A double-blind randomised controlled clinical trial to test the effects of Attention Control Training for infants at increased familial risk of ADHD

Statistical Analysis Plan

Version 2.0

Version date: 16/07/2019

ISRCTN: 37683928

This SAP has been written based on Protocol V3

INTER-STAARS Statistical Analysis Plan

Contents

1. Quantitative Analysis Plan 3
1.1. Brief description of the trial 3
1.1.1. Principal research objectives to be addressed 4
1.1.2. Trial design 4
1.1.3. Method of allocation of groups 6
1.1.4. Duration of the treatment period 6
1.1.5. Frequency and duration of follow-up 6
1.1.6. Visit windows 6
1.1.7. Eligibility screening 7
1.1.8. Measures 7
1.1.9. Sample size estimation (including clinical significance) 10
1.1.10. Brief description of proposed analyses 10
1.2. Data Analysis Plan - Data Description 11
1.2.1. Recruitment, eligibility and representativeness of patients 11
1.2.2. Baseline comparability of randomised groups 13
1.2.3. Adherence to allocated treatment, treatment fidelity and Complier Effect Estimation 13
1.2.4. Loss to follow-up and other missing data 13
1.2.5. Assessing quality of outcome measures 14
1.2.6. Adverse event reporting 14
1.2.7. Descriptive statistics for outcomes measures 14
1.3. Data Analysis Plan – Inferential analysis 15
1.3.1. Aims of formal inferences (overview) 15
1.3.2. Analysis of the primary outcome 15
1.3.3. Primary aim hypothesis test 16
1.3.4. Secondary aim hypothesis tests 16
1.3.5. Missing Data 17
1.3.6. Non-adherence and CACE estimator 17
1.3.7. Sensitivity analysis 17
1.3.8. Analysis of Secondary Outcomes 17
1.4. References 18

INTER-STAARS Statistical Analysis Plan

1. Quantitative Analysis Plan

This document details the presentation and analysis strategy for the primary paper reporting results from the INTER-STAARS trial. It is intended that the results reported in these papers will follow the strategy set out herein; subsequent papers of a more exploratory nature will not be bound by this analysis plan but will be expected to follow the broad principles laid down for the principle paper(s). The principles are not intended to curtail exploratory analysis or to prohibit sensible statistical and reporting practices but they are intended to establish the strategy that will be followed as closely as possible, when analysing and reporting the trial. Reference was made to the trial protocol version 2, ICH guidelines on Statistical Principles (ICH E9 (1998)) reference and CONSORT guidelines (Schultz et al. 2010)

Investigators
Prof Patrick Bolton
Prof Tony Charman
Prof Mark Johnson
Prof Emily Jones
Prof Andrew Pickles
Dr Tim Smith
Prof Edmund Sonuga-Barke
Dr Sam Wass

Principal Investigator
Prof Mark Johnson

Trial Manager
Dr Simona Salomone/Dr Amy Goodwin

Trial Statistician
Professor Andrew Pickles
Emily Robinson/George Vamvakas

1.1. Brief description of the trial

Attention deficit/hyperactivity disorder (ADHD) is a developmental condition that significantly impacts on education, social and life outcomes. Medication is often used as a treatment but has significant limitations. Psychological approaches, although more acceptable to many, have proved disappointing in their efficacy, perhaps due to them being administered too late in development. The best time to intervene is likely in infancy when brains are more amenable to positive environmental effects. To test this idea we will undertake a trial of a novel computer based attention training intervention for infants who are at increased familial risk of ADHD, because they have either a parent or older sibling with the disorder.

1.1.1. Principal research objectives to be addressed

The hypothesis is that we will demonstrate an improvement in attention control in infants at familial risk of ADHD who have completed our Attention Control Training programme.

Primary objectives

To conduct a Phase II RCT of an Attention Control Training programme in infants at familial risk for later ADHD; and to assess our primary composite measure of attention control following study intervention.

Secondary objectives

1. Assess the effects of our training intervention on several secondary outcomes such as neurophysiological measures of attention and other cognitive functions;
2. To test whether these effects will be mediated by changes in the executive attention network and will transfer to a range of testing contexts in the short and medium term; and
3. To test whether there is long-term amelioration of emerging ADHD symptoms.

We will use information collected to plan a final stage RCT if appropriate.

