Statistical Analysis Plan

Promoting physical activity through group self-management support for those with multimorbidity: a randomised controlled trial

(MAP – Movement through Active Personalised engagement)

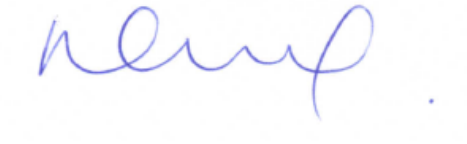
SAP Version: v 0.1 17/07/2019
Based on protocol Version: V6 Dated: 20/06/2018

Sponsor Name:	University of Leicester
Sponsor Address:	University Road, Leicester, LE1 7RH
Funder:	National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care East Midlands (CLAHRC-EM)
Senior Statisticians:	Name and contact details: Prof Laura Gray Department of Health Sciences, University of Leicester
Trial Statistician:	Name and contact details: Mrs Ghazala Waheed Diabetes Research Centre, University of Leicester
Chief Investigator:	Name and contact details: Prof Kamlesh Khunti Diabetes Research Centre University of Leicester

Revision History

Version	Date	Author(s) and Role	Summary of Changes/Comments
0.1	17-07-2019	Trial statistician	Information from protocol included
0.2	19-07-2019	Trial statistician	Changes made for senior statistician's comments on v 0.1
0.3	12-08-2019	Trial statistician	Changes made for chief investigator's comments on v 0.2
1.0	10-09-2019	Trial statistician	Final version

SAP approval for finalised versions:

Chief Investigator	***Prof Kamlesh Khunti***	
		
	Signature	**04.10.2019** **Date**
Supervising Statistician	***Prof Laura Gray***	
		
	Signature	**02/10/2019** **Date**
Trial Statistician	***Mrs Ghazala Waheed***	
		10/09/2019 **Date**
	Signature	

LIST OF ABBREVIATIONS

AE	Adverse event
AR	Adverse reaction
BMI	Body Mass Index
BP	Blood pressure
CHD	Coronary heart disease
CV	Cardiovascular
CVD	Cardiovascular disease
EQ-5D	EuroQol five dimensions
HADS	Hospital Anxiety and Depression Scale
MI	Myocardial Infarction
MVPA	Moderate to Vigorous Physical Activity
PA	Physical Activity
RCT	Randomised Controlled Trial
REC	Research Ethics Committee
RPAQ	Recent Physical Activity Questionnaire
SAE	Serious Adverse Event
SAR	Serious Adverse Reaction
SOP	Standard Operating Procedure
SUSAR	Suspected Unexpected Serious Adverse Reactions
UHL	University Hospitals of Leicester

Contents

1 Introduction

This Statistical Analysis Plan (SAP) outlines the proposed statistical analyses for the MAP study, where a detailed and comprehensive description of the statistical methods is provided. The MAP Study is a randomised controlled trial that is being conducted to assess the effectiveness of a low-cost pragmatic intervention to increase daily physical activity at 12 months in participants with multimorbidity.

The structure and content of this SAP provides sufficient detail to meet the requirements identified by the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH): Guidance on Statistical Principles in Clinical Trials [1]. All work planned and reported for this SAP will follow nationally and internationally accepted guidelines, published by the American Statistical Association and the Royal Statistical Society for statistical practice [2,3].

The purpose of this SAP is to outline the planned analyses that are to be performed on the data to support the completion of the study publication for protocol V6 20_06_18. The SAP will be amended if substantial changes are made to the planned analyses, and in any case, will be finalized before the database lock for this study. Exploratory post-hoc or unplanned analyses not necessarily identified in this SAP may be performed on these data as required. These analyses will be clearly identified in the study publication.

The reader of this SAP is encouraged also to read the clinical protocol for details on the conduct of this study, and the operational aspects of clinical assessments and timing for the process of completing a patient in this study.

