Statistical Analysis Plan, v0.7
Assessing iodine status and associated health outcomes in British women during pregnancy

Darren C Greenwood, Claire Keeble & Diane E Threapleton

Nov 22nd, 2018

1 Introduction

Iodine is a key micronutrient which is essential for healthy growth and particularly important during pregnancy and breastfeeding when demands increase to support the developing baby. Iodine deficiency complications are potentially serious for both mother and child, and can lead to thyroid enlargement, decreased thyroid-hormone production, pregnancy complications, impaired growth and developmental problems. This research aims to provide an estimate of iodine status in pregnant women in the UK plus how iodine may affect health during pregnancy and during the development of the child.

This study utilises existing urine samples where birth and developmental outcomes are known, as well as collects new data and samples for analysis. Three sources of data will be incorporated in this project; data from the pre-existing Born in Bradford project, data from the pre-existing SCOPE study, and newly collected data from the HIBA study.

2 Born in Bradford

The first dataset will use participant data and urine samples from the pre-existing Born in Bradford cohort.

2.1 Linking data

The urine samples are labelled with an aliquot identifier which corresponds to a study identifier, allowing confirmation of the participant and permitting linkage with participant data in the Born in Bradford database. Data from both sources will be monitored to ensure

the format is suitable for data linkage and linking will be tested as the data are collected to ensure there are no problems at the end of data collection.

A sample of the laboratory data from Leeds has been sent to Bradford to ensure the results of the laboratory analyses can be linked back into the Born in Bradford database at the end of the study.

2.2 Data cleaning

Initial data cleaning has already been conducted through the Born in Bradford project, but additional cleaning will be completed where necessary following exploratory data analysis. Any linked variables will be checked to ensure their combination is reasonable (for example, no diabetes treatment recorded for non-diabetic participants). Typographical errors will be corrected and any duplication of the same term will be combined (for example, Male/MALE/male/M). Any outliers will be identified using methods such as distance from the mean and excluded if outside a range considered possible, taking into account that some valid data are anticipated in the tails of the distributions. Cross-tabulation will be used to ensure consistency of characteristics that should change little over time. Missing data will be coded consistently through the dataset and all changes made through data cleaning will be recorded for transparancy and to create an audit trail.

Laboratory data entry has been restricted where possible to ensure values entered are clinically plausible and to minimise typographical errors. Input from clinical members of the team has been used to determine whether a continuous value is plausible and which categories should be assigned to binary or categorical variables. Drop-down boxes have also been utilised to ensure categorical data are recorded consistently and again to reduce the number of typographical errors. Laboratory measures are checked in an ongoing manner and questionable results arising from laboratory analyses will be rerun where possible, or excluded if the run has failed.

2.3 Primary outcomes

2.3.1 Cognitive outcomes

The main cognitive outcomes derive from Early Years Foundation Stage (EYFS) profiles and Key Stage 1 (KS1) educational outcomes. For EYFS, both the EYFS average points total over 17 domains (score range 17 to 41) and achieving Good Level of Development (GLD) (binary) will be presented, with pre-2013 results excluded as incomparable. For KS1, the Year 1

phonics screening test (score range 0 to 40) and KS1 SATS scores in Reading, Writing, Mathematics and Science (working securely at level 2b or beyond pre-2016 or working at the expected standard or beyond for 2016 onwards, binary) will be presented.

2.3.2 Birth and pregnancy outcomes

The main birth outcome of interest is birth weight centile, calculated using the version 8.0.1 of the Bulk centile calculator provided by the Gestation Network based at the Perinatal Institute for Maternal and Child Health [1].

2.3.3 Other outcomes

None.

2.4 Secondary outcomes

2.4.1 Cognitive outcomes

Letter identification (stdandardised), Receptive vocabulary using the British Picture Vocabulary Scale (BPVS)(standardised), Social functioning using the Social Difficulties Questionnaire (SDQ) total difficulty score and split by domains (e.g. internalising and externalising), Fine motor skills using the Clinical Kinematic Assessment Tool (CKAT) overall battery score, along with Tracking, Aiming and Tracing CKAT sub-scores. Key Stage 2 (KS2) SATS scores in English grammar, punctuation and spelling, English reading, and Mathematics.

