{11} < 85$  than other treatment regimes. Under-treatment had no effect on development scores. Initial L- $T_4$  dosages  $>10 \mu\text{g}/\text{kg}$  resulted in a 3.7-fold higher risk of over-treatment than lower dosages.

**CONCLUSIONS:** Data suggest that asRR, compared to  $fT_4$  SSC's, signal over-treatment insufficiently. Using  $fT_4$  SSC's and avoiding over-treatment may optimize cognitive outcome. Lowered  $IQ_{11}$ 's are usually a late complication of severe early over-treatment.

**C**ongenital hypothyroidism (CH) is basically a treatable disorder, with as major goal a normal cognitive development. Even after 40 y of experience, several questions regarding optimal treatment remain, such as how  $fT_4$  and TSH values should be evaluated (1), to what extent over-treatment (OT) and under-treatment (UT) impede normal cognitive development, how these conditions should be defined, and what the optimal dosages of *levothyroxine* (L- $T_4$ ) should be. Recent data demonstrated that high L- $T_4$  treatment may lead to lowered

cognitive outcome at 14 y, both in mild and severe CH (2). In contrast, several longitudinal studies during the 90's show that low L- $T_4$  dosages, in spite of late initiation of therapy and long periods of TSH elevation, may lead to a normal cognitive outcome (3–5). In a recent study, we found that OT during the first 24 mo is a greater risk for an adverse cognitive outcome at age 11 than UT (6). In that and a previous study, we also demonstrated that in CH  $fT_4$  is a better biomarker for euthyroidism than TSH (6,7). Treatment can be individualized by determining  $fT_4$  and TSH steady-state concentrations (SSC) that can provide individual target values and ranges (6,7). In the present study, we investigated whether the individual  $fT_4$  SSC's are more useful than age-specific  $fT_4$  and TSH reference ranges for the evaluation of follow-up values. Thereby, we evaluated the effect of OT during the first 5 mo, the period of the maximal cortical growth spurt, on eventual cognitive outcome at ages 1.8, 6, and 11 y. In addition we analyzed the impact of high initial L- $T_4$  dosages on initial  $fT_4$  concentrations.

## RESULTS

### Sample Classification by Individual $fT_4$ SSC Range and Age-Specific Reference Ranges

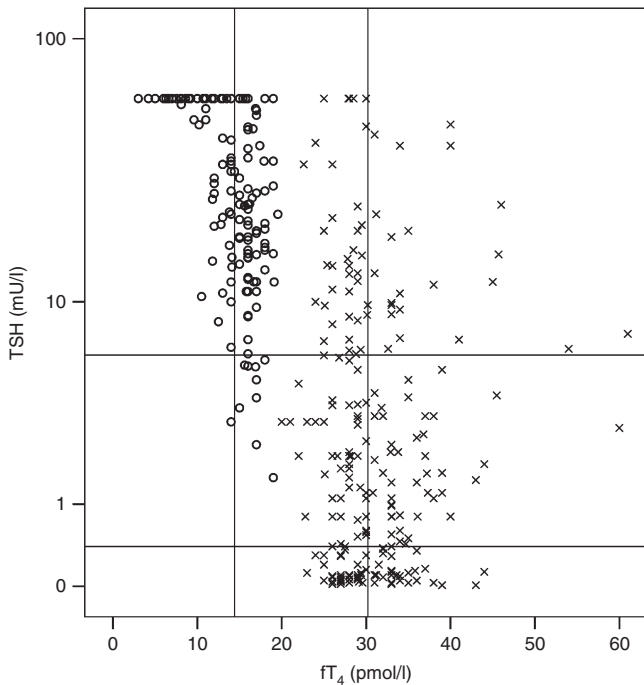
The 95% CI of  $fT_4$  and TSH age-specific reference ranges were 14.4–30.2 pmol/l and 0.42–6.02 mU/l, respectively. The proportion of euthyroid, OT, and UT follow-up samples of the in total 1170 samples from period 0–24 mo considerably differed for the three methods that were applied. The proportions for  $fT_4$  SSC's were 771, 219, and 180, respectively; for the age specific  $fT_4$  reference range 957, 110, and 103, respectively; and for the age specific TSH reference range 481, 130, and 559, respectively ( $P < 0.001$ ). In **Figure 1**, OT and UT control samples, relative to the individual  $fT_4$  SSC range, are shown and compared with the  $fT_4$  and TSH age-specific reference ranges. The  $fT_4$  and TSH age-specific reference ranges missed 50 and 41%, respectively, of the OT episodes, found by the  $fT_4$  SSC's. Consequently, the reference ranges recognized fewer OT patients than did the  $fT_4$  SSC's. Group OT/Over (**Table 1**)

<sup>1</sup>Department of Pediatrics, Sophia Children's Hospital Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands; <sup>2</sup>Department of Developmental and Educational Psychology, Faculty of Social Sciences, University of Leiden, Leiden, the Netherlands; <sup>3</sup>Department of Pediatrics, Leiden University Medical Center, Leiden, the Netherlands; <sup>4</sup>Department of Clinical Chemistry, Division of Endocrinology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands; <sup>5</sup>Department of Internal Medicine, Division of Endocrinology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands. Correspondence: Jacoba J. Bongers-Schokking ([abongers@worldonline.nl](mailto:abongers@worldonline.nl))

Received 5 March 2016; accepted 7 June 2016; advance online publication 21 September 2016. doi:10.1038/pr.2016.159

contained 38 patients, the group OT/No Over contained 11 patients, indicating that the  $ft_4$  age-specific reference range method missed 11 of the 49 over-treated patients (22%). In the repeated measurements ANOVA (Table 1), the interaction term  $age \times ft_4$  SSC was highly significant ( $P = 0.007$ ), while the interaction term  $age \times ft_4$  age-specific reference range was

insignificant ( $P = 0.190$ ). The analysis yielded lower intelligence quotients at age 11 ( $IQ_{11}$ 's) and significantly larger decreases in development scores in the period 1.8–11 y for the groups OT/Over and OT/No Over than for the groups No OT/Over and No OT/No Over.



