STATISTICAL ANALYSIS PLAN FOR SNAP
INCORPORATING FOLLOW-UP ANALYSES POST PRIMARY OUTCOME

Final Version no. 3.0

Date: 12th December 2012

NOTE: This version of the statistical analysis plan was prepared after all primary outcome analyses were completed. Previous Version 2.2 was used for all analyses conducted at the primary outcome point (Section One).

This version is for all analyses conducted on follow-up data, which was collected post-primary outcome. The information for this part of the analysis plan is in Section Two.

Full trial title: Double-blind, randomised, placebo-controlled trial of nicotine replacement therapy in pregnancy

Acronym: Smoking, Nicotine and Pregnancy (SNAP) trial

EudraCT number: 2004-002621-46

International Standardised Randomised Controlled trial Number – ISRCTN (if funded):

ISRCTN07249128

Trial sponsor: *University of Nottingham*

Chief (Principal) investigator: Tim Coleman MD, MRCGP
Reader in General Practice
Division of Primary Care

Analysis Plan prepared by: Sarah Lewis MSc PhD, Matthew Grainge MSc PhD

Reviewed by: _____

Nominated statistician for analysis: Sarah Lewis MSc PhD, Matthew Grainge MSc PhD

Projected start date: May 2007

Expected completion date: November 2010 (primary endpoints), November 2012 (secondary endpoints)

Published Trial protocol: http://www.biomedcentral.com/1472-6963/7/2

SECTION ONE

1. SYNOPSIS OF STUDY DESIGN AND PROCEDURES

1.1. Summary of study background & context in development plan, results from other trials

Maternal smoking during pregnancy harms unborn children and, as up to 30% of pregnancy women smoke, it is a significant public health problem. Currently only around 25% of pregnancy smokers stop for even part of their pregnancy (Owen L, 1999, HDA). Effective methods for promoting smoking cessation by pregnancy women are required. The most effective smoking cessation therapy in non-pregnant smokers involves a combination of behavioural support and pharmacotherapy with either NRT (Silagy C, Cochrane Review) or bupropion. Behavioural support alone can increase smoking cessation rates by up to 7% (Lumley J, Cochrane Review 2002), and the addition of pharmacotherapy increases this further by 1.5 to 2 fold. To date evidence on the effectiveness of NRT in pregnancy comes from 3 studies and is inconclusive (Wisborg K, 2000; Kapur B 2001; Hegaard HK 2003). The biggest of these studies randomised 250 women and produced no clear evidence that NRT was effective since the odds ratio using NRT versus placebo was 1.1 with a 95% CI of 0.7 to 1.8. This odds ratio is much lower than that obtained from meta-analysis of trials using NRT patches in non-pregnant women (OR 1.74). It may be that conventional doses of nicotine contained in NRT may be insufficient for pregnant women because the metabolic clearance of nicotine and cotinine increases by 60% and 140% respectively. However unless the effectiveness of current conventional dose is established it is difficult to justify trials of higher ones.

1.2. Objectives and aims (from protocol)

The *SNAP* trial will investigate whether or not NRT is more effective than placebo in achieving smoking cessation for women who are between 12 and 24 weeks pregnant, who currently smoke 5 or more cigarettes daily and who smoked 10 or more cigarettes daily before pregnancy.

We will also investigate whether there is improvement in pregnancy, child and maternal health outcomes up to 2 years in NRT versus placebo treatment arms, and determine the cost-effectiveness of the intervention.

1.3. Patient population studied

Women between 12 and 24 weeks of pregnancy, who report smoking at least 10 cigarettes per day prior to pregnancy and who still smoke at least 5 cigarettes per day. They must also have an exhaled CO reading of at least 8 ppm.

1.4. Trial configuration: Multicentre, parallel group with 1:1 allocation between NRT and placebo.

1.5. Randomisation procedures

1.5.1. Points of randomisation and the baseline visit (Schematic diagram of trial design, procedures and stages in Appendix A)

After collecting pre-randomisation baseline data, exhaled carbon monoxide readings will be taken from women and assuming that readings indicate that women do smoke

[cut off 8 ppm], informed consent for trial entry will be sought. After consenting to trial entry, women will receive an initial behavioural support session before being randomised.

Randomisation will be via the Nottingham Trials Unit web-based database and randomisation service. In each centre the recruiting research midwife (RM) will have a username and password. (S)he will log on to the trial website that hosts the trial database (http://ctsu.nottingham.ac.uk/snap/login.asp), confirm that the patient eligibility criteria are all met and enter an agreed minimum amount of *registration* data about the participant and centre before randomisation is possible. The computer will then issue a trial number which will be the unique identifier for the trial participant and a trial pack number which will reflect the treatment allocated.

1.5.2. Specify block size, whether randomly varied,

Block size randomly varied from 2 to 4

1.5.3. Stratified allocation, or post-stratified analysis

Randomisation will be stratified by trial centre only.

1.5.4. Minimisation: specify procedure

Not applicable

1.6. Allocation concealment:

1.6.1. Implementation of the random allocation sequence

Numbered packs of active and placebo patches will be distributed by Queens Medical Centre pharmacy and stored in the local site pharmacies. After randomisation, a prescription with a container number will be generated by the database. The local pharmacy will select the patch pack with the appropriate number and issue this to the participant. The research midwife and the trial participant will both be blind to group allocation and NRT / placebo will be prescribed under the supervision local principal investigators. When research midwives visit women at home to enrol them into the trial, immediate internet randomisation will not be possible. In this circumstance the research midwife will return to her / his hospital base to randomise the enrolled woman and the appropriate trial pack will be posted to the trial participant.

1.6.2. Blinding: who were blinded to group assignment

research midwife and research team, participants

1.7. Any safety, data monitoring or special steering or evaluation committees

Trial Steering Committee (Peter Brocklehurst, Oxford; Peter Hajek, Barts and London; Carol Coupland, Nottingham; Sue Maguire Lay member; Michael Murphy, Oxford)

Data monitoring and ethics committee (Janet Peacock, Brunel; Khalid Khan, Birmingham – replaced by Christopher Butler, Cardiff to maintain quorate, after inability to attend; David Field, Leicester)

1.8. Any interim analyses

1.8.1. Reports for Data Monitoring and Ethics Committee (DMEC)

The data monitoring committee will see a comparison of baseline data, efficacy outcomes (overall smoking cessation rates) and safety outcomes (serious adverse events, mean birth weight and gestation at birth) by treatment group at each 6 monthly meeting of the committee. Treatment groups will be blinded by the Nottingham Trials Unit and labelled as A and B for the analysis conducted by the trial statistician and presented to members of the DMEC. Reports to the DMEC will also include details of overall recruitment rates in the trial and list key process variables.

1.8.2. Stopping rules determined as part of the protocol

Stopping rules

1 The DMEC will consider stopping the trial if quit rates in the whole sample fall below 4%, but a final decision on this will only be taken after consideration of any impact of NRT on birth outcomes.

2 A graph of projected recruitment against time will be drawn and should trial recruitment fall below 25% of that which is expected at any time point without reasonable explanation or remedial action being possible, then the DMEC will consider recommending that the trial be stopped.

Explanation: The Trial is powered on the basis of finding quit rates of 16% and 25% in intervention and control groups respectively. Consequently, an overall quit rate of 4% averaged across both groups would not be consistent with a clinically worthwhile treatment effect, and would have only 45% power to detect a difference between treatment arms.

