CATS II long-term anthropometric and metabolic effects of maternal sub-optimal thyroid function in offspring and mothers

Ilaria Muller[1,2], Peter N Taylor[1], Rhian M Daniel[3], Charlotte Hales[1], Anna Scholz[1], Toby Candler[4], Rebecca J Pettit[5], William D Evans[5], Dionne Shillabeer[1], Mohd S Draman[1,6], Colin M Dayan[1], Hiu K C Tang[1,7], Onyebuchi Okosieme[1], John W Gregory[3], John H Lazarus[1], D Aled Rees[8], Marian E Ludgate[1]

[1] Thyroid Research Group, Division of Infection & Immunity, School of Medicine, Cardiff University, UK

[2] Department of Endocrinology, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy

[3] Division of Population Medicine, School of Medicine, Cardiff University, UK

[4] MRC The Gambia at the London School of Hygiene and Tropical Medicine, London, UK

[5] Radiology, Medical Physics and Clinical Engineering Directorate, University Hospital of Wales, Cardiff, UK

[6] Faculty of Medicine, University Sultan Zainal Abidin, Terengganu, Malaysia

[7] Department of Oncology, Nottingham University NHS Trust, Nottingham, UK

[8] Neuroscience and Mental Health Research Institute, School of Medicine, Cardiff University, UK

SHORT TITLE: Maternal thyroid function effects on metabolism

KEYWORDS: thyroid, pregnancy, metabolism, obesity, hypothyroidism, BMI

DISCLOSURE SUMMARY: The authors have nothing to disclose.

CLINICAL TRIAL REGISTRATION NUMBER (CATS-I): ISRCTN46178175

FUNDED BY: The Charles Wolfson Trust (grant to ML), Action Medical Research (project code GN2033, grant to ML), The Henry Smith Charity (project code 20122759 GN 2033, grant to ML) and The American Thyroid Association (project code ATA-2014-033.R1, grant to PNT).

ABSTRACT

CONTEXT & OBJECTIVES. The Controlled Antenatal Thyroid Screening study-I (CATS-I) was a randomized controlled trial investigating the effects of levothyroxine therapy for suboptimal gestational thyroid function (SGTF), comparing outcomes in children of treated (SGTF-T) with untreated (SGTF-U) women during pregnancy. This follow-up study CATS-II reports the long-term effects on anthropometric, bone and cardiometabolic outcomes in mothers and offspring, and includes a group with normal gestational thyroid function (NGTF).

DESIGN & PARTICIPANTS. 332 mothers (197 NGTF, 56 SGTF-U, 79 SGTF-T) aged 41.2±5.3 years (mean±SD) and 326 paired children assessed 9.3±1.0 years after birth for: 1) body mass index (BMI); 2) lean, fat and bone mass by dual-energy x-ray absorptiometry; 3) blood pressure, augmentation index and aortic pulse-wave-velocity; 4) thyroid function, lipids, insulin and adiponectin. The difference between group means was compared using linear regression.

RESULTS. Offspring's measurements were similar between groups. Although maternal BMI was similar between groups at CATS-I, after 9-years (at CATS-II) SGTF-U mothers showed higher BMI (median[IQR] 28.3[24.6-32.6] kg/m^2) compared with NGTF (25.8[22.9-30.0] kg/m^2, p=0.029), driven by fat mass increase. At CATS-II SGTF-U mothers also had higher TSH values (median[IQR] 2.45[1.43-3.50] mU/L) than NGTF (1.54[1.12-2.07] mU/L; p=0.015), since 64% had never received levothyroxine. At CATS-II SGTF-T mothers had BMI (25.8[23.1-29.8] kg/m^2, p=0.672) and TSH (1.68[0.89-2.96] mU/L; p=0.474) values similar to NGTF mothers.

CONCLUSIONS. Levothyroxine supplementation of women with SGTF did not affect long-term offspring anthropometric, bone and cardiometabolic measurements. However, absence

of treatment was associated with sustained long-term increase in BMI and fat mass in women with SGTF.

PRÉCIS

Levothyroxine for suboptimal gestational thyroid function did not affect offspring's anthropometric, bone and cardiometabolic outcomes, but prevented an increase in maternal fat mass over 9 years.

INTRODUCTION

Pregnancy induces physiological changes in the maternal hypothalamic-pituitary-thyroid axis, such that interpretation of thyroid function during pregnancy must take into account trimester-specific reference ranges (1, 2). Suboptimal gestational thyroid function (SGTF), defined as a low free thyroxine (FT4) concentration and/or raised thyroid-stimulating hormone (TSH), is common and associated with adverse pregnancy and offspring outcomes (1, 3, 4).

Overt hypothyroidism is well-known to be associated with impaired fetal neurodevelopment, particularly during the first trimester of pregnancy when the fetus is entirely dependent on maternal thyroid hormone production for optimal brain development (1, 2, 5). However, the effects of mild maternal thyroid dysfunction are less clear. Whereas isolated hypothyroxinemia has been found to be associated with impaired developmental and neurobehavioral outcomes (1, 6), including verbal delay (7), autism (8, 9) and attention-deficit/hyperactivity disorder (ADHD) (10-14) in offspring, such effects have not been described for maternal subclinical hypothyroidism (1, 15-17).

Whilst many studies have examined effects on neuro-intellectual and behavioral outcomes, very little is known about the effects of maternal gestational thyroid dysfunction on offspring anthropometric, bone and cardiometabolic outcomes. Both high TSH levels and low FT4 levels during gestation are known to be associated with maternal weight gain and adverse metabolic pregnancy outcomes (18-20); accordingly, pregnancy-specific reference ranges for thyroid function assessments are influenced by body mass index (BMI) (21-23). In turn, maternal overweight and obesity during gestation and an excessive gestational weight gain, especially during early gestation, are associated with long-term adverse anthropometric and cardiometabolic outcomes in offspring (24-26). Overt hypothyroidism and hyperthyroidism negatively affect bone mass maintenance (27, 28); such effects have also

been observed in long-term subclinical hyperthyroidism, but not hypothyroidism (29-31), with only isolated exceptions (32). Thyroid hormones are also a determinant for bone development and skeletal maturation; it is well known that children with impaired and untreated thyroid function have short stature (28).

The Controlled Antenatal Thyroid Screening (CATS) study I (CATS-I) was the first randomized controlled trial to investigate the effects of antenatal screening and treatment of SGTF on offspring cognitive function (15). No effects on intelligence quotient (IQ) of 3 year-old offspring were found, and a follow-up study (CATS-II) repeating IQ assessment at age 9 (33) confirmed similar findings (17). CATS-II also allowed us to extend the phenotyping of children to other outcomes known to be influenced by thyroid status. We recently reported adverse behavioral outcomes in children born to mothers exposed to excess levothyroxine replacement (34).

Here we present for the first time the effects of treatment for SGTF on maternal and child anthropometric, bone and cardiometabolic outcomes, evaluated 9 years after delivery. We included the maternal analysis as a reference population for their paired children, and as a valuable cohort followed long-term to better evaluate the association between mild/subclinical thyroid dysfunction and overweight/obesity among adult women (35-37), as well as adverse cardiovascular events (38-41).

MATERIALS AND METHODS

STUDY DESIGN AND POPULATION

As previously reported, CATS-II (17, 33, 34) is the follow-up study of CATS-I (ISRCTN 46178175) (15). Briefly, in CATS-I a total of 21,846 pregnant women (median gestation of 12 weeks 3 days) were randomized to have their thyroid function measured at

recruitment (screening group) or at the end of pregnancy (control group). Those in the screening group diagnosed with SGTF, defined as FT4 <2.5th percentile and/or TSH >97.5th percentile of the cohort, were commenced on levothyroxine treatment (SGTF-T) for the duration of their pregnancy, while the women with SGTF in the control group were left untreated (SGTF-U) and subsequently referred to their general practitioner for further management following delivery. In the SGTF-T group, the starting dose of levothyroxine was 150μg daily; TSH and FT4 measurements were repeated after 6 weeks from commencing treatment and at 30 weeks of gestation, with adjustment of levothyroxine dose where necessary, aiming to maintain TSH levels in the 0.1-1.0mlU/L range. In CATS-I the IQ of children of SGTF-T and SGTF-U mothers was measured at age 3 years (15). The follow-up study (CATS-II) repeated IQ assessment in 9-year-old offspring (17, 33), included UK participants only and also evaluated children of women with normal gestational thyroid function (NGTF) during CATS-I. In addition to IQ assessment (17) other outcomes were also evaluated, including child behavior (34) and the anthropometric, bone and cardiometabolic measurements presented here. At CATS-II recruitment, clinical data including history of levothyroxine treatment following CATS-I were collected. The overall study design and population is summarized in **Figure 1**.

BMI AND BMI SDS

Standing heights were measured to the nearest 0.1 cm using a Harpenden stadiometer (Holtain Ltd, Crymych, UK). Participants were weighed to the nearest 0.1 kg in lightweight clothing without shoes, using the weighing function of a Body Fat Analyzer (TBF-305; Tanita, Tokyo, Japan). Body mass index (BMI) was calculated as weight (kg) divided by height (m) squared. Children's BMI standard deviation scores (SDS) were calculated using the UK reference population established in 1990 (42, 43).

DXA SCAN ANALYSIS

Whole body less head (WBLH), total lumbar spine (L1-L4) and left hip measurements of bone and/or lean and fat mass were made in 327/332 (98.5%) mothers and 323/326 (99.1%) children using a Hologic QDR Explorer fan-beam dual-energy x-ray absorptiometry (DXA) scanner (Hologic Inc., Marlborough, USA). Subjects were assessed in the supine position, in the fasted state and after bladder emptying. Whole body scans were acquired in explorer (e) mode, equivalent to array mode, while spine and hip scans were acquired in survey (s) mode, equivalent to fast array. Scans were analyzed using Hologic software V 13.3.0.1:3 with the auto low-density option being applied to the spine scans where appropriate.