**Figure 1.** Visualization of differences in sample classification between steady-state concentrations (SSC) method and age-specific reference range method. Symbols represent CH follow-up  $ft_4$  and TSH determinations of period 0–24 mo after initiation of therapy. 1,170 blood samples were drawn. Euthyroid determinations are omitted for better visualization: plotted are only those determinations that were classified relative to the  $ft_4$  SSC's as over-treated (OT, indicated with crosses) and under-treated (UT, indicated with open circles), defined as  $ft_4$  concentrations above/below the individual +2 SD/–2 SD range, respectively. Horizontal (log TSH) and vertical ( $ft_4$ ) solid lines represent age-specific reference lines, based on a cohort of 208 healthy infants aged 9 d–24 mo. The striking feature of Figure 1 is the large number of high TSH values in the OT samples, i.e., the failure of high  $ft_4$  levels to suppress TSH, illustrating the dysfunction in CH newborns of the TSH feedback system by high  $ft_4$  levels in the first 24 mo. Only 34% ( $n = 51$ ) of the total 152 OT samples with a TSH value > –2 SD originated from the initial period.

**Influence of Over and Under-Treatment on Development Scores**

The characteristics of the three OT 0–5 mo groups are shown in Table 2. Mean  $ft_4$  concentration and total OT duration were in the severe OT group higher than in the two other groups. The analysis of the developmental scores with the groups OT/UT 0–5 mo and OT/UT 6 mo to 6 y, as well as their interaction with age as predictors (Table 3), showed lower  $IQ_{11}$  scores for the five OT groups than for the No OT group. All five OT groups displayed, in contrast to the No OT group, decreasing cognitive scores in the period 1.8–11 y; statistically significant were the decreasing scores of the two severe OT 0–5 mo groups and that of the mild OT 0–5 mo with OT 6 mo to 6 y group. The lowest  $IQ_{11}$ 's were found for severe OT 0–5 mo without OT 6 mo to 6 y (73.1 (13.6)), the highest  $IQ_{11}$ 's for no OT in the period 0–6 y (112.5 (15.4)). Patients with severe OT 0–5 mo had a 5.5 (95% CI: 1.63; 18.21) fold higher risk of an  $IQ_{11} < 85$  than did the other patients ( $P = 0.006$ ). OT 0–5 mo was strongly associated with the decrease in development scores in the period 1.8–11 y ( $r = 0.433$ ,  $P < 0.001$ ), OT 6 mo to 6 y with the decrease in period 6–11 y ( $r = 0.340$ ,  $P = 0.007$ ) (Table 3). The same analysis for UT (Table 3) yielded differences in  $IQ_{11}$  scores and significant changes in development scores from 1.8 to 11 y, but they were not related to the severity of UT 0–5 mo. At the moment of testing at ages 1.8, 6, and 11 y 5, 3, and one in “n” patients, respectively, were hypothyroid, relative to the SSC method, and 1, 5, and 2 patients, respectively, hyperthyroid. Thyroid status at testing did not significantly influence overall test results (data not shown).

**Effect of Initial High and Low L-T<sub>4</sub> Dosages on  $ft_4$  Concentrations**

In Figure 2, the consequences of high and low initial L-T<sub>4</sub> substitution for  $ft_4$  concentrations in the initial period are illustrated. With low initial dosages,  $ft_4$  concentrations in most cases rose slowly to their individual SSC range; with high initial dosages, they often exceeded this range. The risk of OT in the first three follow-up samples ( $n = 183$ ) was 3.7 (95% CI: 1.4; 9.3) ( $P = 0.007$ )

**Table 1.** Effect of two different sample classification methods (SSC method and age specific reference range method) on eventual cognitive outcome (mean (SEM))\*<sup>a</sup>

Classification method			Development scores			Change in development scores	
SSC <sup>b</sup>	asRR <sup>c</sup>	N	1.8 y	6 y	11 y	1.8–11 y	P <sup>d</sup>
OT	Over	38	104.5 (1.8)	102.2 (2.9)	88.6 (2.5)	–15.8 (3.0) (–21.8; –9.9)	<b>&lt;0.001</b>
OT	No Over	11	102.8 (3.2)	100.5 (5.1)	89.8 (4.2)	–13.0 (5.1) (–23.2; –2.7)	<b>0.014</b>
No OT	Over	3	102.2 (5.8)	108.1 (7.1)	106.6 (6.8)	+4.4 (5.4) (–6.4; +15.2)	0.417
No OT	No Over	9	96.3 (3.6)	102.3 (5.3)	100.8 (4.8)	+4.5 (5.4) (–6.4; +15.3)	0.414
P <sup>d</sup>			0.257	0.798	0.068	<b>0.007</b>	

\*Significance for the whole model:  $P = 0.036$ ; significance for  $age \times SSC$ :  $P = 0.007$ ; significance for  $age \times asRR$ :  $P = 0.190$ .  
<sup>a</sup>Development scores are corrected for type of CH (effect size –11.6 ( $P < 0.001$ )) and early/late start of treatment (effect size –7.4 ( $P \leq 0.004$ )). <sup>b</sup>SSC =  $ft_4$  SSC method.  
<sup>c</sup>asRR =  $ft_4$  age-specific reference range method. <sup>d</sup>Bold P values are significant.  
 CH, congenital hypothyroidism; OT, over-treated; SSC, steady-state concentrations; UT, under-treated.

**Table 2.** Characteristics of the OT 0–5 mo groups (mean (SD))

	OT 0–5 mo groups			P <sup>a</sup>
	Severe OT	Mild OT	No OT	
N	18	25	18	
Severe/mild CH	7/11	11/14	8/10	n.s.
Pretreatment fT <sub>4</sub> (pmol/l)	7.0 (5.6)	5.4 (3.7)	5.6 (3.6)	n.s.
Pretreatment TSH mUI/l	54.4 (12.8)	59.8 (1.2)	60.0 (0.0)	n.s.
fT <sub>4</sub> SSC (pmol/l)	20.1 (2.0)	21.8 (2.7)	22.8 (2.3)	n.s.
Initial L-T <sub>4</sub> dose (µg/kg)	8.6 (2.4)	9.1 (2.0)	7.8 (2.1)	n.s.
Maximal initial L-T <sub>4</sub> dose (µg/kg)	9.4 (1.8)	9.5 (2.1)	8.2 (2.0)	n.s.
Mean fT <sub>4</sub> 0–5 mo (pmol/l)	26.8 (9.4)	23.3 (6.5)	18.7 (5.0)	<b>&lt;0.001</b>
Mean TSH 0–5 mo (mUI/l)	17.4 (22.0)	16.0 (20.0)	29.9 (22.6)	n.s.
Mean L-T <sub>4</sub> dose 0–5 mo (µg/kg)	6.9 (1.3)	6.8 (1.4)	6.9 (1.9)	n.s.
Total duration OT 0–5 mo (months)	2.4 (0.9)	0.9 (0.5)	0	<b>&lt;0.001</b>
Total duration UT 0–5 mo (months)	0.2 (0.4)	0.4 (0.7)	0.8 (0.9)	n.s.
OT Severity factor 0–5 mo (SDS*months)	6.4 (3.3)	0.6 (0.5)	0	<b>&lt;0.001</b>
OT Severity factor 6 mo to 6 y (SDS*months)	6.5 (11.3)	1.6 (4.1)	3.6 (7.1)	n.s.
UT Severity factor 0–5 mo (SDS*months)	0.8 (1.6)	0.6 (0.7)	1.0 (1.4)	n.s.
UT Severity factor 6 mo to 6 y (SDS*months)	1.9 (7.4)	1.9 (3.7)	1.3 (3.5)	n.s.