2.4.2 Birth and pregnancy outcomes

Secondary outcomes relating to birth and pregnancy include birth weight (continuous), low birth weight (<2.5kg, binary), small for gestational age (<10^{th} centile, binary), pre-term birth (<37 weeks, binary), crown-to-rump length (continuous), head circumference (continuous), spontaneous abortion/miscarriage (binary), stillbirth (binary), congenital anomalies from linked congenital anomalies records (binary), estimated crown-to-rump length and head circumference from linked data from ultrasound scans during pregnancy (continuous) and APGAR score (ordinal).

2.4.3 Other outcomes

Other secondary outcomes of interest relate to childhood growth trajectories (height and weight, continuous), and maternal health during and after pregnancy, including gestational diabetes (binary), preeclampsia (binary), blood pressure (continuous), length of gestation (continuous), mortality (binary) and maternal mental health (PHQ-9 Depression Test Questionnaire, ordinal). Confirmed diagnosis of autism, Attention Deficit Hyperactivity Disorder (ADHD), based on Read codes in linked General Practitioner records.

2.5 Primary exposure

The main exposure of interest is iodine status during pregnancy, measured through urinary iodine concentration (UIC) of urine samples taken at 26–28 weeks gestation as a ratio to urinary creatinine concentration (continuous, μg/g), which takes into account urinary dilution. Absolute UIC will also be used (continuous, μg/L).

2.6 Sample size

All maternal urine samples collected from the Born in Bradford cohort will be analysed (6971 women), and their associated participant data will be retrieved. The Born in Bradford cohort has detailed outcome measures available relating to complications during pregnancy, birth outcomes and measures of child development which, in conjunction with the size of the cohort, lead to sufficient power to detect any associations that are large enough to be of clinical relevance.

2.6.1 Birth outcome

Per ▮ standard deviation difference in UIC (~100 μg/L) there is approximately 90% power to detect as statistically significant ($p<0.05$) a relative risk of small-for-gestational-age of 1.05, and a difference in birthweight of 10g, plus approximately 80% power for a small-forgestational-age relative risk of 1.04 and a birthweight difference of just 8g. For comparison, mothers who smoke have babies with 200-250g lower birthweight than non-smokers, mothers with a high of caffeine have babies with 70g lower birthweight than those without [2, 3], and mothers exposed to secondhand smoke have 33g lower birthweight than those not [4]. Hence representing a very modest sized association that can be detected.

2.6.2 Early years foundation stage outcome

Per ▮ standard deviation difference in UIC (~100 g/L) there is approximately 90% power to detect a 0.042 standard deviation difference in fine motor skills, and approximately 80% power to detect a difference of 0.037 standard deviations. That is, for example, less than a fifth of the difference between boys and girls, and hence represents a very modest size difference which can be detected.

2.6.3 Key stage 1 outcomes

Per ▮ standard deviation difference in UIC (~100 g/L) there is approximately 90% power to detect a relative risk of not meeting the expected standard of phonic decoding of 1.04, and a relative risk of 1.04 for not achieving level 2b or above in maths key stage 1 results, assuming 75% meet the phonics targets and 80% achieve meet the maths key stage 1 level. There is approximately 80% power for respective relative risks of 1.03 and 1.04. Again, these represent very modest sized associations in absolute terms, equivalent to just 0.4% of the total phonics mark, and about one fifth of the difference in reading accuracy observed in the ALSPAC cohort for UIC<150 μg/L compared to higher concentrations [5]. Hence the study is powered to detect any effects of an achievable change in UIC, equivalent to the effect of an intervention such as introducing iodised salt.

2.7 Response rates

The response rates for the Born in Bradford cohort can be taken from the initial recruitment figures. All participants of the Born in Bradford study who provided urine samples will be included in these analyses.

The Born in Bradford study receives all assessment data that schools routinely report to the local authority. It contains the Early Years Foundation Stage Profile, phonics screening checks (year 1), key stage 1 results in reading, writing, speaking and listening, maths and science (year 2), plus some contextual data (school location, attendance, free school meals, SEN, EAL). Whilst national reporting requirements vary, educational data is recorded for 3000 children from the first two academic years (representing 85% data linkage), and it is anticipated that the same proportion will be present in the final two academic years, which should report on approximately 5900 year 1 children and 5000 year 2 children by the end of 2017. In addition, the Born in Bradford study has CKAT scores capturing fine motor skills

and literacy, and SDQ scores for mental health, for approximately 1500 reception year children.

