Thyroid function in preterm infants and neurodevelopment at 2 years

Fiona L R Williams,1 Alice Lindgren,2 Jennifer Watson,1 Anita Boelen,3 Timothy Cheetham4

ABSTRACT

Objectives Postnatal thyroid dysfunction is common in preterm infants but the relationship between mild dysfunction and neurodevelopment is unclear. Our aim is to describe the relationship between thyroid function and neurodevelopment.

Design Cohort analysis.

Patients 1275 infants born under 31 weeks’ gestation; there were no exclusion criteria.

Setting The infants were part of a UK daily iodine supplementation trial.

Main outcomes Thyroid-stimulating hormone, thyroid-binding globulin and total thyroxine levels were measured in dried blood spots on postnatal days 7, 14, 28 and the equivalent of 34 weeks’ gestation. Neurodevelopment was measured using the Bayley-III Scales of infant development at 2 years of age.

Results No infant was identified as hypothyroid through routine screening. The 3% of infants consistently in the top decile of gestationally age-adjusted thyroid-stimulating hormone levels had a reduction in cognitive score of 7 Bayley units when compared with those not in the top decile (95% CI −13 to −1). A reduction in motor composite score of 6 units (95% CI −12 to −0) and fine motor score of 1 unit (95% CI −2 to –0.1) was also identified. The 0.7% of infants consistently in the bottom decile of age-adjusted thyroxine levels had a reduction in motor composite score of 14 units (95% CI −25 to −2) and its two subset scores, fine and gross motor, of 2 units (95% CI respectively −4.5 to < −0.1 and −4.3 to −0.3).

Conclusions Preterm infants with consistent ‘mild’ thyroid dysfunction score less on neurodevelopmental tests at 2 years of age. Many of these infants will not be detected by current clinical protocols or screening programmes.

INTRODUCTION

The natural history of transiently raised levels of thyroid-stimulating hormone (TSH), with or without associated low thyroxine (T4) levels, in preterm infants is complex and affected by factors such as maternal iodine status,1 developmental immaturity of endocrine systems,2–4 critical illness5–7 and drug exposure.8 The incidence of mildly raised TSH levels depends on the definition, denominator and age-group studied, but in population studies it has been reported as 3.3%8 and 6%,9 and in selected clinical studies, as high as 22.3%.10 There is no consensus on whether thyroid dysfunction should be treated11 12 and its long-term consequences are unclear, as studies are typically methodologically restricted. It has been suggested that infants with transiently raised TSH levels may have a high risk of subclinical hypothyroidism in childhood8 13 and possibly developmental delay,9 although this is not a consistent observation.5 14

The most common neonatal thyroid dysfunction in preterm infants is hypothyroxinaemia (low T4 with normal TSH levels), with an incidence of between 7% and 100%15 in infants <28 weeks’ gestation; the wide range of reported incidence reflects varying definitions and methodologies.

A recently completed trial of daily iodide supplementation in infants born <31 weeks’ gestation with neurodevelopmental outcome measured at 2 years of age16 provided us the opportunity to describe T4, TSH and thyroid-binding globulin (TBG) distributions in a preterm neonatal cohort. We also examined whether abnormal postnatal thyroid function, characterised by a subtle increase in TSH and/or a reduction in T4, or by a pattern of hypothyroxinaemia, was related to neurodevelopmental outcomes.
METHODS
Infants <31 weeks’ gestation were recruited to a UK trial of iodine supplementation (30 μg/kg/day from birth to equivalent of 34 weeks’ gestation) with neurodevelopmental assessment at 2 years. There were no significant differences in neurodevelopmental outcome, or levels of TSH, TBG and T4 between the arms of the trial and so the data were combined. All other regulatory bodies, and all families provided informed consent. The trial protocol is available at https://www.npeu.ox.ac.uk/downloads/files/i2s2/protocol/I2S2-Protocol.pdf

Data collected
Infant blood was collected on dried blood spot cards at postnatal days 7, 14, 28 and 34 weeks’ gestational age (had the fetus remained in utero, referred to hereafter as equivalent gestational age), ±1 day. Cards were sent to the Amsterdam Neonatal Screening Laboratory for measurement of TSH, T4 and TBG (for details see online supplementary appendix). T4 is expressed as absolute levels and as a SD of the daily mean, which is calculated from the results of approximately 150 cards of term and preterm infants; levels ≤−3.0 SD are considered abnormal (in term infants). TBG concentrations <40 nmol/L is indicative of TBG deficiency.

TSH, TBG and T4 whole-blood levels reported in this paper were taken specifically for the trial. TSH levels were considered raised if ≥6 mU/L because many UK laboratories at the time of the trial used 6mU/L as their cut-off for recall (0.2% of all UK infants screened have a blood spot TSH >6 mU/L). (A whole-blood TSH level of 6mU/L is approximately 15 mU/L when measured in a serum sample.) We recorded the outcome of thyroid function tests, whether or not levothyroxine therapy was started, and type of thyroid dysfunction.