<sup>a</sup>Bold P values are significant; n.s., not significant. CH, congenital hypothyroidism; OT, over-treated; SSC, steady-state concentrations; UT, under-treated.

**Table 3.** Mean development scores (SEM) at three ages and mean changes in development scores from 1.8 to 11 y (SEM) (95% CI) for the OT and UT groups 0–5 mo and 6 mo to 6 y\*

OT 0–5 mo	OT 6 mo to 6 y	N	Development scores			Change in development scores <sup>a</sup>	
			1.8 y	6 y	11 y	1.8–11 y	P <sup>b</sup>
Severe	Yes	13	104.1 (3.0)	105.1 (4.8)	90.3 (4.4)	-14.9 (4.8) (-25.4; -4.5)	<b>0.009</b>
Severe	No	5	114.3 (4.8)	97.0 (7.2)	73.1 (6.1)	-41.2 (6.8) (-35.6; -26.8)	<b>&lt;0.001</b>
Mild	Yes	19	105.0 (2.4)	101.5 (4.3)	89.8 (3.1)	-15.2 (3.3) (-21.9; -8.5)	<b>&lt;0.001</b>
Mild	No	6	97.9 (4.4)	94.6 (7.0)	96.1 (5.6)	-1.8 (6.2) (-21.0; +17.3)	0.789
No OT	Yes	13	97.8 (2.9)	101.3 (5.2)	90.6 (3.9)	-7.0 (4.1) (-15.0; +1.3)	0.094
No OT	No	5	99.4 (4.9)	102.9 (7.3)	112.5 (6.9)	+13.2 (6.9) (-0.9; +27.3)	0.066
<i>P**<sup>b</sup></i>			0.061	0.849	<b>0.004</b>	<b>&lt;0.001</b>	
UT 0–5 mo	UT 6 mo to 6 y	N	1.8 y	6 y	11 y	1.8–11 y	P <sup>b</sup>
Severe	Yes	5	101.2 (4.6)	108.8 (6.5)	90.2 (7.5)	-11.1 (8.7) (-28.5; +6.4)	0.210
Severe	No	4	104.1 (5.0)	97.5 (7.2)	83.9 (8.3)	-13.4 (9.2) (-33.0; +6.1)	0.163
Mild	Yes	25	97.2 (2.1)	98.8 (3.2)	91.0 (3.2)	-6.3 (3.7) (-14.0; +1.4)	0.103
Mild	No	7	104.2 (3.9)	115.0 (6.8)	97.4 (5.7)	-6.8 (7.2) (-21.8; +8.1)	0.355
No UT	Yes	11	111.2 (3.2)	96.2 (5.0)	91.8 (4.1)	-19.5 (5.8) (-31.4; -7.5)	<b>0.003</b>
No UT	No	9	110.3 (6.6)	108.6 (5.2)	93.1 (5.2)	-17.2 (6.7) (-31.0; -3.4)	<b>0.017</b>
<i>P**<sup>b</sup></i>			0.072	0.131	0.803	0.165	

\*P = 0.001 for interaction term age\*OT 0–5 mo\*OT 6 mo to 6 y; P = 0.077 for the interaction term age\*UT 0–5 mo\*UT 6 mo to 6 y. \*\*P-value for test of equal mean development score or change in development score in all groups.

<sup>a</sup>Development scores are corrected for type of congenital hypothyroidism (effect size -11.6 (P < 0.001) and early/late start of treatment (effect size -7.4 (P ≤ 0.004)). <sup>b</sup>Bold P values are significant.