In addition, data relating to mental health (GHQ28 questionnaire encompassing - anxiety, depression, insomnia and somatic profiling) are available during pregnancy and for a subset of women post-partum (N= approx 500 at 6, 12, 18 months).

2.8 Descriptive statistics

Exploratory data analyses will be conducted, including summary statistics of each variable to ensure all values recorded are plausible, and to ensure any linked variables are possible (for example, there is no medication for depression recorded where the participant is recorded as not having depression).

Estimated iodine status of the population will be presented (median and geometric mean UIC), with % of individuals below 50 and 150 μg/L.

Differences in iodine status relating to ethnicity, obesity, socioeconomic status, maternal age, parity and other variables will be described, as the Born in Bradford cohort contains approximately 45% South Asian mothers, participants from a range of socioeconomic status', and the obesity rate in the cohort is around 30%, allowing such comparisons to be possible. Multiple testing will be accounted for where necessary. Any associations found between variables in the dataset will be verified with expert advisors to ensure the data structure and any associated variables are understood before statistical modelling commences.

The characteristics of (urine and data) participants will be compared with the characteristics of the overall cohort and with regional data where possible, to investigate whether the participants who provided urine samples were not systematically different to those who did not.

2.9 Statistical modelling

All analyses will be completed using an annotated script file which will act as a record of data cleaning and statistical analysis. Transformation of urinary iodine concentration may be required and will be considered during analyses if necessary. Variables will be categorised if beneficial to the interpretation of the analysis and if the categories can be justified clinically or otherwise. Where possible, continuous variables will not be categorised, so information is not lost, and no assumptions made about the functional relevance of previously suggested thresholds.

Unadjusted regression analyses will be described in the text only. Multiple ordinal/logistic regression and multiple linear regression will be used to ascertain potential associations between maternal iodine status and health and developmental outcomes. Cubic splines will be used to model any nonlinear dose-response associations, avoiding any assumptions of pre-defined categories based on different clinical outcomes, e.g. the World Health Organisation (WHO) criteria. Instead, any potential different thresholds relevant to birth outcomes or cognitive outcomes, or otherwise can be identified. This will help inform discussion on levels of iodine status requiring potential intervention for these particular outcomes. From adjusted regression models, the following will be reported;

Variable	Regression coefficient[a]	95% confidence interval	p-value
Variable 1 Variable 2 Variable 3 ...			

[a] where logistic regression is used, odds ratio will be reported in place of the regression coefficient

All model assumptions will be checked, as will the model fit, and graphical representations of the data will be provided where useful. Subgroup analyses will be conducted by participant characteristics such as ethnicity, obesity and child's sex.

2.10 Potential confounders

Directed acyclic graphs (DAGs) will be used to inform the list of covariates included in the models, but are likely to include age (continuous variable), ethnicity (categorical variable), smoking (binary variable) and the gender of the child (binary variable). Over-adjustment will be avoided to maintain precision. Consideration will be given to any variables which already adjust for other variables.

2.11 Missing data

Any data missing that can be derived from other variables will be, such as the number of days between events A and B where the dates of each event are provided. Any other missing data will remain missing and techniques such as multiple imputation will not be adopted, to maintain clinical and demographic accuracy. Where a participant has missing data in a variable required for a given analysis, the given participant will be excluded for that analysis,

but will remain in the dataset for further analyses, hence maximising the number of participants for each analysis.

2.12 Exclusions and sensitivity analyses

The aim is to include all participants, but sensitivity analyses will be conducted to test the robustness of the conclusions when iodine values which fall outside three standard deviations of the mean UIC are excluded. Sensitivity analyses will also be conducted where mothers with complications relating to other conditions are removed. Any clinically implausible values remaining after data cleaning will be excluded from the analysis, since they could be due to a typographical error. However, typographical errors should be minimised by the ability to trace the UIC values back to the raw iodine values.

2.13 Comparison across studies

The iodine status of pregnant women in the Born in Bradford cohort (Bradford) will be compared with those from each geographical site in the SCOPE study (Leeds, Manchester, Kings), to determine whether Bradford is broadly representative of the UK iodine status during pregnancy, or whether there are differences nationally. More detailed demographic comparisons will also be made across the sites.

2.14 Timescales

Target analysis completion date: August 2018.

3 SCOPE

The second dataset will be taken from the pre-existing SCOPE study. This includes data and urine samples from the Leeds, Manchester and Kings sites.

