Statistical Analysis Plan - DIRECT

Study Name: A randomized controlled trial comparing daily calorie restriction versus intermittent fasting to improve glycaemia in individuals at increased risk of developing type 2 diabetes

Trial registration ClinicalTrials.gov ID: NCT03689608

SAP Author: A Vincent

SAP Version: 2

SAP Date: 20/12/2021

Approved by Trial Steering Committee

Leonie Heilbronn Leonie Heilbronn Version: 2 Date: 20/12/2021

Christine Feinle-Bisset Version: 2 Date: 20/12/2021

Andrew Vincent Version: 2 Date: 20/12/2021

Acronyms

CR	Calorie restriction
IF	Intermittent fasting
SC	Standard care
AUSDRISK	Australian type 2 diabetes risk assessment tool
PAS	Primary analysis set
FAS	Full analysis set
FinAS	Final (week 76) analysis set
SAS	Safety analysis set

Hypothesis

Intermittent fasting (IF) would produce superior glycemic benefits than daily calorie restriction (CR) at 6 months post-intervention.

Study Design Change

Changes made prior to 2nd Interim Analysis – accepted by the DMC. Primary investigator (LH) was blinded to the results of interim analysis 1.

Given the interruption due to COVID-19 it was considered ill advised to continue accrual. Regarding study design, the study team chose to consider change in glucose control as the single primary outcome and change in HbA1c as a secondary outcome, with no further interim analyses (for efficacy).

The original design had HbA1c and glucose AUC as dual primary endpoints and envisaged 3 interim analyses (at N/6, N/3 and N/2), with a multiple testing adjusted two-sided alpha=0.05 (ie two-sided alpha=0.025 for each endpoint). Assuming 25% attrition a N=260 would provide at least 80% power for the HbA1c endpoint and >90% power for the glucose AUC endpoint.

With this new design and assuming 25% attrition and a pre-post correlation of 0.4, then a sample size of N=208 (83:83:42 in the three arms) provides at least 85% power to detect a mean difference of 0.4 mmol/h (SD=0.8) in glucose AUC between the two active interventions (two-sided alpha=0.04999992).

At the first interim analysis (N/6) negligible alpha was spent. Assuming trial could stop for a Type I error in either outcome 6.3e-07 was spent (gsDesign).

Primary Outcome Measures

- (Planned) Dual HbA1c and post-prandial glucose AUC
- (Final) Post-prandial glucose AUC

Primary comparison

- Between CR and IF at week 24

Randomization

- Stratified blocked randomisation according to a 2:2:1 ratio (CR:IF:SC)
- Block length 3 or 6
- Stratification factors: Sex (F vs M) & AUSDRISK (12-19 vs 20+)

Cohorts

Primary Analysis Set (PAS) includes all individuals with a week 24 post-prandial glucose assessment. Individuals are included in the groups they were allocated to at randomization.

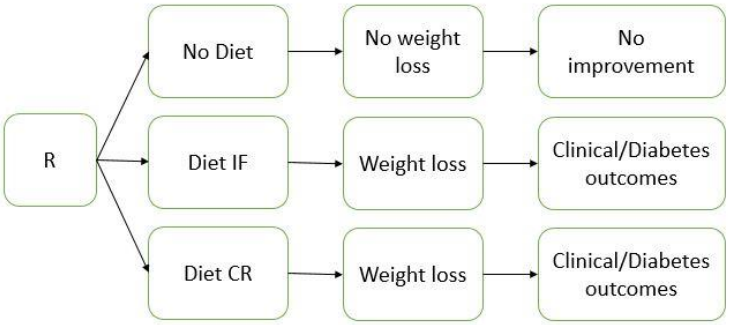
Full Analysis Set (FAS) includes all individuals randomized in the groups they were allocated to at randomization with at least one post baseline week 8 or week 24 assessment.

Final Analysis Set (FinAS) includes all individuals randomized in the groups they were allocated to at randomization with the week 76 assessment.

Safety Analysis Set (SAS) includes all individuals randomized in the groups they were allocated to.

