

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/109041/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Hales, Charlotte , Taylor, Peter N. , Channon, Sue , Paradise, Ruth, McEwan, Kirsten, Zhang, Lei , Gyedu, Michael, Bakhsh, Ameen, Okosieme, Onyebuchi, Muller, Ilaria , Draman, Mohd S., Gregory, John W. , Dayan, Colin , Lazarus, John H., Rees, D. Aled and Ludgate, Marian 2018. Controlled antenatal thyroid screening II: effect of treating maternal sub-optimal thyroid function on child cognition. *Journal of Clinical Endocrinology and Metabolism* 103 (4) , pp. 1583-1591. 10.1210/jc.2017-02378

Publishers page: <http://dx.doi.org/10.1210/jc.2017-02378>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



1 **Controlled Antenatal Thyroid Screening II: effect of treating maternal sub-optimal**
2 **thyroid function on child cognition.**

3 Charlotte Hales¹, Peter N Taylor¹, Sue Channon², Ruth Paradise³, Kirsten McEwan²,
4 Lei Zhang¹, Michael Gyedu¹, Ameen Bakhsh¹, Onyebuchi Okosieme¹, Ilaria Muller¹,
5 Mohd S Draman¹, John W Gregory¹, Colin Dayan¹, John H Lazarus¹, D Aled Rees⁴,
6 and Marian Ludgate¹

7 ¹School of Medicine, Cardiff University, UK. ²Centre for Trials Research, Cardiff
8 University, UK. ³St David's Hospital, Cardiff & Vale University Health Board, UK.
9 ⁴Neuroscience and Mental Health Research Institute, Cardiff University, UK.

10 **Page header:** CATS follow-up assessment of child cognition

11 **Key words:** pregnancy; cognition; subclinical hypothyroidism; subclinical
12 hyperthyroidism; hypothyroidism; hyperthyroidism.

13 **Correspondence & requests for reprints** to: Prof Marian Ludgate, Division of
14 Infection and Immunity, School of Medicine, Cardiff University, Cardiff, Wales UK,
15 ludgate@cardiff.ac.uk phone +44 2920 745457

16 **Funded by** The Charles Wolfson Trust, Action Medical Research (project code
17 GN2033)/The Henry Smith Charity (20122759 GN 2033, grants to ML) and American
18 Thyroid Association (grant to PT).

19 **Declaration of interests:** All authors declare no competing interests.

20 **Word Count 4042**

21 **Abstract**

22 *Context & Objective:* The Controlled Antenatal Thyroid Screening (CATS) study
23 investigated treatment for suboptimal gestational thyroid function (SGTF) on childhood
24 cognition and found no difference in IQ at 3 years between children of treated and
25 untreated SGTF mothers. We have **measured** IQ in the same children at age 9.5-years
26 and included children from normal-GTF mothers.

27 *Design, Setting & Participants:* One examiner, blinded to participant group, assessed
28 children's IQ (WISC-IV), long-term memory and motor function (NEPSY-II) from
29 children of 119 treated and 98 untreated SGTF mothers plus children of 232 mothers
30 with normal-GTF. Logistic regression explored the odds and percentages of IQ<85 in
31 the groups.

32 *Results:* There was no difference in IQ<85 between children of mothers with normal-
33 GTF and combined SGTF i.e. treated and untreated (fully adjusted OR=1.15 (95% CI
34 0.52, 2.51) p=0.731). Furthermore, there was no significant effect of treatment
35 (untreated OR=1.33 (95% CI 0.53, 3.34), treated OR=0.75 (95% CI 0.27, 2.06)
36 p=0.576). IQ< 85 was 6.03% in normal-GTF, 7.56% in treated and 11.22% in untreated
37 groups. Analyses accounting for treated-SGTF women with FT4 >97.5th centile of the
38 entire CATS-I cohort revealed no significant effect on child's IQ<85 in CATS-II. IQ at
39 age 3 predicted IQ at age 9.5 (p<0.0001) and accounted for 45% of the variation.

40 *Conclusions:* Maternal thyroxine during pregnancy did not improve child cognition at
41 age 9.5 years. Our findings confirmed CATS-I and suggest that the lack of treatment
42 effect may be due to the similar proportion of IQ<85 in children of women with normal-
43 GTF and SGTF.

44 **Précis**

45 Cognitive assessments of children aged 9 from the first CATS study confirms no effect
46 of treatment for maternal SGTF on IQ<85 and no IQ difference when compared with
47 children from euthyroid mothers.

48

49 **Introduction**

50 Triiodothyronine (T3) and thyroxine (T4) are essential for early brain development, and
51 maternal thyroid hormones are required by the fetus until its own thyroid starts to
52 function, which can be as late as 18 weeks gestation. (1-3). Prior to this thyroid
53 hormones in the fetal brain are solely of maternal origin (4,5). Thyroid dysfunction
54 occurs in around 2.5% of pregnancies (6) and severe hypothyroidism during the first
55 two trimesters may result in irreversible neurological deficits, although the effect of
56 more modest variation in thyroid hormone levels is unclear. Later in pregnancy the
57 fetus may be better able to compensate for any lack of maternal thyroid hormones but
58 compensation is likely to be incomplete until the fetal thyroid is fully functional at term.
59 (7).

60 Several studies reported that higher levels of maternal thyroid stimulating hormone
61 (TSH) during pregnancy may be associated with a negative impact on the child's
62 intelligence (8-11), but this was not confirmed by others (12,13). Likewise, findings for
63 low maternal T4 levels are contradictory with some (9,13-17), but not all (10,18-21)
64 studies providing evidence of lowered intelligence in the children. As well as
65 intelligence quotient (IQ) and general cognition, further deficits for offspring following
66 exposure to underactive maternal thyroid function have been identified; including
67 memory (15,22-25) and motor difficulties (8,9,16,26,27), amongst others.