Definitions
Two approaches were used when evaluating the hypothalamo-pituitary-thyroid axis and subsequent neurodevelopmental outcome. First, we used deciles, corrected to gestational age subgroup (ie, ≤25, 26–27, 28–30 weeks), to investigate the relationships between TSH, TBG and T4 (top TSH decile and bottom deciles for T4, TBG and T4/TBG ratio vs the remaining deciles) with neurodevelopmental outcome. Secondly, we categorised infants into mutually exclusive groups as: new born screening diagnosis of hypothyroidism, or to study definition of hypothyroid (TSH ≥6 and T4 ≤−3 SD), hyperthyrotopinaemic (TSH ≥6 mU/L and T4 >−3 SD), hypothyroxinaemic (TSH <6 mU/L and T4 ≤−3 SD) or euthyroid (TSH <6 mU/L and T4 >−3 SD). This strategy addresses the function of the hypothalamo-pituitary axis in its entirety rather than TSH or thyroid hormone secretion in isolation.

Detailed clinical data were collected from birth until discharge home or until the equivalent of 36 weeks’ gestational age. In a previous study investigating thyroid hormones and neurodevelopment, we identified 26 factors encompassing maternal, intrapartum, neonatal, lifestyle and the assessor, which contributed to the variation in neurodevelopment. The factors with the strongest statistical associations and for which we had data in this study were gender, gestation at birth, the assessor performing the neurodevelopmental test, the hospital of birth and the level of nursing care reported on the day of blood sampling. The level of nursing care was categorised as high (level 1), medium (level 2) or low (level 3) and was used as a proxy for illness severity. To this list of factors potentially associated with neurodevelopment, we included whether or not the infant received iodine supplementation (ie, trial intervention or placebo) in case it had some unexpected subtle impact, whether or not the infant received levothyroxine therapy and T4/TBG ratio as clinically these were also relevant confounders of neurodevelopment in this group of infants. The T4/TBG ratio is used as a proxy for FT4 concentrations.

Outcome measures
Infants were assessed using the Bayley-III Scales at 2 years of age corrected for prematurity (±1 month). The Bayley-III Scales provide cognitive, motor composite and language composite scores, each with a population mean=100 and SD=15. The latter two scores have two subtests with mean=10 and SD=3. Infants were assessed using trial personnel specifically trained to use Bayley-III; random performances were video-recorded and audited.

Statistics
Infants were included in the analyses for this paper if they had at least one study blood recorded. Data were imputed for the primary outcome of the trial (ie, Bayley-III score) following a 5-point protocol.

General linear models, specifically univariate analysis of variance were constructed to evaluate the adjusted association of thyroid function, reported on days 7, 14, 28 and at 34 weeks’ equivalent gestational age, on Bayley-III scores. The indices of thyroid function were classified as described earlier in the ‘Methods’ section and the outcomes were adjusted for the seven cofactors described previously. Significance was specified at p<0.05.

We predicted that some infants would switch between thyroid categories over the sampling period, while others would not. We therefore constructed general linear models for the main and subset domains of the Bayley-III, which classified thyroid status over the sampling period as consistently within the bottom (T4, T4/TBG ratio) or top (TSH) deciles, or consistently as: euthyroid, hypothyroid, hyperthyrotopinaemic, hypothyroxinaemic or as mixed thyroid function where the infant moves between categories. All bloods recorded for an infant (at day 7, 14, 28 and equivalent of 34 weeks’ gestation) had to meet the definition for either hypothyroid, hyperthyrotopinaemia, hypothyroxinaemia or euthyroid to be classified as ‘consistent’. To be included in these analyses infants had to have two or more study bloods recorded. The Bayley-III outcomes were adjusted for the seven variables described in the ‘Methods’ section and, if appropriate, the T4/TBG ratio. All analyses were undertaken on SPSS V.22.

RESULTS
One thousand two hundred seventy-five infants were recruited to the trial (online supplementary figure 1). One or more bloods were recorded for 1222/1259 (97%) infants, 94% (1185/1259) had ≥2 bloods recorded and 89% (1126/1259) had ≥3 bloods recorded. Bayley-III assessments were available for 997 infants. No infant was identified as hypothyroid as part of the routine neonatal blood spot screening programme, either at 5 days post-delivery or as part of repeat screening conducted at 36 weeks/ time of discharge. Thus, all references to hypothyroid in this paper from this point refer to the study definition of hypothyroid (ie, TSH=≥6 and T4 =<−3 SD).

