'HelpMeDoIt!' a web, app and text based intervention to facilitate social support to achieve and maintain health related change in physical activity and dietary behaviour

STATISTICAL ANALYSIS PLAN

Study Title:	'HelpMeDoIt!' a web, app and text based intervention to facilitate social support to achieve and maintain health related change in physical activity and dietary behavior
Short Title:	HMDI
Sponsor:	NHS Greater Glasgow and Clyde Research and Development Management Office
Sponsor ref:	GN14CP094
Funded by:	National Institute of Health Research Public Health Research Program
Funder ref:	12/180/20
REC ref:	University of Glasgow 200140108
ISRCTN ref:	ISRCTN85615983

Protocol Version:	v5.0	Date:	28 February 2017
SAP Version:	v1.0	Date:	18 January 2018

		Signature	Date
Prepared by:	Sarah Barry Consultant Biostatistician Robertson Centre for Biostatistics University of Glasgow		18/01/2018
Approved by:	Dr Alex McConnachie Assistant Director Biostatistics Robertson Centre for Biostatistics University of Glasgow		18/01/2018
	Prof Sharon Simpson MRC/CSO Social and Public Health Sciences Unit University of Glasgow		
	Dr Mark Kelson College of Engineering, Mathematics and Physical Sciences University of Exeter		

CONTENTS

1. INTRODUCTION

This study is planned to evaluate a web, app and text based intervention, which is aimed to facilitate social support to achieve and maintain health related change in physical activity and dietary behaviour. It is carried out in two main stages, of which stage one is to assess the feasibility of the study and stage 2 is the evaluation phase. This SAP relates to stage 2 (evaluation phase) of the HelpMeDoIt! (HMDI) study.

1.1. STUDY BACKGROUND

Poor diet, physical inactivity and high BMI have been highlighted in the top ten risk factors for burden of disease worldwide. Health related behaviours are significant contributors to diseases such as diabetes, cancer, heart disease, hypertension and stroke. Preventative interventions which are accessible, engaging and which successfully improve health behaviours are necessary to reverse current trends particularly since interventions to date have had limited effectiveness and approaches known to work are not always adopted.

The internet and social media can be effective in influencing behaviour, and can reach large numbers of people. Previous research shows that setting goals, making plans and monitoring how well you are doing is important to facilitate behaviour change. The support of family, friends and others is also crucial in helping people to achieve and sustain behaviour change and healthier lifestyles.

1.2. STUDY OBJECTIVES

The primary aim of the evaluation phase of the HMDI study is to evaluate the feasibility of conducting a larger, definitive trial of the HMDI intervention, and to estimate parameters necessary to design such a study.

1.3. STUDY DESIGN

Randomised controlled trial.

Participants will be randomised in a 2:1 ratio (Intervention: Control).

1.4. SAMPLE SIZE AND POWER

120 participants will be randomised (80 vs. 40). The expected drop out is 30%. This sample size of 84 (after drop out) for analysis is not powered for effectiveness but will provide enough precision to estimate any proportion to within 11 percentage points using a 95% confidence interval. This would also allow for the estimation of a continuous outcome (i.e. BMI) in the intervention arm to within 0.262 of a standard deviation.

1.5. STUDY POPULATION

Participants will be recruited from GP practices, weight management clinics, and by community advertising.

1.5.1. INCLUSION CRITERIA

Adults aged 18 to 70 years old with BMI>30 kg/m^2 who are trying to lose weight and have access to a mobile telephone and the internet.

1.5.2. EXCLUSION CRITERIA

All participants with any of these conditions are excluded from the trial:

- Terminal illness
- Previous bariatric surgery
- Dementia
- Pregnancy
- Poor competence in English (resulting in an inability to complete study materials)
- Contraindications to physical activity
- Being a nominated helper in the trial

1.6. STATISTICAL ANALYSIS PLAN (SAP)

1.6.1. SAP OBJECTIVES

The objective of this SAP is to describe the statistical analyses to be carried out for the evaluation phase of HMDI Study.

1.6.2. GENERAL PRINCIPLES

Continuous variables will be summarised based on the number of observations, number of missing values, mean, standard deviation, median, inter-quartile range, minimum and maximum. Categorical variables will be summarised with the number of observations, number of missing values, and the number and percentage of total non-missing data falling into each category.

Where the two randomised groups are to be compared, the intention-to-treat principle will be applied. Appropriate two-sample tests or regression methods (depending on the distribution of the outcome) will be applied to estimate the intervention effect with a 95% confidence interval and p-value. Where appropriate, regression models will be adjusted for the minimisation factor (BMI <40 kg/m^2, or ≥40 kg/m^2) and other important baseline characteristics. No adjustments for clustering will be required for any analyses.

