# Statistical Analysis Plan

**A randomised controlled trial evaluation of the effectiveness of three minimal human contact interventions to promote fitness and physical activity in an occupational health setting.**

**20 February 2014**

# 1 Introduction

This is the plan for the main trial analyses of the primary and secondary outcome measures from a randomised controlled trial evaluating the effectiveness of three minimal human contact interventions to promote fitness and physical activity in an occupational health setting.

Other analyses of data arising from this trial will be the subject of future analysis plans.

# 2 Study endpoints

## 2.1 Primary efficacy outcomes (measured at baseline and follow-up)

There are 2 co-primary outcomes: PAEE measured using the ActiHeart, and fitness measured using sub-maximal graded treadmill exercise test.

## 2.2 Secondary efficacy outcomes (measured at baseline and follow-up)

- BMI.
- Body fat %.
- Waist.
- Systolic blood pressure.
- Diastolic blood pressure.
- Plasma vitamin C.
- HbA1c.
- Fructosamine.
- Total cholesterol.
- HDL cholesterol.
- LDL cholesterol.
- Total:HDL cholesterol ratio.
- Trigylcerides.
- Quality of life derived from SF-8.
- Perceived stress derived from the validated Perceived Stress Scale.

## 2.3 Intermediate outcomes (measured at baseline and follow-up)

- Summary measures from Theory of Planned Behaviour physical activity questionnaire.

## 2.4 Process measures at follow-up

- Use of intervention website/diary.
- Absence from work due to illness.

## 2.5 Safety outcomes

- Adverse events.

# 3 Analysis populations

The primary analysis of efficacy, intermediate and safety outcomes will use an **Intention To Treat (ITT) population**, which includes all participants in the group to which they were randomised, regardless of the intervention actually received.

A secondary analysis of efficacy and intermediate outcomes will use a **Per Protocol (PP) population.** Inclusion in the PP population will be based on the degree of usage of the intervention website/ completion of diary and will be defined once clean data are available (but before the start of any trial analyses), when the distributions of the degree of usage of the intervention website/completion of diary can be inspected.

# 4 Descriptive analyses

Baseline characteristics of the study population will be summarised separately within each randomised group.

For continuous variables, means and standard deviations will be presented, unless the variable has a highly skewed distribution, in which case medians, 25th and 75th percentiles will be presented. For categorical variables, the number and percentage of participants within each category will be presented. For each variable (continuous or categorical), the % of missing values will be reported.

No p-values will be calculated for these tables.

# 5 Analyses of study outcomes

## 5.1 Primary efficacy outcomes

The co-primary efficacy outcomes, PAEE and fitness, will each be analysed using analysis of covariance (ANCOVA). The outcome in the ANCOVA model will be change (follow-up minus baseline) in PAEE (or fitness), with the baseline value included as a covariate in the model. For each outcome a 3 degrees of freedom test will be performed of the null hypothesis that there is no difference between the 4 randomised groups. If the p-value from this test is <0.05, this will imply that there are differences between the groups. The ANCOVA model will also be used to derive estimates of the differences in mean change and 95% confidence intervals for each of the 6 pairwise comparisons: Intervention 1 vs Control, Intervention 2 vs Control, Intervention 3 vs Control, Intervention 2 vs Intervention 1, Intervention 3 vs Intervention 1, Intervention 3 vs Intervention 2. p-values for each of the pairwise comparisons will not be calculated.

Where baseline values of the outcome are missing, the missing indicator method will be used to enable these participants to be included in the analysis.

An analysis will be performed to check whether adjusting for age, sex and BMI (the randomisation stratifiers) in the ANCOVA model has any impact on the conclusions; if it has no impact, then they will not be included in the model.

## 5.2 Secondary efficacy and intermediate outcomes

For each continuous secondary and intermediate outcome, the 6 pairwise differences between randomised groups will be estimated, together with 95% confidence intervals, using ANCOVA as described in section 5.1. No p-values will be calculated. Any continuous endpoints whose distribution is skewed will be log transformed prior to analysis, in which case a ratio of geometric means (and confidence interval) will be reported.

### 5.3 Process measures

Each process measure reported at follow-up will be summarised separately within each randomised group, using means and standard deviations, or medians and 25th,75th percentiles, or numbers and percentages as appropriate depending on the distribution of the variable. No p-values or confidence intervals will be calculated.

### 5.4 Safety outcomes

The numbers and types of adverse events within each randomised group will be reported. No p-values or confidence intervals will be calculated.

## 6 Considerations for analysis

### 6.1 Missing data

*Missing values of outcomes*

If an individual has a missing value for an efficacy outcome, they will be excluded from the analysis. The pattern of missing data will be described. Levels of missing data are expected to be low, but if this is not the case, the potential impact of missing data will be explored in sensitivity analyses using a pattern mixture model (White 2012).

*Missing baseline values of endpoints*

For continuous efficacy endpoints, those participants with a missing baseline value of the variable will be included in the analysis using the missing indicator method (White 2005), which is a valid method for pre-randomisation measures in trials ensuring that no further participants are excluded while maintaining the advantage of improved precision.

### 6.2 Subgroup analyses

Subgroup analyses by men/women and BMI (below/above median value) will be investigated for the 2 co-primary efficacy outcomes only. The first step will be to perform a 3 degrees of freedom test of the null hypothesis that there is no interaction between any of the 3 intervention groups and (1) gender, (2) BMI (below/above median). If the p-value from this test is ≥0.05 no further analyses will be performed. If the p-value is <0.05, the effect of each intervention (i.e. difference between intervention vs control and 95% confidence interval) will be estimated within each subgroup.

### 6.3 Multiplicity

Given the number of outcome variables, randomised groups and therefore comparisons, the focus of the results will be on estimated differences and 95% confidence intervals. p-values will only be calculated as described in sections 5.1 and 6.2.

# 7 References

White IR, Thompson SG (2005) Adjusting for partially missing baseline measurements in randomized trials. Stat.Med 24: 993-1007.

White IR, Carpenter J, Horton NJ (2012) Including all individuals is not enough: lessons for intention-to-treat analysis. Clinical Trials 9: 396-407.