68 The Controlled Antenatal Thyroid Screening (CATS) study commenced in 2002
69 (CATS-I) and was the first randomised controlled trial (RCT) to investigate the effect
70 of screening and treatment for hypothyroidism during pregnancy on child cognition
71 (28). Women (n=21,846) were recruited at a median gestation of 12 weeks, 3 days;
72 (ten UK centres and one in Turin, Italy). Mothers were defined as having suboptimal

73 gestational thyroid function (SGTF) if their FT4 was <2.5th percentile and/or TSH
74 >97.5th percentile as assessed during the CATS study and half were treated with
75 150µg thyroxine daily. Offspring born to SGTF mothers had their IQ assessed at age
76 3 years no difference was found between those whose mothers were treated (mean
77 IQ 99.2) or untreated (100.0) during pregnancy ($p=0.40$). Similar results were obtained
78 in a recent study from Casey and colleagues who reported no beneficial effect, on
79 offspring cognition up to age 5, of treating mothers with subclinical hypothyroidism or
80 hypothyroxinemia at 16.7 or 17.8 weeks mean gestation respectively (29). The young
81 age of the children when tested in these large RCTs might explain the reported lack
82 of treatment effect. IQ evaluations below age 5 may serve as a general indicator of
83 cognitive function but may not be best suited as a longer term measure of cognitive
84 function (30). Therefore the primary aim of CATS II was to measure the children's
85 cognitive function at age 9 years using a more in-depth battery of tests. Furthermore,
86 neither of these trials compared the IQs of children from euthyroid mothers with those
87 of SGTF mothers to elucidate whether there is a deficit requiring treatment. Our
88 second aim addressed this point by assessing cognitive function in children from
89 mothers with normal gestational thyroid function (normal-GTF). The dose of thyroxine
90 used in the CATS study was relatively high and recent reports suggest adverse effects
91 of cognition from both too much and too little thyroid hormone (31). Consequently we
92 explored a possible effect of 'over-treatment' (defined as maternal FT4 above the
93 97.5th percentile of the CATS-I UK cohort) on IQ scores. Finally we analysed the
94 correlation between cognitive assessments undertaken at age 3 and 9 years as this
95 will be invaluable when designing future studies.

96 **Methods**

97 **Study Design and Population**

98 The original CATS study was previously described in detail (28). Briefly CATS-I
99 recruited 21,846 women (excluding history of thyroid disease, twin pregnancies,
100 maternal age <18 years or gestational age >15 weeks and 6 days), predominantly in
101 the UK, at their first antenatal hospital appointment. Participants were randomized
102 either to screen (treated) or control (untreated) groups; the former having their thyroid
103 function tested immediately and the latter after their child was born. If the mother's
104 FT4 was <2.5th percentile and/or TSH >97.5th percentile, they were classified as
105 having SGTF; percentiles being calculated from the CATS cohort. Women in the
106 screen group with SGTF were treated with levothyroxine (starting dosage 150µg) for
107 the remainder of their pregnancies. The primary outcome was children's IQ at age 3
108 from the screen and control groups.

109 CATS-II included only UK participants for logistical reasons (n=16,346). The target
110 sample size was informed by prior power calculations (see below). All CATS mothers
111 from the UK SGTF treated and untreated groups (n=609) were invited to participate
112 by letter. The Welsh Demographics Service and Patient Data Registrar provided
113 current addresses. Those without SGTF in the control and screen branches of the
114 RCT, were pooled (UK n=15 737), and named 'normal-GTF'; a random sample of
115 4,000 from this group was also invited to participate, again by letter (figure 1).

116 **Cognitive Assessments**

117 CATS-II IQ and additional cognitive assessments were conducted when children were
118 aged 7.00 to 10.92 years (32); either in the research centre or in their homes. One
119 psychologist (CH) undertook all of the CATS-II assessments to allow good consistency

120 and was unaware of participant group. Ten percent of assessments were double
121 scored (RP) to ensure accuracy (32). IQ was measured using the Wechsler
122 Intelligence Scale for Children (WISC) fourth edition UK version which generated a
123 full-scale IQ (FSIQ) calculated equally from four sub-domains: verbal comprehension
124 IQ (VCIQ), perceptual reasoning IQ (PRIQ), working memory IQ (WMIQ) and
125 processing speed IQ (PSIQ). Additional cognitive assessments (8,22) were
126 administered to some children (those not too tired following WISC administration)
127 using the Developmental Neuropsychological Assessment (NEPSY) second edition,
128 details can be found in the supplemental information. These assessments tested long-
129 term memory (memory for designs delayed- MDD, and list memory- LM), working
130 memory (memory for designs- MD, and narrative memory- NM) and fine motor
131 coordination (fingertip tapping dominant hand- FTDH, and fingertip tapping non-
132 dominant hand, FTNDH). As the normal-GTF group means for both assessments were
133 close to the anticipated values (WISC-IV IQ:100, additional NEPSY assessments:10),
134 the authors conclude there was no selection bias in which children completed all
135 assessments in CATS-II.

136 CATS-II was approved by the Wales Research Ethics Committee 2 (reference
137 10/WSE03/33) and Cardiff & Vale University Health Board. Written and informed
138 consent was obtained from all mothers both in CATS-II and initially during their
139 pregnancies; child assent was obtained during the research centre visits. Missing data
140 were largely due to non-response to invitation.

141 **Sample Size Justification**

142 Samples of 120 participants from the treated (CATS-I screen) and untreated SGTF
143 (CATS-I control) groups would have 90% power to detect a difference of 6 points in
144 mean IQ (31) or 80% power with a 5% two-sided significance level to detect a 1.97

145 increase in odds of IQ < 85 in untreated SGTF assuming mean IQ to be 100 with a SD
146 of 15 (32). 240 participants (1.5%) from the normal-GTF group (CATS-I normal thyroid
147 function in test and screen groups) were required to assess whether maternal SGTF
148 influenced her child's IQ.

149 **Analyses**

150 The data were analysed in SPSS version 20 and STATA version 12 in accordance
151 with the pre-specified statistical plan (32).

