Statistical Analysis Plan

SGLT-2 Inhibitor Empagliflozin effects on appetite and weight regulation: a randomised double-blind placebo-controlled trial (SEESAW)

Trial registration

MREC No: 16/EM/0040
ISRCTN: 82062639
ClinicalTrials.gov: NCT02798744
EudraCT: 2015-001594-40

SAP revision history

Date	Version	Justification for SAP version
12 April 2019	0.1	First draft for review
31 May 2019	0.2	Revised version of SAP based on study design
26 June 2019	0.3	Changes made for collaborators' and senior statistician's comments
28 June 2019	0.4	Changes made for collaborators' and senior statistician's comments
05 July 2019	0.5	Changes made for collaborators' and senior statistician's comments
12 August 2019	0.7	Changes made according to 3 arm parallel design allowing for multiple testing
26 September 2019	0.8	Changes made for collaborators' and senior statistician's comments
19 November 2019	1.0	Final version for review
12 December 2019	2.0	Final version

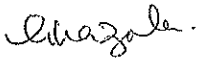
SAP responsibilities

Role in SAP development	Name, affiliation	Role in trial
SAP author	Ghazala Waheed, University of Leicester	Study Statistician
SAP reviewer 1	Laura Gray, University of Leicester	Study Senior Statistician
SAP reviewer 2	Melanie Davies, University of Leicester	Study Principal Investigator
SAP reviewer 3	Tom Yates, University of Leicester	Co-Investigator
SAP reviewer 4	Danielle Bodicoat, Danielle Bodicoat Ltd	Consultant Statistician

SAP signatures

Role	Name, affiliation	Date	Signature
Principal Investigator	Melanie Davies, University of Leicester	16/12/2019	
SAP authors			
Trial Statistician	Ghazala Waheed, University of Leicester	19/11/2019	
Co-investigator	Tom Yates, University of Leicester	16-12-19	
Supervising Statistician	Laura Gray, University of Leicester		

SAP signatures

Role	Name, affiliation	Date	Signature
Principal Investigator	Melanie Davies, University of Leicester		
SAP authors Trial Statistician Co-investigator	Ghazala Waheed, University of Leicester Tom Yates, University of Leicester	19/11/2019	
Supervising Statistician	Laura Gray, University of Leicester	16/12/2019	

1 Introduction

1.1 Trial background and rationale

Empagliflozin (Jardiance™) is one of several SGLT-2 inhibitors commercially available and licensed in the UK as a treatment for Type 2 diabetes. All of the SGLT-2 inhibitors have been shown to lower HbA1c and result in weight loss. However, observed weight loss is less than predicted for these agents, and the underlying mechanisms for this are unknown.

It is hypothesised that appetite hormones may play a role in this weight loss deficit. Those likely to be involved are peptide Y-Y (PYY) and ghrelin, as both are important in the control of appetite and weight regulation. They are secreted from intestinal L cells which are found in the distal small intestine. Ghrelin, which stimulates hunger, has been shown to increase with weight loss and energy restriction diets [1]. Glucagon-like peptide 1 (GLP-1) is another hormone which is involved in the regulation of appetite and satiety along with other glucose homeostatic effects [2]. There are no studies that have investigated the impact of SGLT-2 inhibitors on appetite hormones and effect on body composition. It is essential to understand these mechanisms in order to maximize the weight loss achievable with these agents. They may need to be combined with appropriate dietary measures such as energy restriction diets and weight-lowering or weight-neutral hypoglycaemic therapies for optimal clinical benefit.

The aim of the study is to explore the relationship between appetite hormones and Empagliflozin (Jardiance™) in order to understand the underlying mechanisms for observed weight loss which does not equate with that predicted for these agents.

1.2 Trial objectives/hypotheses

- The primary objective is to investigate the cause for the discrepancy in predicted and observed weight loss with Empagliflozin (Jardiance™) by measuring PYY, as a marker of appetite regulation.
- The secondary objectives are to determine the effects of Empagliflozin (Jardiance™) on other markers of appetite regulation, resting energy expenditure, total body weight and body composition.

2 Methods

2.1 Trial design

The trial is a randomised, double-blind placebo-controlled 4 arm parallel trial, in which participants are randomised (1:1:1:1) to one of the four arms at baseline: Empagliflozin 25mg once daily; Placebo only; Empagliflozin 25mg once daily plus energy restriction diet; Placebo once daily plus energy restriction diet. Participants are followed up for 24 weeks, with intermediate assessments after 2 weeks, 6 weeks and 12 weeks.