1.9. Discuss any statistical implications of the study design.

None

1.10. Efficacy and Safety Variables

Primary end point: The primary outcome is defined as self-reported, prolonged and total abstinence from smoking between quit date and delivery, validated by exhaled CO and salivary cotinine at childbirth. Occasional minor lapses (no more than 5 cigarettes in total) will not be counted as a return to smoking.

The primary outcome is derived from responses at 1 month and delivery as follows.

Abbreviation: CO = exhaled carbon monoxide breath test, COT = saliva cotinine measurement. At the outset of the trial only CO was obtained at delivery but, at DMEC/TSC request this was changed at COT was added. Consequently, for most participants, both CO and COT are available at delivery.

Positive response (i.e. abstinent from smoking)

At 1 month:
'smoked since quit date' = 'no' or 'missing'
OR 'how often have you smoked = 5 times or less'
OR 'at least weekly but less than daily'
OR missing

(i.e. any response other than on most days or frequently).
AND

At delivery:

'smoked in last 24 hrs = no'
AND 'smoked since quit date = no'
AND CO result is between 0 and 8.
AND/OR COT[1] less than 10ng/ml[2]

OR
'how often have you smoked' = '5 times of less'
AND CO result is between 0 and 8.
AND/OR COT less than 10ng/ml

Negative response:

Any other response at 1 month and delivery, including those withdrawn from, or refused follow-up at 1 month or delivery. The only exception is where consent for using any data is withdrawn.

Permitted timing and rules of data collection:

Self reported smoking data will be used if this is collected within i) eight weeks of the one month follow up point and ii) within one month of delivery.

Secondary end points:

a) Smoking

1. Self reported, prolonged abstinence from smoking between quit date and one month.
2. Self reported, prolonged abstinence from smoking between quit date and delivery.
3. Self reported, prolonged abstinence from smoking between quit date and delivery, with biochemical validation of this at both one month follow up and delivery.

[1] Some participants will only have CO measurements and, for these women, readings in the stated reference range are defined as a positive primary outcome (even without COT). Most trial participants will have both CO and COT measurements and, for these women, BOTH readings must fall within defined ranges to count as having a positive outcome.

[2] If a new normative value becomes available before the trial ends (when saliva cotinine samples are analysed), this will be used. The cited value is recommended by a subcommittee on biochemical verification of smoking status, convened by the SRNT.

4. Self reported smoking cessation for previous 24hr period at delivery validated by exhaled CO and saliva cotinine estimation.
5. Self reported, prolonged abstinence from smoking between quit date and 6 months after delivery.
6. Self reported smoking cessation for previous 7 day period at 6 months after delivery (point prevalence).
7. Self reported, prolonged abstinence from smoking between quit date and 2 years after delivery.
8. Self reported smoking cessation for previous 7 day period at 2 years after delivery (point prevalence).

b) Fetal loss and morbidity

1. Miscarriage (non-live birth prior to 24 weeks gestation) and stillbirth (non-live birth at 24 weeks gestation or later)
2. Neonatal death (i.e. from live birth to 28 days)
3. Post-neonatal death (29 days to 2 years)
4. Individualized birth weight Z score (i.e. birth weight adjust for gestational age, maternal height, maternal weight at booking and ethnic group).
5. Unadjusted birth weight and birth weight as Z-score
6. Apgar score
7. Cord blood pH
8. Gestational age at birth
9. Intraventricular haemorrhage
10. Neonatal enterocolitis
11. Neonatal convulsions
12. Congenital abnormality
13. Neonatal intensive care unit (NICU) admission
14. Infant ventilated > 24 hrs
15. Elective termination
16. Elective termination undertaken for fetal morbidity judged incompatible with fetal / infant survival

c) Maternal morbidity and mortality

1. Maternal mortality
2. Mode of delivery
3. Proteinuria
4. Hypertension in pregnancy

d) Early childhood outcomes

1. Behaviour and development at 2 years
2. Disability at 2 years
3. Respiratory symptoms at 2 years

e) Health economic data

1. Duration of maternal hospital admission for childbirth
2. Duration of any admission (of baby) to special care
3. Health status at 6 months (EQ5D)

1.11. Determination of Sample Size (from protocol)

Sample size: We need to recruit 525 women into each arm of the study. A trial with 500 women in each arm would detect an absolute difference of 9% in smoking cessation rates between the two groups immediately before childbirth with a two-sided significance level of 5% and a power of 93%. We anticipate that up to 5% of women will be lost to follow up and inflate our sample size (of 500) by a factor of 1.05 to allow for this. This size of study would allow us to detect smaller treatment effects with lower power. For example, we would have 80% power to detect an absolute difference in cessation rates of 7%.

Justification: A Cochrane review has shown that approximately 10% of women who are still smoking at the time of their first antenatal visit will stop smoking with usual care and a further 6% to 7% will stop as a result of a formal smoking cessation program using intensive behavioural counselling[15]. This means that in our control group (*placebo plus intensive behavioural counselling)* we can expect a smoking cessation rate of around 16%. The most recent Cochrane review of NRT, reports a treatment effect (odds ratio) for transdermal patches of 1.74 95%CI (1.57-1.93)[5]. Consequently, if we were to find NRT as effective in pregnancy as it is generally, we could expect a smoking cessation rate of approximately 25% in our treatment group (*NRT plus intensive behavioural counselling*).

The trial would have lower power to detect lesser treatment effects and table 1 illustrates that with our sample size of 1050 we have around 80% power to detect an OR of 1.6 or greater. We felt that this was adequate power.

If the quit rate (overall) is lower than 16% in the placebo group, the consequences to study power are shown in table 2. (Stopping rule for overall quit rate of 5%?)

Table 1 Sample size versus power to detect a range of odds ratios for the effectiveness of NRT

Odds ratio for effectiveness of NRT	Sample size required to detect this level of effectiveness at 80% power (alpha = 0.05)	Sample size required to detect this level of effectiveness at 90% power (alpha = 0.05)
1.3	3383	4491
1.4	2031	2688
1.5	1382	1824
1.6	1016	1341
1.7	790	1041

Table 2 Sample size versus power to detect an OR of 1.74 for various quit rates on placebo

Quit rate in the placebo group (average quit rate)	Power to detect OR of 1.74 at sample size 500 per group
16% (20%)	92%
14% (18%)	90%
12% (16%)	86%
10% (13%)	80%
8% (11%)	72%
6% (8%)	60%
4% (5%)	43%

1.12. Changes in the Conduct of the Study or Planned Analysis

None

1.13. Protocol amendments that have statistical implications should be described.

We have amended the primary outcome from that specified in the original protocol. This is because since commencement of the trial ambiguities in the way that the primary outcome was described in the trial protocol became apparent. Furthermore, recently proposed standards for the measurement of smoking outcomes in clinical trials came into existence (the Russell criteria) which we were able to incorporate into the revision. The revised primary outcome was approved by both the trial TSC and DMEC committees.

1.14. Any changes made to the planned analyses in the Study Protocol should be described in this section giving justification for the changes.

For outcome measures where the offspring is the unit of analysis (e.g. birth weight), non-independence of observations would need to be taken account of. Our strategy for dealing with this was not explained in great detail in the study protocol so has been expanded in this updated version of this analysis plan (see section 2.1.2 below).

2. ANALYSIS CONSIDERATIONS

2.1. Types of analysis

2.1.1. General methods

We will compare our binary primary outcome, prolonged and total abstinence from smoking at time of birth, between treatment arms.