For bone analysis, bone area (cm^2), bone mineral content (BMC, g) and the BMC per unit area of bone (bone mineral density [BMD], g/cm^2) were assessed. In particular, for mothers the BMD values of femoral neck (FN-BMD) and total lumbar spine (LS-BMD) were compared across the subject groups. For children DXA bone mineral measurements are strongly influenced by several factors including sex, chronological and skeletal age, height, weight, ethnicity and pubertal development (44-48). In order to reduce such biases, the bone mineral apparent density (BMAD) was also considered; it was calculated for the femoral neck (FN) using the method of Lu et al. (49), and for the lumbar spine (LS) using the geometric assumptions made by Carter et al (50), but estimating the volume of each individual vertebral body (L1–L4) from its bone area and summing the result rather than the total volume of L2–L4 (51). Depending on the body region studied (LS, FN, WBLH), BMD and BMAD were expressed as standard deviation scores (SDS) compared to the UK population (cohorts of the Ward study [W] or the Alphabet study [A]) (51, 52) and the US population (Hologic manufacturer [H]), for a total of 8 measurements.

For body composition, absolute WBLH fat and lean mass (kg), as well as relative fat

mass (%), were measured.

BIOCHEMICAL ANALYSIS

Fasting blood samples were collected from 294/332 (88.5%) mothers but only 83/326 (25.5%) children, since most of them refused phlebotomy. Serum was prepared by centrifugation at 4500 rpm for 10 minutes at +4°C, and stored at -80°C until analysis.

Thyroid-stimulating hormone (TSH), free-thyroxine (FT4), free-triiodothyronine (FT3), autoantibodies to thyroid peroxidase (TPOAb), triglyceride, total cholesterol, high-density lipoprotein (HDL) cholesterol and insulin were measured by Chemiluminescent Microparticle Immunoassay (Architect® System, Abbott Laboratories, USA). Normal reference ranges for thyroid function (females aged above 18 years) were 0.30 – 4.4 mIU/L for TSH, 9.0 – 19.1 pmol/L for FT4 and 2.6 – 5.7 pmol/L for FT3. According to the assay cut-off, TPOAb values were considered positive if ≥6 IU/ml and negative if <6 IU/ml.

High molecular weight adiponectin (APN) was measured by ELISA (EMD Millipore, Billerica, MA, USA).

CARDIOVASCULAR FUNCTION

The Vicorder device (Skidmore Medical, UK), a non-invasive cuff-based oscillometric technique that simultaneously measures the upstroke of femoral and carotid pulsations, was used to calculate the aortic pulse wave velocity (PWV), a measure of arterial stiffness (53), in addition to other measurements of peripheral and central blood pressure. This technique is reproducible (54), validated in both adults (55, 56) and children (57, 58) and agrees closely with invasive measures of central blood pressure (59). All measurements were performed by a single trained operator (DS). Measurements were taken with subjects relaxed in a quiet room, and with the head raised to 30°. Aortic PWV was measured by cuffs placed over the right carotid and the right thigh, with the length between the two arteries determined using a

tape measure placed over the suprasternal notch and the mid-point of the thigh cuff. Measurements were recorded when the pressure waveforms were reproducible over both arteries. Additional measurements were undertaken with the cuff placed on the right upper arm, including the systolic blood pressure (SP), the diastolic blood pressure (DP) and the aortic augmentation index (AI), a measure of the pulse wave reflection influenced by vessel stiffness (53).

DATA ANALYSIS

Histograms were performed to assess the distribution of variables. Variables were summarized using the sample mean ± standard deviation (SD) if approximately normally distributed, or using the sample median and interquartile range (IQR) otherwise. In addition, variables judged to be right-skewed were log-transformed for inclusion in analyses. Categorical variables (i.e. TPOAb positivity and child sex) were summarized using percentages, and the statistical significance of associations between them calculated using Fisher's exact test.

The effects of SGTF were first estimated comparing NGTF versus SGTF-U (p1U) and NGTF versus SGTF-T (p1T) in an unadjusted linear regression model (Model 1) and then adjusting for age, sex (children analysis only), ethnicity, socioeconomic status (defined in (17)) and smoking during pregnancy (Model 2). The same analysis was repeated comparing SGTF-U versus SGTF-T (p2) using linear regression adjusting for all the variables included in Model 2 and additionally baseline (at CATS-I) values of TSH and FT4 (Model 3).

In the CATS-I study the FT4 levels of SGTF-T mothers repeated at either 20 or 30 weeks of gestation were classified as optimal (SGTF-Topt) or suggestive for possible levothyroxine overtreatment (SGTF-Tover), if respectively below or above the threshold of

17.7 pmol/l, calculated as the top 2.5th percentile of the entire CATS-I UK population at recruitment (15, 34). The analysis of Model 3 was then repeated for the SGTF-Topt and SGTF-Tover subgroups. In the children's analysis, adjustment was additionally made for the corresponding variable in the mother, where available.

Supplemental Table 1 summarizes the Models used for each analysis (60). The data were analyzed in STATA, version 12. Obtained p values <0.05 were considered statistically significant.

RESULTS

CHILDREN

The female:male ratio was similar among the NGTF, SGTF-U and SGTF-T groups, whereas the children of NGTF women were slightly older (by a few months only) than SGTF-U and SGTF-T as previously reported (17, 34) (**Table 1**); thus age and sex were included in all regression models.

No differences were observed between NGTF, SGTF-U and SGTF-T groups in BMI SDS or any of the DXA or cardiovascular measurements (**Table 1**). Of note, the BMI SDS scores were approximately 0.5 SD higher than the UK 1990 reference population (42, 43).

All biochemical measurements were similar among groups except for HDL cholesterol (**Table 1**), which was marginally lower in the SGTF-T group compared with SGTF-U (p=0.048).

MOTHERS

Anthropometric, bone and cardiometabolic outcomes

Table 2 summarizes the results obtained among mothers: SGTF-T were slightly younger (39.7 ± 4.8 years) compared with NGTF (41.2 ± 5.5 years; p=0.002), but had a similar age to SGTF-U (40.9 ± 4.7 years; p=0.144). Untreated mothers (SGTF-U) had a higher BMI (median: 28.3, IQR: [24.6 - 32.6] kg/m^2) compared with NGTF mothers (25.8 [22.9 - 30.0] kg/m^2; p=0.029); in contrast, the BMI of treated mothers (SGTF-T) was similar to NGTF mothers (25.8 [23.1 - 29.8] kg/m^2; p=0.672). Importantly, BMI at CATS-I did not differ across the three groups (**Figure 2**): NGTF 25.0 [22.4 - 28.3] kg/m^2, SGTF-U 26.0 [23.4 - 30.1] kg/m^2 (p=0.111), SGTF-T 25.6 [23.0 - 29.2] kg/m^2 (p=0.112). When additionally adjusted for BMI at entry into CATS-I, the difference in BMI at CATS-II between SGTF-U and NGTF mothers remained significant (p=0.040). DXA analysis showed that the BMI increase was attributable to an increase in fat, but not lean mass (**Table 2, Figure 2**). DXA analysis showed no difference in BMD calculated at the femoral neck, total lumbar spine or whole body less head.

Among the metabolic measurements, SGTF-U mothers had higher triglyceride levels (median: 1.01, IQR: [0.78-1.40] mmol/L) than NGTF mothers (0.90 [0.70-1.10] mmol/L; p=0.041). Fasting insulin levels were also higher in SGTF-U mothers (6.30 [4.40-9.15] μIU/ml) than SGTF-T mothers (5.50 [3.85-7.15] μIU/ml; p=0.046). However, when BMI was included in the regression models the between-group significance was lost (p=0.212 and p=0.169 respectively).

There was no difference in any of the cardiovascular measurements between groups.

Thyroid function

As expected, a higher percentage of SGTF-U (20/50, 40%) and SGTF-T (38/71, 53%) mothers were TPOAb positive, compared with NGTF mothers (21/173, 12%; p<0.001; **Table**

2). Importantly, SGTF-U mothers at recruitment to CATS-II had significantly higher TSH levels (median [IQR]: 2.45 [1.43-3.50] mU/L) compared with NGTF mothers (1.54 [1.12-2.07] mU/L; p=0.015); TSH levels were also higher compared with SGTF-T mothers, albeit not quite reaching statistical significance (1.68 [0.89-2.96] mU/L; p=0.070). However, TSH concentrations were not different between SGTF-U and SGTF-T mothers at CATS-I recruitment (3.37 [1.22-4.45] mU/L and 4.21 [2.33-5.23] mU/L respectively, p=0.098). SGTF-T mothers had FT4 levels within the normal range (15.0 ± 2.9) in the CATS-II study, although these were significantly higher than in the NGTF group (13.6 ± 1.7; p<0.001).

The clinical history revealed that the majority of SGTF-U women had never been treated with levothyroxine, in contrast to those from the SGTF-T group, half of whom at recruitment to CATS-II were still on treatment commenced during CATS-I (**Table 3**). Of note, when TSH values were included in the regression model for BMI, the significance between SGTF-U and NGTF was lost (p=0.131).

Furthermore, TSH and BMI correlated positively in mothers in CATS-I (p=0.037), but only when excluding the SGTF-T group in CATS-II (p=0.027; **Supplemental Figure 1**) (60).

LEVOTHYROXINE OVER-TREATMENT SUBANALYSIS

Mothers

As shown in **Supplemental Table 2** (60), at CATS-II SGTF-T_{opt} (n= 58) and SGTF-T_{over} (n= 21) women had similar thyroid function (p=0.962 for TSH levels), however SGTF-T_{over} had lower BMI (p=0.002), absolute fat mass (p=0.007), lean mass (p<0.001), systolic blood pressure (p=0.036) and higher HDL cholesterol levels (p=0.002) compared with SGTF-T_{opt} (60). When including BMI in the regression models, the between-group significance was lost

for fat mass (p=0.882) and systolic blood pressure (p=0.074), but not for HDL cholesterol (p=0.022) and lean mass (p=0.003).