2.2 Randomisation

Randomisation is stratified for age [1.≤50years 2.>50years] and body mass index (BMI; [1. BMI 25.0-29.9kg/m2 2. BMI ≥30.0kg/m2]).

2.3 Sample size

The power calculation is based on the primary outcome of total PYY with a standard deviation of area under the curve (AUC) of 96.2pg/ml, based on a study looking at the influence of resistance and aerobic exercise on hunger and circulating levels of total PYY in healthy males[3]. To detect a minimum clinically significant difference of 120pg/ml in AUC of total PYY between groups, we will require 15 patients (60 in total) in each of the four arms to complete the trial with 80% power and 2-sided alpha of 1.7%. This will allow three comparisons between arms.

2.4 Framework

This is a 4 arm parallel trial with two interventions (Empagliflozin and an energy restriction diet). There will be three co-primary analyses: 1) Empagliflozin only versus placebo control, 2) energy restricted diet only versus placebo control, and 3) Empagliflozin & energy restricted diet vs placebo control.

2.5 Interim analyses and stopping guidance

None

2.6 Timing of final analysis

Analyses described in this SAP will be performed following completion of the trial and database lock.

2.7 Timing of outcome assessments

Outcomes are assessed at 24 weeks, with intermediate assessments at 2 weeks, 6 weeks and 12 weeks.

2.8 Blinding

This is a double-blind study. The treatment that is double-blinded is Empagliflozin. Both participants and study investigators are blinded to Empagliflozin allocation by way of a placebo medication that is identical in appearance to Empagliflozin. All measurements, including the primary outcome, are recorded with blinding to treatment allocation in place. Diet allocation is not blinded because it is not possible to blind this. The statistician performing the analyses will not be blinded. This is because the statistician needs to be able to identify the placebo group to know which of the four arms is the reference group for the three primary comparisons with placebo.

3 Statistical principles

3.1 Adherence

Adherence to randomised treatment was assessed at every study visit and documented in the CRFs. Adherence to Empagliflozin/placebo was checked by the study nurse counting in tablets that had not been taken. Participants are deemed to be compliant if they had taken 80% or more of their randomised medication.

Adherence to the energy restriction diet was assessed at visits 2-4 by checking the food diary. Participants are deemed to be compliant if they met their individual diet energy intake restriction threshold.

3.2 Analysis populations

3.2.1 Complete Case Population

All tests of the effect of treatment on outcomes will be conducted on a complete case (CC) population, using intention to treat principles. Patients will be excluded from analyses if they have missing data on any of the variables required for the analysis in question. Those with missing data at baseline will be included in the analysis using the missing indicator method. Detail is given in section 5.3.2. This population will be used for the main analysis of the primary outcome as well as all analyses of secondary outcomes.

3.2.2 Full analysis data set using multiple imputation for missing data

A sensitivity analysis using multiple imputation will be used to impute missing data, to assess the robustness of the conclusions to missing data. More detail is in section 5.3.2. This population will be used for sensitivity analyses of the primary outcome. It will not be used for secondary outcomes.

3.2.3 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 a 'by analysis' basis. The PP population will be defined as follows for each arm:

- Empagliflozin 25mg once daily: Participants must take at least 80% of the prescribed treatment for the duration of the study.
- Placebo only: Participants must take at least 80% of the prescribed (placebo) treatment for the duration of the study, and not be prescribed Empagliflozin by their GP. This will be determined by the concomitant medications recorded for each patient at each visit.
- Empagliflozin 25mg once daily plus energy restriction diet: Participants must take at least 80% of the prescribed treatment for the duration of the study, and meet their individual diet energy intake restriction threshold, following food diary checks, at visits 2-4. .
- Placebo once daily plus energy restriction diet: Participants must take at least 80% of the prescribed (placebo) treatment for the duration of the study, not be prescribed Empagliflozin by their GP, and meet their individual diet energy intake restriction threshold, following food diary checks, at visits 2-4.

Participants with a substantial protocol deviation logged, which in the opinion of the investigators might affect the primary outcome, will also be removed from the PP analysis. This population will be used for sensitivity analyses of the primary outcome. It will not be used for secondary outcomes.

4 Trial population

4.1 Eligibility criteria

Eligibility criteria are described in SEESAW_Protocol_v7.0_ 20.03.2018.

4.2 Recruitment

A CONSORT chart will display the flow of patients through the trial.