Differences in smoking cessation rates between NRT and placebo arms will be reported as odds ratios and 95% confidence intervals obtained using logistic regression, adjusted for centre as a stratification variable. The primary analysis will adjust for no further variables for the reason that in a multivariate analysis results and therefore overall conclusions are liable to be sensitive to decisions concerning what variables to adjust for and how these are specified. Nevertheless, the adjustment for baseline covariates is often advised firstly to correct for any chance imbalances in important prognostic variables following randomisation and secondly, because adjusting for highly important prognostic variables in an RCT can improve the precision of treatment effect estimates even when the outcome measure is binary (Robinson LD, 1991). Statistical testing for baseline imbalances is not advised and instead key covariates should be selected prior to analysis based on the likely magnitude of the association with the outcome measure (European Agency for the Evaluation of Medicinal Products, 2003). Therefore as a sensitivity analysis for our data will report treatment effects adjusting for the following terms in addition to centre:

I. Salivary cotinine taken at baseline which we believe to be the most accurate indicator of levels of baseline nicotine addiction.

II. Partner's smoking status as a two level category (Partner smokes vs. partner does not smoke/ no partner reported)

III. Age of finishing full time education (in years)

Of the above variables, salivary cotinine and age of leaving full time education will be treated as linear terms in the absence of knowledge from either pregnant or non-pregnant populations indicating a relationship of a different nature between these variables and the probability of cessation. For a small number of women still in full time education at the time of enrolment the participant's current age will be used instead of age of finishing education.

2.1.2 Unit of analysis considerations

For the primary outcome measures relating to smoking cessation the women randomised will represent the unit of analysis. For some secondary outcomes (e.g. birth weight) however the offspring will be the unit of analysis instead. A small number of children will be born as multiple births (e.g. twins) and data for these cases will be clustered rather than independent. Therefore outcomes where the offspring is the unit of analysis will comprise singleton births only to allow for the fact that observations will be non-independent and that non-singleton births are likely to have very different birth outcomes in any case. In a subsidiary analysis multiple births will be included and clustering accounted for using an approach previously published (Gates S & Brocklehurst P, 2001). This adapts methodology previously created for use with cluster randomised RCTs, assuming that each women is regarded as the 'cluster' and her number of offspring the cluster size.

2.1.3 Effect modification and sub group analyses

We will look for effect modification by pre-treatment salivary cotinine levels and by age at leaving full time education as a proxy for socioeconomic status. Our multiple logistic regression models (both centre only and multivariate) will therefore be augmented with appropriate interaction terms. Initially, both salivary cotinine and age at leaving education will be fitted as continuous terms to maximise power when testing for an interaction. If evidence of an interaction is present (taken as a p-value of less than 0.05) then further subgroup analyses will dichotomise these variables (at the median) for ease of interpretation. The purpose of these models is to establish whether or not smoking cessation is constant across all levels of pre-treatment salivary cotinine in the NRT group, or reduces with increasing pre-treatment saliva cotinine, which could be indicative of inadequate replacement of nicotine. Similarly, we would be interested to know whether women with low or high levels of education could benefit preferentially from NRT use.

If there is evidence of interaction, we will perform subgroup analysis of the efficacy of NRT compared to placebo in subgroups defined by levels of pre-treatment plasma cotinine and by levels of socio economic status.

2.1.4 Timing of analyses

There will be two analyses. The first will be conducted upon data obtained around delivery. The second will be conducted at 2 years after delivery, using data obtained between delivery and this time point. Data collected for secondary outcomes will not be analysed until the trial has ended with respect to the primary outcome measure.

2.2. Analysis populations

The primary efficacy analysis will be performed using the Full Analysis (Intention to Treat) set i.e. we will include all those women randomised to NRT or placebo, presuming that those who do not provide data at follow-up are continuing to smoke. In the analysis of secondary and safety outcomes, we will again conduct our analysis on the intention to treat population; this will entail obtaining information (eg fetal and maternal morbidity etc) on all participants regardless of whether they return for treatment at one month, and whether they participate at delivery.

We will additionally analyse the safety data in the population of women who used at least one patch.

To prevent the 'double counting' of fetal demise outcomes, comparisons of adverse outcomes between groups will require the denominators listed below.

i) Low birth weight and preterm births: number randomised minus number of elective terminations + miscarriages
ii) Miscarriages & post randomisation fetal deaths: number randomised minus number of elective terminations for fetal morbidity judged incompatible with fetal infant / life + 'missed abortions' (a type of miscarriage) in which fetal death is documented to have occurred prior to randomisation]
iii) Perinatal deaths: number randomised minus number ofelective terminations (+ all miscarriages (inc. missed abortions) + stillbirths

2.3. Protocol Deviations

Failure to use the allocated treatment (whether NRT or placebo) will not constitute a protocol deviation. Possible protocol deviations include:

1) Women who choose to withdraw from the trial, and choose not to consent for the use of their data for primary or secondary outcomes.

2) Women provided with the 'wrong' treatment either at randomisation or 1 month follow-up.

We will give a by-patient listing of protocol deviations.

2.4. Derived variables: derivation of the primary outcome is described in section 1.10. Derivation of birth weight z score?
Low birth weight – births of <2500g

Preterm birth – births of < 37 weeks gestation

Post-randomisation fetal death (a composite measure of all fetal deaths after randomisation which could reflect death due to trial interventions) – ***defined as*** – all [miscarriages + stillbirths + neonatal deaths + elective terminations conducted for fetal abnormalities judged inconsistent with fetal / infant life] NB: 'missed abortion' miscarriages with documented fetal death prior to trial enrolment are not included

Perinatal deaths (a composite measure of all infant deaths following live births) – ***defined as*** – all [stillbirths + neonatal deaths]

2.5. Missing data conventions:

Women with missing smoking data at any point will be presumed to be continuing to smoke.

2.6. Treatment Compliance and per protocol analysis

We will compare compliance in terms of patches used at 1 month (number using them for a certain period of time?) and number of intensive cessation sessions attended between NRT and placebo groups.

Initially we will tabulate compliance data and use of non-trial NRT by treatment group. This will be done separately for data obtained at one month and prior to delivery.

In a subsequent analysis, we will create a variable describing the actual amount of NRT taken throughout the trial period for use in a per protocol analysis. This variable will depict the percentage of trial days where NRT was received (0 for the placebo group who reported taking no non-trial NRT and 100 for NRT group who reported full compliance with treatment), incorporating both the level of compliance with trial patches (NRT group only) and the use of non-trial NRT (placebo group only). Logistic regression analysis adjusting for centre will be carried out, with the above smoking cessation measures as outcomes. Women who do not provide primary outcome data will be excluded from this analysis. Only compliance data taken at delivery will be considered.

2.7. How to pool centres: We do not anticipate differences in treatment efficacy between centres, and we will check for study centre effects by fitting a centre-by-treatment interaction. If there is significant heterogeneity, we will combine results from different centres using a random effects model.

2.8. Documentation and other considerations

2.9. Software used

We will use SAS version 9.1.3 for all analysis.

2.10. Levels of significance

All tests will be two-tailed, using a P value of < 0.05 to indicate statistical significance, and 95% confidence intervals will be calculated.