1.1.2. Trial design

The trial is designed as a Phase II double-blind multi-centre randomised control trial. Fifty 10 months old infants at increased familial risk for ADHD will be randomised to either an attention training or "passive viewing" control arm. Both arms consist of up to nine weekly 'in home' training sessions conducted over three months. Participants will be assessed in the laboratory and at home immediately pre- and post-training, with additional laboratory follow-up at 24 and 36 months. The primary endpoint will be a composite measure of performance on untrained attention control tasks measured immediately post treatment.

INTER-STAARS Statistical Analysis Plan

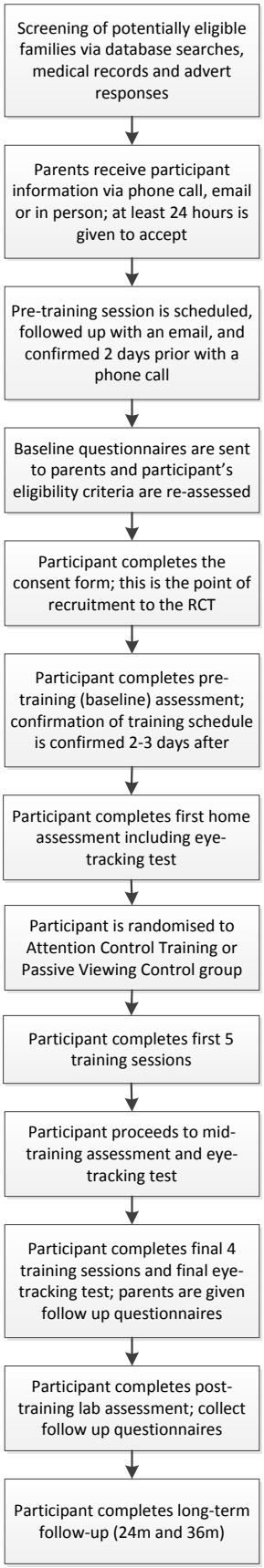

Figure 1. Trial design flow diagram

1.1.3. Method of allocation of groups

Randomisation will take place via a web-based service hosted at the King's Clinical Trial Unit (KCTU), that can be accessed at www.ctu.co.uk by clicking 'randomisation – advanced' on the lower right hand side of the page. This system can only be accessed by trial staff that are trained and have previously been allocated with a username and password. Requests for passwords are via the trial manager to the King's CTU.

Randomisation will be at the level of the individual patient, minimising on gender and study site. The allocation sequence will be generated dynamically, with an 80% probability of allocating to the arm which reduces the imbalance.

Once randomised, the system will automatically generate an unblinded email confirmation which will be sent to the study staff who made the request; researchers who will administer pre- and post- training lab assessments will be blind to participants' group allocation throughout the trial, and research assistants running the training will not be blind because they will need to deliver the intervention.

Every effort will be made to ensure that parents/carers are kept blind to participants' group allocation throughout the trial, and will be instructed not to reveal any information about the training programme during lab assessments. Further information about blinding can be found in the protocol, section 4.3.

1.1.4. Duration of the treatment period

Study treatment is described in detail in the protocol, section 4. As mentioned in section 1.1.2 of this document, the active intervention consists of up to nine weekly 'in home' training sessions over a three month period.

1.1.5. Frequency and duration of follow-up

Participants will be assessed in the laboratory and at home immediately pre- and post-training, with additional home-based assessment after the first five sessions of training, and laboratory follow-up at 24 and 36 months.

1.1.6. Visit windows

The first assessment task will take place prior to randomisation at the first home visit, and the last follow-up task will take place at the last home visit at the end of the training schedule. The study team will be flexible in scheduling all interim visits, and some measures may need to be continued at a different date, or skipped. Therefore, a minimum of 3 days and a maximum of 2 weeks will be given between each study visit to collect relevant measures. The training session number will be recorded as well as the home visit number to ensure that the correct training video is administered depending on the participant's training schedule progress.

1.1.7. Eligibility screening

Participants must give consent to participate in the trial and meet the eligibility criteria prior to randomisation. They will have at least 24 hours after being given the patient information sheet to agree to participate, and the formal consent will be taken at the first pre-training lab assessment. Inclusion and exclusion criteria are described in section 4.1 of the protocol.