SGTF-T_{over} women had lower BMI (p=0.001) and height (p=0.032) already at CATS-I, such that the difference in BMI at CATS-II between SGTF-T_{opt} and SGTF-T_{over} lost significance when adjusted for baseline BMI at CATS-I (p=0.267).

Children

As shown in **Supplemental Table 3** (60), there were no differences in terms of age (p=0.516) and sex (p=0.260) between SGTF-T_{opt} (n= 57) and SGTF-T_{over} (n= 21) offspring. However, similarly to their mothers, SGTF-T_{over} children had lower height (p=0.016), BMI SDS (p=0.001) and lean mass (p=0.004) compared with SGTF-T_{opt} children, as well as FN-BMD-H (p=0.037), WBLH-BMD-H (p=0.002) and WBLH-BMD-A (p=0.002). Of note, when adjusting for the corresponding maternal measurement (p^5 column, **Supplemental Table 3**), only BMI SDS (p=0.006) and height (p=0.044) remained significantly lower in SGTF-T_{over} children compared with SGTF-T_{opt} children (60). When also adjusting for paternal height, the difference in height between SGTF-T_{over} and SGTF-T_{opt} children lost significance (p=0.298).

DISCUSSION

To our knowledge, the present study is the first to evaluate several long-term anthropometric, bone and cardiometabolic outcomes in children and mothers from a large cohort of women with SGTF randomized to receive levothyroxine treatment during pregnancy.

No significant effects were observed on offspring outcomes evaluated at 9 years of age. Only a slight reduction in HDL cholesterol levels was observed among children of treated mothers compared with those who were untreated, albeit this was of marginal clinical

significance. Considering the additional limitation of the low number of children consenting to phlebotomy, further studies in larger cohorts are needed before any firm conclusions can be drawn in this context. It was noteworthy that the children's BMI was higher compared with the UK children reference population established 30 years ago; this is in line with the global secular trends in rates of childhood overweight and obesity observed over the last three decades, as a likely consequence of unhealthy lifestyle (61, 62).

Long-term maternal bone and cardiovascular outcomes were also unaffected by SGTF, whether treated or not. However, BMI was significantly greater at 9 years follow-up only in the group of mothers with SGTF who had not been randomized during CATS-I to receive levothyroxine replacement (SGTF-U), with DXA analysis showing that this weight gain was predominantly attributable to an increase in fat rather than lean mass. On the contrary, the group of mothers with SGTF who were started on levothyroxine replacement during CATS-I (SGTF-T), 9 years later had similar BMI and fat mass values to women with NGTF. Of note, the baseline BMI at enrolment into the CATS-I study was similar among all groups and, when included in the regression model, did not influence the BMI change observed at CATS-II. These observations suggest that the BMI increase had occurred in the 9-year time window from recruitment into CATS-I and CATS-II, and only in the group of women with untreated SGTF at CATS-I. In line with the higher prevalence of overweight and obesity in the SGTF-U group, these women also had significantly higher triglyceride and insulin levels; however, these differences lost significance when adjusted for BMI, suggesting that these metabolic alterations were driven by overweight/obesity as expected (63-65). Untreated SGTF women also had current higher TSH concentrations compared with the other two groups, since the majority of them had not been commenced on levothyroxine treatment during CATS-I or the following 9 years. On the contrary, nearly half of women with SGTF commencing levothyroxine at CATS-I (SGTF-T) were still on treatment 9 years later; their higher FT4

levels likely reflected the measurement of exogenous T4 (66), especially if blood was withdrawn after taking levothyroxine (67). The fact that the differences in BMI among groups lost significance when adjusted for current TSH levels, suggests that the increase in fat mass was largely driven by untreated suboptimal thyroid function. In this study other BMI-influencing factors such as physical activity, dietary habits and family history for obesity were not evaluated, thus their hypothetical influence on observed BMI differences among groups cannot be excluded. However SGTF-U and SGTF-T women had been randomised at CATS-I, thus we assume that confounding factors, other than those related to thyroid function, should not differ between these groups. SGTF-T and NGTF women at CATS-II presented similar values of thyroid and metabolic outcomes (TSH, BMI, fat mass, triglyceride, insulin), parameters that were different only in the SGTF-U group. This difference was statistically significant only when comparing SGTF-U and NGTF women but not SGTF-U and SGTF-T women, with the sole exception of insulin levels, likely due to the small size of SGTF-T group (N=79) compared with NGTF group (N=197).

Small variation in TSH levels even within the normal reference range is positively associated with higher BMI and weight gain in other cohorts (68-71), and such a relationship is bilateral (35-37). Obesity, likely acting via leptin, activates the hypothalamic-pituitary-thyroid axis and induces a consequent rise in TSH levels (72, 73). On the other hand, impaired thyroid function favors weight gain due to the consequent myxoedema and reduction of resting energy expenditure (REE), particularly in overt hypothyroidism (37). However, even smaller variations of thyroid function, usually considered clinically insignificant, induce measurable REE modifications, and therefore if sustained have the potential to affect body weight (74). Accordingly, in our study women with long-term suboptimal thyroid function, if left untreated experienced increased body weight and fat

mass, while those who were treated did not. This suggests that in our cohort thyroid function determined BMI and not vice versa.

A recent meta-analysis did not highlight any benefits on several outcomes, including BMI, of levothyroxine treatment for subclinical hypothyroidism (75). However, this analysis included a very large trial evaluating subjects above 65 years of age (76), which represents a population that may not be comparable with younger individuals such as those analyzed in our study. With respect to specific effects on BMI, the majority of studies were small-scale and based on shorter follow-up periods (75), therefore not allowing definitive conclusions. The study of Zhao et al (77) was one of the largest randomized trials of levothyroxine replacement, enrolling 369 middle-aged males and females affected with subclinical hypothyroidism who were followed-up for 15 months. In line with the results of our study, subjects receiving levothyroxine showed a significant BMI reduction at the end of the follow-up period, while untreated subjects did not. Clearly, further randomized trials are needed to clarify the benefits of levothyroxine replacement on BMI in individuals affected with subclinical hypothyroidism below 65 years of age.

Our exploratory sub-analysis showed that women exposed to over-replacement with levothyroxine during pregnancy (SGTF-T_{over}), at CATS-II displayed lower BMI, height, absolute fat mass, lean mass, SP and higher HDL values compared with those with optimal gestational FT4 levels (SGTF-T_{opt}); fat mass and blood pressure seemed to be driven by BMI. However, SGTF-T_{over} women were noted to have been thinner and shorter from baseline, before commencing levothyroxine treatment during CATS-I; in fact when including baseline BMI in the regression model, the BMI change observed at CATS-II lost significance. Furthermore, levothyroxine doses were promptly reduced during CATS-I to correct the raised FT4 levels, such that thyroid function of the SGTF-T_{over} group at CATS-II was normal and similar to SGTF-T_{opt}. Considering that all women were commenced on a standard 150 µg

dose of levothyroxine during pregnancy, it is appropriate to conclude that this likely induced an excessive increase of FT4 levels in this subgroup of thinner women, since they would have required a smaller dose. This study further highlights the importance of adjusting levothyroxine treatment for body weight, especially during pregnancy where high FT4 levels have been associated with several negative outcomes (1, 2, 6), including a higher prevalence of behavioral difficulties in the offspring of this cohort (34). Similar to their mothers, SGTF-T_{over} children were thinner and shorter compared to SGTF-T_{opt} children. However, correction for the corresponding maternal and paternal measurements, where available, reduced or totally eliminated the between-group significance, indicating a genetic component rather than an effect of levothyroxine overtreatment on anthropometric outcomes.

The strengths of our study include the large sample size, baseline randomization, analysis of several anthropometric, bone and cardiometabolic outcomes, and longitudinal design with one of the longest available follow-up periods. Our study has limitations, however, including a lack of detailed information about the levothyroxine doses used and the length of drug withdrawal periods during the 9 years between CATS-I and CATS-II, as well as a lack of correction for other BMI-influencing factors, such as dietary habits, physical activity and family history for obesity.

In conclusion, for the first time we evaluated the long-term effects of SGTF and treatment with levothyroxine during pregnancy on a series of offspring anthropometric, bone and cardiometabolic measurements, finding no significant evidence for benefit or harm. Women with long-term untreated mild suboptimal thyroid function persisting after pregnancy showed a significant increase in BMI, fat mass, triglyceride and insulin levels, that were absent in the group of women treated with levothyroxine. Our study also emphasizes the need for careful adjustment of levothyroxine dose for bodyweight to avoid overtreatment, especially during pregnancy. Our findings thus highlight the need for dedicated large-scale randomized trials to

investigate the long-term benefits of levothyroxine treatment in young and middle-aged individuals with suboptimal thyroid function, whether in relation to pregnancy or not. If our observations were confirmed, the current indications to such treatment may need to be revised.

ACKNOWLEDGMENTS

CATS-II was funded by The Charles Wolfson Trust (grant to ML), Action Medical Research (project code GN2033, grant to ML), The Henry Smith Charity (project code 20122759 GN 2033, grant to ML) and The American Thyroid Association (project code ATA-2014-033.R1, grant to PNT). We are extremely grateful to the children, parents and families who participated in the study. Special thanks are extended to Julie Pell, Julie Evans, Sophie Fuller and Beverley Carey for their dedicated support of the CATS project.