152 The primary analysis assessed the odds of FSIQ <85 in the normal-GTF and the
153 merged SGTF group; an interaction term for treatment of SGTF was then added, all
154 using logistic regression. Mean IQ differences and percentages with FSIQ<85 were
155 also compared between the three groups. Univariate analysis was followed by
156 multivariate analysis to adjust for key potential covariates in four models:

157 Model 1; Crude

158 Model 2; adjusted for child sex

159 Model 3; adjusted for model 2, and age of mothers at birth of offspring and whether
160 the child was breastfed.

161 Model 4; adjusted for model 3, and schooling (Welsh- or English-medium school
162 attended), place of assessment (home or research centre) and socioeconomic status
163 (calculated from postcode social deprivation scores obtained from
164 [https://statswales.wales.gov.uk/Catalogue/Community-Safety-and-Social-
165 Inclusion/Welsh-Index-of-Multiple-Deprivation/WIMD-2014](https://statswales.wales.gov.uk/Catalogue/Community-Safety-and-Social-Inclusion/Welsh-Index-of-Multiple-Deprivation/WIMD-2014) for Wales and
166 <http://apps.opendatacommunities.org/showcase/deprivation> for England. A score of 1
167 signifies most deprived and 5 least deprived.)

168 Step-wise analysis of covariates was performed only for binary outcomes but all six
169 covariates were included in continuous analyses.

170 Secondary analyses explored several aspects. We assessed using likelihood ratio
171 tests whether response to treatment best fitted a proportional or non-proportional
172 model. Using the normal-GTF to the untreated SGTF data we could investigate
173 whether maternal TSH influenced FSIQ.

174 We also compared subdomain-IQs (VCIQ, PRIQ, WMIQ, and PSIQ) in the treated and
175 untreated SGTF groups to explore the effect of treatment; initially by logistic
176 regressions for scores <85, then a multivariate analysis of covariance (MANCOVA,
177 adjusted for the six covariates) for mean scores. The additional cognitive assessments
178 were also compared by a MANCOVA, and an analysis of covariance (ANCOVA) for
179 the LM subtest (reduced dataset due to late introduction).

180 Sensitivity analyses comprised comparison of CATS-I and CATS-II VCIQ, PRIQ and
181 FSIQs using Pearson correlations.

182 As exploratory analyses, within the broad term of SGTF, we investigated subclinical
183 hypothyroidism (FT4 >2.5th and TSH >97.5th percentiles), isolated hypothyroxinaemia
184 (FT4 <2.5th and TSH <97.5th percentiles), and overt hypothyroidism (FT4 <2.5th and
185 TSH >97.5th percentiles). These were calculated by MANCOVAs (IQs, additional
186 cognitive assessments, and LM) to include interactions between the three groups;
187 normal-GTF, and whether maternal SGTF was treated or not.

188 Finally, we explored differences between participants taking account of those we
189 defined as 'overtreated' i.e the treated SGTF group whose FT4 values were above the
190 17.7pmol/L threshold established by the 97.5th percentile at recruitment in the UK

191 CATS sample. We compared over-supplementation to child FSIQ <85 first, followed
192 by analyses for mean scores, all adjusted for the same covariates detailed above.

193 Supplemental exploratory analyses can be found in the supplemental information:
194 Subclinical hypothyroidism, isolated hypothyroxinemia and overt hypothyroidism.

195 **Results**

196 **Group characteristics**

197 In CATS-I, 16 346 women were UK-based and provided the prospective cohort for
198 CATS-II. There were 382 treated and 371 untreated for SGTF; of these, 303 treated
199 and 306 untreated SGTF offspring completed IQ testing at age 3.2 years. No data
200 were collected from the normal-GTF group.

201 In CATS-II, IQ assessment occurred in a total of 449 children at a mean age of 9.5
202 years; 119 treated SGTF, 98 untreated SGTF, and 232 from the normal-GTF group
203 (figure 1). Smaller groups completed the additional cognitive assessments (please see
204 supplementary data for explanations); 110 treated SGTF, 85 untreated SGTF, and
205 215 normal-GTF.

206 At recruitment into CATS-I, CATS-II mothers from normal-GTF, treated and untreated
207 SGTF groups had median TSHs of 1.16, 4.09, and 3.57mU/L, respectively and mean
208 FT4s were 14.12, 11.92, and 11.79pmol/L, respectively (table 1). The CATS-I and
209 CATS-II SGTF samples were largely unbiased (statistics presented in supplementary
210 table 1).

211 Significant differences between the CATS-II participant groups are detailed in table 1.
212 As anticipated, maternal FT4 and TSH at recruitment into CATS-I were higher (FT4)
213 and lower (TSH), in the normal-GTF compared with both SGTF groups. Maternal TSH

214 was higher in the treated compared with untreated SGTF CATS-II mothers. Mean
215 maternal age at consent into CATS-I was higher in the normal-GTF compared to the
216 treated SGTF group, though only by 0.8 years. Similarly, a difference between the
217 groups was seen in those from the SGTF groups being more likely to opt for
218 participation from their home rather than attending the research clinic. The children in
219 the normal-GTF group were significantly older (by just 4 months) than the SGTF
220 groups.

221 **Primary analysis**

222 There was no significant difference for odds of FSIQ <85 between the normal-GTF
223 and merged SGTF groups (fully adjusted OR=1.15 (95% CI 0.52, 2.51) p=0.731). This
224 non-significant finding was sustained when an interaction term for treatment was
225 included, although treatment improved FSIQ (untreated fully adjusted OR=1.33 (95%
226 CI 0.53, 3.34), treatment OR=0.75 (95% CI 0.27, 2.06) p=0.576). Table 2 displays the
227 FSIQ regression models.

228 The percentages with IQ<85 were 6.03% in normal-GTF, 7.56% in treated and 11.22%
229 in untreated SGTF groups (table 3, Chi p for the trend = 0.11).

230 **Secondary analyses**

231 *Do data fit a proportional or non-proportional model?*

232 Mean child FSIQs per group were 103.10 (SD 11.68), 101.76 (12.04), and 102.31
233 (13.27), for the normal-GTF, treated and untreated SGTF groups, respectively (table
234 3). There was no difference between the mean FSIQ scores of the three participant
235 groups (p=0.678). There was no significant difference for odds of the normal-GTF
236 children having higher FSIQs compared to the treated SGTF children (OR=0.99 (95%

237 CI 0.38, 2.52) $p=0.98$); this was due to a mean IQ difference of only 0.79 between the
238 groups.

239 *Does maternal TSH predict FSIQ?*

240 Analysis of the relationship between FSIQ and thyroid status in normal-GTF and
241 untreated SGTF revealed no clear association between TSH ($B=0.43$ (95% CI -0.68,
242 1.56) $p=0.442$) and FT4 ($B= 0.33$ (95% CI -0.25, 0.91) $p=0.270$) on FSIQ in the fully
243 adjusted model.