Missing data will not be imputed unless specified below. Where appropriate, multiple imputation will be used. Ten imputed datasets will be generated for each analysis. For each imputed variable, a parsimonious regression model will be

developed, based on data collected at the same or earlier study visits. Randomised group will not be included as a predictor in these models. Where multiple imputation has been applied, full details of the models used to impute data will be provided as part of the final statistical outputs.

1.6.3. CURRENT PROTOCOL

The current study protocol at the time of writing is version 5.0, dated 28/02/2017. Future amendments to the protocol will be reviewed for their impact on this SAP, which will be updated only if necessary. If no changes are required to this SAP following future amendments to the study protocol, this will be documented as part of the Robertson Centre Change Impact Assessment processes.

1.6.4. DEVIATION POLICY

No deviations from the analyses specified in the study protocol are expected.

Any deviations from this SAP in the final analysis will be documented and justified in the final report.

1.6.5. SOFTWARE

Analysis will be carried out using R for Windows v3.2.2 or SAS for Windows v9.3, or higher versions of these programs.

2. ANALYSIS

2.1. STUDY POPULATIONS

The analysis population will consist of all subjects who are randomised. The number randomised, and cumulative number randomised each month during the study will be reported. This information will also be reported graphically.

2.2. BASELINE CHARACTERISTICS

Baseline characteristics will be summarised overall and by randomised group as specified in Section 1.6.2. No statistical comparison will be made between randomised groups. The following characteristics will be summarised:

Socio-demographic, lifestyle and occupational information:

- Age
- Gender
- SIMD
- Source of hearing about HMDI
- Marital status
- Ethnicity
- Education status

- Employment status
- Annual household income
- Computer and phone use

Health status and quality of life variables:

- Height (m), weight (kg) and BMI (kg/m2)
- Waist circumference (cm)
- Hip circumference (cm)
- Change of weight in the last 3 months
- Weight loss activity
- Weight loss medication
- Health problems (i.e. heart disease, diabetes, etc.)
- Health professional visit
- Physical activity (7 Day PAR and accelerometer)
- Diet (DINE and RECALL, 4x repeat 24 hour multiple pass recall)
- Social support (Exercise & Eating Social Support Scales)
- Self-efficacy (Weight & Exercise Efficacy Lifestyle Scales)
- Motivation (Treatment self-regulation Questionnaire)
- Smoking status
- Health related quality of life
    - Mental health (GHQ12)
    - ICECAP-A
    - EQ-5D
    - EQ-VAS

2.3. FEASIBILITY OUTCOMES

Follow-up rates at 12 months will be reported overall and by randomised group, with 95% confidence intervals. Reasons for failure to follow up will be reported. Baseline data will be summarised for those with and without follow-up data. The association between baseline factors and follow-up will be assessed using logistic regression, with follow-up (yes/no) as the response variable. For each baseline factor, the odds ratio for follow-up will be reported with a 95% CI and p-value.

Use of the intervention will be summarised for the intervention group, overall and in relation to selected baseline characteristics. The availability and utility of data relating to data usage for the app will be explored, and a range of summary measures will be presented in the final statistical outputs. For example, the following measures may be presented:

- Participant
    - Number of logins
    - Number of views of main dashboard
    - Number of views of progress charts
    - Number of ‘enter your weight’ updates
    - Number of views of ‘smiles’ feature
    - Number of views of ‘rewards’ feature

  - Number of 'view profile'
  - Number of template goals created
  - Number of custom goals created
  - Number of goals deleted
  - Number of goals completed
  - Number of goals updated
  - Number of helpers invited
  - Number of helpers denominated
  - Number of uses of 'contact helper' feature
  - Number of smiles sent to helper
  - Number of times read summary email
  - Number who unsubscribed from summary email
  - Number of times used forgotten password feature
  - Number of views of notifications (new feature)
  - Number of views of FAQs (new feature)

- Nominated helper
  - Number of logins
  - Number of views of friend's progress charts
  - Number of views of 'smile's feature
  - Number of views of 'rewards' feature
  - Number of 'view profile'
  - Number of views of friend's goals
  - Number of times goals liked
  - Number of uses of 'contact friend' feature
  - Number of 'smiles' sent to friend
  - Number of times used forgotten password feature
  - Number of views of notifications (new feature)
  - Number of views of FAQs (new feature)

Similarly, data relating to website usage will be explored, and a range of summary measures will be presented in the final statistical outputs. For example:

- For both participant and helper
  - Number of logins
  - Number of page views
  - Average duration of each login

Data from the USE questionnaire (App Questions) and follow-up questionnaire data (Questionnaire part 13) relating to web and app usage will also be summarised.