2.11. Format of electronic files for archiving

Excel and SAS

3. ANALYSIS OF PARTICIPANT CHARACTERISTICS

3.1. Describe methods used to summarise data.

Continuous data that are approximately normally distributed will be summarised in terms of the mean, standard deviation, median, minimum, maximum and number of observations. Skewed data will be presented in terms of the maximum, upper quartile, median, lower quartile, minimum and number of observations. Categorical data will be summarised in terms of frequency counts and percentages.

3.2. Disposition

3.2.1. We will summarise the number of patients screened for entry, excluded prior to randomisation by major reason and overall, the number of patients randomised and the number entering and completing each phase of the study by treatment group and overall. We will use CONSORT flow chart for this.

3.3. Baseline

3.3.1. We will summarise demographic variables (e.g. age, daily number of cigarettes prior to delivery and currently, gestational age at randomisation, exhaled CO, ethnic group, education, parity, etc) by treatment group and overall.

4. ANALYSIS OF EFFICACY

4.1. Specify how mis-randomised patients will be handled: (in general, patients will be analysed as randomised).

4.2. Description of response variables

4.2.1. Primary

- Self-reported prolonged and total abstinence
  - Binary

4.2.2. Secondary

- Self reported prolonged abstinence at 1 month, prolonged and total abstinence without CO validation and prolonged and total abstinence with CO validation at 1 month and birth
  - All binary
- Miscarriage, still birth, neonatal death and post natal death
  - All binary
- Unadjusted birth weight and birth weight by z-score
  - normally distributed
- Maternal mortality
  - Binary
- Mode of delivery
  - Categorical
- Apgar score (<7)
  - Binary
- Cord blood pH (<7)
  - Binary
- Gestational age at birth (weeks)
  - continuous
- Intraventricular haemorrhage
  - Binary
- Neonatal enterocolitis
  - Binary
- Neonatal convulsions

Binary
- Neonatal intensive care unit admission
  - Binary
- Infant ventilated (>24 hours)
  - Binary

4.2.3 Post trial follow-up
- Behaviour and development at 2 years
  - ?
- Disability / respiratory symptoms at 2 years
  - All Binary
- Duration of maternal hospital admission for childbirth
  - Continuous, skewed?

- Duration of admission of baby to special care
  - Continuous, skewed?
- Health status at 6 months EQ5D
  - Continuous, not normally distributed

4.3. Analysis of response Variables

4.3.1. Primary
- Null hypothesis: no difference in prolonged abstinence to childbirth
  - Size of the difference between treatments will be expressed as an odds ratio from logistic regression.

4.3.2. Secondary
- State null hypothesis, relate this to statement of objectives and type of data (continuous or binary). Describe presentation of results:
  - size of the difference between treatments (or relevant parameter, eg hazard ratio)
  - the associated confidence interval and the results of the hypothesis testing.

4.4. Exploratory/Other analyses

5. ANALYSIS OF SAFETY & TOLERABILITY

5.1. All patients who are randomised (i.e. intention to treat) will be included in the safety analysis.

5.2. Extent of exposure

5.2.1. We will define this as the number of patches reportedly used by the participant up to 1 month, and up to delivery. This will be summarised by treatment group using median, UQ, LQ.

5.3. Adverse events (AE).

To minimise the likelihood of women or infants being harmed by unexpected effect(s) of nicotine that could not predicted from previous research, the Data Monitoring & Ethics Committee will have access to birth outcome data. These data will be available for the DMEC to analyse as is considered appropriate to investigate whether or not significant or clinically-important differences arise between study groups (e.g. in birth weight).

a) The following will be considered ***adverse events*** **(AEs)**:

Withdrawal from patch treatment due to i) skin reaction or ii) other symptom(s) which are potentially caused by NRT (listed in section 4.10 BNF)

AEs will be reported in an annual safety report to the MHRA, REC and Sponsor.

b) The following will be considered ***Serious Adverse Events*** (SAEs):

Baby: miscarriage, still birth, neonatal and post-neonatal death
Maternal: maternal death, other events requiring hospital admission apart from those related to the underlying pregnancy or a pregnancy related condition* (see footnote for excluded hospital admissions)[3]

Any other serious unexpected event.

All SAEs will be reported on a standard form and assessed by Professor Jim Thornton or a named deputy to determine whether or not they should be considered as being **Suspected Unexpected Serious Adverse Reactions (SUSARs)** which are potentially-related to trial treatments.

Life threatening or fatal SUSARs will be reported to the MHRA and REC within 7 days (follow up report within 15 days) and also to relevant NHS trust R&D office according to local policies.

Non life threatening SUSARs will be reported to the MHRA and REC within 15 days and also to R&D offices, as appropriate.

SUSARs will also be reported to the DMEC chair along with the treatment allocation group of the trial participant and a cumulative count of SAE and SUSAR frequency in each trial arm.

SAEs which are not considered SUSARs will be reported in an unblinded manner to each DMEC meeting and in the annual report to MHRA, REC and Sponsor with AEs.

5.3.1. Adverse event summaries will be based upon the number of patients reporting adverse events and not the number of events reported

5.3.1.1. Any treatment emergent AE by treatment group, body system, and preferred term

5.3.1.2. Most common treatment emergent AE by treatment group, body system, and preferred term

6. LIST OF PROPOSED SUMMARY TABLES

[3] The following hospital admissions are **not** SAEs but will be treated as AEs: delivery (not AE or SAE), recognised pregnancy or postnatal complications, including pre-term delivery before 32 weeks, low birth weight (< 2,500g), birth injury, infection, thrombosis, haemorrhage, hypertensive disease, instrumental delivery (not AE or SAE), caesarean section (not AE or SAE), and antenatal admissions for pregnancy related diseases such as false labour, infection, thrombosis, haemorrhage, hypertensive disease, suspected or confirmed fetal compromise, vaginal bleeding fetal congenital abnormalities, and infant hospital admissions. Incidental hospital admissions for minor, gastrointestinal diseases, respiratory, cardiac, renal skin, psychiatric and neurological problems.

6.2. We will produce a CONSORT flow diagram showing enrollment, dispositions, exclusions, evaluable participants

6.3. Participant Characteristics and Background summary table

Characteristic	Placebo	NRT	Total
Centre			
Nottingham.. N (%)			
Age (Mean, SD)			
(range)			
No of cigs daily before preg (Median, IQR, Range)			
No of cigs daily at randomisation (Median, IQR, Range)			
Gestational age (mean, SD)			
Ethnic group			
White N (%)			
Age left full time education (Mean, SD)			
Parity			
0 N (%)			
Etc			
Time to first cig			
< 30 minutes N (%)			
Etc			
Partners smoking status			
Smoker N (%)			
Height (Mean, SD)			
Weight (Mean, SD)			

6.5. Smoking efficacy summaries

	Placebo	NRT	OR (95% CI)*	P value	OR (95% CI)^	P value
Total	N	N				
Self reported prolonged abstinence to delivery						
No	N (%)	N (%)	1			
Yes	N (%)	N (%)	OR			
Self reported prolonged abstinence to 1 month						
No	N (%)	N (%)	1			
Yes	N (%)	N (%)	OR			
Self reported prolonged abstinence to 6 months						
No	N (%)	N (%)	1			
Yes	N (%)	N (%)	OR			
Self reported 7 day cessation at 6 months						
No	N (%)	N (%)	1			
Yes	N (%)	N (%)	OR			
Self reported prolonged abstinence to 2 years	NB: Follow up to this point is not yet complete					
No	N (%)	N (%)	1			
Yes	N (%)	N (%)	OR			
Self reported 7 day cessation at 2 years	NB: Follow up to this point is not yet complete					
No	N (%)	N (%)	1			
Yes	N (%)	N (%)	OR			

* adjusted for centre only (as a stratification factor)
^ adjusted for centre, salivary cotinine at baseline, partner's smoking status and age at leaving full time education.