The outlined boxes in this SAP represent information from the protocol. The analyses will be carried out by the Trial Statistician (Ghazala Waheed) and supervised by the Senior Statistician (Prof Laura Gray) to ensure the integrity of the data processing at all stages.

1.1 Study Objectives

1.1.1 Primary objective

To assess the effectiveness of a low-cost pragmatic intervention to increase daily physical activity at 12 months in participants with multimorbidity

1.1.2 Secondary objectives

- To assess the effectiveness of a structured self-management education programme on improving health-related behaviours such as adherence to medication or treatment regimes, communication and physical wellbeing;
- To assess the effectiveness of the structured education programme on improving self-reported quality of life;
- To assess accessibility and effectiveness of implementing the programme in a population with multimorbidity.
- To assess whether the intervention affects development of existing long term conditions or incidence of new long term conditions in the 5 years following the study

1.2 Study design

1.2.1 Summary

The study is a randomised controlled trial (RCT) which will test the effectiveness of a tailored structured self-management education programme with text message support on the primary (overall volume of physical activity) and secondary outcomes compared with usual care.

Participants with two or more chronic conditions will be recruited from primary and secondary care and randomised to attend the self-management structured education programme or continue receiving usual care.

1.2.2 Randomisation

Participants will be randomised 1:1; stratified by gender (male; female) and ethnicity (White European; other) to either intervention or control groups.

If two people are taking part from the same household they will be randomised to the same arm to prevent contamination taking place.

Both control and intervention participants will be informed of the outcome of the randomisation. In addition intervention participants will be sent a letter confirming the date and the venue of the education sessions.

1.2.3 Study treatment interventions

This will be a randomised control trial with an intervention group and a control group.

1) Intervention group: Those randomised to the intervention group will be invited to attend four patient-centred education sessions of the structured education programme and to receive regular reminders with Short Message Service (SMS) messages (text messages) after the first session. The programme focuses on promotion of personalised physical activity and health-related self-management activities, which impact on quality of life.

These structured education sessions have been developed during an iterative process in conjunction with PPI work and other stakeholders through the use of existing infrastructure consisting of Diabetes Education and Self-Management in on-going and Newly Diagnosed (DESMOND) groups and a network of trained facilitators.

2) Control group: No change in the usual care of these participants will take place and they will continue with their routine care management in line with their direct care team's current recommendations for people with multimorbidity.

1.2.4 Inclusion Criteria

Patients will be included in the study based on the following criteria:

- Good understanding in written and verbal English;
- Able to give informed consent;
- 40-85 years old inclusive;
- Have two or more chronic conditions (see ***table1*** in MAP protocol v6;20.6.2018[4])
- Have access to a mobile phone for use in potential study activities;
- Able to walk independently;

There are no limitations with disease duration, medication use and/or initiation of new treatment unless there are specific recommendation for bed rest and minimising physical activity for a specific condition.

1.2.5 Exclusion Criteria

Patients with following will be excluded:

- Patients with limited understanding of written and verbal English;
- Patients with: dementia, learning disability, mental health disorders other than depression, epilepsy;
- Patients in palliative care;
- Pregnancy;
- Patients currently participating or participated in another interventional trial in the previous 12 weeks;
- Patients with frailty. The definition of frailty will be at the investigator's discretion and based on patients:
    - Living in care homes or institutions;
    - Having daily support for daily activities such as washing, cooking, household tasks etc;
    - Having had unintentional significant weight loss in the last 3-6 months;
    - Having BMI less than 18.5 kg/cm^2

1.2.6 Sample size calculations

The primary outcome is change from baseline to 12 months average daily physical activity, as quantified by the Euclidean norm minus one (ENMO) method measured in milligravity units (mg). This is the main measure of activity derived from the activity monitor. In order to detect a minimum clinically significant difference of 2.1 mg, which is equivalent to an overall increase in physical activity volume of approximately 30 minutes of light walking at 4km/h, assuming a standard deviation of 5.3 mg [1], a power of 80% and significance level of 5%, the sample size requires 202 participants. To allow for 20% loss to follow-up and 20% non-compliance of the activity monitor, we will therefore need to recruit 338 participants (169 in each group).