1.1.8. Measures

The mechanisms for which all of the following measures will be recorded are described in detail in the protocol, sections 4.6 and 4.7.

Baseline

The following demographics will be reported at baseline for biological mother, father:

- Ethnicity (White; Asian; Black; Mixed; other)
- Age (years)
- Country of birth

The following demographics will be reported at baseline for the infant's primary caregiver:

- Relationship to infant
- Caregiving role (primary; other significant; other)
- Highest level of education – above and below tertiary
- First language of the primary caregiver
- Other languages spoken by the primary caregiver
- Proportion of language that infant hears in English

The following demographics will be reported at baseline for the family:

- Annual household income (<£20K; £20K-£30K; £30K-£40K; £40K-£60K; £60K-£80K; £80K-£100K; £100K-£150K; £150K+)

The following tasks/domains will be reported at baseline, either in the pre-training lab assessment or at the first home visit assessment:

- Parent questionnaires
    - Infant Behaviour Questionnaire-Revised (IBQ-R)*
    - Sleep and Settle Questionnaire (SSQ) – average duration of night sleeps in the past week; average duration of day sleeps in the past week; average number of night wakings in the past week.
    - Gestational age of infant (weeks)
    - Family medical/psychiatric history in the following categories:
        - ASD (0 = no relative, 1 = first degree relative, 2 = second degree relative, 3 = other relative)
        - Depression and/or anxiety (0 = no relative, 1 = first degree relative, 2 = second degree relative, 3 = other relative)

        - Schizophrenia, manic depression bipolar, been admitted to a hospital for a psychiatric illness (0 = no relative, 1 = first degree relative, 2 = second degree relative, 3 = other relative)
    - Genetic syndrome (0 = no relative, 1 = first degree relative, 2 = second degree relative, 3 = other relative)
    - Vineland Adaptive Behaviour Scale (composite score)
    - Visual Media Exposure History (minutes per day of screen exposure (television))
- Behaviour observational measures
    - Toy Play*
    - Mullen Scales of Early Learning (composite score)
    - Early Social Communication Scales (ESCS)*
- Eye-tracking tasks (see "primary outcome measures")
    - Cognitive control
    - Gap-overlap
    - Sustained attention

*Note: please see secondary outcome measures below for further detail

Of note, the majority of infants in the trial also participated in the partner study STAARS. As part of this study, they completed a wider protocol (see Appendix X: STAARS protocol). These measures (beyond those listed above) will not form part of the primary report for INTER-STAARS, but may be used for subsequent analysis.

Primary outcome measures

Our primary outcome measure will be a composite of three eye-tracking components:

1. Cognitive control
    - the percentage of trials on which infants correctly anticipate the location of the target stimulus (average percentage across pre-switch and post-switch phases);
2. Disengagement effect
    - difference between the mean saccadic reaction time in the baseline and overlap conditions for valid trials;
3. Sustained attention
    - mean longest look duration during the interesting stimulus blocks

The descriptive statistics for the primary outcome will include each component together with a composite formed as the simple sum of normalized standardized scores of each component. The effect estimate will be obtained as the common effect across all three standardized components accounting for each baseline component score in a repeated measures ANCOVA.

Secondary outcome measures

Several secondary outcomes will be used to assess secondary objectives, as mentioned in section 1.1.1. Namely, to assess the effects of our training intervention on neurophysiological measures of attention and other cognitive functions, behavioural measures of sustained attention and activity level will be employed at 14 months, and include the following measures:

- Infant Behaviour Questionnaire-Revised (IBQ-R)
    - Effortful control (composite measure comprised of average scores across the soothability, duration of orienting, cuddliness and low intensity pleasure scales);
    - Activity level;
    - Duration of orienting (specifically on its own, because may be more sensitive than overarching effortful control variable).
- Toy play – Laboratory Temperament Assessment Battery (Lab-TAB)
    - Observed during the 'task orientation' in the Lab-TAB (multiple objects), administered in the home following the Home Behavioural SOP for the procedure, using proportion of time that the infant was both manipulating *and* looking at the stimuli simultaneously. We will not include the structured free play episode because it is less suited to coding of this variable.
- Early Social Communication Scales (ESCS)
    - Book task for Responding to Joint Attention, removing trials where the infant was already looking at the beginning of the trial, and including only immediate looks. Total Immediate Looks / Trial N.
    - Early Social Communication Scales (ESCS) adjusted for age and administration.
    - Initiating Joint Attention coded during book and object spectacle tasks; total of the two lower (Eye Contact and Alternates) and two higher (Points and Show) forms as specified on page 50 of the ESCS manual
- Average look duration to the screen across the first 10 minutes of the pre-post battery (shortened from 20 minutes in the protocol so that we can have the same duration for all infants).

