Statistical analysis plan: Study Of Drinks With Artificial Sweeteners in People With Type 2 Diabetes (SODAS)

ClinicalTrials.gov ID: NCT03944616

Sponsor: University of California, Irvine

PI: Andrew Odegaard

Date: 4/30/24

Introduction

Background and rationale:

Diet beverages sweetened with artificial sweeteners occupy a unique category in the food environment as they are a source of intensely sweet taste with no calories. Diet beverages are the single largest contributor to artificial sweetener intake in the U.S. diet, and people with diabetes are the highest consumers of diet beverages, tending to consume them as a replacement for dietary sources of sugar, especially in place of sugar-sweetened beverages. This behavior has been endorsed by dietetic and scientific organizations, and diet beverages are marketed as being synonymous with better health, suitable for weight loss, and thus advantageous for diabetes control. The underlying public health concern is that there are few data to support or refute the benefit or harm of habitual diet beverage consumption by people with diabetes; therefore randomized trials with relevant outcomes must be conducted because they would address many limitations of previous research and inform dietary recommendations on diet beverage intake and primary and secondary prevention of chronic disease.

To address this gap the investigators tested the effect of diet beverage intake on diabetes control parameters in free-living adults with type 2 diabetes in a randomized, two arm parallel trial with a run-in period of 2-weeks and an active intervention period of 24-weeks. This study recruited and randomized 181 participants with type 2 diabetes who were usual consumers of commercial diet beverages and randomized them to receive and consume either: 1) A commercial diet beverage of choice (3 servings or 24 oz. daily); or 2) Unflavored bottled water of choice (sparkling or plain) (3 servings or 24 oz. daily). The primary outcome was the central measure of clinical diabetes control in glycated hemoglobin (HbA1c). The study also measured the effect of the intervention on the nature and magnitude of glycemic excursions via continuous glucose monitors, as well as clinical markers of cardiovascular and metabolic risk. Lastly, the study collected detailed data on measures of plausible mechanisms whereby diet beverage intake may alter risk by assessing the effect of the intervention on overall diet quality, components of dietary intake, and influences on dietary intake. Instruments measuring diabetes related quality of life, sleep, physical activity, and objective measures of physical activity were also collected. The study also collected repeated measures of the gut microbiome and stored biospecimens for metabolomic analysis in the future.

The following material is reproduced from the protocol and reflects further analytic details and updates in the protocol due to the Covid-19 pandemic and the evolution of the evidence base during main conduct of the study, with the intent of documentation prior to locking the dataset and carrying out analyses.

Objectives

The objective of this trial was to test the effect of removing the primary major source of artificial sweeteners in the diet by randomizing eligible persons with controlled type 2 diabetes (HbA1c ≤ 8.5%) who were habitual consumers of artificially sweetened diet beverages to substitute plain bottled water for artificially sweetened commercial diet beverages or to continue the habitual consumption of artificially sweetened beverages of choice. All beverages were provided and caffeine intake was kept constant within individual participants.

The primary clinical outcome was the main clinical measure of diabetes care – HbA1c.

The secondary clinical outcomes were other glycemic, metabolic, and cardiovascular disease risk measures.

Other main secondary outcomes include the effect on overall diet quality, components of dietary intake, and influences and preferences of dietary intake. As well as measures from diabetes-specific instruments assessing quality of life, medication usage (glycemic, lipid, blood pressure), and sleep (Pittsburgh Sleep Quality Index).

Study Methods

Trial Design

Multi-site (Irvine, CA and Minneapolis, MN), parallel, randomized trial, with a 1:1 ratio for allocation to intervention and control arms. The intervention is the substitution of plain bottled water for artificially sweetened diet beverages; the control arm is usual habitual intake of artificially sweetened diet beverages of choice. All beverages for both arms were provided to study participants. The duration of the trial was 24 weeks with a 2-week run-in period.

Participants could not be blinded to intervention status; the project statistician will be blinded to study arm allocation for all participants until analysis is complete for all primary and secondary outcomes (including planned sensitivity analyses as documented in this statistical analysis plan, unless otherwise noted).