3.1 Linking data

The participant data will be collated by the SCOPE Cork site, anonymised, and sent directly to Leeds. Urine samples will be pulled separately in each of the three participating sites (Leeds, Manchester and Kings) and sent directly to Leeds. Data from both sources will be monitored as they are collected to ensure the format is suitable for data linkage. Urine

samples will be linked to the corresponding participant data using a four-digit participant identification number. Linking will be tested as the data are collected to ensure there are no problems at the end of data collection.

3.2 Data cleaning

Initial data cleaning has been conducted during the initial SCOPE study. Summary statistics and plots will be used to ensure all values recorded are plausible and data will be cleaned where necessary. Input from clinical members of the team was used to determine whether a continuous value is plausible and which categories should be assigned to binary or categorical variables. Any linked variables will be checked the ensure their combination is reasonable (for example, no diabetes treatment recorded for non-diabetic participants). Typographical errors will be corrected and any duplication of the same term will be combined (for example, Male/MALE/male/M). Any outliers will be identified using methods such as distance from the mean and excluded if outside the range considered possible, taking into account that some valid data are anticipated in the tails of the distributions. Cross-tabulation will be used to ensure consistency of characteristics that should change little over time. Missing data will be coded consistently through the dataset and all changes made through data cleaning will be recorded for transparency and to create an audit trail.

Laboratory data entry has been restricted where possible to ensure values entered are clinically plausible and to minimise typographical errors. Drop-down boxes have also been utilised to ensure categorical data are recorded consistently and again to reduce the number of typographical errors. Laboratory measures are checked in an ongoing manner and any questionable results arising from laboratory analyses will be rerun where possible, or excluded if the run has failed.

3.3 Primary outcome

The main outcome of interest is birth weight centile (continuous) [1], with secondary outcomes including small for gestational age (<10^{th} centile, binary), birth weight (continuous), low birth weight (<2.5kg, binary), preterm birth (<37 weeks, binary), spontaneous abortion/miscarriage (binary) and stillbirths (binary).

3.4 Primary exposure

The main exposure of interest is iodine status during pregnancy, measured through both the urinary iodine concentration (UIC) and the iodine to creatinine ratio, at 15 weeks' and 20 weeks' gestation (continuous).

3.5 Sample size

The requested sample size is all samples from all SCOPE sites in England (Leeds, Manchester and Kings). SCOPE blood samples were not requested, but urine samples were requested for all participants. SCOPE protocol requires three 0.9ml aliquots to remain in storage for each participant for each timepoint (15 weeks and 20 weeks' gestation). Provided this condition could be met, 2 × 0.9ml urine samples were provided for laboratory analysis, accompanied by the corresponding participant data. All participant variables relevant to the study interests were requested - a total of 271 variables were provided for 641 participants, along with 2351 aliquots (urine).

3.6 Response rates

SCOPE response rates can be taken from the initial recruitment figures and compared with the number of participants provided for analysis in this study. A small number of women were excluded from the SCOPE study due to specified underlying medical conditions unrelated to iodine status, that are known to lead to preeclampsia, SGA and pre-term birth, and excludes known major foetal anomalies and abnormal karyotypes, but otherwise the cohort is thought to be broadly representative of UK mothers, including socioeconomic features.

3.7 Descriptive statistics

Summary statistics and plots will be generated to ensure that all variables have plausible values and any values outside the expected range will be checked. Differences in iodine status relating to ethnicity, obesity, socioeconomic status, maternal age, parity and other variables will be described. Any associations found between variables in the dataset will be verified with expert advisors to ensure the data structure and any associated variables are understood before statistical modelling commences.

3.8 Statistical modelling

All analyses will be completed in an annotated script file which will act as a record of data cleaning and statistical analysis. Transformation of urinary iodine concentration may be required and will be considered during analyses if needed. Variables will be categorised if beneficial to the interpretation of the analysis and if the categories can be justified clinically or otherwise. Where possible, continuous variables will not be categorised, so information is not lost.

Unadjusted regression analyses will be described in the text only. Multiple logistic regression will be used to estimate associations between iodine status and binary outcomes such as maternal complications and small for gestational age, adjusting for known confounders. Multiple linear regression will be used to investigate continuous outcomes such as birthweight, again adjusting for confounders. From adjusted regression models, the following will be reported;

Variable	Regression coefficient[a]	95% confidence interval	p-value
Variable 1 Variable 2 Variable 3 ...			

