**NUDGE: STATISTICAL ANALYSIS PLAN**

Study Name: A micro-randomized pilot study to examine the impact of just-in-time nudging on evening snacking in adults with type 2 diabetes

Trial registration number: ACTRN12622000531718

SAP Author: Andrew Vincent

SAP Date: 2023-03-31

**CONTENTS**

Preface

Study Objectives

* Primary
* Secondary
* Exploratory

Study Outcomes

* Primary outcomes
* Secondary outcomes

Study Methods

* Study design
* Study population
* Treatment assignment
* Sample size

Sequence of planned analyses

* Analysis cohorts
* Analyses

Statistical Considerations

* Imputation of missing data
* Protocol violations and deviations
* Covariate adjustment
* Multiple comparisons
* Statistical modelling

Descriptive Statists

* Flow chart (CONSORT flow diagram)
* Accrual/allocation issues
* Baseline characteristics

Statistical Methods

* Primary outcome
* Secondary efficacy outcomes
* Efficacy outcomes
* Safety outcomes

References

**ABBREVIATIONS**

|  |  |
| --- | --- |
| AHI | Apnoea-hypopnea index |
| AUC | Area under the curve |
| CGM | Continuous glucose monitoring |
| hr | Hour |
| iAUC | Incremental area under the curve |
| GEE | Generalized estimating equations |
| MRT | Micro randomized trial |
| NTNV | Night-to-night variability  |
| SAP | Statistical analysis plan |
| SES | Social economic status  |

**PREFACE**

This statistical analysis plan (SAP) describes the planned analyses and reporting for the micro-randomized pilot study to examine the impact of just-in-time nudging on evening snacking in adults with type 2 diabetes**.** The purpose of this SAP is to outline the considerations and the pre-specified analyses for the NUDGE study.

This project is funded by a Wellbeing SA Chronic Disease Integrated Partnership Grants (Round 1 2021) and Diabetes Australia (Y23G-WITG).

This study has been approved by the University of Adelaide Human Research Ethics Committee (H-2022-002).

**Study Objectives**

**Primary**

The primary objective of this study is to determine whether a pictographic digitally delivered nudge soon after dinner leads to a reduction in after-dinner snacking events and as measured objectively by a reduction in glucose after dinner as measured by continuous glucose monitoring (CGM).

**Secondary**

Secondary aims of this study are to assess:

1. The feasibility of the just-in-time nudge delivery platform.
2. The acceptability of the co-designed picto-graphic nudges.
3. The effect of just-in-time digital nudging on behaviour change with respect to the frequency of after-dinner snacking.

**Exploratory**

The exploratory aims of this study are to assess the relationships between:

1. After-dinner glucose levels and both mean and night-to-night variation (N2NV) in sleep quality factors.
2. Digital nudging and both mean and N2NV sleep quality factors.

### **Study Outcomes**

**Primary outcome**

Difference in incremental area under CGM curve (iAUC/hr) between nudging days and non-nudging days during the period 90 minutes post-dinner until 4am.

**Secondary outcomes**

*Glycaemic/Snacking outcomes*

1. Duration of the nudging intervention will be assessed by estimating both intervention-lag effects and the time-by-treatment interaction over the two weeks.
2. Influence of baseline characteristics (age, sex and social economic status (SES)) on treatment intervention will be assessed by inclusion of pairwise interactions with treatment.
3. Peaks in CGM glucose during the period 90 minutes post-dinner until 4am are assumed to reflect evening snacking behaviour, as such the frequency and magnitude of CGM peaks during the 2-week intervention period on nudging days will be compared to those on non-nudging days.
4. Nudge specific differences between the seven picto-graphic nudges will be assessed for both glycaemic response (iAUC/hr) and snacking behaviour (CGM peaks) over the 2-week intervention.
5. Mean within-individual differences in iAUC/hr and glucose peaks (number and magnitude) between the lead-in and intervention 2-week periods.