Randomization

Participants were randomly assigned in a 1:1 ratio to the control (continue habitual diet beverage, 24 oz / day) group or the intervention group (consume plain bottled water in place of diet beverages, 24 oz / day). Randomization was performed with the use of permuted blocks and was stratified by site and sex. The randomization list was created prior to the start of the study and maintained by an independent statistician during the conduct of the study. The randomization list was not accessible to any other research team members.

Sample size estimation

The conservative power estimates were based on detecting a clinically significant 0.3 or 0.4 change in HbA1c percentage units for the primary aim. Based upon the study population HbA1c data from a pilot study, and two previously published studies of nutritional interventions in a population with T2D with highly similar inclusion criteria, we conservatively estimated the standard deviation of the change of HbA1c to be 1.0%. Based on this evidence, we hypothesized HbA1c reduction would be significantly greater among participants in the water consuming group than those in the diet beverage group with an expected effect size of 0.3-0.4 on the primary outcome. We set power to 80% and the 2-tailed alpha-level to .05. We designed the study for a planned repeated measures regression model (0, 12, 24 weeks) testing the between arm intervention effect on HbA1c change over 24 weeks (treatment x time interaction), with the baseline to 24 week HbA1c correlation estimated at r=0.6 to 0.8. These assumptions provide an estimated sample size requirement of 160 participants completing the study when incorporating the more conservative assumptions on the correlation of the outcome measures over time. We designed the study for an anticipated participant attrition rate of 20% and aimed to recruit 100 participants to each intervention (200 total) to achieve ≥80% power or higher for the primary aim. Overall, we randomized 181 participants, and 179 participants provided complete data on all measures.

For the secondary outcomes, we will have reasonable power if the intervention has a medium to large effect size on those variables (i.e., effect size ≥ 0.4), but will have limited power for those that have a relatively small effect size (i.e., effect size < 0.4). Specifically, making similar assumptions as noted above, our estimated sample size will have 80% or higher power to detect a between-group difference of 4.4 μmol in Fructosamine, 5.0 ml/min in eGFR, 3.2 mmHg in SBP, 2.3 mmHg in DBP, 20 mg/dl in triglycerides, 9 mg/dL in LDL, 2.6 mg/dL in HDL, 9 mg/dL in total cholesterol, 4.8 units in diet quality score, and 0.8 kg in weight at 24 weeks. We expect to have similar high-powered analyses related to the primary and secondary CGM metrics of interest – time in range (time in hyperglycemia), and CV, and secondary clinical cardiovascular (Apo-B, ApoA1, Fibrinogen) and metabolic (liver function panel, TSH).

Framework

Superiority comparisons, for both primary and secondary outcomes.

Statistical interim analysis and stopping guidance

No interim analysis was planned or conducted, and stopping rules were not used in the study.

Timing of final analysis

Final analysis of the trial outcomes (as described in this plan) will be conducted once all follow-up data has been collected, processed, and collated in the REDCap data management system. Analysis of primary and secondary outcomes (including specified sensitivity analyses) will be completed prior to unblinding the data.

Timing of outcome assessments

Outcomes were assessed for all measures at baseline (0), 12 weeks, and 24 weeks post baseline. Measures which are sensitive to changes in shorter time periods were also measured at interim study visits 6 and 18 weeks post baseline (e.g., blood pressure, Fructosamine, weight).

Statistical Principles

Confidence intervals and P values

All tests will be conducted using two-sided tests, at a 5% significance level; corresponding confidence intervals for effect sizes will be 95% CI (again two-sided). No adjustment will be made for multiple testing.

Adherence and protocol deviations

Self-reported adherence will be defined as a percentage of total intervention days during which participants reported adhering to their randomized intervention (H2O: consuming at least 24 ounces / day plain water and avoiding artificially sweetened diet beverages) and (Diet beverages: consuming at least 24 ounces / day artificially sweetened diet beverages of choice). This was assessed by participants completing a weekly survey recording whether they have consumed the full provided amount of beverages, and if they didn't, the % of beverages provided or if randomized to water, any consumption of diet beverages. A significant deviation is defined as a participant who reported consuming the assigned intervention less than 80% of the time. Similarly, we will also leverage the 5, unannounced 24-hour dietary recalls during the active 24-week period. A significant deviation of protocol is defined as not consuming the assigned intervention in < 80% of the recalls.