2 Outcome Measures

2.1 Primary outcome measures

The primary outcome measure is change in overall volume of physical activity measured using accelerometer between baseline and 12 months.

A complete case population will be used for the primary analysis. This will be followed by sensitivity analyses which will include an intention to treat (ITT) population by imputing missing data to assess the robustness of the findings to missing data.

2.2 Secondary outcome measures

Below is a list of all of the secondary outcomes collected at baseline and 12 months.

Ambulatory activity:

Patients will be asked to wear an accelerometer 24 hours/day for up to eight days to record total volume of physical activity, which includes light, moderate, vigorous and moderate to vigorous physical activities (MVPA).

Anthropometric measures:

- Weight
- Body mass index (kg/m^2)
- Waist circumference (cm)
- Hip circumference (cm)
- Waist to hip ratio

Clinical data:

- Blood pressure (mmHg)
- Muscle strength (measured by dynamometer)
- Resting heart rate (bpm)
- Total cholesterol (mmol/l)
- HDL cholesterol (mmol/l)
- LDL cholesterol (mmol/l)
- Kidney function (sodium, potassium, urea, creatinine and eGFR)
- HbA1c (mmol/mol, %)
- Medical history, details of any history of disease

Questionnaire Measures

A number of different questionnaire measures will be completed by participants at baseline and 12 months. Details of these questionnaire measures are as follows:

i. Lifestyle behaviours (self-reported): measured by dietary questions, sleeping pattern, smoking status and physical activity measured by RPAQ[5];
ii. Medication adherence: measured by ASK-12 (Adherence Starts with Knowledge (12 items survey)[6]
iii. Quality of life**:** measured by EQ-5D-5[7,8];
iv. Psychological health, i.e. depression and anxiety**:** measured by HADS[9];
v. Chronic Disease Self-Efficacy: measured by Managing Chronic Disease Scale[10];
vi. Self-efficacy for exercise: measured by Self-Efficacy for Exercise (SEE) Scale[11];

Smoking status, sleeping behaviours

No attempt will be made to impute the missing values for the secondary endpoints.

3 Analysis Sets/Populations

3.1 Missing values and outliers

Missing values represent a potential source of bias in clinical trials. Therefore, every effort will be made to ensure the data is fully explored for potential problems and have been correctly entered into the database. Participants who did not attend the baseline assessment after consent will be excluded from all analyses, but will be noted on the CONSORT flow chart. For the primary and secondary analyses, participants who do not have an outcome measurement will be excluded; as discussed in section 3.2. Sensitivity analyses will be performed to assess the robustness of the findings to missing data.
Outliers will be identified by examining graph plots such as normality plots, box plots and histograms. Values that visually appear outside of the main distribution (i.e. outliers) will be further investigated and if this is the correct observation it will remain in the analysis, however further assumptions will be made to assess possible influences on the results using a sensitivity analysis; by removing the outlier or outliers to compare results. Where results from the two analyses are discrepant, this fact will be reported and discussed in the study publication.

3.2 Complete Case Population

The level of missing data for each outcome will be assessed. If missing outcome data is present the initial analysis will be based on the complete cases. All tests of the effect of treatment on outcomes will be conducted on a complete case (CC) population. That is, all the recruited participants will be included in the analysis, with the exception of those with missing outcome data. This will be done on a 'by analysis' basis, i.e. only those participants with missing data for variables required for a specific analysis will be removed.

3.3 Intention-to-treat Population/Full analysis set

The intention-to-treat (ITT) analysis population will consist of all participants randomised. Participants will be analysed in the group to which they were randomly allocated and regardless of any protocol deviations or violations. Multiple imputation methods will be used to impute missing outcome data. This will be a sensitivity analysis.

