

Study Title: Investigating Non-invasive brain Stimulation To Enhance fluency in People who stutter

Short title: INSTEP Trial
University of Oxford, Medical Sciences Division
IDREC Ethics Ref: R52173/RE001

Document title: INSTEP trial statistical analysis plan
Date and Version No: 23rd October 2017, version 1.0

Chief Investigator: Professor Kate Watkins, Department of Experimental Psychology, University of Oxford

Investigators: Dr. Jennifer Chesters, Department of Experimental Psychology, University of Oxford

Ms. Charlie Wiltshire, Department of Experimental Psychology, University of Oxford

Ms. Louisa Needham, Department of Experimental Psychology, University of Oxford

All the investigators declare that there are no conflicts of interest associated with the proposed study.

INSTEP trial statistical analysis plan

We plan to carry out analyses on the 'intention to treat' sample, i.e. including all participants who complete the study. If any participant has data that are group outliers, we will also complete sensitivity analyses (Thabane et al., 2013) by re-running all analyses excluding the relevant participant(s), to evaluate the robustness of the treatment effect. Speech data in people who stutter tends to significantly deviate from a normal distribution. Transformation of these data is problematic for reporting confidence intervals in interpretable units, therefore we will use change scores (baseline subtracted) in the analyses to overcome this limitation. For consistency, we will use this approach of baseline subtraction and analysis of change scores in analysis of other variables even if they do not suffer from the same skewness as the speech data. Variables that are not normally distributed will be analysed using nonparametric methods.

Primary outcome measure

The primary outcome measure for the trial is the change in stuttering severity, measured using the Stuttering Severity Instrument, version 4 (SSI-4, Riley, 2009). Change from baseline score will be calculated (by subtraction of the baseline) at one week, 6 weeks and 12 weeks following the intervention. Change scores will be entered into a mixed model analysis of variance (ANOVA), with stimulation group (tDCS, sham) as the between subjects factor and speaking task and time point (1 week, 6 weeks, 12 weeks) as the within subjects factor. If a significant interaction between stimulation group and time point is found, separate repeated measures ANOVAs will be performed for each stimulation group or with three planned comparisons between groups at each time point.

Secondary outcome measures

We have three secondary outcome measures:

- i. Change from baseline score in percentage of disfluent syllables, during reading.
- ii. Change from baseline score in percentage of disfluent syllables, during conversation.
- iii. Change in total score on Overall Assessment of a Speaker's Experience of Stuttering (OASES).

Change in percentage of disfluent syllables will be calculated by subtraction of the baseline percentage for reading and conversation separately at one week, 6 weeks and 12 weeks following the intervention. Data will be analysed as described above for the primary outcome measure. Anticipating a significant effect of group in both measures, we will also compare the changes in reading with those in conversation at each time point separately for the tDCS group.

To test for the effects on the OASES measure, we will use a mixed model ANOVA with stimulation group (tDCS, sham) as the between subjects factor and time point (6 weeks, 12 weeks). We will not collect data at the 1-week post-intervention time-point to avoid violating test-retest validity of the scale. If a significant

interaction between stimulation group and time point is found, separate planned comparisons between groups will be performed for each time point.

Other pre-specified measures

We are including the Beck Anxiety Inventory to monitor for any differences in anxiety between the groups, which should be minimized at baseline by the randomization procedure. Any effect of the intervention of stuttering severity may also have consequences for anxiety. We will assess the change in anxiety scores with the same analysis as planned for the primary outcome measure. We will also examine the correlation of the change in anxiety score and the change in (i) stuttering severity (SSI-4), and percentage of disfluent syllables on (ii) reading and (iii) conversation tasks

The Premonitory Awareness in Stuttering Scale (PAiS) tests for anticipation of stuttering events. We have included this as an exploratory measure. We also include subjective ratings of stuttering severity and speech naturalness, and an objective measure of speech naturalness. For each scale (PAiS, subjective severity, subjective naturalness, objective naturalness) change scores will be analysed as described for the primary outcome measure.

In addition, we include exploratory analysis of the effects of tDCS during the intervention period. Effects of tDCS immediately after intervention will be analysed using change in percentage of disfluent syllables (baseline subtracted) measured post-intervention on each day, for both reading and conversation. A mixed model ANOVA will be used with stimulation group (tDCS, sham) as the between subjects factor and time point (day 1 to 5) and task (reading, conversation) as the within-subjects factors. Carry-over of effects following a period including sleep will be analysed using change in percentage of disfluent syllables (baseline subtracted) measured immediately before the intervention, for both reading and conversation. A mixed model ANOVA will be used with stimulation group (tDCS, sham) as the between subjects factor and time point (day 2 to 5) and task (reading, conversation) as the within-subjects factors. The Day 1 time point will be omitted in this analysis, as no intervention will have been performed prior to that measurement.

Finally, in further exploratory analyses, we wish to examine the three submeasures that contribute to the Stuttering Severity Instrument total score. These are: (i) frequency, expressed in percent syllables stuttered and converted to scale scores of 2-18; (ii) duration is the average length of the three longest stuttering events timed to the nearest one tenth of a second and converted to scaled scores of 2-18; and, (iii) physical concomitants for which the four types of physical concomitants are rated on a scale of 0-5 and totaled. We will compare changes from baseline in these variables as described for the analysis of the primary outcome measure.