Objective adherence was also assessed by urinalysis of spot urine samples obtained in the morning of primary clinical data collection visits after an overnight fast. Urinalysis for the presence and levels of the spectrum of FDA approved artificial sweeteners was performed, where presence would indicate intake within previous 24 hours. Measures were collected at weeks 0, 12, 24. Because levels of artificial sweeteners detected in urine do not necessarily reflect the amount of intake of diet beverages sweetened with artificial sweeteners and may also represent passive intake via foods with artificial sweeteners, there are not reference levels as defined thresholds. Participants randomly assigned to consume diet beverages will be defined as adherent with the presence of any artificial sweetener in the urine sample. Participants randomly assigned to substitute bottled water for diet beverages will be defined as adherent if there is no presence of any artificial sweetener in the urine sample. These objective measures will be integrated with the self-reported measures so a more robust per-protocol analysis will be able to be conducted.

Analysis populations

Intention-To-Treat (ITT)

The main analyses for all primary and secondary outcomes will be conducted on an intention-to-treat basis, using all randomized participants. Randomization was predicated on meeting all eligibility criteria for the study (clinically measured and self-reported data, completion of 2-week run-in data collection). In this population, treatment will be assigned based on the arm to which participants were randomized, regardless of adherence or follow up visit attendance.

Per-Protocol

The Per-Protocol (PP) population is the subset of the ITT population and will consist of all participants with no major protocol violations and follow up data. This analysis will also exclude the participants who were impacted by a Covid-19 infection and had to alter participation for data collection.

Trial population

Eligibility

We included men, women and non-binary participants with T2D, age 35 years and older, able to provide informed consent, otherwise healthy, who met the following criteria:

- Physician diagnosed type 2 diabetes ≥ 6 months prior to screening
- HbA1c 6.5-8.5% at participant screening
- Current treatment with lifestyle changes or stable diabetes-related medication levels for the past 3 months
- Willingness to provide consent to contact treating physician and physician agreement to refrain from changing diabetes-related medications during the trial (change defined as > 2 fold change in dose of any 1 hyperglycemic agent or addition or subtraction of an agent)
- No physician-directed medication change for 3 months if prescribed medication for lipids or blood pressure
- Usual consumers of diet beverages (≥ 3 servings/ week (24 oz.) and the willingness to maintain fidelity of the intervention, and participate in all aspects of the intervention
- Not actively looking to make major lifestyle alterations during the study period with stable weight for 2 months (within 3%).

Exclusion Criteria:

- Type 1 diabetes or suspected type 1 diabetes (lean with polyuria, polydipsia, and weight loss with little response to metformin)
- "Secondary" diabetes due to specific causes (e.g. monogenic syndromes, pancreatic surgery, and pancreatitis)
- Diabetic Ketoacidosis hospitalization within last 6 months
- Severe/major hypoglycemia in the last 3 months-severe/major hypoglycemia is defined as a hypoglycemic event in which patient requires assistance of another person to manage the episode
- Glucocorticoid use (prednisone 2.5 mg/d or more or its equivalent)
- History of intolerance or allergy to diet beverages or AS or phenylketonuria
- Any condition that is known to affect the validity of the glycemic measures (Hba1c)
- Major cardiovascular disease event or surgery within past 6 months
- Gastrointestinal disease
- Renal or liver disease
- Current treatment for cancer
- Those with major surgery planned or history of bariatric surgery
- Antibiotic treatment (> 6 days) within past 6 months
- Currently pregnant (via self-report) or planning to become pregnant during study period; <1 year postpartum and breast feeding
- Current participation in another interventional clinical trial
- Previous randomization in this study,
- Heavy alcohol consumption (on average >2 drinks/day for women and >3 drinks/day for men)
- Habitual consumer of SSB ≥ 1 serving / day (8 oz.)
- Does not drink diet beverages
- BMI < 20.0 kg/m2

Recruitment

Participants were recruited via multiple avenues – medical records, physician referral, advertisements, recruitment lists at the respective institutions.