Sixty-two per cent of infants had one or more TSH or T4 values outwith the reference range during the study. T4 and TBG levels and T4/TBG ratios increase as gestation rises and also over time (table 1). TSH levels also increase as gestation rises, but peak and flatten at day 14 for infants born ≤25 weeks, and peak

ORIGINAL RESEARCH
### Table 1  Distributions of T4, TSH, TBG and T4/TBG ratio by day of sampling and gestational age group; mean (SD) number

| Gestational weeks | Day 7 blood | Day 14 blood | Day 28 blood | Week 34 blood |
|-------------------|-------------|--------------|--------------|---------------|
| T4 nmol/L         |             |              |              |               |
| ≤25 weeks         | 18.0 (8.6)  | 22.4 (11.6)  | 31.9 (13.3)  | 45.1 (14.0)   |
| 26–27 weeks       | 26.4 (10.2) | 32.8 (12.7)  | 39.9 (14.6)  | 44.0 (14.8)   |
| 28–30 weeks       | 36.5 (13.6) | 43.2 (14.2)  | 47.3 (15.8)  | 49.3 (14.7)   |
| TSH mU/L          |             |              |              |               |
| ≤25 weeks         | 1.3 (0.8)   | 1.6 (1.2)    | 1.6 (1.3)    | 1.6 (1.1)     |
| 26–27 weeks       | 1.4 (1.4)   | 1.9 (2.5)    | 1.7 (1.4)    | 1.5 (1.1)     |
| 28–30 weeks       | 1.8 (1.4)   | 1.9 (1.8)    | 1.6 (1.3)    | 1.6 (1.2)     |
| TBG nmol/L        |             |              |              |               |
| ≤25 weeks         | 115.6 (40.1)| 144.0 (46.6)| 169.7 (54.0)| 198.0 (58.4)  |
| 26–27 weeks       | 131.9 (47.6)| 155.1 (51.2)| 170.1 (51.2)| 179.3 (59.9)  |
| 28–30 weeks       | 146.5 (45.7)| 162.9 (48.9)| 174.0 (53.2)| 176.2 (52.5)  |
| T4/TBG ratio      |             |              |              |               |
| ≤25 weeks         | 14.5 (5.5)  | 13.2 (4.6)   | 14.7 (8.2)   | 16.2 (4.5)    |
| 26–27 weeks       | 16.1 (5.3)  | 16.0 (5.3)   | 17.0 (4.9)   | 17.7 (5.7)    |
| 28–30 weeks       | 18.3 (5.1)  | 18.9 (6.6)   | 18.9 (5.0)   | 19.4 (5.1)    |

T4, total thyroxine; TBG, thyroid-binding globulin; TSH, thyroid-stimulating hormone.

### Table 2  Adjusted* linear regression models of TSH levels which were consistent over at least two study bloods and the Bayley-III main and subtest domains

| Bayley-III main domains | N   | Effect estimate | T statistic | P value  | 95% CI Lower | 95% CI Higher |
|------------------------|-----|-----------------|-------------|----------|--------------|---------------|
| Cognitive score        |     |                 |             |          |              |               |
| Consistently top decile TSH | 27  | −6.9            | −2.266      | 0.024    | −12.9        | −0.9          |
| Mixed dysfunction      | 474 | −1.5            | −1.521      | 0.129    | −3.5         | 0.4           |
| Never in top decile    | 489 | Reference       |             |          |              |               |
| Motor composite score  |     |                 |             |          |              |               |
| Consistently top decile TSH | 27  | −5.8            | −1.967      | 0.050    | −11.7        | −0.1†         |
| Mixed dysfunction      | 474 | −0.6            | −0.583      | 0.560    | −2.5         | 1.3           |
| Never in top decile    | 488 | Reference       |             |          |              |               |
| Language composite score|      |                 |             |          |              |               |
| Consistently top decile TSH | 27  | −5.3            | −1.528      | 0.127    | −12.0        | 1.5           |
| Mixed dysfunction      | 473 | −2.1            | −1.912      | 0.056    | −4.3         | 0.1           |
| Never in top decile    | 488 | Reference       |             |          |              |               |
| Receptive language subtest |      |                 |             |          |              |               |
| Consistently top decile TSH | 27  | −0.7            | −1.179      | 0.239    | −1.9         | 0.5           |
| Mixed dysfunction      | 473 | −0.2            | −1.143      | 0.253    | −0.6         | 0.2           |
| Never in top decile    | 488 | Reference       |             |          |              |               |
| Expressive language subtest |      |                 |             |          |              |               |
| Consistently top decile TSH | 27  | −1.1            | −1.719      | 0.086    | −2.3         | 0.2           |
| Mixed dysfunction      | 473 | −0.5            | −2.066      | 0.009    | −0.9         | −0.1          |
| Never in top decile    | 487 | Reference       |             |          |              |               |
| Fine motor subtest     |     |                 |             |          |              |               |
| Consistently top decile TSH | 27  | −1.3            | −2.165      | 0.031    | −2.4         | −0.1          |
| Mixed dysfunction      | 474 | −0.3            | −1.436      | 0.151    | −0.6         | 0.1           |
| Never in top decile    | 487 | Reference       |             |          |              |               |
| Gross motor subtest    |     |                 |             |          |              |               |
| Consistently top decile TSH | 27  | −0.7            | −1.215      | 0.224    | −1.7         | 0.4           |
| Mixed dysfunction      | 473 | 0.1             | 0.554       | 0.580    | −0.2         | 0.4           |
| Never in top decile    | 489 | Reference       |             |          |              |               |

*adjusted for gender, gestational age, level of nursing care required on day 14, whether or not in receipt of iodine supplement, whether or not treated with L-thyroxine during the neonatal period or at 2 years of age, assessor and hospital of birth.