*Feasibility/Acceptability outcomes*

1. Feasibility of “just-in-time” messaging will be assessed as the time between nudge being sent and time it is observed, and the difference between dinner time and time observed.
2. Nudge acceptability is assessed by:
	* The mean responses to the two Nudge Content Questionnaire items and the frequency of categories of responses to the open-ended item.
	* Mean responses to the four sub-scales and the overall engagement score of the Nudge User Engagement Scale Questionnaire

*Sleep association outcomes*

1. Sleep quality measures include sleep duration and depth, timing of sleep onset, waking times, AHI and heart rate as measured by under-mattress Withing sleep analyser.
2. N2NV in sleep quality measures is assessed as the standard deviation of each measure.

**Study Methods**

### **Study design**

This is a single site micro-randomised trial (MRT) (Bidargaddi et al. 2018).

**Schedule of assessments**

|  |  |  |  |
| --- | --- | --- | --- |
| **Assessment items** | **Online screening****and consent** | **Lead-in** | **MRT** |
| Clinic visit 1 (Week -2) | Clinic visit 2 (Week 0) | Clinic visit 3 (Week 2) |
| Informed consent | √ |  |  |  |
| Eligibility screen | √ |  | √ |  |
| Medical history | √ |  |  |  |
| Demographics  | √ |  |  |  |
| Height,waist/hip/neck circumference, neck length |  | √ |  |  |
| Body weight, blood pressure |  | √ | √ | √ |
| Continuous glucose monitoring* iAUC 90minutes after dinner to 4 am
* Number and magnitude of episodic increments in CGM glucose
 |  | √ | √ |  |
| Sleep monitoring* sleep duration and depth
* timing of sleep onset
* number of awakenings
* AHI
* heart rate
 |  | √ | √ |  |
| Dinner timing |  | √ | √ |  |
| Food frequency questionnaire |  | √ |  |  |
| Eating habits questionnaire |  | √ |  | √ |
| Epworth Sleepiness Scale |  | √ |  | √ |
| Sleep Condition Indicators |  | √ |  | √ |
| Stanford Brief Activity Survey |  | √ |  | √ |
| Questionnaire-based exit interview* Nudge specific questionnaire
* User engagement score
 |  |  |  | √ |

**Study population**

|  |
| --- |
| ***Participant inclusion criteria***  |
| * Men and women with type 2 diabetes.
* Age 18–75 years
* Managed with diet or a stable dose of oral antidiabetic medications for at least 3 months
* Report habitual snacking after dinner most nights (3 or more per week)
* Resident in the greater Adelaide area
* In possession of and uses a smartphone
* Capacity to provide written informed consent and willingness to participate and adhere to the study protocol
 |
| ***Participant exclusive criteria***  |
| A personal history/diagnosis (self-reported) of:* Type 1 diabetes
* Individuals with advanced diabetes complications (renal dialysis, above ankle amputation, registered partially blind)
* Major psychiatric disorders (schizophrenia, major depressive disorder, bipolar disorder, borderline personality disorder)
* Eating disorder or restrictive eating pattern
* Any gastrointestinal condition or medication causing vomiting or affecting absorption.
* Current, planned or recently completed (within 6 months) treatment for cancer (excluding non-melanoma skin cancer)
* Insomnia
* Significant liver, kidney, cardiovascular or any other medical conditions that would in the opinion of the study physician adversely affect participation in the study (e.g., renal dialysis, liver cirrhosis, cardiac failure limiting activities of daily living, degenerative neurological disorder)
* Any other condition deemed unstable by the study physician
* Habitual evening consumption of alcohol other than one standard drink with dinner

Currently taking the following medications: * Use of insulin or sulfonylurea
* Glucocorticoids
* Antipsychotic medications
* Opioid medications unless combined with paracetamol in a single formulation and used occasionally on a pro re nata basis

Additional exclusion criteria include:* Shift-workers
* Pregnant or currently breastfeeding
* Weight stable at least 2 weeks prior to study participation
* Anyone unable to comprehend the study protocol or provide informed consent (i.e., due to English language or cognitive difficulties)
* Those who have not had a COVID-19 vaccination
 |

**Treatment assignment**

Within the two-week intervention period, the days for which a participant is available for nudging are block randomized with block size two to “nudge” or “no nudge” (equal allocation ratio). The seven co-designed picto-graphic nudges are then randomly allocated to each block (without repetition). If a participant indicates they are unavailable for nudging for some days then that participant will receive some, but not all of the seven nudges.