Withdrawal/follow-up

Participants were able to withdraw/disengage from the intervention but still remain in the trial for follow-up measurement. Any participants withdrawing their consent for data to be used in analysis will be removed from

analysis. Participants will be separately enumerated as “withdrawals” if notification is given that the participant no longer wishes to participate in the trial; and “lost to follow up” if the participant could not be contacted or did not provide data (after the research team has completed the follow-up contact protocol). Since randomization occurs immediately following baseline measurements (but prior to engaging with the intervention) withdrawal/loss to follow-up will be reported at the following time-points: between baseline and first primary follow-up (12 weeks); and between first primary and final follow-up (24 weeks).

Analysis

Baseline characteristics

Participant baseline characteristics will be presented for both study arms, for the factors used in the stratified randomization (study site and sex) and for baseline sociodemographic variables (age, gender, race/ethnic group, education) and clinical characteristics, (medication effect score (glycemia), duration of type 2 diabetes, N (%) using insulin therapy, therapeutic intensity score (blood pressure), lipid treatment (intensity category-none, low, moderate, high), % Using medication for managing other condition(s), % with history of infection of COVID at baseline, BMI (kg/m^2), weight, systolic and diastolic blood pressure, LDL, HDL, triglycerides, eGFR, HbA1c, fructosamine, fasting glucose, fasting insulin, Diet Quality (HEI estimate from baseline/run-in period of unannounced 24-hr recalls), Energy intake (baseline/run-in period), Usual artificial sweetened beverage intake (servings / week), Class of habitual intake sweetener at baseline (%: aspartame/Ace-K, Sucralose/Ace-K, Aspartame/Sucralose/Ace-K, Sucralose, Stevia, other), % consuming caffeinated ASB and caffeine levels, physical activity levels, Pittsburgh Sleep quality Index.

These baseline results will not be compared with any formal statistical tests. Descriptive statistics will be used to present appropriate summary statistics by study group for continuous measures (mean, standard deviation if variable is approximately normally distributed; median and interquartile range if data are skewed) and for categorical measures (frequency and percentages).

Outcome definitions

Primary: HbA1c at 24 weeks. Measures occurred at baseline (0), 12-, and 24-weeks post randomization.

Secondary:

1) continuous glucose monitoring (CGM) metrics (see below) Measures occurred over 2-week run-in (baseline), weeks 11-12, weeks 23-24 post randomization.

***Inclusion*:**

The range of CGM data for inclusion in this study will be 5 to 14 days for each period, consistent with manufacturer's recommendations. All CGM data that meets this criterion should be included in the final analysis, but the proportion of participants who meet 70% data-obtainment (10 days) during the 14 days should also be reported as part of the data completeness. We will also report proportion of participants who meet 50% (7 days) data-obtainment.

Outcomes: Use week 24 as endpoint, and adjust for baseline, (and include week 12) for any repeated measures/mixed models

-Time in range indicates the amount of time that glucose readings are within a defined target glucose range of 70–180 mg/dL (3·9–10·0 mmol/L (*This is primary CGM metric)*

-Time in tight range is defined as the percentage of time that glucose readings are within 70–140 mg/dL (3·9–7·8 mmol/L)

-Time below range refers to the amount of time that glucose readings are below the target glucose range of less than 70 mg/dl (3·9 mmol/L

-Time above range refers to the amount of time that glucose readings are above the target range of more than 180 mg/dL (10·0 mmol/L

-The coefficient of variation is a measure of glucose variability that is correlated with time below range and is calculated as 100 × (SD divided by mean glucose)

-Mean glucose: A measure of the mean 24 h glucose concentration calculated across all recorded glucose readings during a wear period

- The SD of mean glucose values is a measure of dynamic glucose variability; SD is strongly correlated with mean glucose

Further reporting:

The change in time in range will be reported separately for each study group from beginning to end of the study, and the difference between the study groups should be compared statistically with adjustment for the baseline value, and repeated measures and noted covariates.

Interpretation: A difference of ≥3% (absolute percentage points) in time in range is considered clinically meaningful for a treatment group difference in mean time in range

2) Clinical: Fructosamine, fasting glucose, fasting insulin, weight, blood pressure, fasting lipid panel (total cholesterol, HDL, LDL, triglycerides), Apo-B, Apo-a1, Fibrinogen, kidney function (eGFR via albumin and cystatin-C), liver function panel, TSH, C-reactive protein.