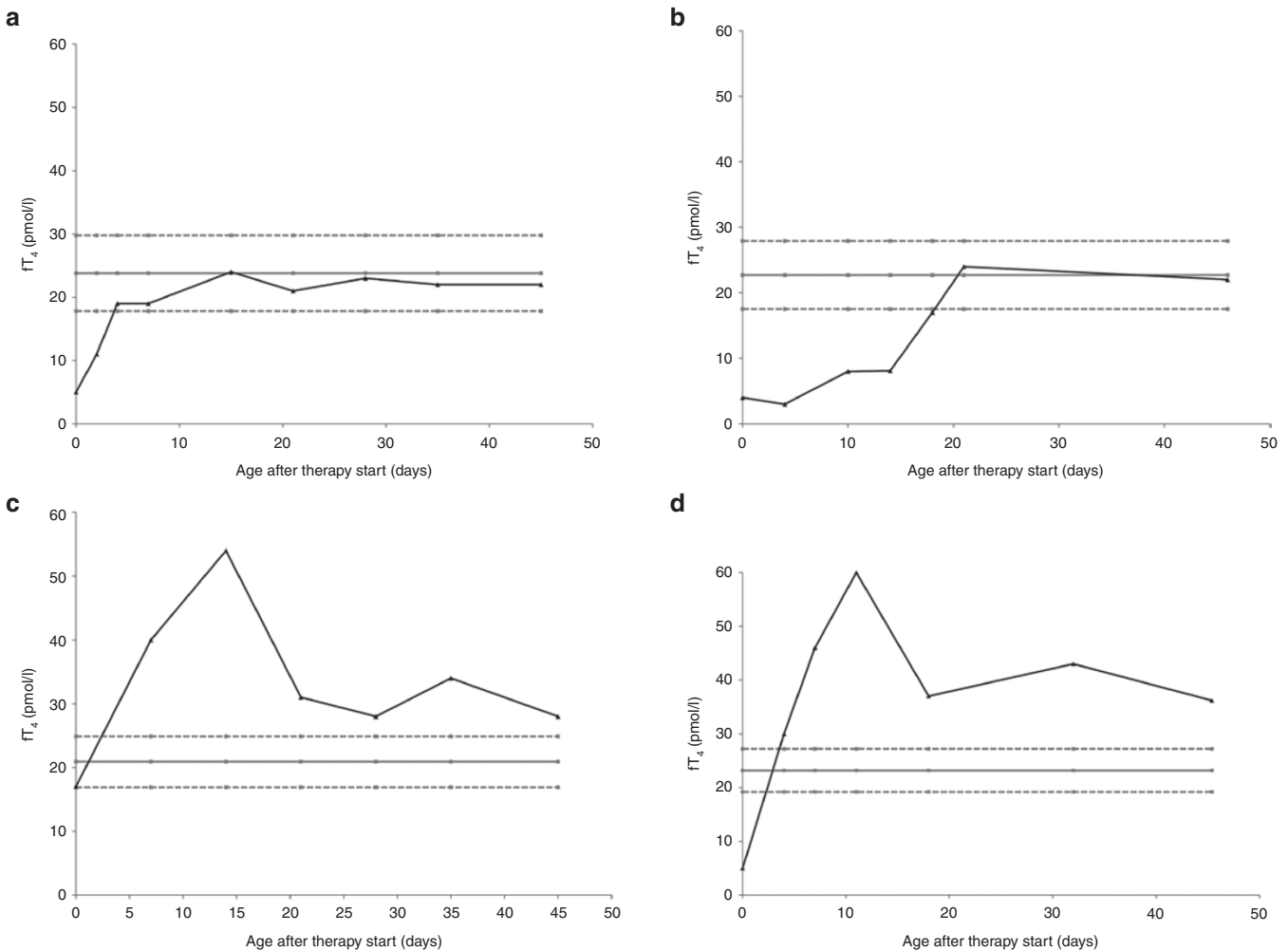
fold higher for initial L-T<sub>4</sub> dosages > 10 µg/kg than for dosages ≤ 10 µg/kg. Free T<sub>4</sub> concentrations in these samples, collected at ages 4.4 (2.1), 9.2 (3.9), and 17.0 (10.3) days after therapy initiation, showed a positive relation with initial L-T<sub>4</sub> dose (P < 0.001)

and a negative relation with minimal initial fT<sub>4</sub> deficit (P < 0.001). **Table 4** shows the mean fT<sub>4</sub> concentrations (in SDS) (95% CI) relative to the individual SSC's, obtained by substituting ascending initial L-T<sub>4</sub> dose values (µg/kg) and minimal

**Table 4.** Free  $T_4$  concentrations (in SDS) of the first three follow-up samples, at 4.4, 9.2, and 17.0 days, by initial L- $T_4$  dose and minimal initial  $fT_4$  deficit\*

Initial L- $T_4$ dose ( $\mu\text{g}/\text{kg}$ )	Minimal initial $fT_4$ deficit (pmol/l)		
	5	10	15
5	-0.7 (-2.4; 1.1)	-2.3 (-3.9; -0.7)	-4.0 (-5.4; -2.5)
7.5	<b>1.0 (-0.7; 2.7)</b>	0.7 (-2.1; 0.8)	-2.3 (-3.7; -0.9)
10	<b>2.7 (1.2; 4.1)</b>	<b>1.0 (-0.1; 2.0)</b>	0.7 (-1.7; 0.5)
12.5	<b>4.3 (1.9; 6.7)</b>	<b>2.7 (0.3; 5.1)</b>	<b>1.0 (-1.0; 3.0)</b>
15	<b>5.9 (3.6; 8.4)</b>	<b>4.3 (1.9; 6.7)</b>	<b>2.7 (0.7; 4.7)</b>

\*Fields with bold figures: increased risk of over-treatment, i.e.,  $fT_4$  concentration > +2 SDS, relative to the individual SSC's of the patients. SSC, steady-state concentrations.



**Figure 2.** Examples of  $fT_4$  concentrations after low and high initial L- $T_4$  dosages. Four examples of individual initial  $fT_4$  curves, two examples of cases with low initial L- $T_4$  dosages and without initial OT (panel **a**: mild CH, initial L- $T_4$  dose 7.2  $\mu\text{g}/\text{kg}$ , panel **b**: severe CH, initial L- $T_4$  dose 6.0  $\mu\text{g}/\text{kg}$ ) and two examples of cases with high initial L- $T_4$  dosages and with initial OT (panel **c**: mild CH, initial L- $T_4$  dose 10.3  $\mu\text{g}/\text{kg}$ , panel **d**: severe CH, initial L- $T_4$  dose 11.2  $\mu\text{g}/\text{kg}$ ). Black lines: individual  $fT_4$  concentration, gray lines: mean individual  $fT_4$  SSC  $\pm$  2 SD, determined with  $fT_4$  control values after the initial period (i.e., > 1.5 mo until 11 y). SSC, steady-state concentrations.

initial  $fT_4$  deficit values (pmol/l) in the regression equation:  $fT_4$  (in SDS) =  $-2.30 + 0.66 \cdot \text{initial L-}T_4 \text{ dose} - 0.33 \cdot \text{minimal initial } fT_4 \text{ deficit}$ . **Table 4** shows which combinations of initial L- $T_4$  dosage and minimal initial  $fT_4$  deficit (bold figures) may lead to initial over-treatment, i.e., to  $fT_4$  values > +2 SDS, relative to the individual SSC's.