244 Analysis of women with SGTF by dividing FSIQ score into quintiles did not reveal any
245 benefit of treatment in the fully adjusted model ($p=0.98$) with no evidence of a non-
246 proportional effect ($p=0.75$) (data not shown).

247 *Does treatment for SGTF affect any subdomain?*

248 No differences were found between subdomain-IQ scores <85 (see table 2 for sub-IQ
249 regression models) or for mean subdomain-IQ scores for VCIQ, PRIQ, WMIQ, and
250 PSIQ between the groups ($p=0.193$). The mean scores of the additional cognitive
251 assessments were also compared, with no difference identified between the three
252 participant groups ($p=0.732$, LM $p=0.266$, table 3).

253 **Sensitivity Analysis**

254 As CATS-II followed the UK sample we analysed the CATS-I UK only cohort ($n=609$)
255 and revealed $IQ<85$ in 14% treated and 17% untreated, the difference was not
256 significant. Furthermore there was no significant difference in percentage $IQ<85$
257 treated versus untreated in the CATS-II subset of CATS-I ($n=212$).

258 Pearson correlations to assess how associated the scores were from the WPPSI-III
259 and the WISC-IV for the treated and untreated SGTF groups found that all scores were

260 positively correlated ($p < 0.0001$). Furthermore age 3 IQ predicts 45% of the variation
261 in age 9 IQ with other variables such as breast feeding contributing only an additional
262 1%.

263 **Exploratory analyses**

264 Different types of abnormal thyroid function (subclinical hypothyroidism, isolated
265 hypothyroxinemia) were also explored using MANCOVA. No significant differences
266 were found in the mean IQ scores, IQ<85 or additional assessments between children
267 of treated and untreated mothers. Similar results were obtained in the offspring of a
268 small number of women with overt hypothyroidism identified during participation in
269 CATS, although IQ<85 was apparent in 0% of the treated but 10% of the untreated
270 groups. These analyses are presented in supplementary table 2.

271

272 *Over-supplementation*

273 Finally, we explored differences between participants, taking account of those in the
274 treated SGTF group with raised FT4 values (20 weeks mean FT4 16.19 (2.83), TSH
275 median 0.33 (0.08-0.99); 30 weeks mean FT4 15.56 (2.50), median TSH 0.27 (0.03-
276 0.84). The threshold for high FT4 was established by the 97.5th percentile recruitment
277 in the UK CATS sample (17.7pmol/L); one-third of the treated SGTF had FT4
278 >17.7pmol/L.

279 We compared over-supplementation to child FSIQ <85 first, followed by analyses for
280 mean scores, all adjusted for the same covariates detailed above. There was no
281 significant effect on child's IQ<85 and no difference between mean IQ scores of the
282 groups or additional cognitive assessments, including the LM subtest ($p=0.875$,
283 $p=0.765$, $p=0.951$, respectively), data not shown.

284 Of note, we observed no detrimental effect of over-supplementation on IQ<85 in
285 children of such women in CATS-I for whom we had information on FT4 levels after
286 therapy was initiated (UK cohort, n=609).

287 **Discussion**

288 We revisited the effects of treatment for SGTF on cognition in the CATS children at an
289 average age of 9.5 years. Our results confirm those of CATS-I, in that we saw no
290 significant differences in FSIQ<85 or mean IQ scores in the children of treated and
291 untreated women. Our results also confirm those of Casey and colleagues who
292 reported no beneficial effect, on offspring cognition up to age 5, of treating mothers
293 with subclinical hypothyroidism or hypothyroxinemia at 16.7 or 17.8 weeks mean
294 gestation respectively (29). Of interest Haddow et al (8) reported that mean FSIQ
295 scores and FSIQ scores <85 were not significantly different comparing children born
296 to mothers who were treated or not ($p=0.20$ and $p=0.90$, respectively), although the
297 study was retrospective and the treatment groups were small. In contrast to our
298 findings however the study by Haddow et al showed that the IQ of children born to
299 untreated mothers was significantly lower than those of control children.

300

301 One criticism of CATS-I was that cognitive assessments were conducted in children
302 at too young an age for differences to be evident Our current findings indicate that this
303 may not be the case as we found that IQ scores at age 3 and 9 were strongly correlated
304 in the two CATS studies with FSIQ at age 3 predicting 45% of the variability in FSIQ
305 at age 9 and with other factors contributing very little.

306

307 The design of the CATS-I study has also been questioned in relation to the timing of
308 initiation of levothyroxine therapy. The fetus relies wholly on maternal thyroxine
309 delivery up until about 14-18 weeks gestation when its own thyroid gland becomes
310 functional (7). Fetal brain development begins immediately after conception and
311 therefore treatment initiated at 12-13 weeks may have missed the early critical phase
312 of brain development. The CATS study participants were recruited during their first
313 scheduled visit to the antenatal clinic which generally fell towards the end of the first
314 trimester (median of 12 weeks and 3 days). (33) Similarly, thyroxine supplementation
315 in the study by Casey et al (29) was started even later and thus future trials would
316 benefit from recruiting women at a much earlier stage of pregnancy in order to
317 overcome these limitations.

318 A further consideration in the CATS study design is that the starting dose of
319 levothyroxine administered may have been too high and therefore adverse outcomes
320 in women who were over-treated may have masked any benefits of treatment. The
321 CATS-I study was the first RCT to investigate the effects of treatment for SGTF in
322 pregnancy and hence there were no previous studies for guidance. Furthermore, there
323 is no universal consensus on thyroxine supplementation dose even for the treatment
324 of women with overt hypothyroidism who become pregnant. Of note, guidelines for the
325 management of thyroid function during pregnancy recommend assay of TSH alone
326 and indeed treatment in CATS-I was monitored and adjusted based on TSH levels. As
327 a result, approximately one third of the treated mothers achieved a high FT4 which
328 was accompanied by a switch from a positive correlation between FT4 and age 9
329 cognition at recruitment to a negative correlation after treatment (supplemental
330 information). However, in contrast to a study illustrating a bi-phasic effect of FT4 on
331 cognition, with children of women with both low and high thyroxine levels displaying

332 lower IQs and smaller grey matter and cortex volumes (35)₍₃₁₎, we observed no
333 significant difference in the proportion of IQ<85 at age 9 in children of over-treated
334 mothers compared with the rest. Furthermore we did not find any detrimental effect on
335 IQ<85 in children of such women when we analysed the age 3 cognition data in CATS-
336 I (UK only cohort).