4.3 Baseline characteristics

The demographic and clinical characteristics and medical history will be tabulated and summarised by treatment groups and in total. This will include the stratification variables.

Numbers (with percentages) for binary and categorical variables and medians (with lower and upper quartiles and minimums and maximums) for continuous variables will be presented. There will be no formal comparison of baseline variables between treatment groups.

5 Analysis

5.1 Primary outcome definitions

The primary outcome is change in 3-hour time averaged AUC of PYY from baseline to week 24.

The PYY outcome is based on the response to a standardised mixed breakfast with data collected pre- and postprandially, specifically fasted (pre-standardised breakfast), 30 minutes post-breakfast and then 60 minutes, 90 minutes, 120 minutes, 150 minutes and 180 minutes post-breakfast. PYY AUC will be derived using the trapezium rule and then divided by time period to calculate the time averaged AUC. This represents the average PYY response across the postprandial period.

PYY 3-hour time averaged AUC responses at 2 weeks, 6 weeks, and 12 weeks follow-up will be analysed in the same way as the primary outcome and investigated as secondary outcomes.

5.2 Secondary outcome definitions

Below is a list of all of the secondary outcomes collected at baseline and at least one follow-up visit. Outcomes measured at 2, 6 and 12 weeks will be treated as secondary outcomes.

(i) Appetite hormones ghrelin and GLP-1 and appetite perception
Ghrelin and GLP-1 will each be recorded in the same manner as the primary outcome, i.e. as time averaged AUC.

(ii) Resting energy expenditure

(iii) Change in blood and urine biochemical parameters.

- Insulin time averaged AUC
- Glucose time averaged AUC
- Glucagon time averaged AUC
- C-peptide time averaged AUC

The above biochemical parameters will be measured at 7 time points specifically, fasted (pre-standardised breakfast), 30 minutes post-breakfast and then 60 minutes, 90

minutes, 120 minutes, 150 minutes and 180 minutes post-breakfast. Time averaged AUC values will be calculated as per that for PYY.

- Leptin
- HbA1c
- C-reactive protein
- Fasting plasma glucose
- Lipids (total HDL and LDL cholesterol, FFAs)
- eGFR
- urinary glucose excretion
- ALT
- Leptin

(iv) Physical activity measures:

- ActiGraph for assessment of daily step count and daily time spent sedentary and in light-intensity and moderate-to-vigorous intensity physical activity (including only those people with at least 3 days of data)

(v) Study questionnaires

- Three-Factor Eating Questionnaire – this will assess the patient's eating traits (restraint, disinhibition, hunger)
- International Physical Activity Questionnaire (IPAQ) –to assess walking, moderate and vigorous activities, along with time spent sedentary.
- Visual Analogue Scale (VAS) - which will measure perceived hunger, fullness, satisfaction and prospective food consumption ratings. VAS is measured at every time point as per the blood samples and so will also be analysed as an AUC measure, as per that for PYY.

(vi) Anthropometric measures:

- Body weight (kg)
- Body fat (%)
- Fat mass (kg)
- Bone mass (kg) and density (g/m^2)
- Muscle mass (kg)
- Waist circumference (cm)
- Hip circumference (cm)

(vii) Cardiovascular measures:

- Blood pressure (mmHg)
- Resting heart rate (bpm)

5.3 Analysis methods

5.3.1 Analysis of primary and secondary outcomes

A generalised linear model will be used with PYY 3-hour time averaged AUC at 24 weeks entered as the dependant variable and randomisation group entered as a categorical independent variable with the placebo control group as the reference. The effect of the intervention will be adjusted for baseline value of PYY and stratification variables (Age & BMI grouped as two categories as per the stratification method). Data will be checked for normality with an appropriate distribution (i.e. gamma) or transformation used if time averaged AUC PYY is not normally distributed, although we anticipate data will be normally distributed. The overall p value for the effect of treatment will be reported.

5.3.2 Confidence intervals and p-values

The threshold for a significant treatment effect for each intervention group vs placebo control (3 comparisons) will be determined through post-hoc tests using Holm's Sequential Bonferroni procedure with thresholds set at 0.017 for the most significant comparison, 0.025 for the next most significant comparison and 0.05 for the least significant comparison. Tests for significance will stop once any of these thresholds are met or exceeded. For consistency and to allow for a comparison of variance across groups, data will be reported as the mean adjusted intervention effect (change in the intervention of interest minus change in the comparator; mean, 95% CI). Inferences of significance will not be made based on the 95% CI.