†Actual figure −0.012.

TSH, thyroid-stimulating hormone.
Table 3 Adjusted* linear regression models of T4 levels which were consistent over at least two study bloods and the Bayley-III main and subtest domains

| Bayley-III main domains | N   | Effect estimate | T statistic | P value | 95% CI Lower | 95% CI Higher |
|-------------------------|-----|-----------------|-------------|---------|--------------|---------------|
| Cognitive score         |     |                 |             |         |              |               |
| Consistently bottom decile T4 | 7   | −2.2            | −1.988      | 0.046   | −4.5         | <−0.1†        |
| Mixed dysfunction       | 239 |                 | −2.229      | 0.026   | −4.3         | −0.3          |
| Never in bottom decile  | 742 |                 | −2.229      | 0.026   | −4.3         | −0.3          |
| Motor composite score   |     |                 |             |         |              |               |
| Consistently bottom decile T4 | 7   | −1.8            | −1.495      | 0.135   | −4.1         | 0.6           |
| Mixed dysfunction       | 239 |                 | −0.8        | −0.675  | −3.0         | 1.5           |
| Never in bottom decile  | 744 |                 | −0.8        | −0.675  | −3.0         | 1.5           |
| Language composite score|     |                 |             |         |              |               |
| Consistently bottom decile T4 | 7   | −1.34           | −2.336      | 0.020   | −24.6        | −2.1          |
| Mixed dysfunction       | 239 |                 | −0.8        | −0.675  | −3.0         | 1.5           |
| Never in bottom decile  | 744 |                 | −0.8        | −0.675  | −3.0         | 1.5           |
| Receptive language subtest |    |                 |             |         |              |               |
| Consistently bottom decile T4 | 7   | −0.8            | −0.814      | 0.416   | −3.7         | 1.5           |
| Mixed dysfunction       | 239 |                 | −0.8        | −0.814  | −3.7         | 1.5           |
| Never in bottom decile  | 744 |                 | −0.8        | −0.814  | −3.7         | 1.5           |
| Expressive language subtest |    |                 |             |         |              |               |
| Consistently bottom decile T4 | 7   | −0.4            | −1.512      | 0.131   | −0.8         | 0.1           |
| Mixed dysfunction       | 239 |                 | −0.4        | −1.512  | −0.8         | 0.1           |
| Never in bottom decile  | 744 |                 | −0.4        | −1.512  | −0.8         | 0.1           |
| Fine motor subtest      |     |                 |             |         |              |               |
| Consistently bottom decile T4 | 7   | −1.8            | −1.445      | 0.149   | −4.1         | 0.6           |
| Mixed dysfunction       | 239 |                 | −0.8        | −0.814  | −3.7         | 1.5           |
| Gross motor subtest     |     |                 |             |         |              |               |
| Consistently bottom decile T4 | 7   | −1.8            | −1.445      | 0.149   | −4.1         | 0.6           |
| Mixed dysfunction       | 239 |                 | −0.8        | −0.814  | −3.7         | 1.5           |
| Never in bottom decile  | 744 |                 | −0.8        | −0.814  | −3.7         | 1.5           |

*Adjusted for gender, gestational age, level of nursing care required on day 14, whether or not in receipt of iodine supplement, whether or not treated with L-thyroxine during the neonatal period or at 2 years of age, assessor and hospital of birth.
†Actual figure −0.04.
T4, total thyroxine.

and then fall for infants born ≥26 weeks (table 1). None of the infants had TBG concentrations <40 nmol/L.

Classification of thyroid function that was consistent over ≥2 bloods showed that 3% (27/969) were consistently in the top decile of TSH levels (table 2); 0.7% were consistently in the bottom decile for T4 levels (7 infants) (table 3) and 0.2% were consistently in the bottom decile for T4/TBG ratio (2 infants).

Classification of thyroid function that was consistent over ≥2 study bloods by thyroid category showed that the majority (670/1185) of infants switch categories (which represents 4–11 weeks in this cohort of infants). No infant was consistently hypothyroid or hyperthyrotropinaemic, 5% (65/1185) were consistently hypothyroxinaemic and 38% (450/1185) were consistently euthyroid (online supplementary figure 1); 0.3% (4/1185) were consistently in the bottom decile for T4/TBG ratio.