These measures tap behaviours that are relevant in everyday contexts for infants. For example, the IBQ asks about a child's ability to pay attention for extended periods during book reading or interaction with a parent. Further, performance on these measures has been linked to levels of later ADHD symptoms (Friedman et al., 2005; Lawson & Ruff, 2004) Thus, these measures will allow us to assess the generalisability of training effects to naturalistic contexts.

The three components of the primary outcome composite will also be considered as separate secondary outcome measures. See section 1.3 for how these will be analysed.

1.1.9. Sample size estimation (including clinical significance)

Based on the treatment effect size reported for the primary outcome in Wass et al. (2011) of 0.69, a total of 50 participants (25 in each group) are required. Allowing for 10% attrition this will give 82% power for two-tailed significance of 0.05 using analysis of covariance with baseline-outcome correlation of 0.6 (Stata sampsi procedure).

Since this original calculation, the primary outcome has been modified to be a composite that would include additional components of sustained attention and disengaging visual attention. These produced similar effect sizes to the cognitive control in our previous work. Including these in a composite score is expected to provide a more stable and broader measure of treatment success and increased power. For instance, a composite formed by the addition of a single repeated measurement at each of baseline and outcome, conservatively assuming a similar correlation of 0.6 as between baseline and outcome measures, allows power to exceed 80% for an effect size down to 0.5 (Stata sampsi procedure).

1.1.10. Brief description of proposed analyses

Analyses will be carried out by the trial statistician (GV) once the database has been locked. Data will be analysed with an intention-to-treat approach (i.e. analyse all those with data in groups as randomised irrespective of treatment received).

There will be descriptive statistics reported on the measures mentioned in 1.1.8, with an aim to comparing the treatment arms, and to review the patient demographics.

For the primary analysis, to test the treatment effect of the attention control training group compared to the passive viewing control group, analysis of covariance will be used.

Further details of the analyses are given later on in this document.

Data summaries and analyses will be carried out in Stata 15.0.

1.2. Data Analysis Plan - Data Description

1.2.1. Recruitment, eligibility and representativeness of patients

A CONSORT flow chart will be constructed – see Figure 2. The number of patients will be summarised using the following categories: total number of patients contacted; eligible; consenting; and randomised.

Then by treatment arm: patients compliant and non-compliant with intervention; continuing through the trial; withdrawing; lost to follow-up; and excluded or analysed.

Compliance (adherence) is defined as having attended at least 6 of the training programme sessions, for both treatment groups. A summary of the number of training sessions attended, and the number of patients compliant with training sessions will be provided.