[a] where logistic regression is used, odds ratio will be reported in place of the regression coefficient

All model assumptions will be checked, as will the model fit and graphical representations of the data will be provided where useful.

3.9 Potential confounders

The same covariates as used in Born in Bradford models will be included where possible. Over-adjustment will be avoided to maintain precision and consideration will be given to any variables which already adjust for other variables.

3.10 Missing data

Any data missing which can be derived from other variables will be, such as the number of days between events A and B where the dates of each event are provided. Any other missing

data will remain missing and techniques such as multiple imputation will not be adopted, to maintain clinical and demographic accuracy. Where a participant has missing data in a variable required for a given analysis, the given participant will be excluded for that analysis, but will remain in the dataset for further analyses, hence maximising the number of participants for each analysis.

3.11 Exclusions and sensitivity analyses

The aim is to include all participants, but sensitivity analyses will be conducted to test the robustness of the conclusions when iodine values which fall outside three standard deviations of the mean UIC are excluded. Any clinically implausible values remaining after data cleaning will be excluded from the analysis, since they could be due to a typographical error.

3.12 Assessment of between-site variability

Subgroup analyses will be conducted for the Leeds, Manchester and Kings data to investigate whether associations differ nationally. Site will also be included in the regression models as a covariate when the data are combined.

3.13 Comparison across studies

The iodine status of pregnant women in the SCOPE study (Leeds, Manchester and Kings), will be compared with the iodine status of pregnant women in the Born in Bradford study (Bradford) and the newly recruited pregnant women in the HIBA study (also Bradford), to report similarities or differences over time and nationally.

3.14 Timescales

Target analysis completion date: June 2017.

4 Health and Iodine status in Babies (HIBA)

The final dataset will be generated using participants recruited specifically for this study. HIBA (Health and Iodine status in BAbies) is a new longitudinal substudy to ascertain how iodine status varies between trimesters and during lactation. Women will be recruited at

their 12 week dating scan and the following data collected: baseline characteristics, urine samples for iodine analysis, blood for thyroid stimulating hormone (TSH), free thyroxine (fT4), triodothyronine (fT3) and thyroglobulin, Glomerular Filtration Rate (GFR) based on serum creatinine, visual inspection of the thyroid using standard methods, dietary intake using a validated online 24 hour recall tool (myfood24). Data collection will be repeated at 26 and 36 weeks gestation, and 6, 18 and 30 weeks postpartum.

4.1 Timeline for data collection - dietary, samples and questionnaire

Target data collection timepoints are at 12, 26 and 36 weeks' gestation, and 6, 18 and 30 weeks postpartum. However if the target week is not possible, the meeting may be held up to three weeks before or up to three weeks after the target week:

1. From 9 weeks gestation + 0 days to 15 weeks gestation + 6 days
2. From 23 weeks gestation + 0 days to 29 weeks gestation + 6 days
3. From 33 weeks gestation + 0 days to 39 weeks gestation + 6 days
4. From 3 weeks postpartum + 0 days to 9 weeks postpartum + 6 days
5. From 15 weeks postpartum + 0 days to 21 weeks postpartum + 6 days
6. From 27 weeks postpartum + 0 days to 33 weeks postpartum + 6 days

If a meeting is not possible within the permitted timeframe, the meeting is abandoned and recorded as missing data. The next meeting is booked as usual, unless the participant wishes to withdraw from the study. Exact dates will be incorporated into the modelling.

4.2 Linking data

The urine and blood samples will be labelled with the unique participant identification number and this number will also be used on the questionnaires. The myfood24 food diaries will be labelled with a different identifier in the format HIBA01234Visit. A 32 bit token and mapping table will be used to link these two identifiers. Data from all sources will be monitored as they are collected to ensure the format is suitable for data linkage. Linking will be tested as the data are collected to ensure there are no problems at the end of data collection.

4.3 Data cleaning

Summary statistics and plots will be used to ensure all values recorded are plausible and data will be cleaned where necessary. Input from clinical members of the team was used to determine whether a continuous value is plausible and which categories should be assigned to binary or categorical variables. Any linked variables will be checked the ensure their combination is reasonable and any outliers will be identified using methods such as distance from the mean and excluded if outside the range considered possible, taking into account that some valid data are anticipated in the tails of the distributions. Cross-tabulation will be used to ensure consistency of characteristics that should change little over time. Missing data will be coded consistently through the dataset and all changes made through data cleaning will be recorded for transparancy and to create an audit trail.