2.4. EFFICACY OUTCOMES

2.4.1. PRIMARY OUTCOME

The three primary outcomes are physical activity, diet and BMI at 12 months follow up to explore which one is most responsive and sensitive to change for the trial. Outcome measures will be derived from the following sources:

- Physical activity
  - 7-day PAR
  - Accelerometer

- Diet
  - Dietary Instrument for Nutrition Education (DINE)

- BMI (kg/m^2)
  - calculated as $weight\ (kg) \div [height\ (m)]^2$

The primary outcome measures will be summarised overall and by randomised group, and compared using linear regression models, with randomised group, the baseline measurement of the outcome, age and gender as predictor variables. With the exception of BMI, regression models will also adjust for the minimisation factor. The regression coefficient for randomised group will be reported with a 95% CI and p-value. A standardised effect size will be reported with a 95% CI, by dividing the intervention effect estimate and confidence limits by the standard deviation of the outcome measure in the whole randomised population at baseline. The residuals from each regression model will be assessed for Normality; where necessary, the outcome measure (at follow-up and at baseline) will be transformed to improve model fit.

Missing data will not be imputed, unless more than 20% of cases are lost due to missing data, in which case multiple imputation will be performed. Details of the imputation procedure used will be provided in the final statistical report. Results of both the complete case and multiple imputation analyses will then be reported.

Analyses of weight and BMI will be repeated after excluding participants who self-reported weight by telephone, as sensitivity analyses.

2.4.2. SECONDARY OUTCOMES

The following secondary outcome measures will be analysed:

- Weight (kg)
- Waist circumference (cm)
- Waist-to-hip ratio
- Health related quality of life (EQ-5D, ICECAP_A, EQ-VAS, GHQ12)
- Smoking and alcohol drinking status
- Social support (Exercise & Eating Social Support Scales) -
- Self-efficacy (Weight & Exercise Efficacy Lifestyle Scales)
- Motivation (Treatment self-regulation Questionnaire)

These will be summarised overall and by randomised group, and compared using linear or logistic regression models, with randomised group, the baseline measurement of the outcome (where available), age and gender as predictor variables. The regression coefficient, or odds ratio, for randomised group will be reported, with a 95% CI and p-value. A standardised effect size will be reported for

continuous outcome measures, with a 95% CI, by dividing the intervention effect estimate and confidence limits by the standard deviation of the outcome measure in the whole randomised population at baseline. The residuals from each regression model will be assessed for Normality; where necessary, the outcome measure (at follow-up and at baseline) will be transformed to improve model fit.

Missing data will not be imputed for analyses of secondary outcomes.

2.4.3. ADDITIONAL ANALYSES

For each primary outcome measure, a per-protocol analysis will be performed. Protocol violators (PVs) will be identified based on adherence to the intervention. The PV definition, and full list of PVs, will be approved by the chief investigator prior to viewing any unblinded results. The primary analysis regression models will be fitted in the per-protocol population. For each model, the regression coefficient for randomised group will be reported with a 95% CI and p-value.

For each primary outcome measure, intervention effect moderation may be assessed with respect to subgroups defined by age, gender, socioeconomic status and the baseline measurement of the outcome. Other potential moderator variables may also be considered. This will be achieved by fitting regression models with main effects for randomised group and the moderator variable of interest, plus their interaction. All models will be adjusted for the baseline measurement of the outcome, age and gender (if not already included in the model). The intervention effect estimate at each level of categorical moderator variables will be reported with a 95% CI and p-value, and the interaction p-value will be reported. The intervention effect estimate at the 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} percentile of continuous moderator variables will be reported with a 95% CI and p-value, and the interaction p-value will be reported. All such intervention effect estimates will be reported graphically with a forest plot.

For each primary outcome measure, a multivariable regression model will be developed based on baseline factors. The process used to derive this model will be described in the final statistical report, but will be based on a mixture of clinical and statistical judgement. All univariable associations considered in the modelling process will be reported with a 95% CI and p-value. For the final model for each outcome, the regression coefficients will be reported with 95% CIs and p-values. Randomised group, the baseline value of the outcome variable and (except for BMI) the minimisation variable will be forced into all models, regardless of statistical significance.

Within the intervention group, for each primary outcome measure, associations with measures of app and website usage will be assessed, to explore whether any intervention effects may be mediated by adherence. Given the large number of potential measures of app and website usage, this will be restricted to reporting of (Spearman) correlation coefficients. P-values will be reported but will be interpreted cautiously; results will be interpreted in terms of the magnitude of associations and

patterns of types of adherence measures that show the strongest associations with outcomes.

2.5. Safety Outcomes

2.5.1. Adverse Events

Only serious adverse events (SAEs) will be collected. Characteristics of events will be summarised overall and by randomised group, without statistical comparison. The number and percentage of participants experiencing at least one SAE, or at least one SAE of specific types, will be reported overall and by randomised group. All SAEs will be listed.

3. Document History

This is version 1.0 of the SAP for the HelpMeDOIt Study, dated 18/01/18. It is the initial version of this document.

4. Tables

A full set of tables will be prepared using dummy randomisation codes and approved prior to the final analysis.

5. Figures

Other than those figures noted in this SAP, additional figures will be prepared following the final analysis as required for publication and presentation purposes.

6. Listings

Formal listings of SAEs and withdraws will be provided.