INTER-STAARS Statistical Analysis Plan

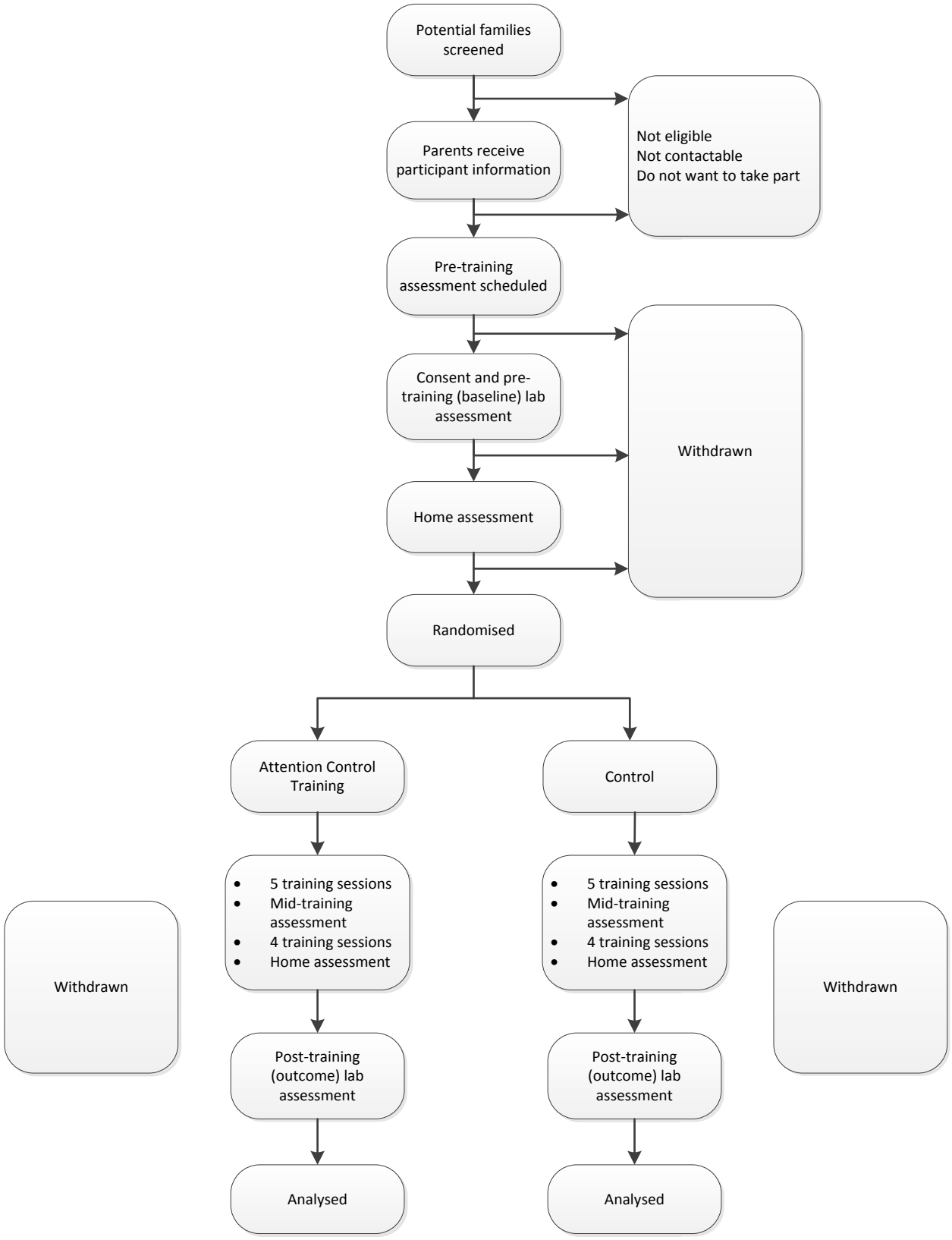

Figure 2: INTER-STAARS CONSORT

1.2.2. Baseline comparability of randomised groups

Table 1 will present, for all available cases, means and standard deviations (proportions and frequencies for categorical variables) disaggregated by group for baseline values of variables contributing to the primary outcome, the secondary outcomes and background infant and family socio-demographic variables. No statistical significance tests or confidence interval will be calculated for the difference between randomised groups on any participant level baseline variables. The randomisation of intervention groups to participants should have ensured that any imbalance over all measured and unmeasured baseline characteristics is due to chance (Altman and Dore 1991).

1.2.3. Adherence to allocated treatment, treatment fidelity and Complier Effect Estimation

Any departures from intended treatment assignment will be described by the number of training sessions completed (minimum 6) and the average duration of game play/session (minimum 5 mins).

Treatment fidelity will be measured by the level of contingency, i.e. percentage of time that the child was looking at the screen that their eyes were detected (amount of eye-tracking data). There is no pre-specified minimum requirement.

We will review compliance with the treatment protocol, defined as failing two of the following three criteria:

- More than 21 days between the last training/control visit and the post test or intermediate test
- Difference between the pre and post test more than 114 days
- More than 70 days between the pre and intermediate tests OR between the intermediate and post tests.

If this exceeds 20% in the treatment arm a complier effect estimate will be calculated. Although the treatment allocation is masked, it is possible that treatment may influence longer-term interest and compliance of the infant. Such selective compliance makes standard per-protocol analysis inappropriate as a causal effect estimator. We will estimate a complier average causal effect estimator (CACE) using an instrumental variable approach.