Laboratory data entry has been restricted where possible to ensure values entered are clinically plausible and to minimise typographical errors. Drop-down boxes have also been utilised to ensure categorical data are recorded consistently and again to reduce the number of typographical errors. Laboratory measures are checked in an ongoing manner and any questionable results arising from laboratory analyses will be rerun where possible, or excluded if the run has failed.

4.4 Primary outcome

The main outcome of interest is a summary of iodine status changes during pregnancy and lactation (continuous), informed by urine samples at three timepoints during pregnancy (12, 26 and 36 weeks' gestation) and three timepoints during lactation (6, 18 and 30 weeks' lactation).

4.5 Secondary outcomes

Secondary outcomes of interest concern changes in thyroid hormones (thyroid stimulating hormone (TSH), free thyroxine (fT4), triodothyronine (fT3) and thyroglobulin, continuous) and size (from palpation, categorical) during gestation and lactation.

Identification of key sources of iodine in the diet as measured over six timepoints using the myfood24 tool, and dietary patterns associated with iodine status.

The final outcomes of interest correspond to birth and pregnancy outcomes in the Born in Bradford and SCOPE analyses; including birth weight centile (continuous), small for gestational age (<10th centile, binary), birth weight (continuous), low birth weight (<2.5kg,

binary), preterm birth (<37 weeks gestation, binary), spontaneous abortion/miscarriage (binary) and stillbirth (binary), child growth (height and weight) during lactation (continuous).

4.6 Primary exposure

The main exposures of interest in relation to iodine status is dietary intake: Iodine intake (continuous) as measured at 6 time points using the myfood24 tool, food groups and dietary patterns predicting iodine status.
In relation to thyroid hormones (thyroid stimulating hormone (TSH), free thyroxine (fT4), triodothyronine (fT3) and thyroglobulin, continuous) and size (from palpation, categorical) during gestation and lactation, a key exposure of interest is urinary iodine status.

4.7 Sample size

Between-individual and within-individual variances of urinary iodine concentration from spot urines were approximately derived [6]. Three measurements (at 12, 26 and 36 weeks' gestation or 6, 18 and 30 weeks' postpartum) from 200 participants would allow estimation of the geometric mean UIC to within ±4%, and the estimation of mean iodine status at a single timepoint to ±6%, allowing the precision required to investigate changes during and after pregnancy. Iodine status by ethnic group will be reported provided participants are sufficiently distributed across different ethnic groups. Recruiting 100 white women and 100 South Asian women would allow estimation of the geometric mean UIC in the subgroups to within ±7%, and similar power calculations would apply for comparisons between obese and non-obese women. To allow for anticipated attrition of around 30% during the study, the initial recruitment target will be 300 women. Recruiting 200 women and recording three time points during pregnancy would result in approximately 80% power to detect a 15% difference in UIC (equating to approximately 20μg/L for an average woman), hence representing good power to detect small differences.

4.8 Response rates

For each month of data collection, response rates will be recorded as follows:

- The number of women approached in the scan department;
- The number of women declining information or further contact;

- The number of women not eligible;
  - By age;
  - By thyroid condition;
- The number of women to be contacted; • The number of women booked for an appointment;
- The number of women recruited.

The final report will include at least summaries of these response, exclusion and completion rates. Response rates by visit and data source will also be recorded, since there may be individuals who are unable to provide, for example, a urine sample at a particular visit, but who are able to provide a blood sample and answers to the questionnaire and food diary questions. In addition, the number of women withdrawing from the study and with missing visits will be recorded, and a summary of the withdrawal reasons provided if known.

	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5	Visit 6
Questionnaire myfood24 Urine sample Blood sample						

	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5	Visit 6
Missed visit Had withdrawn						

4.9 Descriptive statistics

Exploratory data analyses will be conducted, including summary plots and statistics of each variable to ensure all values recorded are plausible, and any linked variables are possible. Differences in iodine status relating to ethnicity, obesity, socioeconomic and other variables will be described, and multiple testing will be accounted for where necessary. Any associations found between variables in the dataset will be verified with expert advisors to ensure the data structure and any associated variables are understood before statistical

modelling commences. Tables of descriptive statistics will be generated and will include baseline and visit-specific variables such as:

	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5	Visit 6
Mother's age		-	-	-	-	-
Mother's height		-	-	-	-	-
Mother's weight		-	-	-	-	-
Ethnicity		-	-	-	-	-
...		-	-	-	-	-
Iodine intake UIC Iodine supplementation ...						