337 CATS-II included children from normal-GTF women and found no difference in IQ
338 measures between these and children from SGTF mothers, whether treated or not.
339 This confirmed previous studies reporting no effect of low thyroid function on offspring
340 intelligence or cognition (10,12,13,18-21) and may to some extent explain the absence
341 of treatment benefits observed in the trial. However our results contradict many animal
342 studies possibly because the thyroid abnormalities in the CATS mothers are mild when
343 compared with models induced e.g. by thyroidectomy. The lack of agreement on the
344 effects of FT4 on cognition in observational studies is the result of varying definitions
345 of SGTF, the lack of universal pregnancy-specific reference ranges for thyroid function
346 tests and the application of various tools to measure cognition in children across the
347 age spectrum. Hence it is not surprising that the benefits of universal screening during
348 pregnancy on cognition remain hotly debated although other adverse pregnancy
349 outcomes have been well-reported (such as pre-eclampsia, miscarriage, and preterm
350 birth) (34-36).

351 In our protocol paper (32) one of the secondary analyses planned to investigate
352 whether the combination of low maternal FT4 during pregnancy and the presence of
353 an adverse deiodinase 2 (D2) genotype in her child would impact cognition. The
354 hypothesis followed reports that Thr92Ala reduced conversion of thyroxine to tri-
355 iodothyronine (37). We genotyped 426 CATS children finding 73 alanine
356 homozygotes; when a mother had low FT4 during pregnancy and the child had the

357 homozygous alanine D2 genotype, treatment appeared to reduce the odds of
358 FSIQ<85 (reduced OR from 5.72 to 1.85), though this was non-significant and included
359 only a small number of the participants (data not shown).

360 Our study has some limitations although throughout all analyses adjustments were
361 made to control for extraneous effects. 1. The CATS-II power calculation was based
362 on an IQ difference of 6 points, as found by Haddow et al in offspring of women with
363 overt hypothyroidism. We studied women with less severe thyroid dysfunction and
364 thus the study may have been underpowered to detect more subtle cognitive variation.
365 2. This was exacerbated by the recruitment challenges we faced from the outset, with
366 the main problem due to participants having re-located since participating in CATS-I
367 and not responding to invitation. As the study developed, the recruitment process
368 evolved and rates improved but extending the data collection period would have taken
369 the children closer to puberty and its complications. 3. There were some differences
370 noted between the three groups raising the possibility of bias. However, significantly
371 older normal-GTF children than those from the SGTF groups should not have affected
372 the results since both assessment tools used have scores age-corrected in three
373 month intervals. Similarly differences in maternal age at recruitment and place of child
374 assessment were both covariates controlled for in the analyses.

375 In conclusion, results obtained in the current follow-up study have shown no effect of
376 thyroxine supplementation in women with SGTF on child IQ at age 9. These findings
377 support those of the original CATS-I study and a recent large RCT. Our data are
378 consistent with the lack of treatment effect being due to the similar proportion of IQ<85
379 in children of normal-GTF and SGTF mothers rather than the age of cognitive
380 assessment or the relatively high dose of thyroxine supplementation. However, future
381 large randomised trials, with thyroxine interventions at a much earlier stage of

382 pregnancy (or pre-conception), may still be warranted, since the benefits of treatment
383 may not be fully realised unless treatment is commenced early.

384 **Contributors**

385 CH collected the data, was involved in writing the report and analysed the data with
386 PNT. SC, RP, KM, LZ, MG, AB, OO, IM, MSD, JG, CD, JHL, and AR contributed to
387 study design, data analyses and writing the report. ML designed and managed the
388 project, supervised analyses and contributed to the report.

389 **Acknowledgments**

390 We are extremely grateful to the children, parents and families who participated in the
391 study. Special thanks are extended to Dionne Shillabeer, Julie Pell, Julie Evans,
392 Sophie Fuller, and Beverley Carey for their dedicated support to the CATS project.

393 **References**

394

- 395 1. Bernal J, Nunez J. Thyroid hormones and brain development. *European Journal of Endocrinology* 1995;
396 133:390-398
- 397 2. Ahmed OM, El-Gareib AW, El-bakry AM, Abd El-Tawab SM, Ahmed RG. Thyroid hormones states
398 and brain development interactions. *International Journal of Developmental Neuroscience* 2008; 26:147-
399 209
- 400 3. Zimmermann MB. Iodine deficiency. *Endocrine Reviews* 2009; 30:376-408
- 401 4. De Escobar GM. The role of thyroid hormone in fetal neurodevelopment. *Journal of Pediatric*
402 *Endocrinology and Metabolism* 2001; 14:1453-1462
- 403 5. Contempré B, Jauniaux E, Calvo R, Jurkovic D, Campbell S, Morreale De Escobar G. Detection of
404 thyroid hormones in human embryonic cavities during the first trimester of pregnancy. *Journal of Clinical*
405 *Endocrinology and Metabolism* 1993; 77:1719-1722
- 406 6. Lazarus JH, Premawardhana LDKE. Screening for thyroid disease in pregnancy. *Journal of Clinical*
407 *Pathology* 2005; 58:449-452
- 408 7. Rovet JF, Willoughby KA. Maternal thyroid function during pregnancy: Effects on the developing fetal
409 brain. *Maternal Influences on Fetal Neurodevelopment: Clinical and Research Aspects* 2010:55-77.
- 410 8. Haddow JE, Palomaki GE, Allan WC, Williams JR, Knight GJ, Gagnon J, O'Heir CE, Mitchell ML,
411 Hermos RJ, Waisbren SE, Faix JD, Klein RZ. Maternal thyroid deficiency during pregnancy and
412 subsequent neuropsychological development of the child. *The New England journal of medicine* 1999;
413 341:549-555
- 414 9. Li Y, Shan Z, Teng W, Yu X, Li Y, Fan C, Teng X, Guo R, Wang H, Li J, Chen Y, Wang W, Chawinga
415 M, Zhang L, Yang L, Zhao Y, Hua T. Abnormalities of maternal thyroid function during pregnancy