6.6. Secondary efficacy summaries (fetal loss and morbidity)

	Placebo	NRT	OR (95% CI)*	P value	OR (95% CI)^	P value
Total	N	N				
Miscarriage and still birth						
No	N (%)	N (%)	1			
Yes	N (%)	N (%)	OR			
Neo natal death						
No	N (%)	N (%)	1			
Yes	N (%)	N (%)	OR			
Post natal death						
No	N (%)	N (%)	1			
Yes	N (%)	N (%)	OR			
Birth weight (unadjusted)	Mean (SD)	Mean (SD)	Mean difference (95% CI)			
Birth weight (z-score)	Mean (SD)	Mean (SD)	Mean difference (95% CI)			

6.7. Secondary efficacy summaries (maternal morbidity

	Placebo	NRT	OR (95% CI)*	P value	OR (95% CI)^	P value
Total	N	N				
Mode of delivery						
Normal	N (%)	N (%)	1			
Breech	N (%)	N (%)	OR			
Caesarian						

6.8. Adverse events summaries

7. LIST OF PROPOSED APPENDICES (if relevant)

7.1. Trial process

7.2. Adverse events

7.2.1. All adverse events (including non-treatment-emergent events) by patient, centre, age, sex, race, adverse event (body system, preferred

term, reported term), date of onset, date of resolution, duration, severity, seriousness, action taken, outcome and causality.
7.2.2. by-patient listing of all serious adverse events
7.2.3. by-patient listing of all adverse events leading to withdrawal
7.2.4. by-patient listing of all deaths that occurred during the study

7.3. Clinical Laboratory Evaluation
7.3.1. all laboratory data should be provided by treatment group, with abnormal values highlighted, and including centre, patient identifier, age, sex, race, weight and visit. Laboratory reference ranges should also be listed.
7.3.2. Vital Signs, Physical Findings and Other Observations Related to Safety - as for laboratory parameters

SECTION TWO

Measurement of SNAP trial outcomes at two years: combining measures

1. Overview

This section lists the outcomes collected up to the two-year time point. It lists outcomes as per trial protocol and the methods used to measure these.

2. Outcomes included in approved version of trial protocol

Early childhood outcomes
1. Behaviour and development (B&D) at 2 years
2. Disability (Dis) at 2 years
3. Respiratory symptoms (RS) at 2 years

For analysis purposes, items 1 and 2 above will be combined; although these are classed as secondary outcomes for the main trial, for the follow-up, the primary outcome will be 'survival with no impairment'. (See section 'Collation of data for outcomes' below for details of how these will be specified.)

Smoking outcomes
1. Self reported, prolonged abstinence from smoking between quit date and 2 years after delivery.
2. Self reported smoking cessation for previous 7 day period at 2 years after delivery (point prevalence).

3. Methods of outcome measurement

A postal questionnaire (PQ2), completed by women enrolled in the study was the primary method of outcome measurement; non-respondents could also complete this by telephone. The questionnaire included the 30-item Ages and Stages (ASQ-3) questionnaire (see below) which detects developmental delay. When women could not be contacted for follow up, a shorter instrument (excluding the 30 ASQ-3 items) was completed by either participants' GPs or health visitors (HPQ).
Table 3 shows: the distribution of items between the *participant* and *health professional* questionnaires, and how these map on to outcomes above.

ASQ-3 (Ages and Stages) questionnaire
The ASQ-3 24 month questionnaire, used at the 2-year follow up point, is valid for use from 23 months 0 days through to 25 months 15 days and, if used at 24 months. Documentation provided by the publishers of ASQ-3 states that adjustment of infant's age to allow for prematurity is no longer required once they reach 24 months of age, therefore as the questionnaire was sent shortly before the child's 2nd birthday no adjustment to their age was made. Items 1-36 on the participant questionnaire were all taken from ASQ-3.

The ***participant*** questionnaire contained all 30 ASQ-3 items on child development in five domains: communication, gross motor, fine motor, problem solving and personal-social (items 1-5, Table 3). For each domain, a score reflecting infants' development can be obtained (see below). Table 4 shows cut points and 'borderline'

scores for each domain; in clinical practice and some studies (not SNAP), infants who fail one or more domains would be further assessed for developmental delay. There is no global score and infants whose score is below the cut-off in one or more domains would be considered to have not passed the ASQ.

Both the participant and health professional questionnaires included additional items, of which questions 31-36 on the participant questionnaire were taken from ASQ-3, investigating concerns about infant health and development but which do not contribute to domain scores (items 6-12, Table 3). The response could be 'yes' or 'no', with free text for the participant or health professional to describe their concerns.

In addition to the questionnaires at 2 years, parental questionnaires were sent 6 months and 1 year after delivery. Responses from these questionnaires do not form part of the main follow up outcomes, but will be used to help validate smoking and respiratory outcomes which are asked at all three time points. Responses will be summarised in tables to be found at the end of this document. The 6 month questionnaire also included questions that will be used in a separate Health Economics analysis.

Validation of questionnaire responses
For any participants where both PQ2 and HPQ have been completed, the questionnaires will be compared to ensure there are no major differences. Only responses from PQ2 will be used in the analysis unless, on examination by an expert member of the trial team, the HPQ responses are felt to be more reliable (e.g. where there are unexplained inconsistencies in PQ2 responses).

Questions with free text responses will be examined by members of the trial team, and referred to an expert where necessary, and then a judgement made as to whether these are thought to indicate definite or suspected problems with the child's development or health.

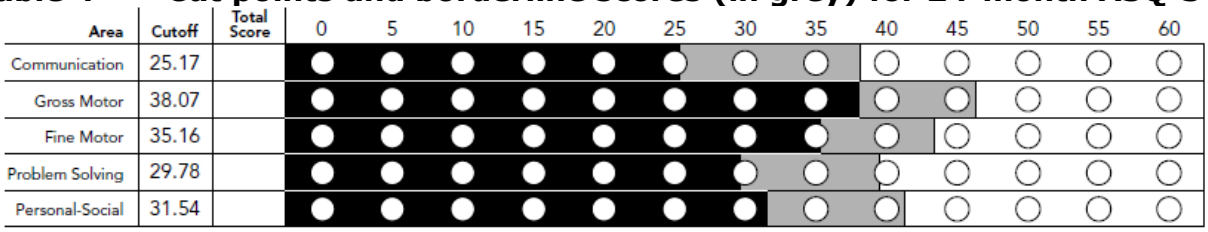
Table 3 Items on participant- and health professional-completed questionnaires distributed at 2 years

Item description	Participant	Health professional	Outcome listed in approved version of protocol
1 Fine motor skills[a]	Y	N	B&D[c]
2 Gross motor skills[a]	Y	N	B&D[c]
3 Communication[a]	Y	N	B&D[c]
4 Problem solving[a]	Y	N	B&D[c]
5 Personal-social[a]	Y	N	B&D[c]
6 Hearing problems[b]	Y	Y	B&D
7 Speech problems[b]	Y	Y	B&D
8 Neuro-motor problems[b]	Y	Y	B&D
9 Vision problems[b]	Y	Y	B&D
10 Behaviour problems[b]	Y	Y	-
11 Feeding problems[b]	Y	Y	-
12 Chest or breathing problems[b]	Y	Y	RS
13 Rating of disability	N	Y	Dis