1.2.4. Loss to follow-up and other missing data

Table 1 will also present, for all available cases, means and standard deviations, proportions and frequencies, disaggregated by assignment group, the primary and secondary outcomes. The numbers with endpoint data within each randomised group will be presented for each outcome. The major known reasons for loss to follow-up will be described and any systematic differences by treatment group in the characteristics of those lost will be described.

1.2.5. Assessing quality of outcome measures

As mentioned in the protocol section 4.7.1, treatment blind analysis will be carried out to assess quality of data collected during testing sessions. Therefore, we will use a home-lab composite primary outcome measure score *unless* the following conditions are not met for at least 75% of the sample:

1. Cognitive control task: a minimum of 2 trials with anticipations (either correct or incorrect) per phase (learning and reversal) per block.
2. Gap-overlap task: a minimum of 5 valid trials per critical condition (baseline and overlap).
3. Sustained attention: a minimum of 4 'looks' per infant during interesting phase (in either one of the two blocks administered).
4. Measurement Lag: Time between last home session and post-training lab-visit less than 4 weeks.

If these conditions are not met for 75% of the sample, then we will use *either* the home measures only (if 1, 2 and 3 are met for 75% of infants for the home battery, but 4 is not met) or the lab measures only, (if 1, 2 and 3 are not met for the home battery but 4 is met) to calculate the final composite score.
If this is not possible (e.g. 1, 2, 3 and 4 are not met for the home battery) then we will select either the lab or the home composite on the basis of which battery has the greater proportion of infants who meet criteria 1, 2 and 3.

1.2.6. Adverse event reporting

Adverse side effects will be described and analysed as part of the primary paper. No continuous monitoring is proposed as we know little about the range of normal variation for these measures. Comparisons will be made of the nights before and after training. They will be assessed by:

- Infant's sleep quantity (duration of mins/night) as assessed by the sleep diary, comparing the night before and night after each training visit
- Fussiness rated on a 1 to 5 scale (average over the 4 hours after the training session, and the 4 hours that is comparable for the previous day, e.g. if training session finishes at 14:00, it would be 14:00-18:00 on that day, and day before)
- Hospitalisation

1.2.7. Descriptive statistics for outcomes measures

The primary and secondary outcomes as listed in section 1.1.8 will be described by treatment group and time point. Means and standard deviations or medians and interquartile ranges will be used for continuous variables, where relevant; this will check whether continuous outcomes can be assumed normally distributed.

1.3. Data Analysis Plan – Inferential analysis

1.3.1. Aims of formal inferences (overview)

The study analysis and publication plan will follow CONSORT guidelines. This statistical analysis plan will be agreed with a Data Monitoring and Ethics Committee before any post randomisation assessment.

We powered our study to ask whether our training program improves attention control in infants at familial risk for ADHD, and whether these effects generalise to other behavioural tasks. The formal statistical analyses will estimate the differences in relevant variables between patients randomised to the Attention Control Training programme compared to the "passive viewing" control group, by intention to treat. Later papers will conduct a number of analyses to examine the mechanisms of change and their long-term consequences.

Group difference estimates and associated 95% confidence intervals will be reported. All data preparation and analysis for the primary paper will be blind to treatment group. If any of the data contain information that may disclose blindness, these data will be re-coded before analysis. The overall significance level will be 5% (two-sided) for the primary and secondary outcomes.

A preliminary analysis of a partial dataset without post-test outcomes is being undertaken as part of a PhD project, and will be supervised by staff that are not part of the study team and kept confidential until after the trial primary paper is complete. There will be no publication nor presentation on this preliminary analysis until the primary paper of the full trial has been published.

Details on the methods for handling missing data are given in sections 1.3.2.

Sensitivity analyses will be used to assess the robustness of conclusions; please refer to section 1.3.4 for details of the planned sensitivity analyses.

No interim analysis is planned.

1.3.2. Analysis of the primary outcome

The analysis population will include all patients randomised with sufficient information to carry out the analysis, i.e. complete primary outcome data and minimisation factors. The primary outcome is a primary composite measure of attention control, see section 1.1.8. See section 1.2.5 for the decision rule as to whether this will be derived from laboratory or home based assessments.

The three components will be analysed together initially for evidence of a common treatment effect, equivalent to a treatment effect for the simple sum of normalized standardized scores of each component. This decision is due to uncertainty in the correlation between the components of the composite; both from a clinical point of view and after reviewing data from a previous study (Wass 2011), which suggested there is little evidence to support the use of a common factor. The data of that study when used to construct a standardized sum as a composite, showed a difference in primary outcome between the two treatment groups.