- 416 affect neuropsychological development of their children at 25-30 months. *Clinical endocrinology* 2010;
417 72:825-829
- 418 **10.** Su PY, Huang K, Hao JH, Xu YQ, Yan SQ, Li T, Xu YH, Tao FB. Maternal thyroid function in the first
419 twenty weeks of pregnancy and subsequent fetal and infant development: A prospective population-
420 based cohort study in China. *Journal of Clinical Endocrinology and Metabolism* 2011; 96:3234-3241
- 421 **11.** Klein RZ, Sargent JD, Larsen PR, Waisbren SE, Haddow JE, Mitchell ML. Relation of severity of
422 maternal hypothyroidism to cognitive development of offspring. *Journal of Medical Screening* 2001;
423 8:18-20
- 424 **12.** Smit BJ, Kok JH, Vulmsa T, Briet JM, Boer K, Wiersinga WM. Neurologic development of the newborn
425 and young child in relation to maternal thyroid function. *Acta Paediatrica, International Journal of*
426 *Paediatrics* 2000; 89:291-295
- 427 **13.** Henrichs J, Bongers-Schokking JJ, Schenk JJ, Ghassabian A, Schmidt HG, Visser TJ, Hooijkaas H, De
428 Muinck Keizer-Schrama SMPF, Hofman A, Jaddoe VVW, Visser W, Steegers EAP, Verhulst FC, De
429 Rijke YB, Tiemeier H. Maternal thyroid function during early pregnancy and cognitive functioning in
430 early childhood: The generation R study. *Journal of Clinical Endocrinology and Metabolism* 2010;
431 95:4227-4234
- 432 **14.** Ghassabian A, El Marroun H, Peeters RP, Jaddoe VW, Hofman A, Verhulst FC, Tiemeier H, White T.
433 Downstream effects of maternal hypothyroxinemia in early pregnancy: Nonverbal IQ and brain
434 morphology in school-age children. *Journal of Clinical Endocrinology and Metabolism* 2014; 99:2383-
435 2390
- 436 **15.** Suárez-Rodríguez M, Azcona-San Julián C, Alzina de Aguilar V. Hypothyroxinemia during pregnancy:
437 The effect on neurodevelopment in the child. *International Journal of Developmental Neuroscience* 2012;
438 30:435-438
- 439 **16.** Pop VJ, Brouwers EP, Vader HL, Vulmsa T, Van Baar AL, De Vijlder JJ. Maternal hypothyroxinaemia
440 during early pregnancy and subsequent child development: A 3-year follow-up study. *Clinical*
441 *endocrinology* 2003; 59:282-288
- 442 **17.** Berbel P, Mestre JL, Santamaría A, Palazón I, Franco A, Graells M, González-Torga A, de Escobar GM.
443 Delayed neurobehavioral development in children born to pregnant women with mild hypothyroxinemia
444 during the first month of gestation: the importance of early iodine supplementation. *Thyroid : official*
445 *journal of the American Thyroid Association* 2009; 19:511-519
- 446 **18.** Craig WY, Allan WC, Kloza EM, Pulkkinen AJ, Waisbren S, Spratt DI, Palomaki GE, Neveux LM,
447 Haddow JE. Mid-gestational maternal free thyroxine concentration and offspring neurocognitive
448 development at age two years. *Journal of Clinical Endocrinology & Metabolism* 2012; 97:E22-28
- 449 **19.** Chevrier J, Harley KG, Kogut K, Holland N, Johnson C, Eskenazi B. Maternal thyroid function during
450 the second half of pregnancy and child neurodevelopment at 6, 12, 24, and 60 months of age. *Journal of*
451 *Thyroid Research* 2011; 2011
- 452 **20.** Grau G, Aguayo A, Vela A, Aniel-Quiroga A, Espada M, Miranda G, Martinez-Indart L, Martul P,
453 Castaño L, Rica I. Normal intellectual development in children born from women with hypothyroxinemia
454 during their pregnancy. *Journal of Trace Elements in Medicine and Biology* 2015; 31:18-24
- 455 **21.** Oken E, Braverman LE, Platek D, Mitchell ML, Lee SL, Pearce EN. Neonatal thyroxine, maternal
456 thyroid function, and child cognition. *The Journal of clinical endocrinology and metabolism* 2009;
457 94:497-503
- 458 **22.** Willoughby KA, McAndrews MP, Rovet JF. Accuracy of episodic autobiographical memory in children
459 with early thyroid hormone deficiency using a staged event. *Developmental Cognitive Neuroscience*
460 2014; 9:1-11
- 461 **23.** Willoughby KA, McAndrews MP, Rovet J. Effects of early thyroid hormone deficiency on children's
462 autobiographical memory performance. *Journal of the International Neuropsychological Society* 2013;
463 19:419-429
- 464 **24.** Willoughby KA, McAndrews MP, Rovet JF. Effects of maternal hypothyroidism on offspring
465 hippocampus and memory. *Thyroid* 2014; 24:576-584
- 466 **25.** Pharoah POD, Connolly KJ. Relationship between maternal thyroxine levels during pregnancy and
467 memory function in childhood. *Early Human Development* 1991; 25:43-51
- 468 **26.** Ishaik G, Asztalos E, Perlman K, Newton S, Frisk V, Rovet J. Hypothyroxinemia of prematurity and
469 infant neurodevelopment: A pilot study. *Journal of Developmental and Behavioral Pediatrics* 2000;
470 21:172-179
- 471 **27.** Henrichs J, Ghassabian A, Peeters RP, Tiemeier H. Maternal hypothyroxinemia and effects on cognitive
472 functioning in childhood: How and why? *Clinical endocrinology* 2013; 79:152-162
- 473 **28.** Lazarus JH, Bestwick JP, Channon S, Paradise R, Maina A, Rees R, Chiusano E, John R, Guaraldo V,
474 George LM, Perona M, Dall'Amico D, Parkes AB, Joomun M, Wald NJ. Antenatal thyroid screening
475 and childhood cognitive function. *The New England journal of medicine* 2012; 366:493-501