5.3.3 Secondary outcomes

PYY time averaged AUC responses at weeks 2, 6 and 12 will also be analysed using the same methods as above. For summary purposes, Generalised Estimating Equations (GEE) taking account of repeated measures (2, 6, 12 and 24 weeks) with an exchangeable correlation matrix adjusted for baseline value and stratification variables (Age & BMI) will be used to derive an average PYY AUC response across time for each of the intervention groups compared to the control group, with significance assigned using the same method as that described for the primary outcome. Objective physical activity measures will be adjusted for ActiGraph wear time at baseline and follow up.

All other secondary outcomes will be analysed using the same approach as that described for the primary outcome. Models of ActiGraph measured outcomes will additionally be adjusted for ActiGraph wear time at baseline and follow-up.

5.4 Missing data

5.4.1 Missing data for AUC values – complete case analysis

This study includes AUC values as the primary and key secondary outcomes, with AUC values calculated using data collected over 7 time points; fasting (pre-standardised

breakfast), and at 30 minutes post-breakfast and then 60 minutes, 90 minutes, 120 minutes, 150 minutes and 180 minutes post-breakfast. Those with at least 4 valid values across the sampling time points will be included in the complete case analysis. Those with 4 or more missing data points will be excluded from the complete case analysis.

Those with at least 4 valid data points will have missing time point data using a regression method. Specifically, for each time-point a regression model will be fitted that includes BMI, age, ethnicity, randomisation group, and fasting PYY (excluded from the regression model for fasting PYY). Fitted estimates of PYY will be produced for each participant at each time-point based on the relevant regression model, and used to impute missing data.

5.4.2 Additional Analysis

For a sensitivity analysis, we will also deal with missing data in the primary outcome (PYY postprandial responses at 24 weeks) by running a Generalised Estimating Equations (GEE) model with an exchangeable correlation matrix to compare groups, allowing for repeated postprandial measurements from the same individual. This method allows all individuals with at least a single measure of PYY to be included. In the GEE model with repeated PYY values at 24 weeks as the response variable, the effect of the intervention will be adjusted for the fasting value at the baseline and stratification variables (Age & BMI). Gaussian family distribution with identity link will be used for the response and robust standard errors to provide consistent parameter and standard error estimates. Post-hoc treatment effects for the mean PYY response over the pre to post-prandial assessment period will be analysed according to the same *a priori* contrasts and correction for multiple testing as those reported for the primary analysis.

5.4.3 Multiple imputation

An additional sensitivity analysis of the primary outcome will be performed using multiple imputation which substitutes predicted values from a regression equation assuming Missing At Random. This will be carried out where AUC values are reported as missing based on the above criteria. The imputation will be carried out by the MI command in Stata [4]. 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 [5].

The following procedure will be followed:

- Multiple imputation will be carried out by chained equations to impute the entire PYY AUC values, using the variables PYY AUC, treatment group, baseline value and the stratification variables.
- The MI will be set as wide
- A regression method will be used, where the registered variables will be individually adjusted for the randomisation and stratification factors (BMI and age), with 100 imputations to avoid biased estimates, seed (2259) and the force options
- MI estimate will provide the final results using the same regression model as the primary analysis and covariate adjustments.

5.4.4 Per-protocol

The primary outcome analysis will be repeated when restricted to a per-protocol dataset (see Section 3.2.3)

5.4.5 Subgroup analyses for primary outcome

No subgroup analyses are planned.

5.5 Safety data

The number (%) of individuals experiencing either an adverse event or a serious adverse event will be summarised by randomised group.

5.6 Statistical software

Analyses will be performed using Stata version 15.0 or later [4].

6 References

1. Sumithran P, Prendergast LA, Delbridge E, Purcell K, Shulkes A, Kriketos A, Proietto J. Long-term persistence of hormonal adaptations to weight loss. New England Journal of Medicine. 2011 Oct 27;365(17):1597-604.

2. Madsbad S. The role of glucagon-like peptide-1 impairment in obesity and potential therapeutic implications. Diabetes, Obesity and Metabolism. 2014 Jan;16(1):9-21.

3. Broom DR, Batterham RL, King JA, Stensel DJ. Influence of resistance and aerobic exercise on hunger, circulating levels of acylated ghrelin, and peptide YY in healthy males. Am J Physiol Regul Integr Comp Physiol. 2009; 296(1):R29-35.

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

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