**DISCUSSION**

The results of our study on CH treatment suggest that adequate signaling and avoidance of over-treatment are of great importance and that this can best be done with  $fT_4$  SSC's. In our study, TSH and  $fT_4$  age-specific reference ranges, recommended in the recent ESPE guidelines (1), missed, relative to

the  $fT_4$  SSC's, 41 and 50%, respectively, of the over-treatment episodes and 22% of the over-treated patients, because reference ranges are too wide to signal all over-treatment episodes (Figure 1). This had consequences for eventual cognitive outcome. The 11 over-treated patients, missed by the  $fT_4$  age-specific reference range, had significantly decreasing development scores from 1.8 to 11 y, suggesting real over-treatment, while the three patients recognized by the  $fT_4$  age-specific reference range as over-treated, but not so by the SSC's, as well as the nine patients classified by both methods as not-over-treated, did not show decreasing development scores in this period. Our conclusion is that reference ranges in CH, with only two fixed values for upper and lower limit, will never be able to properly distinguish between normal or abnormal for the individual patient, whatever values are taken. Our data indicate that severe over-treatment during the first 5 mo leads to normal development scores at 1.8 and 6 y, but to considerably lowered IQ's at 11 y. Patients with severe over-treatment 0–5 mo had a 5.5-fold higher risk of an  $IQ_{11} < 85$  than those with mild over-treatment 0–5 mo or without over-treatment. Over-treatment in period 6 mo to 6 y seemed to contribute to the IQ decrease from 6 to 11 y. For under-treatment, no negative effect on cognitive development scores could be established. Patients treated with initial L-T<sub>4</sub> dosages > 10 µg/kg had a 3.7-fold higher risk of over-treatment within the first 17 d than those with lower initial dosages had. These data suggest that during the first neonatal months, CNS development is very sensitive to over-treatment with thyroid hormone, more sensitive than to later over-treatment. From animal models (8–12), we know that early over-treatment may enhance neural growth and development and, when over-treatment continues, it even may lead to an irreversible arrest in neural development. A similar development was seen in our severely over-treated patients: at age 1.8 y they had normal to above normal development scores, but thereafter decreasing scores. In contrast, under-treatment in this period may lead to a delayed, but reversible neural development with later catch up (13). In CH, initial low-dose treatment seems preferable to initial high-dose treatment, in contrast to the advice in most recent studies (14–18) and ESPE guidelines (1). A slow  $fT_4$  rise with under-treatment seems more preferable than a fast rise with over-treatment (Table 4), because we earlier demonstrated that a fast TSH normalization does not lead to higher  $IQ_{11}$ 's (6). Low-dose treatment allows  $fT_4$  to gradually rise to its SSC value and to remain on that level, if medication is not importantly changed. A further advantage of initial low-dose treatment is that individual  $fT_4$  SSC can be estimated during the initial period (7). This is not possible with initial high-dose treatment, because  $fT_4$  values are often not yet stable then. Initial high-dose treatment was introduced in the 80's, in order to shorten the period of neonatal hypothyroidism (16) and L-T<sub>4</sub> doses between 10 and 14 µg/kg were considered safe and effective to do so (17). The normal to above-normal cognitive development scores in CH infants treated with a higher L-T<sub>4</sub> dose at young age, found by us as well as by others (14,15,18,19), strengthened the conviction,

also emphasized in the ESPE guidelines (1) that high L-T<sub>4</sub> dosages and maintenance of TSH in the normal range are necessary for a normal eventual outcome. The sharp decline in IQ values from 6 to 11 y in our study is unusual and therefore unexpected, because in healthy subjects IQ at age 6 is a good predictor for IQ at ages 11 and 13 (20,21). The longitudinal study of Rovet, showing decreasing development scores in CH children from infancy to adolescence (22), supports our data. A consequence of the changing cognitive scores over time is that no meaningful conclusions on eventual cognitive outcome can be drawn from data that do not include adolescent patients (14,15,18,19). The results of the only so far published transversal study on adolescent CH patients, treated according to the recent guidelines with low TSH values during the first year (mean 1.0 mU/l) and demonstrating lowered cognitive outcomes at age 14, irrespective the type of CH (2), are in concert with our results. Moreover, the longitudinal studies during the 90's (3–5) showing normal cognitive outcomes in initially low treated patients, in spite of long periods of under-treatment and late initiation of therapy, support our normal cognitive outcomes in the not-over-treated patients. These last studies also demonstrate that low initial L-T<sub>4</sub> dosages may result in normal adolescent IQ's.

The most important issues in CH treatment are early initiation of therapy and avoidance of over-treatment. For that purpose, frequent monitoring and L-T<sub>4</sub> dose adaptation and correct interpretation of the  $fT_4$  and TSH concentrations are needed. It is recommended to start with a low L-T<sub>4</sub> dose, taking the estimated minimal initial  $fT_4$  deficit into account (Table 4), thus minimizing the risk of over-treatment, i.e., an  $fT_4$  concentration > 2 SDS, and to take a first follow-up sample at day 3 or 4. L-T<sub>4</sub> dosage can then be decreased, when increase of the initial  $fT_4$  concentration is such that for the next days a rise to above 30 pmol/l can be foreseen. We advise to collect follow-up samples in the period 10–60 d at 10–14 d intervals. Once  $fT_4$  has stabilized,  $fT_4$  SSC can be estimated, also when TSH is still elevated, providing a target value to aim at (7). With five samples, mean  $fT_4$  SSC can be estimated with 95% CI at a precision level of 90% (7). In further follow-up, when more  $fT_4$  determinations, valid for the assessment of the  $fT_4$  SSC, i.e., within TSH interval 0.5–10 mU/l, become available, the individual  $fT_4$  SSC range can be established. Such range is crucial for the signaling of over and under-treatment. A correct assessment of the thyroid status by  $fT_4$  is in CH even more important than in healthy subjects, because TSH in CH is less predictable (6,7) (Figure 1) and thus less informative about the thyroid status than in the normal situation. TSH SSC's are less well defined, with a larger variation around the individual mean SSC (Table 5), and signal overtreatment less well than  $fT_4$  SSC's. Therefore,  $fT_4$  should be the most important factor in guiding decisions on sample evaluation and therapy, more important than TSH. Two-thirds of our CH patients had a TSH SSC above the normal range, i.e., > 2.5 mU/l (7). This means that for the majority of them the lower limit of TSH will be above 0.05 mU/l, the lower limit indicated in the ESPE guidelines (1). We therefore recommend keeping TSH in the

**Table 5.** Mean values (SD) (95% CI) of  $fT_4$  and TSH SSC's of the 61 congenital hypothyroidism patients of the study group

	Age interval <sup>a</sup> (month)	N samples per patient	Mean SSC (SD) (95% CI)	CV <sup>b</sup> (%)
Postinitial $fT_4$ SSC (pmol/l)	1.5–132	21.0	21.7 (2.6) (16.6;28.7)	12.0 (3.9)
Initial $fT_4$ SSC (pmol/l)	0.2–1.5	4.6	22.5 (3.0) (16.2;27.3); $P = 0.07^c$	13.2 (6.3)
Postinitial TSH SSC (mUI/l)	1.5–132	21.0	3.3 (1.9) (0.9; 5.7)	59.9 (13.6)

<sup>a</sup>Age after initiation of therapy. <sup>b</sup>Mean of the coefficients of variation (CV) within the individual SSC's. <sup>c</sup>Difference between initial SSC's and SSC's for interval 1.5–132 mo. SSC, steady-state concentrations.

upper half of the TSH SSC range (3.3–5.9 mUI/l), in order to avoid over-treatment.