Outcomes

Efficacy Measures



	baseline	8A	24A	76A
Diet/Behaviour				
Total energy intake	•	•	•	•
Carbohydrate intake	•	•	•	•
Protein intake	•	•	•	•
Fat intake	•	•	•	•
Fibre intake	•	•	•	•
Physical activity (steps count)	•	•	•	•
Weight Loss Measures				
Body weight (fasting)*	•	•	•	•
Waist circumference	•		•	•
Hip circumference	•		•	•
fat mass	•		•	•
fat free mass	•		•	•
Diabetes Measures				
Postprandial glucose AUC	•		•	•
HbA1c	•	•	•	•
Fasting glucose	•	•	•	•
Fasting insulin	•	•	•	•
Fasting NEFA	•	•	•	•
Fasting triglyceride	•	•	•	•
Postprandial insulin	•		•	•
Postprandial NEFA	•		•	•
Postprandial triglyceride	•		•	•
Clinical Outcomes				
Systolic blood pressure	•	•	•	•
Diastolic blood pressure	•	•	•	•
LDL	•	•	•	•
HDL	•	•	•	•
Cholesterol	•	•	•	•
hsCRP	•	•	•	•

* Non fasting weight also reported every two weeks, with fasting weight assessed as above.

Adverse Event Measures

- The maximum grade of adverse events will be calculated per individual per event category per time period (weeks 1-24 and weeks 25-76).
- Counts and percentages will be presented for each treatment group and overall.
- For each category and assessment period, the number of individuals will be compared between groups when there are at least 4 individuals with an event.
- Summaries and comparisons will also be performed for only adverse events that are considered at least *possibly related* to the intervention.

Data Presentation - Outcomes

All efficacy outcomes will have mean (SD) and median (IQR) reported for all treatment groups at all assessment times and change from baseline at post baseline assessments.

Cohort: FAS

Table: Measured outcomes at four visits (Baseline = Week 0, Visit 1 = Week 8, Visit 2 = Week 24, and Visit 3 = Week 76).

		Baseline*	Visit 1	Visit 2	Visit 3
Outcome 1					
Group 0	Mean (SD)	XX (YY)	XX (YY)	XX (YY)	XX (YY)
	Median [IQR]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]
	N (%)	XX (YY %)	XX (YY %)	XX (YY %)	XX (YY %)
Group 1	Mean (SD)	XX (YY)	XX (YY)	XX (YY)	XX (YY)
	Median [IQR]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]
	N (%)	XX (YY %)	XX (YY %)	XX (YY %)	XX (YY %)
Group 2	Mean (SD)	XX (YY)	XX (YY)	XX (YY)	XX (YY)
	Median [IQR]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]
	N (%)	XX (YY %)	XX (YY %)	XX (YY %)	XX (YY %)
Outcome 2					
Group 0	Mean (SD)	XX (YY)	XX (YY)	XX (YY)	XX (YY)
	Median [IQR]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]
	N (%)	XX (YY %)	XX (YY %)	XX (YY %)	XX (YY %)
Group 1	Mean (SD)	XX (YY)	XX (YY)	XX (YY)	XX (YY)
	Median [IQR]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]
	N (%)	XX (YY %)	XX (YY %)	XX (YY %)	XX (YY %)
Group 2	Mean (SD)	XX (YY)	XX (YY)	XX (YY)	XX (YY)
	Median [IQR]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]
	N (%)	XX (YY %)	XX (YY %)	XX (YY %)	XX (YY %)

*When more than one baseline assessment has occurred for some individuals, the within-individual means will be used.

Table: Change from baseline* outcomes at the three post-baseline visits.

		Visit 1	Visit 2	Visit 3
Outcome 1				
Group 0	Mean (SD)	XX (YY)	XX (YY)	XX (YY)
	Median [IQR]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]
	N (%)	XX (YY %)	XX (YY %)	XX (YY %)
Group 1	Mean (SD)	XX (YY)	XX (YY)	XX (YY)
	Median [IQR]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]
	N (%)	XX (YY %)	XX (YY %)	XX (YY %)
Group 2	Mean (SD)	XX (YY)	XX (YY)	XX (YY)
	Median [IQR]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]
	N (%)	XX (YY %)	XX (YY %)	XX (YY %)
Outcome 2				
Group 0	Mean (SD)	XX (YY)	XX (YY)	XX (YY)
	Median [IQR]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]
	N (%)	XX (YY %)	XX (YY %)	XX (YY %)
Group 1	Mean (SD)	XX (YY)	XX (YY)	XX (YY)
	Median [IQR]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]
	N (%)	XX (YY %)	XX (YY %)	XX (YY %)
Group 2	Mean (SD)	XX (YY)	XX (YY)	XX (YY)
	Median [IQR]	XX [XY, ZZ]	XX [XY, ZZ]	XX [XY, ZZ]
	N (%)	XX (YY %)	XX (YY %)	XX (YY %)

*When more than one baseline assessment has occurred for some individuals, the within-individual means will be used.