3.4 Per-protocol Population

The per protocol (PP) population are those who were compliant with the protocol and have complete data for the analysis concerned on 'by analysis' basis. In the control arm, the PP population will include all participants randomised to that arm with complete outcome data. In the intervention arm, the PP population will be defined as participants who have attended at least one group session of the programme and have outcome with complete data. This will be a sensitivity analysis.

4 General Issues for Statistical Analysis

4.1 Multiplicity, Multiple Comparisons and Interim Analyses

All tests and reported p-values will be two-sided, where a p-value of < 0.05 will be considered to be statistically significant. Estimates will be presented with 95% confidence intervals. There will be no formal adjustment for multiple significance testing. The outcomes however, are clearly categorised by the degree of importance (primary and secondary). No formal interim analyses were planned.

4.2 Analysis Software

All data manipulations, tables, figures and analyses will be performed using Stata version 15.0 [12]. The validated program code (do file) and any other outputs created shall be documented and archived together.

4.3 GENEActiv data processing

All GENEActiv data will be processed by the physical activity team prior to statistical analysis. To be included in an analysis a participant needs to provide at least one valid days of GENEActiv data.

4.4 Demographic and Baseline Characteristics

Demographic and baseline characteristics for individual participants will be summarised by randomisation group. Numbers (with percentages) for binary and categorical variables and means (and standard deviations), or medians (with lower and upper quartiles) for continuous variables will be presented. The number of missing values will be reported in the footnote of the summary table. There will be no tests of statistical significance nor confidence intervals for differences between randomised groups on any of the baseline variables.

4.5 Primary Endpoint Analysis

4.5.1 Primary Analysis of Primary Endpoints

The primary outcome is change in overall volume of physical activity (average acceleration, ENMO) from baseline to 12 months. For the primary outcome treatment arm will be compared using linear regression modelling with

1. a binary indicator for randomisation group as the explanatory variable

2. terms for stratification factors (gender and ethnicity) as confounders

3. adjustment for the change from baseline in accelerometer wear time and baseline overall volume of physical activity.

The primary analysis at 12 months will be based on complete data.

4.5.2 Sensitivity Analyses

4.5.2.1 Per Protocol Population

For the PP analysis, participants who have engaged with at least one group session of the programme will be included. The PP will adhere to the same steps as the primary analysis.

4.5.2.2 Intention-to-treat Population

To allow for full analysis set, missing data will be imputed using a multiple imputation procedure which substitutes predicted values from a regression equation. The imputation will be carried out by the MI command in Stata [13]. MI replaces missing values with multiple sets of simulated values to complete the data, performs standard analysis on each completed dataset, and adjusts the obtained parameter estimates for missing-data uncertainty using Rubin's rules to combine estimates [14].

The following procedure will be followed:

- The MI will be set as wide
- The MI will register imputation of the overall volume of physical activity at 12 months and at baseline
- A regression method will be used , where the registered variables will be individually adjusted for the randomisation and stratification factors (sex and ethnicity), with 100 imputations to avoid biased estimates, rseed (2259) and the force options
- MI estimate will provide the final results using the same regression model as the primary analysis and covariate adjustments.

4.5.2.3 Exploratory analysis of number of sessions attended

We will assess the effect of the number of sessions attended on the primary and secondary outcomes of the physical activity using the same model used in the primary analysis. We will assess three alternative scenarios

I. <1, no sessions attended

II. 1-3 sessions attended

III. All 4 sessions attended

4.5.2.4 Stratified analyses

Interaction effects will be fitted between intervention arm and gender (male vs. female) and ethnicity (White European vs. other). If the interaction term is statistically significant at the 10% level then stratified analyses will be performed for that factor using the same model as the primary analyses.

4.6 Secondary Endpoint Analyses

The analyses of the secondary outcomes will be conducted in a similar manner as the main analysis using the appropriate model type, logistic regression for binary outcomes, linear for continuous and ordinal for ordinal outcome.