The primary outcome measures will be standardized at both baseline and endpoint. To test the treatment effect of the Attention Control Training group compared to the passive viewing control group, a multivariate analysis of covariance (multivariate ANCOVA but fitted in Stata sem) will be used, allowing for an unstructured correlation matrix between components and estimated by full maximum likelihood. We will make use of the repeated measures setup in which both baseline and outcome variables are treated as dependent variables with a design matrix of predictors that allow no group difference at baseline (though a different error variance at baseline and outcome and non-zero covariance). If following section 1.2.5 the primary selected endpoint is home measurement, then an additional midpoint assessment will be available, which will be incorporated into the analysis as an additional equation and on the assumption that half of the full treatment effect will have taken place in the manner of a linear growth curve. This setup is, with complete data two-period data, equivalent to ANCOVA but allows both missing baseline and outcome data to be treated as ignorable under a Missing-at-Random assumption. Dummy variables for the minimisation factors (gender and study site) will also be included. Treatment effect estimates for each of the three components at endpoint will be obtained. Using the baseline standard deviation of each component, an effect size estimate based on the average effect size of each component will be presented together with 95% confidence interval calculated using 1000 boot-strap samples.

Distributional assumptions will be checked by the use of Q-Q plots of residuals, and variables will be transformed where required.

1.3.3. Primary aim hypothesis test

There is a single primary aim hypothesis test based on the composite score. An alpha level of 5% will be used for the primary outcome composite measure to test significance. Confidence intervals will be calculated by bootstrapping 1000 times the sample.

1.3.4. Secondary aim hypothesis tests

As a secondary aim, the effect estimates of the three individual components of the primary outcome composite will also be reported as secondary outcomes. This constitutes three additional tests. Significance will be based on Hochberg's correction for the three single component secondary outcomes to account for multiple testing. The p-values calculated

from the secondary outcome tests will be ordered from lowest to highest and the Hochberg multiplicity corrected p-value will then be computed. Confidence intervals will be calculated by adjusting the alpha level for each subsequent test.

To assess comparability with the pilot study we will also report the treatment effect estimate at the intermediate measurement point.

1.3.5. Missing Data

We will make use of the repeated measures setup in which both baseline and outcomes variables are treated as dependent variables with all valid scores for each component included. Account for missing measures will be made under an assumption of a missing-at-random mechanism. Since all remaining variables in the analysis are required for randomisation there should be no further missing data.

1.3.6. Non-adherence and CACE estimator

If non-adherence to the intervention in the active arm exceeds 20%, a Complier-Average-Causal Effect estimator and its 95% confidence interval will be calculated for the primary outcome using an instrumental variable extension to the primary outcome model of 1.3.2 and in Mplus or Stata gsem (Jo, 2002).

1.3.7. Sensitivity analysis

If the proportion of participants who are missing all components of the primary outcome composite exceeds 10%, a sensitivity analysis will be undertaken in which outcome scores, imputed using Stata mi with all analysis variables included, will be used with the addition of a fixed quantity under four scenarios of missing data bias. These will be based on the observed distribution of change scores in the two groups. In the first those missing will have the score corresponding to the 50% percentile for their group added to their imputed score. This should give results similar (but not identical) to those of the standard ANCOVA estimated under the MAR assumption. In the second and third the added scores would correspond to 25th and 75th percentiles. In the 4th scenario the 25th percentile for the contingent (active treatment) group and the 75th percentile for the non-contingent group (implying a bias towards drop out of the less progressing in the contingent group and the more progressing in the non-contingent group) would be used.

1.3.8. Analysis of Secondary Outcomes

All secondary outcomes are continuously scored and will be analysed using an SEM setup that assumes bivariate normality of baseline and endpoint measures.

The list of secondary outcome measures can be found on page 8/9.

1.4. References

International Harmonised Tripartite Guidelines : Statistical Principles for Clinical Trials E9 (1998). European Medicines Agency.

Jo, B. (2002). Estimation of intervention effects with noncompliance: Alternative model specifications. Journal of Educational and Behavioral Statistics, 27, 385-409

Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010 Mar 24;340:c332.

StataCorp. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC; 2017.