The characteristics of complete participants will be compared with the characteristics of participants with missing data, to investigate whether the participants who completed were systematically different to those who did not.

4.10 Statistical modelling

All analyses will be completed in an annotated script file which will act as a record of data cleaning and statistical analysis. Variables will be categorised if beneficial to the interpretation of the analysis and if the categories can be justified clinically or otherwise. Where possible, continuous variables will not be categorised, so information is not lost. Subgroup analyses will be conducted by participant characteristics such as ethnicity and obesity, provided the participants are sufficiently distributed across different categories of these variables.

Transformation of urinary iodine concentration may be required and will be considered during analyses if needed. Urinary iodine concentration will be used in analyses as six timepoints if changes over time are of interest, or the mean of the first three and final three values will be taken if comparisons between gestation and postpartum are being conducted.

Unadjusted regression analyses will be described in the text only. The association between dietary iodine intake, dietary goitrogens such as brassicas, and subsequent iodine status will be assessed using multiple regression and multilevel models (where appropriate), adjusting for known confounders. The association between thyroid hormone concentrations during pregnancy and postpartum with the iodine measures and thyroid size will be assessed using multilevel models that take into account the repeated measures. All model assumptions will be checked, as will the model fit and graphical representations of the data will be provided where useful.

4.11 Potential confounders

Confounders will be selected using directed acyclic graphs (DAGs) and specialist knowledge will be sought where necessary. Over-adjustment will be avoided to maintain precision. The potential for confounding of associations between iodine status and pregnancy outcomes by renal function will be assessed using a plasma creatinine-based approximation of the Glomerular Filtration Rate (GFR). Consideration will be given to any variables that already adjust for other variables, such as socioeconomic status.

4.12 Missing data

Any data missing which can be derived from other variables will be, such as the number of days between events A and B where the dates of each event are provided. Any other missing data will remain missing and techniques such as multiple imputation will not be adopted, to maintain clinical and demographic accuracy. Where a participant has missing data in a variable required for a given analysis, the given participant will be excluded for that analysis, but will remain in the dataset for further analyses, hence maximising the number of participants for each analysis.

4.13 Exclusions and sensitivity analyses

Sensitivity analyses will be performed to assess the robustness of the conclusions drawn. This will include analyses excluding those who missed visits, those taking iodine supplements, and high risk mothers. Any women later found to have had an undiagnosed thyroid condition will be excluded from the analysis.

4.14 Comparison across studies

The urine analyses from the HIBA study (Bradford) will be compared with the urine analyses from the SCOPE study (Leeds, Manchester and Kings) to determine whether Bradford is representative of, or differs from, other locations nationally. Changes over time will also be compared for HIBA and SCOPE, since both datasets include repeated urine measures.

4.16 Timescales

Target analysis completion date: February 2019.

References

[1] Gardosi J, Francis A. Customised Centile Calculator. GROW version 8.0.1, 2018. Gestation Network (www.gestation.net).

[2] LW Chen, Y Wu, N Neelakantan, M Chong, A Pan, and RM van Dam. Maternal caffeine intake during pregnancy is associated with risk of low birth weight: A systematic review and dose-response meta-analysis. *BMC Medicine*, 12(1):174, 2014.

[3] DC Greenwood, NJ Thatcher, J Ye, L Garrard, G Keogh, LG King, and JE Cade. Caffeine intake during pregnancy and adverse birth outcomes: A systematic review and dose-response meta-analysis. *European Journal of Epidemiology*, 29(10):725–734, 2014.

[4] J Leonardi-Bee, A Smyth, J Britton, and T Coleman. Environmental tobacco smoke and fetal health: Systematic review and meta-analysis. *Archives of Disease in Childhood Fetal and Neonatal Edition*, 93(5):F351–F361, 2008.

[5] SC Bath, CD Steer, J Golding, P Emmett, and MP Rayman. Effect of inadequate iodine status in uk pregnant women on cognitive outcomes in their children: Results from the avon longitudinal study of parents and children (alspac). *The Lancet*, 382:331–337, 2013.

[6] F Konig, M Andersson, K Hotz, I Aeberli, and MB Zimmermann. Ten repeat collections for urinary iodine from spot samples or 24-hour samples are needed to estimate individual iodine status in women. *Journal of Nutrition*, 141:2049–2054, 2011.