Table 4 Number of infants with a raised TSH (≥6 mU/L) levels by day of blood sampling and subsequent raised TSH levels*

| Blood sampling day | One isolated raised TSH level | Two raised TSH levels | Three raised TSH levels | Four raised TSH levels | Total number of infants with a first raised TSH level per day: incidence data |
|--------------------|------------------------------|-----------------------|------------------------|------------------------|---------------------------------------------------------------|
| 7                  | 13                           | 5                     | 2                      | 0                      | 20                                                            |
| 14                 | 19                           | 8                     | 1                      | n/a†                   | 28                                                            |
| 28                 | 11                           | 4                     | n/a†                   | n/a†                   | 15                                                            |
| 34 weeks' gestation| 6                            | n/a†                  | n/a†                   | n/a†                   | 6                                                             |

*On day 7, there were 20 infants with raised TSH levels of whom 13 had an isolated raised level on day 7, 5 infants had two raised levels (on day 7 and either day 14/28 or 34 weeks), and 2 infants had three raised levels (on day 7, and two others from days 14/28 or 34 weeks).
†Not applicable as it is not possible to obtain this many tests.
TSH, thyroid-stimulating hormone.

Bayley-III outcomes

The cognitive score for infants consistently in the top decile for TSH levels was 7 Bayley units lower (95% CI −13 to −1) compared with those who were never in the top decile. There was also a reduction in the motor composite score of 6 units (95% CI −12 to −0.012) and the fine motor subset score of 1 unit (95% CI −2 to −0.1) (table 2). For infants consistently in the bottom decile of T4, the motor composite score was reduced by 13 units (95% CI −25 to −2) compared with those who were never in the bottom decile. This deficit was contributed by both the fine (2 units 95% CI −4.5 to −0.04) and gross (2 units 95% CI −4.3 to −0.3) motor subset scores (table 3).

Infants in the top decile for TSH on day 14 (compared with those not in that decile) had lower cognitive scores, lower expressive language scores on day 7 and day 14 and lower fine motor scores on day 14 (online supplementary tables 2-8). Infants in the bottom decile for T4 (compared with those not in that decile) had significantly lower cognitive, motor composite and gross motor scores on day 14 (online supplementary tables 2-8).

Using adjusted regression models, no infants were consistently hypothyroid or hyperthyrotropinaemic. Infants classified as consistently hypothyroxinaemic, compared with those who were consistently euthyroid, scored 8 units lower on the cognitive score (95% CI −14 to −3), 9 units lower on the motor composite score (95% CI −15 to −4), 11 units lower on the language composite score (95% CI −17 to −4) and lower on all subtest domains (table 3).

Infants classified as hypothyroid on day 7 scored significantly lower on all Bayley-III main, and three of the four subtest domains compared with infants classified as euthyroid. There were two strong significant associations: the classification of hypothyroid on day 7 with the Bayley-III cognitive score (a reduction of 21 units compared with the euthyroid group), and...
the fine motor subtest score (a reduction of 3 units compared with the euthyroid group) (online supplementary tables 9-11). Of the seven infants categorised as hypothyroid on day 7, five had TSH levels <11.0 mU/L, one infant’s TSH was 11.0 mU/L and one infant’s TSH was 22.0 mU/L.

For more clinical characteristics of the infants, see online supplementary appendix.

**DISCUSSION**

This analysis has highlighted two important observations. First, when compared with a euthyroid state, consistent hypothyroxinaemia (ie, measured on at least two occasions), but not isolated episodes of low T4, is associated with significantly lower (8–11 points) Bayley-III scores in all the domains: cognitive, motor and language. Secondly, isolated mildly raised TSH levels (>6 mU/L) are associated with significantly lower Bayley-III scores when compared with euthyroid.

Several studies have reported adverse associations between neurodevelopment and low postnatal T4 levels but it has also been queried whether hypothyroxinaemia is simply an epiphenomenon of prematurity and illness. Observational studies cannot provide the definitive answers of a randomised controlled trial, but designing a trial to answer this question is very challenging. Our data suggest that trials in this field should target only those infants who are consistently hypothyroxinaemic. Pragmatically, such infants are difficult to identify as they represent a small proportion of preterm infants (5% in our cohort) and agreeing a definition for hypothyroxinaemia/low T4 is not straightforward. There is an appreciable gestational-age gradient in levels of T4 and T4 levels also vary according to the postnatal age at sampling and the media used (whole blood, plasma, serum).

Furthermore, to preserve neurodevelopment hypothyroxinaemia needs to be treated quickly as evidence from managing congenital hypothyroidism shows that replacement therapy must start within 2 weeks in order to optimise neurodevelopment. Use of blood spot cards may provide a rapid way to identify hypothyroxinaemia and to facilitate trials and management. (It is important to highlight that hypothyroxinaemia is separate from the congenital hypothyroidism due to dysgenesis or dyshormonogenesis that screening programmes were designed to detect.)

We optimised the strength of our observational design by adjusting for many of the key contributors to neurodevelopment in a preterm population. It is relevant therefore that the association between low T4 levels and Bayley-III scores remained significant following adjustment for gender, gestation at birth and level of sickness. The inclusion of gestation in the regression model was important as it counteracts the impact of using T4 levels derived from term and preterm infants in the definition of hypothyroxinaemia. The data were also adjusted for study relevant factors such as treatment with levothyroxine and daily iodine supplementation, as well as methodological factors such as assessor variability. In this large preterm cohort, 5.7% of infants had mildly raised (≥6 mU/L) TSH levels. This proportion seems low when compared, for example, to the 6% found by Cuestas et al in an exclusively term cohort, but is likely due to differences in the day of testing, and the use of whole-blood rather than serum. Mildly raised TSH levels seem to be important as we found associations with neurodevelopment whether the TSH levels were defined by categories of thyroid function or by deciles. An association between borderline blood spot TSH concentrations and poorer neurodevelopmental outcome has similarly been reported in a large population of primarily term infants.