REFERENCES

1. Korevaar TIM, Medici M, Visser TJ, Peeters RP. Thyroid disease in pregnancy: new insights in diagnosis and clinical management. Nat Rev Endocrinol. 2017;13:610-622.
2. Muller I, Taylor PN, Lazarus JH. Thyroid function in pregnancy. Annals of Thyroid. 2018;3:27.
3. Lazarus J, Brown RS, Daumerie C, Hubalewska-Dydejczyk A, Negro R, Vaidya B. 2014 European thyroid association guidelines for the management of subclinical hypothyroidism in pregnancy and in children. Eur Thyroid J. 2014;3:76-94.
4. Alexander EK, Pearce EN, Brent GA, et al. 2017 Guidelines of the American Thyroid Association for the Diagnosis and Management of Thyroid Disease During Pregnancy and the Postpartum. Thyroid. 2017;27:315-389.
5. Haddow JE, Palomaki GE, Allan WC, et al. Maternal thyroid deficiency during pregnancy and subsequent neuropsychological development of the child. N Engl J Med. 1999;341:549-555.
6. Korevaar TI, Muetzel R, Medici M, et al. Association of maternal thyroid function during early pregnancy with offspring IQ and brain morphology in childhood: a population-based prospective cohort study. Lancet Diabetes Endocrinol. 2016;4:35-43.
7. Henrichs J, Bongers-Schokking JJ, Schenk JJ, et al. Maternal thyroid function during early pregnancy and cognitive functioning in early childhood: the generation R study. J Clin Endocrinol Metab. 2010;95:4227-4234.
8. Roman GC, Ghassabian A, Bongers-Schokking JJ, et al. Association of gestational maternal hypothyroxinemia and increased autism risk. Ann Neurol. 2013;74:733-742.

9. Andersen SL, Laurberg P, Wu CS, Olsen J. Attention deficit hyperactivity disorder and autism spectrum disorder in children born to mothers with thyroid dysfunction: a Danish nationwide cohort study. BJOG. 2014;121:1365-1374.
10. Vermiglio F, Lo Presti VP, Moleti M, et al. Attention deficit and hyperactivity disorders in the offspring of mothers exposed to mild-moderate iodine deficiency: a possible novel iodine deficiency disorder in developed countries. J Clin Endocrinol Metab. 2004;89:6054-6060.
11. Ghassabian A, Bongers-Schokking JJ, Henrichs J, et al. Maternal thyroid function during pregnancy and behavioral problems in the offspring: the generation R study. Pediatr Res. 2011;69:454-459.
12. Pakkila F, Mannisto T, Pouta A, et al. The impact of gestational thyroid hormone concentrations on ADHD symptoms of the child. J Clin Endocrinol Metab. 2014;99:E1-8.
13. Modesto T, Tiemeier H, Peeters RP, et al. Maternal Mild Thyroid Hormone Insufficiency in Early Pregnancy and Attention-Deficit/Hyperactivity Disorder Symptoms in Children. JAMA Pediatr. 2015;169:838-845.
14. Oostenbroek MHW, Kersten RHJ, Tros B, Kunst AE, Vrijkotte TGM, Finken MJJ. Maternal hypothyroxinaemia in early pregnancy and problem behavior in 5-year-old offspring. Psychoneuroendocrinology. 2017;81:29-35.
15. Lazarus JH, Bestwick JP, Channon S, et al. Antenatal thyroid screening and childhood cognitive function. N Engl J Med. 2012;366:493-501.
16. Casey BM, Thom EA, Peaceman AM, et al. Treatment of Subclinical Hypothyroidism or Hypothyroxinemia in Pregnancy. N Engl J Med. 2017;376:815-825.
17. Hales C, Taylor PN, Channon S, et al. Controlled Antenatal Thyroid Screening II: Effect of Treating Maternal Suboptimal Thyroid Function on Child Cognition. J Clin Endocrinol Metab. 2018;103:1583-1591.
18. Han C, Li C, Mao J, et al. High Body Mass Index Is an Indicator of Maternal Hypothyroidism, Hypothyroxinemia, and Thyroid-Peroxidase Antibody Positivity during Early Pregnancy. Biomed Res Int. 2015;2015:351831.
19. Knight BA, Shields BM, Hattersley AT, Vaidya B. Maternal hypothyroxinaemia in pregnancy is associated with obesity and adverse maternal metabolic parameters. Eur J Endocrinol. 2016;174:51-57.
20. Collares FM, Korevaar TIM, Hofman A, et al. Maternal thyroid function, prepregnancy obesity and gestational weight gain-The Generation R Study: A prospective cohort study. Clin Endocrinol (Oxf). 2017;87:799-806.
21. Mannisto T, Surcel HM, Ruokonen A, et al. Early pregnancy reference intervals of thyroid hormone concentrations in a thyroid antibody-negative pregnant population. Thyroid. 2011;21:291-298.
22. Mosso L, Martinez A, Rojas MP, et al. Early pregnancy thyroid hormone reference ranges in Chilean women: the influence of body mass index. Clin Endocrinol (Oxf). 2016;85:942-948.
23. Laurberg P, Andersen SL, Hindersson P, Nohr EA, Olsen J. Dynamics and Predictors of Serum TSH and fT4 Reference Limits in Early Pregnancy: A Study Within the Danish National Birth Cohort. J Clin Endocrinol Metab. 2016;101:2484-2492.
24. Gaillard R, Steegers EA, Duijts L, et al. Childhood cardiometabolic outcomes of maternal obesity during pregnancy: the Generation R Study. Hypertension. 2014;63:683-691.
25. Gaillard R, Steegers EA, Franco OH, Hofman A, Jaddoe VW. Maternal weight gain in different periods of pregnancy and childhood cardio-metabolic outcomes. The Generation R Study. Int J Obes (Lond). 2015;39:677-685.

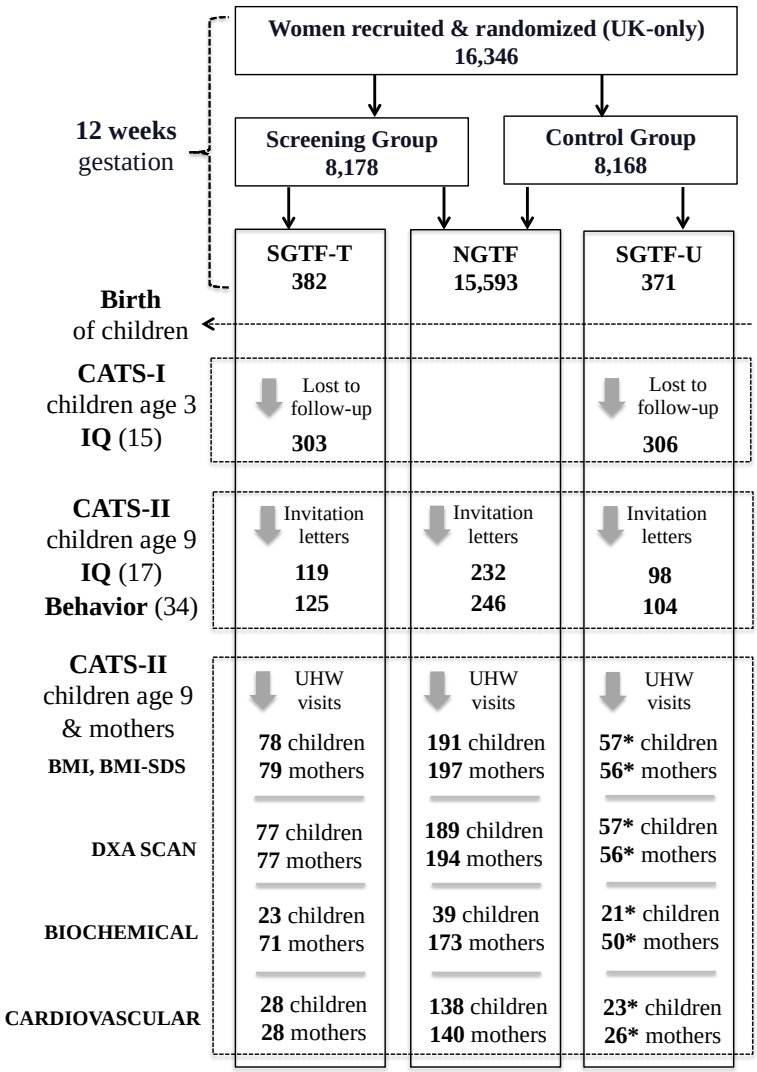
26. Gaillard R, Welten M, Oddy WH, et al. Associations of maternal prepregnancy body mass index and gestational weight gain with cardio-metabolic risk factors in adolescent offspring: a prospective cohort study. BJOG. 2016;123:207-216.
27. Wojcicka A, Bassett JH, Williams GR. Mechanisms of action of thyroid hormones in the skeleton. Biochimica et biophysica acta. 2013;1830:3979-3986.
28. Bassett JH, Williams GR. Role of Thyroid Hormones in Skeletal Development and Bone Maintenance. Endocr Rev. 2016;37:135-187.
29. Waring AC, Harrison S, Fink HA, et al. A prospective study of thyroid function, bone loss, and fractures in older men: The MrOS study. J Bone Miner Res. 2013;28:472-479.
30. Garin MC, Arnold AM, Lee JS, Robbins J, Cappola AR. Subclinical thyroid dysfunction and hip fracture and bone mineral density in older adults: the cardiovascular health study. J Clin Endocrinol Metab. 2014;99:2657-2664.
31. Blum MR, Bauer DC, Collet TH, et al. Subclinical thyroid dysfunction and fracture risk: a meta-analysis. JAMA. 2015;313:2055-2065.
32. Lee JS, Buzkova P, Fink HA, et al. Subclinical thyroid dysfunction and incident hip fracture in older adults. Arch Intern Med. 2010;170:1876-1883.
33. Hales C, Channon S, Taylor PN, et al. The second wave of the Controlled Antenatal Thyroid Screening (CATS II) study: the cognitive assessment protocol. BMC Endocr Disord. 2014;14:95.
34. Hales C, Taylor PN, Channon S, et al. Controlled Antenatal Thyroid Screening II: effect of treating maternal sub-optimal thyroid function on child behaviour. J Clin Endocrinol Metab. 2019;Oct 29;pii: dgz098. Doi: 10.1210/clinem/dgz098. [Epub ahead of print]
35. Reinehr T. Obesity and thyroid function. Mol Cell Endocrinol. 2010;316:165-171.
36. Biondi B. Thyroid and obesity: an intriguing relationship. J Clin Endocrinol Metab. 2010;95:3614-3617.
37. Laurberg P, Knudsen N, Andersen S, Carle A, Pedersen IB, Karmisholt J. Thyroid function and obesity. Eur Thyroid J. 2012;1:159-167.
38. Ortega FB, Lavie CJ, Blair SN. Obesity and Cardiovascular Disease. Circ Res. 2016;118:1752-1770.
39. Piche ME, Poirier P, Lemieux I, Despres JP. Overview of Epidemiology and Contribution of Obesity and Body Fat Distribution to Cardiovascular Disease: An Update. Prog Cardiovasc Dis. 2018;61:103-113.
40. Cooper DS, Biondi B. Subclinical thyroid disease. Lancet. 2012;379:1142-1154.
41. Razvi S, Jabbar A, Pingitore A, et al. Thyroid Hormones and Cardiovascular Function and Diseases. J Am Coll Cardiol. 2018;71:1781-1796.
42. Freeman JV, Cole TJ, Chinn S, Jones PR, White EM, Preece MA. Cross sectional stature and weight reference curves for the UK, 1990. Archives of disease in childhood. 1995;73:17-24.
43. Cole TJ, Freeman JV, Preece MA. Body mass index reference curves for the UK, 1990. Arch Dis Child. 1995;73:25-29.
44. Wang MC, Aguirre M, Bhudhikanok GS, et al. Bone mass and hip axis length in healthy Asian, black, Hispanic, and white American youths. J Bone Miner Res. 1997;12:1922-1935.
45. Nelson DA, Simpson PM, Johnson CC, Barondess DA, Kleerekoper M. The accumulation of whole body skeletal mass in third- and fourth-grade children: effects of age, gender, ethnicity, and body composition. Bone. 1997;20:73-78.
46. Bachrach LK, Hastie T, Wang MC, Narasimhan B, Marcus R. Bone mineral acquisition in healthy Asian, Hispanic, black, and Caucasian youth: a longitudinal study. J Clin Endocrinol Metab. 1999;84:4702-4712.