Primary Outcome Comparison

Objective: Week 24 post-prandial glucose AUC for CR and IF groups.

Cohort: PAS

Linear regression model: Y ~ Group (CR v IF) + Sex (M v F) + AUSDRISK + Y0

- Y = Post prandial glucose; Y0 = Baseline post prandial glucose assessment
- AUSDRISK and Y0 included as a continuous linear variables.
- Missing baseline covariate data are assumed to be infrequent and imputed using cohort means.
- Significance is one-sided $p < 0.04999992$ (two-sided).

	Est [95%CI]	**p-value**
IF v CR	XX [YY, ZZ]	PP

Secondary Analyses (week 24)

The primary outcome and all other efficacy outcomes with week 8 and 24 assessments will be analysed using repeated measures with a random intercept per individual. Efficacy outcomes without a week 8 will be analysed using linear regression of week 24 data only. The FAS cohort will be used, with individuals with no post baseline assessment for each outcome excluded per analysis. Missing baseline covariate data are assumed to be infrequent and imputed using cohort means.

Estimates and 95% CIs will be reported per group for all assessments (restricted maximum likelihood estimates), however pairwise comparisons will only be assessed if the overall effect of Group is significant (likelihood ratio test of nested models with/without Group). Significance will be considered at $p < 0.05$ (two-sided) with no adjustment for multiple testing and conclusions interpreted as exploratory.

Cohort: FAS

Model (wk 8): Y ~ Visit + Group + Visit x Group + Sex (M v F) + AUSDRISK + Y0 + (1|ID)

Model (no wk 8): Y ~ Group + Sex (M v F) + AUSDRISK + Y0

- Y = Week 8 and 24 assessments where available per outcome (see pg 4)
- Y0 = Baseline assessment
- Visit = week 8 and 24.
- Group: CR vs IF vs SC.
- AUSDRISK and Y0 included as a continuous linear variables.

Visual inspection of residual and random effect distributions will be undertaken. If residual distributions appear not satisfactory then log transformations will be applied and if still not satisfactory GLMMs will be considered with different error distributions.

	Group	**Visit (wk 8)**		**Visit (wk 24)**	
	p-value	**Est [95%CI]**	**p-value***	**Est [95%CI]**	**p-value***
Outcome 1 (with wk8)					
CR v SC	PP	XX [YY, ZZ]	PP	XX [YY, ZZ]	PP
IF v SC		XX [YY, ZZ]	PP	XX [YY, ZZ]	PP
IF v CR		XX [YY, ZZ]	PP	XX [YY, ZZ]	PP
Outcome 2 (no wk 8)					
CR v SC	PP			XX [YY, ZZ]	PP
IF v SC				XX [YY, ZZ]	PP
IF v CR				XX [YY, ZZ]	PP

*Only reported if the overall Group test is significant ($p < 0.05$).

Secondary Analyses (week 76)

Individuals are allowed to stop the intervention after week 24, as such the mean group and individual level effects may be quite different at week 76. Consequently we analyse week 76 data separately from the earlier weeks. All efficacy outcomes (see pg 4) will be analysed using linear regression models. The FinAS cohort will be used. Missing baseline covariate data are assumed to be infrequent and imputed using cohort means.

Estimates and 95% CIs will be reported per group for all assessments (restricted maximum likelihood estimates), however pairwise comparisons will only be assessed if the overall effect of Group is significant (likelihood ratio test). Significance will be considered at $p < 0.05$ (two-sided) with no adjustment for multiple testing and conclusions interpreted as exploratory.

Cohort: FinAS

Model: Y ~ Group + Sex (M v F) + AUSDRISK + Y0

- Y = Week 76 assessment
- Y0 = Baseline assessment
- Group: CR vs IF vs SC.
- AUSDRISK and Y0 included as a continuous linear variables.

Visual inspection of residual and random effect distributions will be undertaken. If residual distributions appear not satisfactory then log transformations will be applied and if still not satisfactory GLMMs will be considered with different error distributions.