Evidence of poorer cognitive development and motor development, and in particular fine motor development, with mildly raised levels of postnatal TSH now need to be confirmed in other cohorts with intervention trials a logical next step.

This study has limitations. The use of 6mU/L as a cut-off to categorise hypothyroidism is lower than many national screening programmes which more typically use 10mU/L or higher. We chose 6 mU/L because several UK laboratories at the time of the trial used 6 mU/L as their cut-off for recall. This cut-off corresponded to the 97.3%ile of TSH levels at the Amsterdam Screening Laboratory during the time of the trial and lies around the 99.8%ile on a UK TSH screening programme. Fluctuations in TSH levels in response to iodine (the trial intervention) are well recognised and may be more pronounced in our cohort if the mothers were mildly iodide deficient, which is likely.

The definitions of thyroid function used in this paper are those generally applied to term infants, who do not have the biochemical patterns characteristic of preterm infants. Preterm infants are often critically ill, which may alter thyroid function and thyroid-binding proteins. We included a measure of illness level in the regression models, which we showed previously to be a good summary indicator, and also the T4/TBG ratios in an attempt to control for the preterm infants’ thyroid function. Several other important confounding measures of neurodevelopment were controlled for, which adds weight to our observations.

---

**Table 5** Adjusted* linear regression models of thyroid function which was consistent over at least two study bloods and the Bayley-III main and subtest domains

| Cognitive score | N | Effect estimate T statistic P value Lower Higher 95% CI |
|-----------------|---|----------------|---------------|--------|-------|-------|-------|
| Hypothyroxinaemic | 31 | -8.4 | -2.818 | 0.005 | -14.3 | -2.6 |
| Mixed dysfunction | 561 | -2.1 | -1.864 | 0.063 | -4.3 | 0.1 |
| Euthyroid | 398 | Reference | |

| Motor composite score | N | Effect estimate T statistic P value Lower Higher 95% CI |
|----------------------|---|----------------|---------------|--------|-------|-------|-------|
| Hypothyroxinaemic | 31 | -9.4 | -3.262 | 0.001 | -15.1 | -3.8 |
| Mixed dysfunction | 560 | -1.6 | -1.440 | 0.150 | -3.7 | 0.6 |
| Euthyroid | 398 | Reference | |

| Language composite score | N | Effect estimate T statistic P value Lower Higher 95% CI |
|--------------------------|---|----------------|---------------|--------|-------|-------|-------|
| Hypothyroxinaemic | 31 | -10.7 | -3.175 | 0.002 | -17.2 | -4.1 |
| Mixed dysfunction | 560 | -1.8 | -1.446 | 0.149 | -4.3 | 0.6 |
| Euthyroid | 397 | Reference | |

| Receptive language subtest | N | Effect estimate T statistic P value Lower Higher 95% CI |
|---------------------------|---|----------------|---------------|--------|-------|-------|-------|
| Hypothyroxinaemic | 31 | -1.8 | -2.927 | 0.004 | -2.9 | -0.6 |
| Mixed dysfunction | 560 | 0.4 | 1.788 | 0.074 | -0.9 | <0.1 |
| Euthyroid | 397 | Reference | |

| Expressive language subtest | N | Effect estimate T statistic P value Lower Higher 95% CI |
|----------------------------|---|----------------|---------------|--------|-------|-------|-------|
| Hypothyroxinaemic | 31 | -1.9 | -3.132 | 0.002 | -3.1 | -0.7 |
| Mixed dysfunction | 559 | -0.2 | -1.013 | 0.311 | -0.7 | 0.2 |
| Euthyroid | 397 | Reference | |

| Fine motor subtest | N | Effect estimate T statistic P value Lower Higher 95% CI |
|-------------------|---|----------------|---------------|--------|-------|-------|-------|
| Hypothyroxinaemic | 31 | -1.8 | -3.184 | 0.001 | -2.9 | -0.7 |
| Mixed dysfunction | 559 | -0.3 | -1.213 | 0.226 | -0.7 | 0.2 |
| Euthyroid | 398 | Reference | |

| Gross motor subtest | N | Effect estimate T statistic P value Lower Higher 95% CI |
|---------------------|---|----------------|---------------|--------|-------|-------|-------|
| Hypothyroxinaemic | 31 | -1.3 | -2.587 | 0.010 | -2.4 | -0.3 |
| Mixed dysfunction | 560 | -0.2 | -1.227 | 0.220 | -0.6 | 0.1 |
| Euthyroid | 398 | Reference | |