47. Schoenau E, Neu CM, Rauch F, Manz F. Gender-specific pubertal changes in volumetric cortical bone mineral density at the proximal radius. Bone. 2002;31:110-113.
48. Binkovitz LA, Henwood MJ. Pediatric DXA: technique and interpretation. Pediatr Radiol. 2007;37:21-31.
49. Lu PW, Cowell CT, SA LL-J, Briody JN, Howman-Giles R. Volumetric bone mineral density in normal subjects, aged 5-27 years. J Clin Endocrinol Metab. 1996;81:1586-1590.
50. Carter DR, Bouxsein ML, Marcus R. New approaches for interpreting projected bone densitometry data. J Bone Miner Res. 1992;7:137-145.
51. Ward KA, Ashby RL, Roberts SA, Adams JE, Zulf Mughal M. UK reference data for the Hologic QDR Discovery dual-energy x ray absorptiometry scanner in healthy children and young adults aged 6-17 years. Arch Dis Child. 2007;92:53-59.
52. Crabtree NJ, Shaw NJ, Bishop NJ, et al. Amalgamated Reference Data for Size-Adjusted Bone Densitometry Measurements in 3598 Children and Young Adults-the ALPHABET Study. J Bone Miner Res. 2017;32:172-180.
53. O'Rourke MF, Pauca A, Jiang XJ. Pulse wave analysis. Br J Clin Pharmacol. 2001;51:507-522.
54. Hickson SS, Butlin M, Broad J, Avolio AP, Wilkinson IB, McEniery CM. Validity and repeatability of the Vicorder apparatus: a comparison with the SphygmoCor device. Hypertens Res. 2009;32:1079-1085.
55. Davies JM, Bailey MA, Griffin KJ, Scott DJ. Pulse wave velocity and the non-invasive methods used to assess it: Complior, SphygmoCor, Arteriograph and Vicorder. Vascular. 2012;20:342-349.
56. Shahin Y, Barakat H, Barnes R, Chetter I. The Vicorder device compared with SphygmoCor in the assessment of carotid-femoral pulse wave velocity in patients with peripheral arterial disease. Hypertens Res. 2013;36:208-212.
57. Kis E, Cseprekal O, Kerti A, et al. Measurement of pulse wave velocity in children and young adults: a comparative study using three different devices. Hypertens Res. 2011;34:1197-1202.
58. Kracht D, Shroff R, Baig S, et al. Validating a new oscillometric device for aortic pulse wave velocity measurements in children and adolescents. Am J Hypertens. 2011;24:1294-1299.
59. Pucci G, Cheriyan J, Hubsch A, et al. Evaluation of the Vicorder, a novel cuff-based device for the noninvasive estimation of central blood pressure. J Hypertens. 2013;31:77-85.
60. Muller I, Taylor PN, Daniel RM, et al. Supplemental data of "Controlled Antenatal Thyroid Screening (CATS) study II: Effects of treating maternal sub-optimal gestational thyroid function on anthropometric, bone and cardiometabolic outcomes". Figshare Digital Repository. Deposited 1 January 2020. http://doi.org/10.6084/m9.figshare.11492289.
61. Wang Y, Lobstein T. Worldwide trends in childhood overweight and obesity. Int J Pediatr Obes. 2006;1:11-25.
62. Han JC, Lawlor DA, Kimm SY. Childhood obesity. Lancet. 2010;375:1737-1748.
63. Kahn BB, Flier JS. Obesity and insulin resistance. The Journal of clinical investigation. 2000;106:473-481.
64. Franssen R, Monajemi H, Stroes ES, Kastelein JJ. Obesity and dyslipidemia. Med Clin North Am. 2011;95:893-902.
65. Ye J. Mechanisms of insulin resistance in obesity. Front Med. 2013;7:14-24.
66. Woeber KA. Levothyroxine therapy and serum free thyroxine and free triiodothyronine concentrations. J Endocrinol Invest. 2002;25:106-109.

67. Lips DJ, van Reisen MT, Voigt V, Venekamp W. Diagnosis and treatment of levothyroxine pseudomalabsorption. Neth J Med. 2004;62:114-118.
68. Knudsen N, Laurberg P, Rasmussen LB, et al. Small differences in thyroid function may be important for body mass index and the occurrence of obesity in the population. J Clin Endocrinol Metab. 2005;90:4019-4024.
69. Fox CS, Pencina MJ, D'Agostino RB, et al. Relations of thyroid function to body weight: cross-sectional and longitudinal observations in a community-based sample. Arch Intern Med. 2008;168:587-592.
70. Svare A, Nilsen TI, Bjoro T, Asvold BO, Langhammer A. Serum TSH related to measures of body mass: longitudinal data from the HUNT Study, Norway. Clin Endocrinol (Oxf). 2011;74:769-775.
71. Taylor PN, Razvi S, Pearce SH, et al. Clinical review: A review of the clinical consequences of variation in thyroid function within the reference range. J Clin Endocrinol Metab. 2013;98:3562-3571.
72. Reinehr T, Andler W. Thyroid hormones before and after weight loss in obesity. Arch Dis Child. 2002;87:320-323.
73. Baloch Z, Carayon P, Conte-Devolx B, et al. Laboratory medicine practice guidelines. Laboratory support for the diagnosis and monitoring of thyroid disease. Thyroid. 2003;13:3-126.
74. al-Adsani H, Hoffer LJ, Silva JE. Resting energy expenditure is sensitive to small dose changes in patients on chronic thyroid hormone replacement. J Clin Endocrinol Metab. 1997;82:1118-1125.
75. Feller M, Snel M, Moutzouri E, et al. Association of Thyroid Hormone Therapy With Quality of Life and Thyroid-Related Symptoms in Patients With Subclinical Hypothyroidism: A Systematic Review and Meta-analysis. JAMA. 2018;320:1349-1359.
76. Stott DJ, Gussekloo J, Kearney PM, et al. Study protocol; Thyroid hormone Replacement for Untreated older adults with Subclinical hypothyroidism - a randomised placebo controlled Trial (TRUST). BMC Endocr Disord. 2017;17:6.
77. Zhao M, Liu L, Wang F, et al. A Worthy Finding: Decrease in Total Cholesterol and Low-Density Lipoprotein Cholesterol in Treated Mild Subclinical Hypothyroidism. Thyroid. 2016;26:1019-1029.