**Sample size considerations**

In planning the study we considered that a reduction of less than 10% in the probability of evening snacking would be unlikely to relate to a clinical benefit, and that the evening snacking rate would not be around 60%.

 As such, with fourteen days per individual and 50% chance to nudge/day, a micro-randomized trial with N=59 participants has at least 80% power (alpha=0.05 one-sided) to detect a decrease to 50% evenings with snacking (i.e. RR=1.25 from 40% to 50% snack free evenings) (Qian et al. 2021) (https://tqian.shinyapps.io/mrt\_ss\_binary/)

During the development process, it became apparent that identification of post dinner snacking episodes was more complicated than simply assessing post-dinner glucose peaks via CGM. As such, we decided to use glucose iAUC/hr (local-baseline adjusted; from 90 min post dinner to 4am) as the primary efficacy endpoint of the study as it is assumed that a reduction in evening snacking will result in lower post-dinner glucose levels.

Examination of the 2-week lead-in data indicated that: (i) the mean probability of night time snacking is approximately 85% and a decrease to 75% is a change in iAUC/hr of 0.8 mmol/hr; (ii) the iAUC/hr across individuals SD is about 2.4 and the between-week within-individual correlation is 0.8, suggesting that the SD of within-individual iAUC/hr change is likely to be less than 2.0 mmol/hr.

With these assumptions we note that in a paired t-test where the SD of the within-individual difference of 2.0 mmol/hr, then with N=54 there is >90% power to detect a reduction of 0.8 mmol/hr in iAUC/hr (alpha=0.05 one-sided).

Following Lewis et al 2021 we will use a tri-variate progression criteria (Lewis et al. 2021). (i) When the sample mean of the within individual change in iAUC/hr is increased in nudge vs no-nudge nights we consider the intervention a failure. (ii) When the iAUC/hr change is a >0.8 mmol/hr reduction (and thereby significantly different from zero) we consider the intervention a success. (iii) When the iAUC/hr change is a reduction but not greater than a 0.8 mmol/hr reduction we consider the intervention promising, but in need of further development.

With these criteria, if the true effect is a promising reduction of 0.4 mmol/hr then with N=54 the probability of correctly concluding that it needs further development is 86% and the probability of incorrectly concluding it is either a success or a failure is both 7%.

**Statistical Considerations**

**CGM baseline**

Glucose measured using CGM devices is known to be affected by pressure, e.g., due to lying on the device during sleep. This variation will be accounted for using a baseline estimation algorithm derived from (Brakel 2014). see Appendix A. In brief, this algorithm uses an iterative procedure to identify periods of elevated glucose from the local baseline mean. The algorithm adjustments using CGM data from a different study (Teong et al. 2020), and parameter settings were tuned using the lead-in CGM data.

**Glucose iAUC**

Glucose iAUC is defined as the difference of the CGM measured glucose and the algorithm estimated local CGM baseline.

**Glucose peaks**

During the post-dinner period, glucose peaks are defined as local peaks that occur during periods of elevated glucose (i.e., above the noise-threshold given the estimated local baseline mean). Local peaks are those glucose measures higher than both the previous and subsequent measurement. We note that glucose peaks are only a surrogate of snacking behaviour as during periods of sustained elevated glucose some peaks will occur due to imprecision in the measuring device (i.e., noise).

**Dinner time**

Dinner time is the time of the photo of dinner. If the photo indicates that the dinner has already been started this dinner-time is set to missing. If the photo is taken before 4pm or after midnight the dinner time is set to missing.

For each individual the mean dinner time over the four CGM weeks (lead-in and MRT intervention phases) is calculated and any missing dinner times and their within-block partners are imputed with this within-individual mean dinner time.