- 476 29. Casey BM, Thom EA, Peaceman AM, Varner MW, Sorokin Y, Hirtz DG, Reddy UM, Wapner RJ, Thorp
477 Jr JM, Saade G. Treatment of Subclinical Hypothyroidism or Hypothyroxinemia in Pregnancy. *New*
478 *England Journal of Medicine* 2017; 376:815-825
- 479 30. N M. IQ and human intelligence. Oxford University Press.
- 480 31. Korevaar T, Muetzel R, Medici M, Chaker L, Jaddoe VWV, de Rijke YB, Steegers EAP, Visser TJ,
481 White T, Tiemeier H. Association of maternal thyroid function during early pregnancy with offspring IQ
482 and brain morphology in childhood: a population-based prospective cohort study. *The Lancet Diabetes*
483 *& Endocrinology* 2016; 4:35-43
- 484 32. Hales C, Channon S, Taylor PN, Draman MS, Muller I, Lazarus J, Paradise R, Rees A, Shillabeer D,
485 Gregory JW. The second wave of the Controlled Antenatal Thyroid Screening (CATS II) study: the
486 cognitive assessment protocol. *BMC endocrine disorders* 2014; 14:95
- 487 33. Daniels GH, Dayan CM. *Thyroid Disorders*. Oxford, UK: Health Press.
- 488 34. Van den Boogaard E, Vissenberg R, Land JA, van Wely M, van der Post JAM, Goddijn M, Bisschop
489 PH. Significance of (sub) clinical thyroid dysfunction and thyroid autoimmunity before conception and
490 in early pregnancy: a systematic review. *Human reproduction update* 2011; 17:605-619
- 491 35. Jouyandeh Z, Hasani-Ranjbar S, Qorbani M, Larijani B. Universal screening versus selective case-based
492 screening for thyroid disorders in pregnancy. *Endocrine* 2014; 48:116-123
- 493 36. Reid Sally M, Middleton P, Cossich Mary C, Crowther Caroline A, Bain E. Interventions for clinical and
494 subclinical hypothyroidism pre-pregnancy and during pregnancy. *Cochrane Database of Systematic*
495 *Reviews* 2013(5). <http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD007752.pub3/abstract>.
- 496 37. Castagna MG DM, Cantara S, Ambrosio R, Maino F, Porcelli T, Marzocchi C, Garbi C, Pacini F,
497 Salvatore D. DIO2 Thr92Ala Reduces Deiodinase-2 Activity and Serum-T3 Levels in Thyroid-Deficient
498 Patients. *Journal of Clinical Endocrinology and Metabolism* 2017; 102:1623-1630

499

500

501 **Legends to Figures and Tables**

502 **Table 1: Characteristics of the cohort.** Data are expressed as median (Interquartile
503 range, IQR), mean (standard deviation, SD) or the Number (N) of participants
504 (percentage, %)- Socioeconomic status is based on a social deprivation score with 1
505 being the most deprived. Child's language describes whether the child speaks English
506 (Engl) both at home and in school, Welsh in both locations, a combination of English
507 and Welsh or an additional language. GTF=gestational thyroid function;
508 SGTF=suboptimal gestational thyroid function; CATS=controlled antenatal thyroid
509 screening.

510 **Table 2: Logistic regressions for odds of IQ below 85.** Data are expressed as
511 Odds Ratios (OR) with 95% confidence intervals (95% CI). SGTF=suboptimal
512 gestational thyroid function. FSIQ=full scale intelligence quotient. VCIQ=verbal

513 comprehension intelligence quotient. PRIQ=perceptual reasoning intelligence
514 quotient. WMIQ=working memory intelligence quotient. PSIQ=processing speed
515 intelligence quotient. Model 1=unadjusted. Model 2=adjusted for child gender. Model
516 3=adjusted for model 2 and whether the mother breastfed over one month, and mother
517 age at time of study consent during pregnancy. Model 4=adjusted for model 3 and
518 where the child was assessed, child's language spoken at school and home, and
519 social deprivation [score](#).

520 **Table 3: Mean scores for IQs.** Data expressed as means (standard deviations) of
521 group, or the Number (N) of participants (percentage, %) having IQ<85.

522 GTF=gestational thyroid function; SGTF=suboptimal gestational thyroid function;
523 WISC=Wechsler intelligence scale for children fourth edition UK; VCIQ=verbal
524 comprehension intelligence quotient, PRIQ=perceptual reasoning intelligence
525 quotient; WMIS=working memory intelligence quotient; PSIQ=processing speed
526 intelligence quotient; FSIQ=full scale intelligence quotient; NEPSY=developmental
527 neuropsychological assessment second edition; MD=memory for designs;
528 MDD=memory for designs delayed; FTDH=fingertip tapping dominant hand;
529 FTNDH=fingertip tapping non-dominant hand; NM=narrative memory; LM=list
530 memory. *reduced dataset

531 **Figure 1: Flow-chart of recruitment to the Controlled Antenatal Thyroid**
532 **Screening (CATS) Study** Illustrates initial recruitment for CATS-I, when child IQ was
533 assessed at 3 years of age and the follow-up study CATS-II, in which child IQ was
534 assessed at 9 years of age.

535

536 TABLE 1:

Characteristics	Groups						
	Normal-GTF	Treated SGTF	Untreated	Normal-GTF	Normal-GTF vs.	Treated	SGTF
	N=232	N=119	N=98	vs. Treated SGTF (P)	vs. Treated Untreated SGTF (P)	vs. Treated SGTF (P)	vs. Untreated SGTF (P)
Thyrotropin at CATS-I consent (mIU/L)	1.16 (0.66-1.83)	4.09 (1.79-5.09)	3.57 (1.18-4.49)	0.001	0.001	0.007	
Thyroxine at CATS-I consent (pmol/L)	14.12 (1.76)	11.92 (1.93)	11.79 (1.88)	0.001	0.001	1.000	
Maternal age at CATS-I consent (years)	31.85 (5.16)	30.26 (5.08)	31.05 (4.88)	0.018	0.578	0.767	
Social deprivation/Socio-economic status	4.00 (3-5) (mean 3.71)	4.00 (3-5) (mean 3.78)	4.00 (2-5) (mean 3.37)	0.807	0.359	0.161	
1	26 (11%)	15 (13%)	16 (16%)				
2	27 (11%)	12 (10%)	14 (14%)				