FIGURES AND TABLES

FIGURE 1. Flow chart of study population

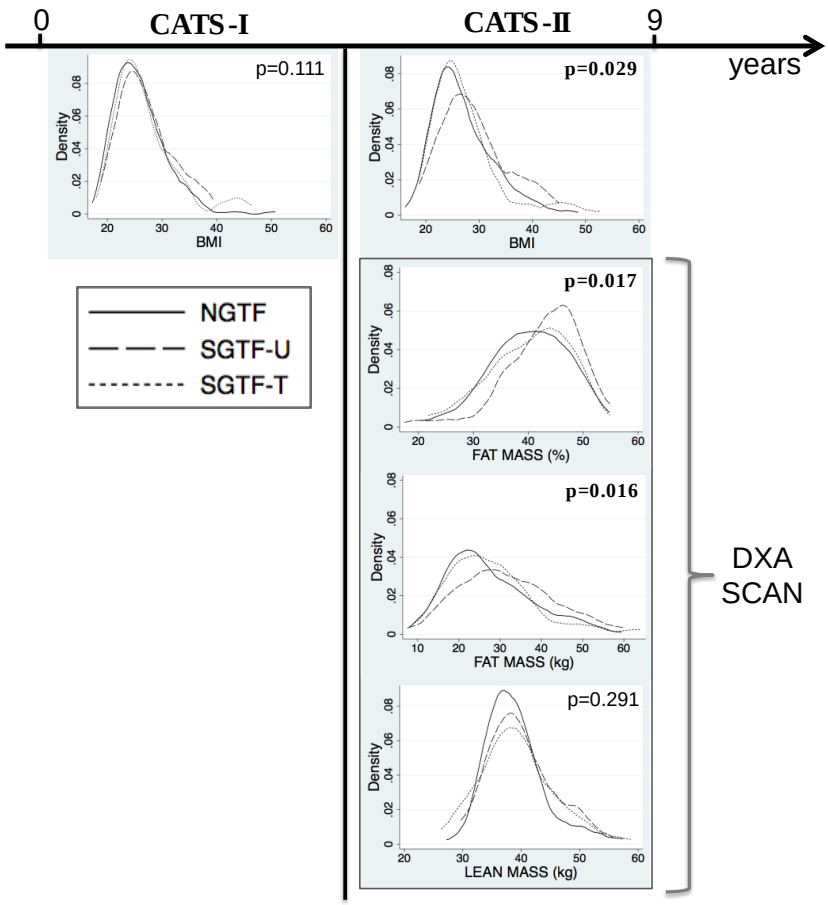


The recruitment of study participants initially started with the CATS-I study, assessing child cognition at 3 years of age (15). A UK-only subset of the original population and an additional third group (NGTF) were involved in the subsequent follow-up CATS-II study, assessing child cognition (17) and behavior (34) at 9 years of age. Only CATS-II children and paired mothers attending visits at the research center were included in the present study and assessed for cardiovascular, metabolic and bone measurements.

***** = Children tended to participate less in this study compared with their mothers, except for one mother in the SGTF-U group, refusing to be included in the study but agreeing for her son to participate.

BMI = body mass index. **CATS** = Controlled Antenatal Thyroid Screening study. **DXA** = dual-energy x-ray absorptiometry. **IQ** = intelligence quotient. **NGTF** = normal gestational thyroid function. **SDS** = standard deviation scores. **SGTF-T** = suboptimal gestational thyroid function treated with levothyroxine during pregnancy (Treated). **SGTF-U** = suboptimal gestational thyroid function not treated with levothyroxine during pregnancy (Untreated). **UHW** = University Hospital of Wales (research center).

FIGURE 2. Comparison of body composition at CATS-I and CATS-II studies among women



BMI = body mass index. **CATS** = Controlled Antenatal Thyroid Screening study. **DXA** = dual-energy x-ray absorptiometry. **NGTF** = normal gestational thyroid function. **SGTF-T** = suboptimal gestational thyroid function treated with levothyroxine during pregnancy (Treated). **SGTF-U** = suboptimal gestational thyroid function not treated with levothyroxine during pregnancy (Untreated).

Reported p values refer to the comparisons between NGTF and SGTF-U.

Table 1. Anthropometric, bone and cardiometabolic outcomes: children

	TOT N = 326	**NGTF N = 191**	**SGTF-U N = 57**	**SGTF-T N = 78**	**p1U**	**p1T**	**p2**
Age (years)	9.3 ± 1.0	9.6 ± 0.7	9.0 ± 1.1	8.8 ± 1.1	**<0.001**	**<0.001**	0.485
Female children N (%)	158 (48.5%)	88 (46.1%)	32 (56.1%)	38 (48.7%)	0.227	0.788	0.486
Height (cm)	136.7 ± 7.9	138.5 ± 7.3	133.5 ± 8.7	134.8 ± 7.4	0.051	0.779	0.089
BMI (kg/m^2)	17.2 [15.8 - 19.2]	17.3 [15.9 - 19.5]	17.0 [15.8 - 19.3]	16.8 [15.5 - 18.7]	0.587	0.822	0.464
BMI-SDS UK1990	0.49 ± 1.15	0.48 ± 1.15	0.57 ± 1.15	0.46 ± 1.15	0.646	0.763	0.481
DXA SCAN[1]							
WBLH Lean Mass (kg)	19.7 ± 3.5	20.3 ± 3.3	18.6 ± 3.6	19.0 ± 3.6	0.395	0.808	0.454
WBLH Fat Mass (kg)	8.36 [6.12 - 11.89]	8.67 [6.27 - 12.90]	8.71 [6.26 - 10.85]	7.46 [5.84 - 10.46]	0.979	0.573	0.553
WBLH Fat Mass (%)	31.8 ± 7.9	31.9 ± 7.8	32.6 ± 8.1	30.8 ± 7.9	0.617	0.475	0.279
LS-BMD-H (SDS)	0.33 ± 1.00	0.28 ± 0.97	0.28 ± 1.00	0.50 ± 1.08	0.721	0.302	0.424
LS-BMD-A (SDS)	0.23 ± 1.01	0.16 ± 0.97	0.19 ± 0.99	0.43 ± 1.10	0.751	0.261	0.403
LS-BMAD-A (SDS)	0.35 ± 1.08	0.30 ± 1.08	0.29 ± 1.04	0.53 ± 1.12	0.950	0.214	0.384
LS-BMAD-W (SDS)	0.27 ± 1.07	0.22 ± 1.07	0.22 ± 1.03	0.43 ± 1.1	0.948	0.223	0.471
FN-BMD-H (SDS)	0.49 ± 1.04	0.00 ± 0.95	0.02 ± 1.17	0.19 ± 1.14	0.900	0.358	0.732
FN-BMAD-W (SDS)	0.21 ± 0.99	0.16 ± 0.88	0.27 ± 1.26	0.29 ± 1.01	0.255	0.314	0.755
WBLH-BMD-H (SDS)	-0.83 ± 0.86	-0.80 ± 0.80	-0.93 ± 0.87	-0.85 ± 0.99	0.459	0.837	0.973
WBLH-BMD-A (SDS)	-0.87 ± 0.96	-0.87 ± 0.90	-0.98 ± 0.95	-0.79 ± 1.10	0.339	0.850	0.546
BIOCHEMICAL[2]							
TSH (mU/L)	1.77 [1.42 - 2.52]	1.64 [1.38 - 2.31]	1.80 [1.41 - 2.29]	1.81 [1.66 - 2.76]	0.511	0.305	0.764
FT4 (pmol/L)	14.69 ± 1.48	14.20 ± 1.38	15.39 ± 1.65	14.82 ± 1.25	0.098	0.597	0.335
FT3 (pmol/L)	5.60 ± 0.64	5.43 ± 0.63	5.87 ± 0.45	5.65 ± 0.73	0.158	0.938	0.290
TPOAb Positive/Total (%)	3/83 (3.6)	1/39 (2.6)	1/21 (4.8)	1/23 (4.3)	1.000	1.000	1.000
Cholesterol TOT (mmol/L)	4.33 ± 0.60	4.35 ± 0.63	4.41 ± 0.62	4.22 ± 0.54	0.398	0.887	0.403
Cholesterol HDL (mmol/L)	1.20 ± 0.26	1.22 ± 0.27	1.26 ± 0.29	1.10 ± 0.21	0.178	0.566	**0.048**
Triglyceride (mmol/L)	0.71 [0.60 - 0.90]	0.70 [0.50 - 1.00]	0.80 [0.62 - 0.88]	0.68 [0.60 - 0.89]	0.789	0.509	0.374
Insulin (μIU/mL)	4.45 [3.30 - 5.95]	4.90 [3.60 - 6.90]	4.40 [3.30 - 5.70]	3.85 [3.3 - 5.1]	0.385	0.272	0.950
Adiponectin (ng/mL)	13.49 ± 5.19	12.36 ± 4.44	14.75 ± 5.41	14.25 ± 5.96	0.368	0.554	0.703
CARDIOVASCULAR[3]							
SP (mmHg)	125.4 ± 12.9	124.8 ± 12.9	126.9 ± 10.4	127.2 ± 14.8	0.330	0.840	0.284
DP (mmHg)	63.0 ± 7.9	63.0 ± 7.7	61.8 ± 9.0	64.1 ± 8.0	0.497	0.682	0.228
AI (%)	7.33 [3.67 - 11.67]	7.33 [3.42 - 11.17]	10.67 [4.00 - 14.00]	7.00 [5.00 - 11.33]	0.363	0.911	0.454
PWV (m/s)	6.22 ± 0.65	6.22 ± 0.68	6.22 ± 0.51	6.17 ± 0.61	0.919	0.903	0.452

The results are presented as mean ± SD or median [IQR], if Normally- or non-Normally distributed, respectively.
[1] = Data available for 323 subjects; [2] = Data available for 83 subjects; [3] = Data available for 189 subjects
A = Alphabet study reference cohort. **AI** = augmentation index. **BMAD** = bone mineral apparent density. **BMD** = bone mineral density. **BMI** = body mass index. **BMI-SDS** = standard deviation score of body mass index. **DP** = diastolic blood pressure. **DXA** = dual-energy x-ray absorptiometry. **FN** = femoral neck. **FT3** = free-triiodothyronine. **FT4** = free-thyroxine. **H** = Hologic manufacturer reference cohort. **LS** = lumbar spine (from L1 to L4). **NGTF** = children of women with normal gestational thyroid function. **PWV** = aortic pulse wave velocity. **SGTF-T** = children of women with suboptimal gestational thyroid function treated with levothyroxine during pregnancy (Treated). **SGTF-U** = children of women with suboptimal gestational thyroid function not treated with levothyroxine during pregnancy (Untreated). **SDS** = standard deviation score. **SP** = systolic blood pressure. **TPOAb** = autoantibodies to thyroid peroxidase. **TSH** = thyrotropin. **W** = Ward study reference cohort. **WBLH** = whole body less head.
p1U = NGTF vs SGTF-U. **p1T** = NGTF vs SGTF-T. **p2** = SGTF-U vs SGTF-T.