	Group	**Visit (wk 76)**	
	p-value	**Est [95%CI]**	**p-value***
Outcome 1			
CR v SC	PP	XX [YY, ZZ]	PP
IF v SC		XX [YY, ZZ]	PP
IF v CR		XX [YY, ZZ]	PP
Outcome 2			
CR v SC	PP	XX [YY, ZZ]	PP
IF v SC		XX [YY, ZZ]	PP
IF v CR		XX [YY, ZZ]	PP

*Only reported if the overall Group test is significant ($p < 0.05$).

Bayesian Joint Secondary Analysis of Postprandial Glucose and HbA1c

The posterior probabilities of the between group difference in HbA1c for CR vs IF, and for SC vs IF+CR combined will be calculated using a sceptical bivariate Gaussian normal prior. This prior will be mean centred and with covariance

$$\Sigma = \begin{bmatrix} 0.14 & 0.08115171 \\ 0.08115171 & 0.096 \end{bmatrix},$$

corresponding to a correlation of 0.7 and variance of 0.14 for HbA1c and 0.15 for post-prandial glucose AUC. (Note there is an error in the protocol paper Teong et al 2020 where the post-prandial glucose variance was reported as 1.5 across standardized mean differences in historical studies. Converting to units of mmol/h we assume a SD of 0.8.)

We assume that observations from individuals in this study are exchangeable under the prior.

This analysis uses the PAS cohort.

To mimic the study design, the posteriors will be calculated in two stages: (A) initially including individuals used in the interim analysis and (B) then the subsequently randomized individuals. Each stage calculates the likelihood for the between group mean difference in the within-individual change from baseline values, assumed to be normally distributed with known variance. The posterior from the first stage is included as the prior for the second stage.

Reported will be

1. The mean and co-variance matrix for the posterior distribution
2. The probability that the change in HbA1c for
    a. IF vs CR is less than -0.3%, and vice versa.
    b. IF+CR vs SC is less than -0.3%.
3. The probability that the change in postprandial glucose AUC for
    a. IF vs CR is less than -0.4 mmol/h and vice versa.
    b. IF+CR vs SC is less than -0.4 mmol/h.
4. The joint probability that for IF+CR vs SC
    a. Either HbA1c is less than -0.3% or postprandial glucose AUC is less than -0.4 mmol/h.
    b. Both HbA1c is less than -0.3% and postprandial glucose AUC is less than -0.4 mmol/h.

Weekly weight assessments

There are non-fasting weight assessments reported every 2 weeks. These data will be analysed similarly to the other secondary analyses, with however piecewise linear effects assumed for the interventions over two time periods: weeks 1-24 and 25-76. The model will include random intercepts and slopes for individuals. Again pairwise comparisons will only be assessed if the interaction tests Time1 x Group and Time2 x Group are significant (likelihood ratio test).

Models: Y ~ (Time1 + Time2) + Group + (Time1 + Time2) x Group + Sex + AUSDRISK + Y0 + (1 + Time1 + Time2|ID)

- Time1 = Week number (1 to 76).
- Time2 = Zero for weeks 1 to 24, and week number minus 24 for weeks 25+.

Int. p-value	Group	Estimate	Coef [95%CI]	p-value*
PP	SC	Slope: wks 1-24	XX [YY, ZZ]	PP
PP	SC	Slope change: wk 25-76	XX [YY, ZZ]	PP
	CR	Slope: wk 1-24	XX [YY, ZZ]	PP
	CR	Slope change: wk 25-76	XX [YY, ZZ]	PP
	IF	Slope: wk 1-24	XX [YY, ZZ]	PP
	IF	Slope change: wk 25-76	XX [YY, ZZ]	PP
	CR v SC	Comparison: slope	XX [YY, ZZ]	PP
	CR v SC	Comparison: slope change	XX [YY, ZZ]	PP
	IF v SC	Comparison: slope	XX [YY, ZZ]	PP
	IF v SC	Comparison: slope change	XX [YY, ZZ]	PP
	IF v CR	Comparison: slope	XX [YY, ZZ]	PP
	IF v CR	Comparison: slope change	XX [YY, ZZ]	PP

*Only reported if the respective interaction p-values for slope and slope-change are significant.

Adverse Event Analyses

It is assumed that the number of AEs will be low, and as such Fisher exact tests will be used to compare: CR vs IF vs SC (ie two degrees of freedom) when the number of individuals with an event is at least 4.