14 If developmental delay, severity	Y[d]	Y	B&D
15 No. of & reasons for hospital admissions	Y	N	Not included in approved version of protocol
16 Respiratory symptoms	Y	Y	RS
17 Asthma diagnosis	Y	N	RS
18 Medications	Y	N	RS
19. Prolonged abstinence from smoking since childbirth?	Y	N	Smoking
20. Smoked in last week?	Y	N	Smoking
21. Smoked any at all in last 2 years?	Y	N	Smoking

[a] complete domains from ASQ-3 questionnaire
[b] Either participant perception of problem or problem documented in medical records
[c] 30 item ASQ-3 domains which measure developmental delay
[d] Measured by outcome domains labelled 'c'

Table 4 Cut points and borderline scores (in grey) for 24-month ASQ-3



Area	Mean score	SD	Mean - 1SD (borderline score)	Mean - 2SD (cut-off point)
Communication	51.23	13.03	38.20	25.17
Gross Motor	54.73	8.33	46.40	38.07
Fine Motor	51.70	8.27	43.43	35.16
Problem Solving	49.40	9.31	39.59	29.78
Personal-Social	51.14	9.80	41.34	31.54

(From Ages & Stages Questionnaires ®, Third Edition (ASQ-3), Squires & Bricker © 2009 Paul H Brookes Publishing Co)

Calculating domain scores for ASQ-3: NB: these items only appear on *participant* (PQ2) questionnaire. For each domain item the responses were 'yes', 'sometimes' or 'not yet' and these are scored 10, 5 and 0 respectively. If there are any unanswered items in a domain then the score can be adjusted as long as *no more than two items have been omitted*. This is calculated by dividing the total score for the domain by the number of questions answered in that domain, and then adding this to the total score once if one question was missed and twice if two were missed. *A domain is not scored if more than two questions are omitted*.

Collation of data for outcomes

NB See previous section for interpretation and use of free text responses from questions for the outcomes listed below.

Behaviour and development, and disability (i.e., combined 'early childhood outcomes' items 1 & 2):

i) ***ASQ-3 scores all normal***. For this study, infants will be classified as having no impairment if scores are above the borderline (i.e. within white section of Table 4 above) for **all** domains of the ASQ-3.

ii) ***Survival with no impairment*** *(Primary outcome for follow up)*. For this study, infants will be classified as having no impairment if scores are above the borderline (i.e. within white section of Table 4 above) for **all** domains of the ASQ-3, ***and*** no problems are reported for items 31-35 in the participant questionnaire. Where there is no participant questionnaire, and a health professional questionnaire has been completed then HPQ questions 1-4, 9 and 10 should all indicate an absence of any problems.

iii) ***Definite developmental impairment.*** For this study, infants will be classified as having developmental impairment if scores are at or below the cut-off point (i.e. within one or more black areas of Table 4 above) in one or more domains of the ASQ-3 ***or***, where no participant questionnaire has been completed, the health professional questionnaire indicates severe problems for any of the questions 1-4 and/or severe disability (Q9) and/or severe development delay (Q10).

iv) ***Suspected developmental impairment***. Items will be scored as indicated above. For this study, infants will be classified as having suspected developmental impairment if **all** ASQ-3 scores are above the cut-off point (i.e. not in the black area of Table 4 above) **but,** scores in one or more ASQ-3 domains are at or below the borderline levels (i.e. within one or more grey areas of Table 4 above). In addition, this classification will be used if there are any *participant* or *health professional* perceptions of developmental impairment. This will include any infant that has a 'Yes' response to one or more of the free text questions in either the participant (Q31-37) or health professional (Q1-6) questionnaires and that, after examination of the responses by the research team, is judged to have mild, moderate or possible impairment, **and/or** any child that is classed as having mild or moderate disability (Q9) and/or mild or moderate development delay (Q10) on the health professional questionnaire. Any infant who has any responses that would class them as having definite developmental impairment would exclude them from being placed in this category.

Respiratory symptoms:

i) ***Respiratory symptoms.*** Infants will be judged to have a respiratory symptoms if, at 2 years, any of the questions 38-42 of the participant questionnaire and/or question 7 of the health professional questionnaire indicate that the infant has respiratory problems. These questions include hospital admissions for respiratory problems, problems with chest or breathing (yes/no, free text), wheeze or whistling in chest (yes/no, frequency), doctor diagnosed asthma (yes/no), asthma medications taken (yes/no, inhaler description free text). Health professional questionnaire: does child have problems with their chest or breathing (yes/no, free text).

Outcomes below are only recorded on participant questionnaire

Smoking behaviour:
Responses to the participant questionnaire Q45-47 will be used in a manner consistent with Russell Criteria and reporting of smoking outcomes at the primary outcome point.

i) *Positive outcome for 'Self-reported, prolonged abstinence from smoking between quit date and 2 years after delivery':*
The participant must have met the criteria for prolonged abstinence at delivery (ie positive primary outcome), PLUS ONE OF THE FOLLOWING RESPONSES
'smoked since two year old was born' = 'No'
OR
'how often have you smoked' = '5 times or less'

If any participant questionnaires have been completed at 6 and/or 12 months, these should all have the same responses as above for a positive outcome.

ii) *Smoked in last week (self-reported smoking cessation for previous 7-day period at 2 years after delivery – point prevalence)*:
'smoked since two year old was born' = 'No'
OR 'smoked in last week' = 'No'

iii) *Smoked in last 2 years (self-reported prolonged abstinence from smoking between delivery and 2 year questionnaire) (NB: this outcome was not listed in the protocol)*:
'smoked since two year old was born' = 'No'
OR
'how often have you smoked' = '5 times or less'

If any participant questionnaires have been completed at 6 and/or 12 months, these should all have the same responses as above for a positive outcome.

4. Statistical analysis

Baseline characteristics will be compared between those participants and infants who did and did not have outcomes ascertained at 2 years after delivery.

<u>*Non-smoking outcomes*</u>: Primary analysis of non-smoking behaviour outcomes will be on an intention-to-treat basis with participants analysed in the treatment groups to which they were randomised. Participants with no live birth (i.e. miscarriage, stillbirth and elective termination) will be excluded from the intention-to-treat analysis, but all others, including those who withdrew consent before delivery and those with no delivery details, will be included.

Development and health outcomes will all be analysed as binary indicators of presence or absence of disability or disease as described above. Odds ratios will be obtained using logistic regression adjusted for centre as the stratification variable. We will use multiple imputation to deal with missing values as described below, under an assumption of missing at random (MAR) and this assumption will be tested as described below. In the case of very small numbers in cells we will analyse using Fishers exact test.