* adjusted for gender, gestational age, level of nursing care required on day 14, whether or not in receipt of iodine supplement, whether or not treated with L-thyroxine during the neonatal period or at 2 years of age, assessor and hospital of birth.
In summary, for the majority of infants in this study cohort, postnatal thyroid dysfunction was restricted to an isolated episode. No infants were consistently hypothyroid or hyperthyrotopinaemic; 38% were euthyroid throughout the sampling period and 5% were consistently hypothyrotopinaemic. Infants with TSH levels consistently in the top decile or categorised as hypothyroid specifically on postnatal day 7 performed worse on the Bayley scores than those categorised with mixed dysfunction or euthyroid. A small cohort of infants was consistently hypothyroid and this was associated with appreciably worse outcomes than euthyroid infants on all Bayley-III domains. Hypothyrotopinaemia in preterm infants does not appear to be a benign disorder on the basis of the associations that we have identified, and the potential role of thyroxine supplementation in this group of infants requires further investigation.

Acknowledgements The authors would like to thank the parents and children who participated in this study and the many NHS staff in UK hospitals who recruited, managed or received infants under the trial protocol. The authors would like to thank Dr Shona Livingston for helpful comments on the final submission.

Contributors FLRW: conceptualised and designed the study, designed and contributed to the data analysis, drafted the initial manuscript and approved the final manuscript as submitted. AL: contributed to the study design, critically reviewed the manuscript and approved the final manuscript as submitted. JW: contributed to the data analysis, reviewed and revised the manuscript and approved the final manuscript as submitted. AB: coordinated and supervised the analysis of blood spot cards, critically reviewed the manuscript and approved the final manuscript as submitted. TC: contributed to the design of the data analysis, reviewed and revised the manuscript and approved the final manuscript as submitted. JW: contributed to the data analysis, drafted the initial manuscript and approved the final manuscript as submitted. AB: coordinated and supervised the analysis of blood spot cards, critically reviewed the manuscript and approved the final manuscript as submitted. LW: contributed to the data analysis, supervised the analysis of blood spot cards, contributed to the design of the data analysis, reviewed and revised the manuscript and approved the final manuscript as submitted. JW: contributed to the design of the data analysis, critically reviewed the manuscript and approved the final manuscript as submitted. AB: coordinated and supervised the analysis of blood spot cards, critically reviewed the manuscript and approved the final manuscript as submitted. TC: contributed to the design of the data analysis, reviewed and revised the manuscript and approved the final manuscript as submitted.

Funding This study was funded by Medical Research Council (UK) and managed by NIHR on behalf of the MRC-NIHR partnership.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The trial was approved by Tayside Committee on Medical Research Ethics (08/SS0501/13).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. De-identified participant data will be shared in accordance with the National Perinatal Epidemiology Unit Data Sharing policy and Nuffield Department of Population Health Data Sharing policy. Requests for access to the data will be considered by the National Perinatal Epidemiology Unit Data Sharing committee from the date of publication. Data will be shared after approval of a proposal with investigator approval and completion of a signed data sharing agreement. Access to the data can be requested from ctu@npeu.ox.ac.uk.