3	36 (15%)	15 (13%)	17 (17%)			
4	43 (18%)	19 (16%)	20 (20%)			
5	100 (43%)	58 (49%)	31 (32%)			
Child breastfed over one month N (%)	150 (65%)	72 (60%)	56 (57%)	0.445	0.198	0.616
Child characteristics						
Male children N (%)	177 (50%)	65 (55%)	49 (50%)	0.457	0.943	0.497
Child age at participation	9.83 (9.00-10.33)	9.58 (9.08-9.94)	9.50 (9.00- 9.94)	0.001	0.024	0.710
Where child was assessed				0.001	0.001	0.554
Home	120 (52%)	92 (77%)	79 (81%)			
Research centre	112 (48%)	27 (23%)	19 (19%)			
Child's language				0.950	0.364	0.541
English school/home	180 (78%)	95 (80%)	85 (87%)			
Welsh school/Engl home	42 (18%)	20 (17%)	11 (11%)			

Welsh school/home	7 (3%)	3 (2%)	1 (1%)
Engl school/other home	2 (1%)	1 (1%)	1 (1%)
Welsh school/other home	1 (1%)	0	0

537

538 Table 2:

IQs	MODELS	Merged SGTF to Normal-GTF OR (95% CI)	P Interaction	OR Untreated (95% CI)	OR treatment (95% CI)	P treatment interaction
FSIQ	1	1.58 (0.78, 3.21)	0.206	1.97 (0.86, 4.50)	0.65 (0.26, 1.63)	0.355
	2	1.57 (0.77, 3.19)	0.217	1.98 (0.86, 4.55)	0.63 (0.25, 1.59)	0.325
	3	1.38 (0.66, 2.86)	0.389	1.77 (0.75, 4.16)	0.61 (0.23, 1.58)	0.308
	4	1.15 (0.52, 2.51)	0.731	1.33 (0.53, 3.34)	0.75 (0.27, 2.06)	0.576
VCIQ	1	1.08 (0.57, 2.03)	0.820	0.89 (0.38, 2.09)	1.38 (0.55, 3.48)	0.491
	2	1.07 (0.57, 2.02)	0.833	0.89 (0.38, 2.09)	1.36 (0.54, 3.44)	0.506
	3	0.99 (0.52, 1.88)	0.968	0.82 (0.34, 1.93)	1.38 (0.54, 3.53)	0.491
	4	0.93 (0.47, 1.83)	0.834	0.70 (0.29, 1.73)	1.62 (0.62, 4.20)	0.317
PRIQ	1	1.82 (0.84, 3.94)	0.130	2.54 (1.06, 6.07)	0.49 (0.18, 1.33)	0.156
	2	1.82 (0.84, 3.94)	0.131	2.54 (1.06, 6.07)	0.49 (0.18, 1.33)	0.156
	3	1.60 (0.73, 3.53)	0.238	2.31 (0.95, 5.62)	0.46 (0.17, 1.28)	0.132
	4	1.35 (0.59, 3.09)	0.482	1.78 (0.69, 4.56)	0.56 (0.19, 1.58)	0.268

WMIQ	1	1.48 (0.78, 2.81)	0.232	1.35 (0.60, 3.04)	1.17 (0.50, 2.77)	0.715
	2	1.47 (0.77, 2.79)	0.241	1.35 (0.60, 3.05)	1.15 (0.49, 2.73)	0.742
	3	1.33 (0.69, 2.57)	0.393	1.21 (0.53, 2.78)	1.18 (0.49, 2.84)	0.713
	4	1.26 (0.63, 2.53)	0.513	1.04 (0.43, 2.50)	1.42 (0.57, 3.52)	0.449
PSIQ	1	0.79 (0.36, 1.71)	0.550	0.88 (0.33, 2.32)	0.81 (0.25, 2.61)	0.729
	2	0.78 (0.36, 1.69)	0.524	0.88 (0.33, 2.33)	0.79 (0.24, 2.53)	0.688
	3	0.75 (0.34, 1.63)	0.463	0.85 (0.32, 2.27)	0.77 (0.24, 2.49)	0.664
	4	0.75 (0.33, 1.68)	0.482	0.82 (0.20, 2.24)	0.85 (0.26, 2.77)	0.783

539

540

541 Table 3:

Groups				
Cognitive assessment	Normal-GTF N=232	Merged SGTF N=217	Treated SGTF N=119	Untreated SGTF N=98
WISC				
VCIQ	99.81 (11.26)	98.60 (11.42)	97.56 (9.95)	99.86 (12.93)
<85	28 (12%)	30 (14%)	19 (16%)	11 (11%)
PRIQ	105.37 (12.30)	104.55 (12.87)	104.49 (12.26)	104.63
	11 (5%)	18 (8%)	7 (6%)	(13.64)
<85				11 (11%)
WMIQ	99.91 (11.24)	99.81 (12.72)	99.73 (13.28)	99.90 (12.07)
<85	18 (8%)	24 (11%)	14 (12%)	10 (10%)
PSIQ	103.66 (12.75)	102.39 (12.73)	103.16 (12.71)	101.45 (12.75)
	22 (9%)	18 (8%)	8 (7%)	10 (10%)
<85				
FSIQ	103.10 (11.68)	102.01 (12.59)	101.76 (12.04)	102.31
	15 (6%)	21 (10%)	10 (8%)	(13.28)
<85				11 (11%)
NEPSY				
MD	10.36 (2.92)	9.69 (3.13)	9.63 (3.27)	9.76 (2.96)
MDD	10.34 (2.65)	9.86 (2.84)	9.77 (2.79)	9.98 (2.92)
FTDH	12.24 (1.60)	11.94 (1.45)	11.90 (1.41)	12.01 (1.52)
FTNDH	12.51 (1.37)	12.24 (1.41)	12.21 (1.39)	12.31 (1.46)
NM	11.56 (2.76)	11.06 (2.76)	11.02 (2.78)	11.12 (2.74)

	N=170	N=146	N=78	N=68
LM*	10.93 (2.84)	10.62 (2.86))	10.63 (3.13)	10.60 (2.54)

542

Figure 1