<u>*Smoking behaviour:*</u> The denominator for smoking behaviour outcomes will be all women randomised at the outset of the trial and those lost to follow up (for any

reason, including fetal death) or who withdraw will included in the denominator and, if their smoking status is not known, will be counted as smokers. Differences in smoking cessation outcomes between NRT and placebo arms will be reported as odds ratios and 95% confidence intervals obtained using logistic regression, adjusted for centre as a stratification variable. Those with missing data at follow-up will be presumed to be continuing smokers – further details of sensitivity analysis of the impact of missing data is described below. The primary analysis will adjust for no further variables for the reason that in a multivariate analysis results and therefore overall conclusions are liable to be sensitive to decisions concerning what variables to adjust for and how these are specified. Nevertheless, the adjustment for baseline covariates is often advised firstly to correct for any chance imbalances in important prognostic variables following randomisation and secondly, because adjusting for highly important prognostic variables in an RCT can improve the precision of treatment effect estimates even when the outcome measure is binary (Robinson LD, 1991). Statistical testing for baseline imbalances is not advised and instead key covariates should be selected prior to analysis based on the likely magnitude of the association with the outcome measure (European Agency for the Evaluation of Medicinal Products, 2003). Therefore as a sensitivity analysis for our data we will report treatment effects adjusting for the following terms in addition to centre:

I. Salivary cotinine taken at baseline which we believe to be the most accurate indicator of levels of baseline nicotine addiction.

II. Partner's smoking status as a two level category (Partner smokes vs. partner does not smoke/ no partner reported)

III. Age of finishing full time education (in years)

Of the above variables, salivary cotinine and age of leaving full time education will be treated as linear terms in the absence of knowledge from either pregnant or non-pregnant populations indicating a relationship of a different nature between these variables and the probability of cessation. For a small number of women still in full time education at the time of enrolment the participant's current age will be used instead of age of finishing education.

4.1 Handling of missing data and sensitivity analysis of assumptions

We will explore the patterns of missingness in the data by crosstabulating baseline characteristics including maternal age, centre, education, ethnic group, parity and heaviness of addiction, and birth outcome variables, including birth weight, gestation, and smoking at delivery in those with and without missing outcomes. We will also explore the association between smoking status at delivery and smoking outcomes at 6 months, 1 and 2 years.

Our primary analysis of smoking outcomes assumes that data are missing not at random (MNAR) i.e. that those missing at follow-up are likely to be smokers. We will test that assumption by using the data at delivery (where data is almost complete) to explore whether smokers were less likely to have data at subsequent follow-ups. We will also conduct a sensitivity analysis to explore alternative associations (odds ratios) for the relationship between smoking status and missingness (using multiple imputation by the method described by Hedeker).

Our analysis of developmental and health outcomes will involve both a complete case analysis and an analysis using multiple imputation by chained equations which assumes that data are missing at random (MAR) i.e. that missingness is associated with baseline characteristics but not with the outcome itself. In our multiple imputation, we will include all of the variables listed above and we will use at least 20 imputations (ensuring that the number of imputations exceeds the % of incomplete cases in the dataset). We will test the MAR assumption by again exploring alternative associations (odds ratios) for the relationship between each outcome and missingness. We will also conduct sensitivity analyses comparing the results of analysis using outcomes based on parental responses only and those based on the combination of parental and health professional responses.

4.2 Unit of analysis considerations

Apart from the exclusions listed above, for the smoking cessation outcomes the women randomised will represent the unit of analysis. For some other outcomes (e.g. child development) the offspring will be the unit of analysis instead. A small number of children will be born as multiple births (e.g. twins) and data for these cases will be clustered rather than independent. Therefore, outcomes where the offspring is the unit of analysis will comprise singleton births only to allow for the fact that observations will be non-independent and that non-singleton births may have very different outcomes in any case. In a subsidiary analysis multiple births will be included and clustering accounted for using an approach previously published (Gates S & Brocklehurst P, 2001). This adapts methodology previously created for use with cluster randomised RCTs, assuming that each women is regarded as the 'cluster' and her number of offspring the cluster size.

4.3 Effect modification and sub group analyses

There are no planned interaction or subgroup analyses.

4.4 Secondary analyses (not necessarily for inclusion in principal secondary outcomes paper)

In secondary analyses, we will explore the association between smoking outcomes at 1 month after quit date and at delivery and developmental and health outcomes at 2 years. In this analysis, we will adjust for treatment group and centre as a priori confounders, and we will explore the potential confounding effects of maternal age, education, ethnicity, parity, and duration of breast feeding, including those which lead to a 10% or greater change in the effect of smoking status when included in the model.

Adherence: We will also explore the dose response effect between amount of NRT used. Adherence will be quantified in three ways, i) as a continuous measure (i.e. reported number of patches used, taken from data collected at delivery), ii) as a categorical measure (i.e. with 'high' and 'low' adherence categories created based on prior literature and / or the distribution of reported patch use within participants and iii) by comparing those women who accepted a second batch of nicotine patches with those who did not. We will investigate the relationship between adherence and principal developmental outcomes: no impairment, suspected impairment and definite impairment.

5. Summary information on data for tables: SNAP two year follow-up period

5.1 Completeness of follow up at different time points (Table 1)

(NB This will be presented in CONSORT diagram format rather than as a table, with numbers and reasons for withdrawal and/or exclusion from analysis given at each stage (delivery, 6 months, 1 year, 2 years)

	NRT N (%)	Placebo N (%)
Randomised and included in ITT analysis at delivery	521	529
Completed to primary outcome		
Number providing birth outcomes at delivery		
Participant questionnaires sent at 6 months		
Participant questionnaires returned at 6 months: i) postal return & ii) telephone completed)		
Participant questionnaires sent at 1 year		
Participant questionnaires returned at 1 year: i) postal return & ii) telephone completed)		
Participant questionnaires sent at 2 years		
Participant questionnaires returned at 2 years: i) postal return & ii) telephone completed)		
Health professional questionnaires sent at 2 years		
Health professional questionnaires completed at 2 years		
Outcome data provided at 2 years (**either** PQ2 ***or*** HPQ completed)		
Outcome data provided at 2 years (**both** PQ2 ***and*** HPQ completed)		

5.2 Baseline comparison of groups providing data at 2 years (Table 2)

	NRT (n=)		**Placebo (n=)**	
	n, mean or median	**%, SD or IQR**	**n, mean or median**	**%, SD or IQR**
Characteristics at enrolment/randomisation				
Age – yr (Mean, SD)				
No of cigs daily before pregnancy – N (Median, IQR)				
Cigarettes smoked daily at enrolment – N (Median, IQR)				
Gestational age at randomisation - weeks (mean, SD)				
Race or ethnic group – N (%) White British Other				
Age left full time education – yr (Mean, SD)				
Index of multiple deprivation				
Parity - N (%) 0-1, 2-3, ≥4				
Salivary cotinine level at enrolment – ng/ml (median, IQR)				
Time from awakening to first cigarette N (%) 0-15min, 16-60min, >60min				

Women with partner who smokes – N/Total N with partner (%)				
Height - cm (Mean, SD)				
Weight - kg (Mean, SD)				
Previous preterm birth – N (%)				
Length of first behavioural support session – N (%) < 30min, 31-60min, >60min				
Use of NRT earlier in pregnancy – N (%)				
Characteristics at delivery				
Met primary outcome criteria (ie abstinence from quit date to delivery, with salivary and/or CO validation) – N (%)				
Birth weight, unadjusted – kg (mean, sd)				
Gestational age – wk (mean, sd)				
Preterm birth – no/total no (%)				
Low birth weight (<2.5kg) no/total no (%)				
NICU admission – no/total no (%)				
Apgar score at 5 mins <7 – no/total no (%)				
Congenital abnormalities – no/total no (%)				
Infant on ventilator >24hr - no/total no (%)				

Assisted vaginal delivery - no/total no (%)				
Delivery by caesarean section - no/total no (%)				

5.3 Comparison of participant and infant characteristics between children who did and did not have outcomes at 2 years (Web Table)