REFERENCES
1. Zimmermann MB. Iodine deficiency. Endocr Rev 2009;30:376–408.
2. American Academy of pediatrics, section on endocrinology and Committee on genetics, rose SR; American thyroid association, public health Committee; Lawson Wilkins pediatric endocrine Society, brown RS. update of newborn screening and therapy for congenital hypothyroidism. Pediatr 2006;117:290–303.
3. Parks JS, Lin M, Grosse SD, et al. The impact of transient hypothyrotopinaemia on the increasing rate of congenital hypothyroidism in the United States. Pediatrics 2010;125:554–63.
4. Cacalacra f, Motta RM, Miscio G, et al. Subclinical hypothyroidism in early childhood: a frequent outcome of transient neonatal hyperthyrotopinaemia. J Clin Endocrinol Metab 2002;87:3209–14.
5. Unuvar T, Demir K, Abacı A, et al. The role of initial clinical and laboratory findings in infants with hyperthyrotopinaemia to predict transient or permanent hypothyroidism. J Clin Res Pediatr Endocrinol 2013;5:170–3.
6. Simpson J, Williams FLR, Delahunty C, et al. With collaboration from the Scottish preterm thyroid group, serum thyroid hormones in preterm infants and relationships to indices of severity of intercurrent illness. J Clin Endocrinol Metab 2005;90:1271–9.
7. Williams FLR, Ogston SA, van Toor H, et al. Serum thyroid hormones in preterm infants: associations with postnatal illnesses and drug usage. J Clin Endocrinol Metab 2005;90:5954–63.
8. Lazar L, Frumkin RB-D, Battat E, et al. Natural history of thyroid function tests over 5 years in a large pediatric cohort. J Clin Endocrinol Metab 2009;94:1678–82.
9. Cuestas E, Guido ML, Capra RH. Transient neonatal hyperthyrotopinaemia is a risk factor for developing persistent hyperthyrotopinaemia in childhood with Repercussion on developmental status. Eur J Endocrinol 2015;172:483–90.
10. Aten O, Wang MK, Brnjac L, et al. Mild neonatal hyperthyrotopinaemia: 10-year experience suggests the condition is increasingly common but often transient. Clin Endocrinol 2013;79:832–7.
11. O’Grady MJ, Cody D. Subclinical hypothyroidism in childhood. Arch Dis Child 2011;96:280–4.
12. Wasnner AJ, Brown RS. Hypothyrodism in the newborn period. Curr Opin Endocrinol Diabetes Obes 2013;20:449–54.
13. Daluvi AL, Linder B, DiMarto-Nardi A, et al. Three-Year follow-up of borderline congenital hypothyroidism. J Pediatr 2000;136:53–6.
14. Demirel F, Bideci A, Camuradan MO, et al. L-Thyroxine treatment in infants with hyperthyrotopinaemia: 4-year experience. Int J Clin Pract 2007;61:1333–6.
15. Williams F, Hume R. The measurement, definition, aetiology and clinical consequences of neonatal transient hypothyroidism. Ann Clin Biochem 2011;48:7–22.
16. Williams FLR, Ogston S, Hume R, et al. And the I2S2 consortium. Supplemental iodide for preterm infants and developmental outcomes at 2 years: an RCT. Pediatrics 2017;139:e20163703.
17. Williams F, Hume R, Ogston S, et al. A summary of the iodine supplementation study protocol (I2S2): a UK multicentre randomised controlled trial in preterm infants. Neonatology 2014;105:282–9.
18. Elvers LH, Loeber JG, Verkerk PH. Thyroxine-Binding globulin determination as an extra parameter in congenital hypothyrotopinaemia screening. Proc. 3rd meeting of ISNS (Levy HL, Mermos RJ, Grady GF eds) Boston, 1996: 250–2.
19. Delahunty C, Falconer S, Hume R, et al. And the Scottish Preterm Thyroid Group. Levels of neonatal thyroid hormone in preterm infants and neurodevelopmental outcome at 5.5 years: millennium cohort study. J Clin Endocrinol Metab 2010;95:4898–908.
20. Khalid R, Williams P, Williams FLR. Do studies reporting infant neurodevelopment adjust for the variability of assessors? Dev Med Child Neurol 2016;58:131–7.
21. Boelen A, van Veen M, Verkerk PH, et al. Measuring free thyroxine levels in neonatal heel-prick samples. Clin Chim Acta 2013;423:51–5.
22. Bayley N. Bayley scales of infant and toddler development®. 3rd edn, 2005(Bayley–III®).
23. van Wassenberga AG, Kok JH, de Vijlder JJM, et al. Effects of thyroxine supplementation on neurologic development in infants born at less than 30 weeks’ gestation. N Engl J Med 1997;336:21–6.
24. Den Ouden AL, Kok JH, Verkerk PH, et al. The relation between neonatal thyroid levels and neurodevelopmental outcome at age 5 and 9 years in a national cohort of very preterm and/or very low birth weight infants. Pediatr Res 1996;39:142–5.
25. Reuss ML, Panetti N, Pinto-Martin JA, et al. The relation of transient hypothyrotopinaemia in preterm infants to neurologic development at two years of age, N Engl J Med 1996;334:821–7.
26. Ng SM, Turner MA, Gamble C, et al. An explanatory randomised placebo controlled trial of levothyroxine supplementation for babies born <28 weeks’ gestation: results of the TIPIT trial. Tract; 2013:14:211–20.
27. Fisher DA. Hypothyrotopinaemia in premature infants: is thyroxine treatment necessary? Thyroid 1999;9:715–20.
28. Williams FLR, Simpson J, Delahunty C, et al. Developmental trends in cord and postpartum serum thyroid hormones in preterm infants. J Clin Endocrinol Metab 2004;89:5314–20.
29. Bongers-Schokking JJ, Koot HM, Wiersma D, et al. Influence of timing and dose of thyroid hormone replacement on development in infants with congenital hypothyroidism. J Pediatr 2000;136:292–7.
30. Lain SJ, Bentley JP, Wiley V, et al. Associations between borderline neonatal thyroid-stimulating hormone concentrations and educational and developmental outcomes: a population-based record-linkage study. Lancet Diabetes Endocrinol 2016.
31. Korada SM, Pearce M, Ward Platt MP, et al. Difficulties in selecting an appropriate neonatal thyroid stimulating hormone (TSH) screening threshold. Arch Dis Child 2010;95:169–73.
32. Combet E, Bouga M, Pan B, et al. Iodine and pregnancy – a UK cross-sectional survey of dietary intake, knowledge and awareness. Br J Nutr 2015;114:108–17.
33. van Wassenber- Leemhuis A, Ares S, Colonem B, et al. Thyroid hormone supplementation in preterm infants born before 28 weeks gestational age and neurodevelopmental outcome at age 36 months. Thyroid 2014;24:1162–70.