A drawback of our study is that  $fT_4$  measurements were performed in 27 different laboratories. The adult reference ranges of the laboratories varied, but the differences were so small that we accepted these differences. Furthermore, the  $fT_4$  values of the patients, with a few exceptions determined in their own local laboratory throughout the study, were expressed as SDS value of the individual SSC's, making them independent of the used  $fT_4$  determination method. Therefore, the various  $fT_4$  determination methods did not change the eventual conclusions. Another drawback of our study is its limited sample size in a cohort with various different treatment schedules, resulting in small treatment groups. The diversity of treatment methods, however, allowed a comparison between the effect of low and high treatment on eventual cognitive outcome at age 11, which would not have been possible in a cohort with a single treatment protocol. A future cohort study with more CH patients, treated with different protocols and repeatedly psychologically tested until adulthood, might reveal the most optimal treatment protocol. The feasibility of such study, however, is uncertain in view of ethical issues.

## METHODS

### Subjects

The study group consisted of 61 children (19 boys) with CH, born between February 1993 and July 1996, and followed from birth until 11 y, the same cohort described previously (6,7,23,24). The study was approved by the privacy committee of the CH Screening Board and by the Erasmus MC review board. The original cohort, recruited by questionnaires sent out by the national CH coordination office to all local pediatricians, consisted of 98 patients, of which 71 patients were selected and 27 were excluded: four infants of mothers with known thyroid abnormalities, five infants at risks for abnormal psychomotor development (asphyxia, meningitis, chromosomal abnormality, severe hypotonia syndrome, prematurity), one infant that died, and 17 patients with mild CH, because testing capacity was limited and we expected more developmental abnormalities in severe CH than in mild CH (6,23,24). After written informed parental consent, the study group consisted of 27 children with severe CH and 34 with mild CH. Severe CH was defined as total incapability to produce  $T_4$ , mild CH as partial incapability (6,8). All patients were psychologically tested at 1.8 (0.4) years (mean (SD)), 46 at 6.0 (0.6) years, and 55 at 11.0 (0.7) years. Reasons for not re-examining all patients at ages 6 and 11 y have been described previously (6,9).

### Treatment and Monitoring

The local pediatrician made all decisions on treatment regimen. Until 1991, an initial daily L- $T_4$  dose of 5–6  $\mu\text{g}/\text{kg}$  with an upper limit for  $T_4$  of 180 nmol/l was recommended; in 1992, the recommendation changed to 9–12.5  $\mu\text{g}/\text{kg}$  for the initial L- $T_4$  dose and 35 pmol/l for the upper limit of  $fT_4$ . Follow-up data on  $fT_4$ , TSH, and L- $T_4$  dosages ( $n = 2,524$ ) were collected. Only 1.1% of the samples were suspicious for bad adherence to medication. On average 31 (range 18–78)

control measurements per patient were performed in the period 0–6 y, of which 10 (range 5–17) were performed in the period 0–5 mo. Two groups were formed based on the timing of treatment start: early  $\leq 12$  d ( $n = 28$ ), and late  $> 12$  d ( $n = 33$ ), and two groups based on initial daily L- $T_4$  dosage: low ( $\leq 10$   $\mu\text{g}/\text{kg}$ ,  $n = 40$ ) and high ( $> 10$   $\mu\text{g}/\text{kg}$ ,  $n = 21$ ). Mean initial L- $T_4$  dosage was 8.6 (2.2)  $\mu\text{g}/\text{kg}$ . In 17 patients, initial L- $T_4$  was increased from 7.5 (2.5) to 9.4 (2.3)  $\mu\text{g}/\text{kg}$  during the initial period,  $\leq 1.5$  mo after therapy start. Maximal initial L- $T_4$  dose was defined as either the increased dose, or the initial dose, when L- $T_4$  dosage was not increased.

### Laboratory Tests

Free  $T_4$  and TSH were measured in 27 different laboratories, whose quality of measurement is surveyed by the National External Quality Assurance Program (the Netherlands, SMLK) (25). Adult reference values were 11–25 pmol/l for  $fT_4$  and for TSH, at the time of this study, 0.4–4.3 mUI/l.

### Individual Steady-State Concentrations (SSC's)

Individual  $fT_4$  and TSH steady-state concentrations (SSC's) were defined as those concentrations at which both  $fT_4$  and TSH are constant, despite ongoing processes that might change them, such as medication, thyroxine clearance, and alterations in the equilibrium between the different deiodinase pathways (26). In the euthyroid situation, individual  $fT_4$  and TSH SSC's are in equilibrium with each other. In healthy, euthyroid subjects and subjects with stable thyroid disease, individual SSC's represent the mean value of a series of determinations under standard conditions, with an individual SSC range that is about half that of the total population (27–29). For the assessment of SSC's in CH, those samples were included that we assumed to represent a stable and euthyroid state, obtained in the postinitial period (i.e.,  $> 1.5$  mo after therapy start until 11 y) that were within the TSH interval 0.5–10 mUI/l. From the thus obtained series of determinations, those samples were excluded, which were outside the 95% of this series. Included were only samples that were inside both the individual  $fT_4$  as well as the individual TSH SSC range (Table 5). As described previously, individual  $fT_4$  and TSH SSC's decreased slightly in the period 0–11 y, and were largely independent of the L- $T_4$  dosage (7). In the initial period,  $fT_4$  SSC's were somewhat higher than in the postinitial period, without reaching significance: 22.5 (3.0) vs. 21.7 (2.6) pmol/l, respectively ( $P = 0.07$ ) (7). Minimal initial  $fT_4$  deficit (pmol/l), determined for each patient, was the difference between  $-2$  SD value of the 95% CI of the  $fT_4$  SSC range of the whole cohort, i.e., 16.6 pmol/l (Table 5), and individual pretreatment  $fT_4$ .