	Followed up: PQ2 (n=)		**Followed up: HPQ (n=)**		**Not followed up (n=)**	
	n, mean or median	**%, SD or IQR**	**n, mean or median**	**%, SD or IQR**	**n, mean or median**	**%, SD or IQR**
Characteristics at enrolment/randomisation						
Age – yr (Mean, SD)						
No of cigs daily before pregnancy – N (Median, IQR)						
Cigarettes smoked daily at enrolment – N (Median, IQR)						
Gestational age at randomisation - weeks (mean, SD)						
Race or ethnic group – N (%) White British Other						
Age left full time education – yr (Mean, SD)						
Index of multiple deprivation						
Parity - N (%) 0-1, 2-3, ≥4						
Salivary cotinine level at enrolment – ng/ml (median, IQR)						

Time from awakening to first cigarette N (%) 0-15min, 16-60min, >60min						
Women with partner who smokes – N/Total N with partner (%)						
Height - cm (Mean, SD)						
Weight - kg (Mean, SD)						
Previous preterm birth – N (%)						
Length of first behavioural support session – N (%) < 30min, 31-60min, >60min						
Use of NRT earlier in pregnancy – N (%)						
Characteristics at delivery						
Met primary outcome criteria (ie abstinence from quit date to delivery, with salivary and/or CO validation) – N (%)						
Birth weight, unadjusted – kg (mean, sd)						
Gestational age – wk (mean, sd)						
Preterm birth – no/total no (%)						
Low birth weight (<2.5kg) no/total no (%)						
NICU admission – no/total no (%)						
Apgar score at 5 mins <7 – no/total no (%)						

Congenital abnormalities – no/total no (%)						
Infant on ventilator >24hr - no/total no (%)						
Assisted vaginal delivery - no/total no (%)						
Delivery by caesarean section - no/total no (%)						

5.4 Behaviour and development, and respiratory symptoms at 2 years (Table 3 and Table 4)

	NRT			Placebo		
	PQ2	HPQ	PQ2 & HPQ	PQ2	HPQ	PQ2 & HPQ
	N (%)					
ASQ-3 domain scores all normal [1]						
Survival with no impairment [2]						
Definite developmental impairment [3]						
Suspected development impairment [4]						
Infant death [5]						
Respiratory symptoms [6]						

PLUS
(this information could be in the text rather than a table)

	Complete case analysis		Multiple imputation	
	PQ2	PQ2 & HPQ	PQ2	PQ2 & HPQ
	OR (95% CI)		OR (95% CI)	
Survival with no impairment [2]				

[1] Score above borderline score in ASQ-3 (white section) for all domains, and no problems reported in additional sections of ASQ-3 (ie any hearing, talking, understanding, neuromotor, vision, behaviour, feeding problems)

[2] ie all ASQ-3 domain scores normal and no other development problems reported in PQ2 and/or HPQ

[3] ie ASQ-3 score equal to or below cut-off (black section) in ≥1 domain, HPQ indicates severe disability and/or severe developmental delay

[4] ie ASQ-3 borderline score (grey section) in ≥ 1 domain, but no scores equal to or below cut-off (black section), and/or judged to have mild/moderate or possible impairment, disability or development delay from the additional questions on the PQ2 and/or HPQ including problems with hearing, speech, neuro-motor or vision.

[5] infant death between birth and 2 year follow-up – information obtained from NHS records

[6] any report of respiratory symptoms, asthma diagnosis, asthma medications at 2 year follow-up

Table 4 for Appendix/Web

	NRT	Placebo	OR (95% CI)*	P value
ASQ-3 domain scores [1]:				
Number of participants providing data for at least one domain				
Fine motor skills (mean, sd)				
Gross motor skills				
Communication				
Problem solving				
Personal-social				
Score below cut-off in ≥2 domains – N (%)				
ASQ-3 supplementary questions (N (%) reporting problem):				
Hearing				
Speech (talking)				
Speech (understanding)				
Neuromotor (walking, running, climbing)				
Vision				
Behaviour				
Feeding				
HPQ				

Number providing data				
N (%) reporting problems with:				
Hearing				
Speech				
Neuromotor				
Vision				
Behaviour				
Feeding				
Current health status N (%): None Mild disability Moderate disability Severe disability				
Concerns about development – N (%)				
(if Yes) Formal development assessment carried out				
Overall development delay N (%): None Mild Moderate Severe				

[1] Footnote to include cut off scores for each domain

5.5 Smoking and respiratory outcomes post-delivery (Table 5)

	NRT	Placebo	OR (95% CI) [1]	P value	OR (95% CI) [2]	P value
Smoking outcomes						
6 months after delivery:						
Number providing data						
Self-reported prolonged abstinence since delivery N (%) [3]						
Self-reported 7 day cessation N (%)						
Prolonged abstinence from smoking between quit date and 6 months after delivery N(%) [4]						
1 year after delivery:						
Number providing data						
Self-reported prolonged abstinence N (%) [3,5]						
Self-reported 7 day cessation N (%) [5]						
Prolonged abstinence from smoking between quit date and 1 year after delivery N(%) [4]						
2 years after delivery:						
Number providing data						
Self-reported prolonged abstinence N (%) [3]						

Self-reported 7 day cessation N (%)						
Prolonged abstinence from smoking between quit date and 2 years after delivery N(%) [4]						
Respiratory outcomes						
1 year after delivery [6]						
Number providing data						
Wheeze or whistling N (%)						
If yes, how many attacks in last year? N (%) 0 1-3 4-12 >12						
How often has sleep been disturbed due to wheezing? N (%) Never <1 night/week ≥1 night/week						
Doctor diagnosed asthma N (%) [7]						
Dry cough at night N (%)						
Seen by paediatrician or chest specialist about chest or breathing problems N (%)						
2 years after delivery (maternal questionnaire)						
Number providing data						
Wheeze or whistling N (%)						

If yes, how frequently? N (%) Every day Every week Once/month or less						
Doctor diagnosed asthma N (%) [7]						
Medicines taken for cough/wheeze/chest problems N (%)						
2 years after delivery (HPQ)						
Number providing data						
Problems with chest or breathing N (%)						

[1] adjusted for centre only (as a stratification factor)

[2] adjusted for centre, salivary cotinine at baseline, partner's smoking status and age at leaving full time education

[3] Self-reported prolonged abstinence in the table = smoked ≤5 times since baby was born

[4] Participant met criteria for prolonged abstinence at delivery (ie positive primary outcome) plus self-reported smoking ≤5 times since baby was born

[5] Cessation information was collected at 1 year but was not listed as an outcome in the protocol

[6] Respiratory symptoms were collected at one year but were not listed as outcomes in the protocol

[7] Has a doctor ever said your child has asthma?

The following is additional information obtained from questionnaires but not included in the tables above

Other outcomes related to smoking:

(NB To possibly report in text only (not listed as outcomes in protocol))

Asked at 6 months only:

In the last 6 months have you used any NRT? Yes / No *(binary)*

In the last 6 months how many times have you met and spoken to a smoking cessation advisor from an NHS stop smoking service?

None / One / Two / Three or more *(categorical – or binary – None / One or more)*

5.6 Maternal morbidity and mortality between birth and 2 years post-delivery [1]

	Placebo	NRT	OR (95% CI) [2]	P value
Total	N	N		
Maternal death				
Yes	N (%)	N (%)	OR	

[1] no maternal deaths reported at primary outcome therefore = total maternal deaths

[2] adjusted for centre only (as a stratification factor)