**Analysis sets**

The analyses of the glycaemic and snacking outcomes (including the primary outcome) will include all individuals that initiated the 2-week intervention phase and who have at least the first two consecutive days of CGM data available (i.e., the first nudge-no-nudge randomized block).

CGM data is considered to be available only if the entire period 90 minutes post dinner to 4am is recorded. Post dinner periods with incomplete CGM data are set to missing, and these blocks are not included in the analyses.

If the CGM data after adjustment of the baseline-algorithm are not considered to be believable, for example recording no change in glucose over many hours, then a per-protocol analysis cohort will be constructed excluding affected individuals.

The analyses of feasibility, acceptability and sleep association outcomes are complete case analyses, analysing all individuals with available data.

**Covariate adjustment**

All comparisons of treatment effect, i.e. nudge vs non-nudge days, of glycaemic and snacking outcomes (iAUC/hr and glucose peaks) will include adjustment for the within-individual mean level observed during the two-week lead-in period. Similarly analyses of sleep meaure outcomes assessed during the intervention period will adjust for that outcome assessed during the lead-in period. For the comparison of the lead-in and intervention periods we will adjust for age, sex and SES.

**Model assumptions**

Linear model assumptions, for example residual error distributions, will be assessed visually and if the assumptions appear violated, simple log-transformation of the outcome will be considered. If the transformation is unable to resolve the distributional issue, then analysis models assuming different error distributions will be considered.

**Statistical Significance**

Given this a pilot study, we prioritise type-II error control over type-I. As such, we use one-sided tests (alpha=0.05) for all glycaemic and snacking outcome analyses, and do not adjust for multiplicity of testing. Further, as noted in the sample size calculation, for the primary outcome we consider non-significant estimates of treatment effect in favour of nudge (vs no-nudging) to be “promising”, with the caveat that the intervention requires further refinement. For all other analyses of secondary outcomes, we use two-sided tests (alpha=0.05) and do not adjust for multiple testing. We consider significant results of all analyses except the primary analysis to be exploratory.

**Descriptive Statistics**

A CONSORT flow diagram will present the number of individuals who (i) participated in online screening, (ii) clinic visit 1, (iii) two-week lead-in phase, (iv) clinic visit 2, (v) enrolled in the two-week MRT phase, and (vi) completed the exit interview (clinic visit 3).



Descriptive summary statistics will be reported for baseline characteristics in individuals who participated in both the lead-in and the MRT two-week intervention phases.

Baseline characteristics that will be reported are listed in the schedule of assessments.

**Statistical methods for the primary outcome**

*Primary analysis*

P1: The primary outcome will be assessed as the within-individual means of the differences of each nudge day with the paired non-nudge day of the glucose iAUC/hr assessed by CGM over the period 90 minutes after dinner until 4am.

The primary analysis of the primary outcome is an adjusted linear regression of these means, adjusting for the within-individual means of iAUC/hr assessed during the two-week lead-in phase.

*Secondary analyses of the primary outcome*

P2: The mean nudge effect will be estimated using mixed effects models of the within-block difference adjusting for fixed a one-day lag factor, a linear time effect and the baseline lead-in AUC/hr and a random intercept and slope per individual.

P3: If there are CGM data that are not believed to genuinely reflect glucose levels, then analyses P1 and P2 will be repeated excluding all CGM measures from affected individuals.

**Statistical methods secondary outcomes**

***Glycaemic/snacking outcomes***

1. The influence of age, sex and SES will be examined by the inclusion of covariates in the P1 analysis.
2. Differences in peak frequency and magnitude (mmol) will be analysed using the same methodology as in P1 and P2 (referred to as S2a and S2b). It is expected that these outcomes will likely require log-transformation or alternative modelling assumptions (e.g., negative binomial for frequency).
3. Multi-day lag effects will be assessed by comparing likelihood ratio tests of nested models with/out the effects.
4. Nudge specific effects will be assessed by extending the analyses (P2) above to include nudge specific estimates. If there is a problem with model convergence, the equivalent mixed-effects model will be constructed with two crossed random effects: individual and nudge.
5. Analyses detailed in P1 and S2a will be repeated, however with the comparison between two-week lead-in and MRT phases of mean within-individual outcomes (iAUC/hr, peak frequency and peak magnitude).