### Comparison of $fT_4$ SSC's With Age-Specific $fT_4$ Reference Ranges

1,170 follow-up samples were collected from 61 patients over the first 24 mo after therapy start and the measured  $fT_4$  values were classified as euthyroid, when  $fT_4$  concentration was within the individual  $fT_4$  SSC  $\pm 2$  SD range, and as OT or UT, when it was above or below this range. This classification was compared with the classification based on the 95% CI of age-specific  $fT_4$  and TSH reference ranges, constructed from 208 samples of healthy infants, aged 9 d to 24 mo, admitted to our hospital for small surgical procedures (30). Over-treatment relative to the age-specific  $fT_4$  reference range was defined as an  $fT_4$  concentration  $> 95\%$  CI. For each classification, two groups were formed, based on whether or not over-treatment had occurred in period 0–24 mo: the groups OT and No OT for the  $fT_4$  SSC's, and the groups Over and No Over for the  $fT_4$  age-specific reference range. The groups were combined in a contingency table: OT/Over, OT/No Over, No OT/Over, and No OT/No Over.

### Severity of Over and Undertreatment

The severity of the OT/UT episodes, defined as the product of the distance between the  $fT_4$  value and the individual  $fT_4$  SSC +2 SD line/−2SD line (in SDS) and the duration of the OT/UT episode (in months), resulting in the unity SDS.mo, was calculated for each episode, for the period 0–6 y. The severity factor (SF) per patient, being the sum of the severity determinations, was calculated for age periods 0–5 mo and 6 mo to 6 y. Three of the following groups were formed based on SF OT/UT 0–5 mo: (i) Severe, when >2 SDS.month; (ii) Mild, when between 0 and 2 SDS.months; and (iii) No OT/UT, when SF was 0. For the period 6 mo to 6 y, two groups were formed: Yes or No OT/UT.

### Development Tests

At age 1.8 y, the infants were tested with the Bayley Development Scale (31); only the mental development index (MDI) scores was used in the analysis. At ages 6 and 11 y, the short version of the Revised Amsterdam Child Intelligence Test (RAKIT) (32,33) was applied, a reliable ( $r > 0.95$ ), normed, Dutch child intelligent test for children aged 4–11.6 y. IQ scores across the ages are stable. The RAKIT-short consists of the following six subtests: Exclusion (visual reasoning); Word Knowledge (passive knowledge of word meaning); Disc Placement (visual-spatial-motor task); Learning Names (verbal-memory-association task); Hidden Figures (visual analysis); and Word Fluency (fast word production). Around all test moments, thyroid function was checked. At all ages, control groups of healthy children were included: 207 at age 1.8 (8), 37 at age 6 (9), and 25 at age 11 (6). Cognitive development scores for study group and control groups were corrected for socio-economic status (SES) and ethnicity, as these characteristics might influence scores (32,34,35). After correction, study group scores were standardized by the scores of the control groups that were set at a mean of 100 and SD of 15.

### Statistical Analysis

Statistical analysis was done with SPSS version 23.0 (SPSS, Chicago, IL). Chi-square tests were used to analyze differences in sample classification between the  $fT_4$  SSC's and the reference ranges. A repeated measurements ANOVA was used to compare the cognitive development scores of the four contingency table groups over time. In this analysis, the four groups and age were the independent variables, cognitive development scores at ages 1.8, 6, and 11 y the dependent variable. This analysis gives unbiased results even in case of (randomly) missing values. In all analyses, development scores were adjusted for type of CH (mild/severe) and start of treatment (early/late). An ANOVA procedure was applied comparing the clinical characteristics of the 0–5 mo treatment groups. A repeated measurements ANOVA was used to estimate the effect of the factors OT/UT 0–5 mo and OT/UT 6 mo to 6 y on the cognitive development scores at three ages. Logistic regressions were applied to determine the risk of an  $IQ_{11} < 85$  for the OT 0–5 mo groups and the risk of OT by high initial dosages of L- $T_4$ . Selected for this last analysis were the first three control samples per patient after initiation of therapy ( $n = 183$ ). Lastly, a linear regression was used to analyze the dependency of  $fT_4$  concentrations on initial L- $T_4$  dosage and minimal initial  $fT_4$  deficit. A two-tailed significance of  $P < 0.05$  was set for all tests.

### STATEMENT OF FINANCIAL SUPPORT

This study was financially supported by unrestricted grants from the Theia Foundation, Jan Dekker & Ludgardine Bouwman Foundation, and Erasmus University Rotterdam Association Trust Fund.

Disclosure: None of the authors have competing financial interests.