The assumptions of each analysis will be assessed and alternate parameterisations will be considered where appropriate.

4.7 Subgroup Analyses

We will conduct the following subgroup analyses to assess if the intervention effect is statistically different between these groups.

- Age
- Sex
- Median number of comorbidities
- Depression
- Arthritic condition
- Cardiovascular condition
- Type 2 diabetes
- Physical activity; low and high

The main analysis for each subgroup will involve using the same analytic principles as the primary analysis, but stratifying by the subgroup to estimate the mean difference in the overall volume of physical activity when compared to the control group. Further subgroup analyses will investigate the interaction effects between the treatment and subgroup, this will include an additional interaction term to the model; 'subgroup × treatment'. Subgroup results will be graphically presented as forest plots, with the estimated difference of overall volume of physical activity between treatment groups with 95% confidence interval for the interaction.

5 Safety Reporting

5.1 Adverse Events (AEs)/ Serious Adverse Events (SAEs)

All AEs/SAEs occurring during the study observed by the investigator or reported by the participant, attributed to the study, will be recorded on the CRF. The relationship of AEs/SAEs to the study will be assessed by a medically qualified investigator.

AEs will be recorded on the AE Record Sheet and periodically discussed by the study steering group committee as required. Any safety concerns arising from the team will be reported to the Sponsor as soon as possible.

All SAE information must be recorded on an SAE form and sent to the Sponsor using the appropriate reporting form and the contact details on there within 24 hours of the study team being made aware of the SAE.

Any adverse event/serious adverse event occurring whilst a participant is continuing in the study, whether or not attributed to the study, will be documented in the study publication.

6 References

1. ICH Topic E9: Statistical Principles for Clinical Trials. CPMP/ICH/363/96. September 1998. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002928.pdf Accessed on 28-September-2016

2. ASA. Ethical Guidelines for Statistical Practice. Prepared by the Committee on Professional Ethics. 1999.

3. RSS. The Royal Statistical Society: Code of Conduct. 1993

4. Dallosso H, Yates T, Mani H, Gray LJ, Dhalwani N, Baldry E, Gillies C, Cradock S, Batt M, Davies MJ, Khunti K. Movement through Active Personalised engagement (MAP)—a self-management programme designed to promote physical activity in people with multimorbidity: study protocol for a randomised controlled trial. Trials. 2018 Dec;19(1):576.

5. Besson, H., et al., *Estimating physical activity energy expenditure, sedentary time, and physical activity intensity by self-report in adults.* Am J Clin Nutr, 2010. **91**(1): p. 106-14.

6. Matza, L.S., et al., *Derivation and validation of the ASK-12 adherence barrier survey.* Ann Pharmacother, 2009. **43**(10): p. 1621-30.

7. *About EQ-5D. http://www.euroqol.org/about-eq-5d.html (accessed 15/09/2016).*

8. Herdman, M., et al., *Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L).* Qual Life Res, 2011. **20**(10): p. 1727-36.

9. Zigmond AS, *Snaith RP. The hospital anxiety and depression scale.* Acta psychiatrica Scandinavica 1983. **67**(6): p. 361-370.

10. Lorig, K.R., et al., *Effect of a self-management program on patients with chronic disease.* Eff Clin Pract, 2001. **4**(6): p. 256-62.

11. Resnick, B. and L.S. Jenkins, *Testing the reliability and validity of the Self-Efficacy for Exercise scale.* Nurs Res, 2000. **49**(3): p. 154-9.

12. StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP.

13. StataCorp. STATA Multiple Imputation Reference Manual: Release 13. 2013 [cited 2016 July 8]; 13:[Available from: https://www.stata.com/manuals13/mi.pdf]

14. Rubin DB. Inference and missing data. Biometrika. 1976 Dec 1;63(3):581-92.