Table 2. Anthropometric, bone and cardiometabolic outcomes: mothers

	TOT N = 332	**NGTF N = 197**	**SGTF-U N = 56**	**SGTF-T N = 79**	**p1U**	**p1T**	**p2**
Age (years)	41.2 ± 5.3	41.8 ± 5.5	40.9 ± 4.7	39.7 ± 4.8	0.252	**0.002**	0.144
Height (cm)	164.0 ± 6.4	164.3 ± 6.1	163.2 ± 5.7	164.1 ± 7.6	0.213	0.580	0.402
BMI CATS-I (kg/m^2)	25.4 [22.7 - 28.7]	25.0 [22.4 - 28.3]	26.0 [23.4 - 30.1]	25.6 [23.0 - 29.2]	0.111	0.112	0.806
BMI (kg/m^2)	26.1 [23.1 - 30.3]	25.8 [22.9 - 30.0]	28.3 [24.6 - 32.6]	25.8 [23.1 - 29.8]	**0.029**	0.672	0.139
DXA SCAN[1]							
WBLH Lean Mass (kg)	39.2 ± 5.6	38.9 ± 5.2	40.0 ± 5.7	39.3 ± 6.4	0.291	0.698	0.661
WBLH Fat Mass (kg)	28.3 ± 10.7	27.5 ± 10.3	31.5 ± 11.4	27.8 ± 10.7	**0.016**	0.791	0.084
WBLH Fat Mass (%)	40.7 ± 7.3	40.2 ± 7.2	42.8 ± 7.2	40.4 ± 7.4	**0.017**	0.784	0.072
LS-BMD (g/cm^2)	1.07 ± 0.12	1.06 ± 0.11	1.06 ± 0.12	1.08 ± 0.12	0.834	0.150	0.541
FN-BMD (g/cm^2)	0.84 ± 0.11	0.83 ± 0.11	0.85 ± 1.24	0.83 ± 0.11	0.297	0.949	0.321
BIOCHEMICAL[2]							
TSH CATS-I (mU/L)	1.66 [0.94 - 3.37]	1.22 [0.77 - 1.79]	3.37 [1.22 - 4.45]	4.21 [2.33 - 5.23]	**<0.001**	**<0.001**	0.098
TSH (mU/L)	1.64 [1.10 - 2.52]	1.54 [1.12 - 2.07]	2.45 [1.43 - 3.50]	1.68 [0.89 - 2.96]	**0.015**	0.474	0.070
FT4 (pmol/L)	14.0 ± 2.3	13.6 ± 1.7	13.9 ± 2.8	15.0 ± 2.9	0.471	**<0.001**	0.200
FT3 (pmol/L)	4.13 ± 0.51	4.15 ± 0.45	4.07 ± 0.53	4.11 ± 0.63	0.239	0.299	0.428
TPOAb Positive/Total (%)	79/294 (26.9)	21/173 (12.1)	20/50 (40.0)	38/71 (53.5)	**<0.001**	**<0.001**	0.196
Cholesterol TOT (mmol/L)	4.93 ± 0.87	4.91 ± 0.83	5.07 ± 0.93	4.88 ± 0.94	0.224	0.708	0.506
Cholesterol HDL (mmol/L)	1.29 ± 0.32	1.32 ± 0.31	1.24 ± 0.33	1.24 ± 0.34	0.125	0.101	0.785
Triglyceride (mmol/L)	0.90 [0.70 - 1.20]	0.90 [0.70 - 1.10]	1.01 [0.78 - 1.40]	0.80 [0.64 - 1.29]	**0.041**	0.730	0.193
Insulin (μIU/mL)	5.90 [4.40 - 7.80]	5.90 [4.60 - 7.80]	6.30 [4.40 - 9.15]	5.50 [3.85 - 7.15]	0.231	0.073	**0.046**
Adiponectin (ng/mL)	10.59 ± 4.80	10.63 ± 4.71	10.54 ± 4.65	10.52 ± 5.19	0.964	0.988	0.872
CARDIOVASCULAR[3]							
SP (mmHg)	125.7 ± 12.8	126.6 ± 12.9	124.6 ± 9.3	122.1 ± 14.4	0.569	0.155	0.220
DP (mmHg)	70.3 ± 8.3	70.6 ± 8.4	70.8 ± 8.0	68.7 ± 7.9	0.966	0.409	0.578
AI (%)	17.5 ± 7.2	18.1 ± 7.5	15.6 ± 4.8	16.5 ± 7.00	0.186	0.588	0.494
PWV (m/s)	8.80 [7.50 - 10.95]	9.28 [7.66 - 11.17]	7.87 [7.37 - 9.47]	7.93 [6.83 - 9.77]	0.491	0.101	0.605

The results are presented as mean ± SD or median [IQR], if Normally- or non-Normally distributed, respectively. All values refer to the Controlled Antenatal Thyroid Screening (CATS) study II analysis, except for BMI CATS-I and TSH CATS-I, relative to CATS study I analysis.

[1] = Data available for 327 subjects; [2] = Data available for 294 subjects; [3] = Data available for 194 subjects.

AI = augmentation index. **BMD** = bone mineral density. **BMI** = body mass index. **DP** = diastolic blood pressure. **DXA** = dual-energy x-ray absorptiometry. **FN** = femoral neck. **FT3** = free-triiodothyronine. **FT4** = free-thyroxine. **LS** = lumbar spine (from L1 to L4). **NGTF** = normal gestational thyroid function. **PWV** = aortic pulse wave velocity. **SGTF-T** = suboptimal gestational thyroid function treated with levothyroxine during pregnancy (Treated). **SGTF-U** = suboptimal gestational thyroid function not treated with levothyroxine during pregnancy (Untreated). **SP** = systolic blood pressure. **TPOAb** = autoantibodies to thyroid peroxidase. **TSH** = thyrotropin. **WBLH** = whole body less head.

p1U = NGTF vs SGTF-U. **p1T** = NGTF vs SGTF-T. **p2** = SGTF-U vs SGTF-T.

Table 3. Levothyroxine treatment

	TOT N = 332	**NGTF N = 197**	**SGTF-U N = 57**	**SGTF-T N = 79**	**p**
Never	176 (53.0%)	139 (70.6%)	38 (66.7%)	0 (0.0%)	**<0.001**
Yes stopped	36 (10.8%)	3 (1.5%)	0 (0.0%)	33 (41.8%)	**<0.001**
Yes current	65 (19.6%)	8 (4.1%)	16 (28.1%)	41 (51.9%)	**0.008**
Unknown	55 (16.6%)	47 (23.8%)	3 (5.2%)	5 (6.3%)	1.000

History of women's levothyroxine treatment collected at recruitment into the Controlled Antenatal Thyroid Screening study II (CATS-II).
NGTF = normal gestational thyroid function. **SGTF-T** = suboptimal gestational thyroid function treated with levothyroxine during pregnancy (Treated). **SGTF-U** = suboptimal gestational thyroid function not treated with levothyroxine during pregnancy (Untreated). **Yes stopped** = taken in the past but not currently. **Yes current** = started in the past and currently still on treatment.
Reported p values refer to the comparisons between SGTF-U and SGTF-T

Supplemental Table 1. Models used for analysis

MODEL	INCLUDED VARIABLES	PARAMETERS EVALUATED	COMPARISON
1	None (unadjusted)	Age	All
2	Age, ethnicity, social class, smoking during pregnancy	All numerical parameters except age	NGTF vs SGTF-U (p1U) NGTF vs SGTF-T (p1T)
3	Age, ethnicity, social class, smoking during pregnancy, baseline (CATS-I) TSH, baseline (CATS-I) FT4	All numerical parameters except age	SGTF-U vs SGTF-T (p2) SGTF-Topt vs SGTF-Tover (p)

CATS-I = Controlled Antenatal Thyroid Screening study I. **FT4** = free-thyroxine. **NGTF** = normal gestational thyroid function. **SGTF-T** = suboptimal gestational thyroid function treated with levothyroxine during pregnancy (Treated). **SGTF-Topt** = SGTF-T optimally treated with levothyroxine. **SGTF-Tover** = SGTF-T over-treated with levothyroxine. **SGTF-U** = suboptimal gestational thyroid function not treated with levothyroxine during pregnancy (Untreated). **TSH** = thyrotropin.