### REFERENCES

- Léger J, Olivieri A, Donaldson M, et al.; on behalf of the ESPE-PES-SLEP-JSPE-APEG-APPES-ISPAE, and the Congenital Hypothyroidism Consensus Conference Group. Consensus guidelines on screening, diagnosis, and management of congenital hypothyroidism. *J Clin Endocrinol Metab* 2014; 99: 363–384.
- Dimitropoulos A, Molinari L, Etter K, et al. Children with congenital hypothyroidism: long-term intellectual outcome after early high-dose treatment. *Pediatr Res* 2009;65:242–8.
- New England Congenital Hypothyroidism Collaborative. Correlation of cognitive test scores and adequacy of treatment in adolescents with congenital hypothyroidism. *J Pediatr* 1994;124:383–387.
- Simons WF, Fuggle PW, Grant DB, Smith I. Intellectual development at 10 years in early treated congenital hypothyroidism. *Arch Dis Child* 1994;71:232–4.
- Salerno M, Militerni R, Di Maio S, Bravaccio C, Gasparini N, Tenore A. Intellectual outcome at 12 years of age in congenital hypothyroidism. *Eur J Endocrinol* 1999;141:105–10.
- Bongers-Schokking JJ, Resing WC, de Rijke YB, de Ridder MA, de Muinck Keizer-Schrama SM. Cognitive development in congenital hypothyroidism: is overtreatment a greater threat than undertreatment? *J Clin Endocrinol Metab* 2013;98:4499–506.
- Bongers-Schokking JJ, de Ridder MA, de Rijke YB, de Muinck Keizer-Schrama SM. Experience in treating congenital hypothyroidism: implications regarding free thyroxine and thyrotropin steady-state concentrations during optimal levothyroxine treatment. *Thyroid* 2013;23:160–5.
- Davenport JW, Gonzalez LM. Neonatal thyroxine stimulation in rats: accelerated behavioral maturation and subsequent learning deficit. *J Comp Physiol Psychol* 1973;85:397–408.
- Stone JM, Greenough WT. Excess neonatal thyroxine: effects on learning in infant and adolescent rats. *Dev Psychobiol* 1975;8:479–88.
- Marta CB, Adamo AM, Soto EF, Pasquini JM. Sustained neonatal hyperthyroidism in the rat affects myelination in the central nervous system. *J Neurosci Res* 1998;53:251–9.
- Yilmazer-Hanke DM, Hantsch M, Hanke J, Schulz C, Faber-Zuschratter H, Schwegler H. Neonatal thyroxine treatment: changes in the number of corticotropin-releasing-factor (CRF) and neuropeptide Y (NPY) containing neurons and density of tyrosine hydroxylase positive fibers (TH) in the amygdala correlate with anxiety-related behavior of wistar rats. *Neuroscience* 2004;124:283–97.
- Fernandez M, Pironi S, Manservigi M, Giardino L, Calzà L. Thyroid hormone participates in the regulation of neural stem cells and oligodendrocyte precursor cells in the central nervous system of adult rat. *Eur J Neurosci* 2004;20:2059–70.
- Rodriguez-Peña A, Ibarrola N, Iñiguez MA, Muñoz A, Bernal J. Neonatal hypothyroidism affects the timely expression of myelin-associated glycoprotein in the rat brain. *J Clin Invest* 1993;91:812–8.
- Selva KA, Harper A, Downs A, Blasco PA, Lafranchi SH. Neurodevelopmental outcomes in congenital hypothyroidism: comparison of initial T4 dose and time to reach target T4 and TSH. *J Pediatr* 2005;147:775–80.
- Albert BB, Heather N, Derraik JG, et al. Neurodevelopmental and body composition outcomes in children with congenital hypothyroidism treated with high-dose initial replacement and close monitoring. *J Clin Endocrinol Metab* 2013;98:3663–70.
- Fisher DA, Foley BL. Early treatment of congenital hypothyroidism. *Pediatrics* 1989;83:785–9.
- Germak JA, Foley TP Jr. Longitudinal assessment of L-thyroxine therapy for congenital hypothyroidism. *J Pediatr* 1990;117(2 Pt 1):211–9.
- Soliman AT, Azzam S, Elawwa A, Saleem W, Sabt A. Linear growth and neurodevelopmental outcome of children with congenital hypothyroidism detected by neonatal screening: A controlled study. *Indian J Endocrinol Metab* 2012;16:565–8.
- Salerno M, Militerni R, Bravaccio C, et al. Effect of different starting doses of levothyroxine on growth and intellectual outcome at four years of age in congenital hypothyroidism. *Thyroid* 2002;12:45–52.
- Moffitt TE, Caspi A, Harkness AR, Silva PA. The natural history of change in intellectual performance: who changes? How much? Is it meaningful? *J Child Psychol Psychiatry* 1993;34:455–506.
- Marcus Jenkins JV, Woolley DP, Hooper SR, De Bellis MD. Direct and Indirect Effects of Brain Volume, Socioeconomic Status and Family Stress on Child IQ. *J Child Adolesc Behav* 2013;1:1.
- Rovet JF. Congenital hypothyroidism: long-term outcome. *Thyroid* 1999;9:741–8.
- Bongers-Schokking JJ, Koot HM, Wiersma D, Verkerk PH, de Muinck-Keizer-Schrama SM. Influence of timing and dose of thyroid hormone replacement on development in infants with congenital hypothyroidism. *J Pediatr* 2000;136:292–7.
- Bongers-Schokking JJ, de Muinck Keizer-Schrama SM. Influence of timing and dose of thyroid hormone replacement on mental, psychomotor,

- and behavioral development in children with congenital hypothyroidism. *J Pediatr* 2005;147:768–74.
25. Friedrichsen S, Christ S, Heuer H, et al. Regulation of iodothyronine deiodinases in the Pax8<sup>-/-</sup> mouse model of congenital hypothyroidism. *Endocrinology* 2003;144:777–84.
  26. Bons JAP, Vogt MHJ, Janssen MJW, et al. A multicenter evaluation of dys-thyroxinemia in a defined patient cohort results in regional harmonization of reference ranges. *Ned Tijdschr Clin Chem Labgeneesk* 2013;3866–68.
  27. Andersen S, Pedersen KM, Bruun NH, Laurberg P. Narrow individual variations in serum T(4) and T(3) in normal subjects: a clue to the understanding of subclinical thyroid disease. *J Clin Endocrinol Metab* 2002;87:1068–72.
  28. Biersack HJ, Hartmann F, Rödel R, Reinhardt M. Long term changes in serum T4, T3, and TSH in benign thyroid disease: proof of a narrow individual variation. *Nuklearmedizin* 2004;43:158–60; quiz 162–3.
  29. Karmisholt J, Andersen S, Laurberg P. Variation in thyroid function tests in patients with stable untreated subclinical hypothyroidism. *Thyroid* 2008;18:303–8.
  30. Lem AJ, de Rijke YB, van Toor H, de Ridder MA, Visser TJ, Hokken-Koelega AC. Serum thyroid hormone levels in healthy children from birth to adulthood and in short children born small for gestational age. *J Clin Endocrinol Metab* 2012;97:3170–8.
  31. Van der Meulen BF, Smrkovski M. Manual for the Bayley Developmental Scales (BOS). Swets en Zeitlinger, Amsterdam, the Netherlands; 1982.
  32. Bleichrodt N, Resing WCM, Drenth PJD, Zaal JN. [Intelligence measurement in children]. Zwets en Zeitlinger, Lisse, the Netherlands; 1987.
  33. Bleichrodt N, Drenth PJD, Zaal JN, Resing WCM. Revisie Amsterdamse Intelligentie Test: Manual. Swets and Zeitlinger, Lisse, the Netherlands; 1987.
  34. Resing WCM, Bleichrodt N, Drenth PJD. Het gebruik van de RAKIT bij allochtoon etnische groepen. [The use of the RAKIT for non-ethnic Dutch groups]. *Nederlands Tijdschrift voor de Psychologie* 1986;4:179–188.
  35. Georgas J, Weiss L, Van de Vijver FJR, Saklofske DH. Cultures and Children's intelligence: A cross-cultural analysis of the WISC-III. Academic Press, New York; 2003.