***Feasibility/Acceptability outcomes***

1. A mixed effects model will be used to estimate the mean time differences across individuals between the time a nudge is sent and opened. This model will include a linear time fixed effect and random intercepts and slopes per individual. It is expected that this outcome will not be linear and likely need transformation.
2. Linear mixed effects models will be constructed of the two nudge content items. Each model will have random intercepts per individual and fixed effects for the seven nudges. Pairwise comparisons of differences between nudge acceptability will be performed only if the overall likelihood ratio tests of these two models with their nested sub-models without the nudge fixed effects are significant. Finally a nudge will be considered acceptable if both estimates are at least 50%.
3. Responses to the third open-ended nudge content item will be coded into categories and frequency of each category will be reported for each nudge.
4. The overall nudge user engagement score will be analysed by (generalized) linear regression models. The four subscales will be modelled using linear mixed effects models with random intercepts per individual.

***Sleep association outcomes***

1. The effect of glucose control on sleep quality will be assessed during the two-week lead-in phase, with a mixed model of sleep quality regressed onto the prior evening’s post-diner glucose iAUC/hr, adjusting for age, sex and SES with a random intercept per individual.
2. The effect of sleep quality on glucose control will be assessed during the two-week lead-in phase, with a mixed model of the daily iAUC/hr regressed onto the prior evening’s measures of sleep quality, adjusting for age, sex and SES with a random intercept per individual.
3. Associations between mean glucose control and N2NV in sleep quality will be assessed during the two-week lead-in phase with a regression of the mean iAUC/hr regressed on the standard deviation of the sleep measure, adjusting for age, sex and SES.
4. The effect of nudging on sleep quality is assessed during the two-week intervention period, with regressions of mean and N2NV sleep quality measures to determine if they differ between nights after a nudge day verses a no-nudge day. These regressions adjust for the sleep measure variable assessed during the lead-in period.
5. The effect of sleep quality on nudge effectiveness will be assessed by inclusion of both mean and N2NV in sleep quality measures as covariates in the P1 analysis.

**REFERENCES**

Bidargaddi, N., D. Almirall, S. Murphy, I. Nahum-Shani, M. Kovalcik, T. Pituch, H. Maaieh, and V. Strecher. 2018. 'To Prompt or Not to Prompt? A Microrandomized Trial of Time-Varying Push Notifications to Increase Proximal Engagement With a Mobile Health App', *JMIR Mhealth Uhealth*, 6: e10123.

Brakel, J.P.G. van. 2014. 'Robust peak detection algorithm using z-scores'.

Lewis, M., K. Bromley, C. J. Sutton, G. McCray, H. L. Myers, and G. A. Lancaster. 2021. 'Determining sample size for progression criteria for pragmatic pilot RCTs: the hypothesis test strikes back!', *Pilot Feasibility Stud*, 7: 40.

Qian, T., H. Yoo, P. Klasnja, D. Almirall, and S. A. Murphy. 2021. 'Rejoinder: 'Estimating time-varying causal excursion effects in mobile health with binary outcomes'', *Biometrika*, 108: 551-55.

Teong, X. T., K. Liu, A. T. Hutchison, B. Liu, C. Feinle-Bisset, G. A. Wittert, K. Lange, A. D. Vincent, and L. K. Heilbronn. 2020. 'Rationale and protocol for a randomized controlled trial comparing daily calorie restriction versus intermittent fasting to improve glycaemia in individuals at increased risk of developing type 2 diabetes', *Obes Res Clin Pract*, 14: 176-83.

**Appendix A:** CGM baseline and peak detection algorithm, and example of three consecutive nights from an individual during the lead-in phase.



****