Supplemental Table 2. Levothyroxine over-treatment sub-analysis: mothers

	SGTF-T_{opt} N = 58	**SGTF-T_{over} N = 21**	**p**
Age (years)	40.3 ± 4.8	38.1 ± 4.7	0.069
Height (cm)	164.9 ± 7.4	161.8 ± 7.7	**0.032**
BMI CATS-I (kg/m^2)	26.4 [23.8 - 32.0]	23.2 [21.3-25.7]	**0.001**
BMI (kg/m^2)	26.7 [23.9 - 30.9]	23.5 [22.2 - 25.3]	**0.002**
DXA SCAN[1]			
WBLH Lean Mass (kg)	40.9 ± 6.3	34.8 ± 4.3	**<0.001**
WBLH Fat Mass (kg)	30.1 ± 11.1	21.9 ± 7.0	**0.007**
WBLH Fat Mass (%)	41.3 ± 7.2	37.9 ± 7.6	0.180
LS-BMD (g/cm^2)	1.10 ± 0.12	1.05 ± 0.11	0.198
FN-BMD (g/cm^2)	0.85 ± 0.12	0.80 ± 0.73	0.111
BIOCHEMICAL[2]			
TSH CATS-I (mU/L)	4.09 [1.94 - 5.30]	4.40 [3.75 - 4.62]	0.379
TSH (mU/L)	1.52 [0.86 - 2.96]	1.76 [1.22 - 3.32]	0.962
FT4 (pmol/L)	15.06 ± 2.79	14.73 ± 3.31	0.762
FT3 (pmol/L)	4.13 ± 0.66	4.05 ± 0.56	0.867
TPOAb Positive/Total (%)	23/51	10/20	0.794
Cholesterol TOT (mmol/L)	4.81 ± 0.85	5.07 ± 1.16	0.073
Cholesterol HDL (mmol/L)	1.15 ± 0.33	1.44 ± 0.27	**0.002**
Triglyceride (mmol/L)	0.91 [0.68 - 1.29]	0.80 [0.60 - 1.21]	0.562
Insulin (μIU/mL)	5.50 [3.90 - 7.50]	5.65 [3.75 - 6.85]	0.772
Adiponectin (ng/mL)	10.16 ± 5.70	11.44 ± 3.53	0.727
CARDIOVASCULAR[3]			
SP (mmHg)	127.6 ± 13.5	113.7 ± 11.7	**0.036**
DP (mmHg)	70.4 ± 8.0	66.0 ± 7.3	0.767
AI (%)	17.8 ± 6.3	14.6 ± 7.9	0.746
PWV (m/s)	8.05 [6.83 - 13.03]	7.13 [6.80 - 8.17]	0.855

The results are presented as mean ± SD or median [IQR], if having a normal or non-normal distribution, respectively. All values refer to the Controlled Antenatal Thyroid Screening (CATS) study II analysis, except for BMI CATS-I and TSH CATS-I, relative to CATS study I analysis.

[1] = Data available for 56 SGTF-T_{opt} and 21 SGTF-T_{over} subjects;

[2] = Data available for 51 SGTF-T_{opt} and 20 SGTF-T_{over} subjects;

[3] = Data available for 17 SGTF-T_{opt} and 11 SGTF-T_{over} subjects.

AI = augmentation index. **BMD** = bone mineral density. **BMI** = body mass index. **DP** = diastolic blood pressure. **DXA** = dual-energy x-ray absorptiometry. **FN** = femoral neck. **FT3** = free-triiodothyronine. **FT4** = free-thyroxine. **LS** = lumbar spine (from L1 to L4). **PWV** = aortic pulse wave velocity. **SGTF-Topt** = suboptimal gestational thyroid function optimally treated with levothyroxine during pregnancy. **SGTF-Tover** = suboptimal gestational thyroid function over-treated with levothyroxine during pregnancy. **SP** = systolic blood pressure. **TPOAb** = autoantibodies to thyroid peroxidase. **TSH** = thyrotropin. **WBLH** = Whole body less head.

Supplemental Table 3. Levothyroxine over-treatment sub-analysis: children

	SGTF-Topt N = 57	**SGTF-Tover N = 21**	**p**	**p[4]**	**p[5]**
Age (years)	8.8 ± 1.2	9.0 ± 0.9	0.516	NA	NA
Female children N (%)	30 (52.6%)	8 (38.1%)	0.260	NA	NA
Height (cm)	135.1 ± 7.7	132.9 ± 8.1	**0.016**	NA	**0.044**
BMI (kg/m^2)	17.6 [16.0 - 19.1]	15.2 [14.9 - 16.7]	**0.005**	NA	**0.041**
BMI-SDS UK1990 (SDS)	0.70 ± 1.07	-0.19 ± 1.14	**0.001**	NA	**0.006**
DXA SCAN[1]					
WBLH Lean Mass (kg)	19.6 ± 3.3	17.6 ± 3.9	**0.004**	NA	0.109
WBLH Fat Mass (kg)	9.66 ± 4.64	7.22 ± 4.19	0.087	NA	0.298
WBLH Fat Mass (%)	31.9 ± 8.0	28.0 ± 7.1	0.234	NA	0.511
LS-BMD-H (SDS)	0.59 ± 1.00	0.26 ± 1.24	0.174	0.394	0.603
LS-BMD-A (SDS)	0.51 ± 1.07	0.23 ± 1.21	0.217	0.449	0.669
LS-BMAD-A (SDS)	0.57 ± 1.07	0.42 ± 1.25	0.405	0.423	0.679
LS-BMAD-W (SDS)	0.46 ± 1.05	0.33 ± 1.27	0.455	0.467	0.702
FN-BMD-H (SDS)	0.33 ± 1.05	-0.19 ± 1.30	**0.037**	0.164	0.272
FN-BMAD-W (SDS)	0.31 ± 1.03	0.23 ± 0.99	0.952	0.803	0.988
WBLH-BMD-H (SDS)	-0.62 ± 0.80	-1.44 ± 1.19	**0.002**	**0.036**	0.216
WBLH-BMD-A (SDS)	-0.57 ± 0.96	-1.39 ± 1.26	**0.002**	**0.042**	0.193
BIOCHEMICAL[2]					
TSH (mU/L)	1.81 [1.66 - 2.76]	1.82 [1.55 - 2.41]	0.540	NA	0.526
FT4 (pmol/L)	14.91 ± 1.24	14.40 ± 1.40	0.251	NA	0.280
FT3 (pmol/L)	5.83 ± 0.55	4.77 ± 0.92	**0.036**	NA	0.051
TPOAb Positive/Total (%)	1/19 (5.3)	0/4 (0.0)	1.000	NA	NA
Cholesterol TOT (mmol/L)	4.24 ± 0.53	4.13 ± 0.68	0.197	NA	0.218
Cholesterol HDL (mmol/L)	1.08 ± 0.21	1.20 ± 0.14	0.717	NA	0.332
Triglyceride (mmol/L)	0.68 [0.60 - 0.90]	0.65 [0.55 - 0.75]	0.484	NA	0.505
Insulin (μIU/mL)	3.70 [3.30 - 5.60]	4.00 [2.80 - 4.60]	0.824	NA	0.544
Adiponectin (ng/mL)	13.91 ± 5.92	15.87 ± 6.76	0.292	NA	0.298
CARDIOVASCULAR[3]					
SP (mmHg)	125.3 ± 16.9	130.5 ± 10.2	0.768	NA	0.260
DP (mmHg)	64.2 ± 8.4	63.9 ± 7.8	0.586	NA	0.637
AI (%)	6.67 [3.67 - 11.33]	7.12 [7.00 - 9.00]	0.367	NA	0.412
PWV (m/s)	6.03 ± 0.46	6.37 ± 0.77	**0.041**	NA	**0.035**

The results are presented as mean ± SD or median [IQR], if having a normal or non-normal distribution, respectively.

[1] = Data available for 56 SGTF-T_{opt} and 21 SGTF-T_{over} subjects;

[2] = Data available for 19 SGTF-T_{opt} and 4 SGTF-T_{over} subjects;

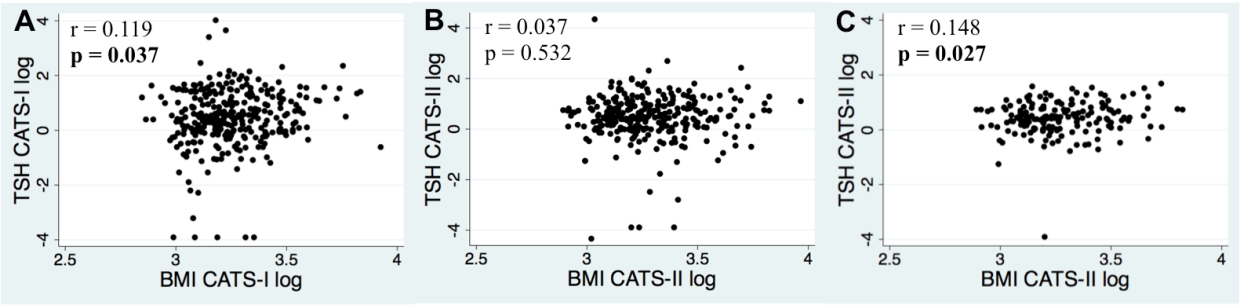
[3] = Data available for 18 SGTF-T_{opt} and 10 SGTF-T_{over} subjects;

[4] = linear regression adjusted also for child's height;

[5] = linear regression adjusted also for the corresponding maternal parameter. For BMI SDS used mother's BMI. For bone DXA scan standard deviation scores used the maternal BMD of the same area.

A = Alphabet study reference cohort. **AI** = augmentation index. **BMAD** = bone mineral apparent density. **BMD** = bone mineral density. **BMI** = body mass index. **BMI-SDS** = standard deviation score of body mass index. **DP** = diastolic blood pressure. **DXA** = dual-energy x-ray absorptiometry. **FN** = femoral neck. **FT3** = free-triiodothyronine. **FT4** = free-thyroxine. **LS** = lumbar spine (from L1 to L4). **NA** = Not Applicable. **PWV** = aortic pulse wave velocity. **SGTF-Topt** = children of women with suboptimal gestational thyroid function optimally treated with levothyroxine during pregnancy. **SGTF-Tover** = children of women with suboptimal gestational thyroid function over-treated with levothyroxine during pregnancy. **SDS** = standard deviation score. **SP** = systolic blood pressure. **TPOAb** = autoantibodies to thyroid peroxidase. **TSH** = thyrotropin. **W** = Ward study reference cohort. **WBLH** = whole body less head.

Supplemental Figure 1. Correlation between BMI and TSH: mothers



Panel A: entire women cohort at the Controlled Antenatal Thyroid Screening study I (CATS-I). **Panel B**: entire women cohort at the Controlled Antenatal Thyroid Screening study II (CATS-II). **Panel C**: women cohort at CATS-II excluding those with suboptimal gestational thyroid function treated with levothyroxine during pregnancy (SGTF-T).