3) clinical related: Outcome is score at 24 weeks (MES and TIS) Medication effect score (MES, glycemia, 0, 6, 12, 18, 24 weeks), Therapeutic intensity score (blood pressure, 0, 6, 12, 18, 24 weeks). For any changes in lipid-intensity therapy we will score participants according to therapeutic intensity category (0-none, 1- low intensity, 2- moderate intensity, 3- high intensity), and examine changes in mean over course of study (0, 6, 12, 18, 24 weeks).

4) participant-reported measures:

*Data from unannounced 24-hour dietary recalls (7 total)**:*** Recalls during run-in (2) represent baseline diet (time 0), Recalls post-randomization until week 12 (2-3 recalls) summarized as week 12, recalls > week 12 – 24 (2-3 recalls) summarized as week 24. Primary dietary intake metrics for baseline, week 12, week 24: Overall diet quality (HEI), total calories, artificial sweetener intake, added sugars. Other carbohydrate-related measures: total CHO (% energy), CHO quality (amount of whole grains, Glycemic Index/Glycemic Load (GI/GL), fiber. For descriptive purposes, we will also report overall macronutrient composition (protein, fat).

Other metrics derived from instruments:

-Food Craving Inventory

-Dietary Practices: modified UK Diabetes and Diet Questionnaire

-Pittsburgh Sleep Quality Index

-Diabetes-related Quality of Life (DHP-18 + ADDQol)

-Medication Effect Score (glycemia), Therapeutic Intensity Score (Blood pressure), Lipid therapy intensity (lipids)

Analysis methods

All primary analyses will be based on intention-to-treat principles (with individuals analyzed in the group to which they were randomized). Demographic and baseline characteristics will be summarized using descriptive statistics. All outcomes will be analyzed as continuous variables.

Analysis will use linear mixed models, with randomization group, categorical time points, and the interaction between randomization group and time to estimate mean outcome levels between intervention and control arms at the 24-week endpoint, adjusted for baseline outcome level (accounting for baseline differences in outcome) and for other important baseline covariates (sex, study site, + prespecified covariates). This analysis will include the 0,12 and 24-week measurements available for all the participants, taking advantage of all available data (mixed models analysis treats missing outcome measurements at certain timepoint as missing

at random conditional on the other outcome levels at other time points and baseline covariates, i.e. these individuals are expected on average to have outcomes similar to other people with the same baseline covariates and outcome trajectories). As the study involves repeated measures data, these models will include random effects for individuals to account for intraclass correlation between measurements from the same person at different follow-up times.

A-priori specification of adjustment covariates beyond randomization stratification criteria (sex, site)*:*

-For primary glycemic, CGM and weight outcomes: Medication effect score (MES).

-For secondary outcomes of clinical CVD risk– blood pressure, lipids, apo-b, Apoa1, fibrinogen: (MES, blood pressure therapeutic intensity score (TIS) for blood pressure), (MES, intensity of lipid therapy for lipid/lipoprotein and fibrinogen)

-For secondary outcomes -metabolic and inflammatory related – kidney and liver function, TSH, c-reactive protein: (MES, and TIS for kidney function), (MES, TIS, intensity of lipid therapy for liver function), and (MES, TIS, intensity of lipid therapy for TSH and c-reactive protein)

-For secondary outcomes of Diet Quality and related components + food craving inventory: MES, TIS, intensity of lipid therapy and COVID-19 infection

-For secondary instrument-based outcomes of DHP-18 + ADDQol, MES, sleep (PSQI): Adjust for MES, TIS, intensity of lipid therapy, age and race/ethnic group when DHP-18 and ADDQol are outcome measures. When PSQI is outcome measure, adjust for MES, TIS, and caffeine. When MES is outcome, adjust for age, race/ethnic group, duration of diabetes

Missing data

Missing data will be reported in both the summary of baseline data and for follow-up data. At this point we do not expect to be missing covariate data at baseline, but any missing covariate data will be filled in with single-mean imputation for all analysis models, which allows inclusion of all participants. This will be implemented as part of the main analysis

It is assumed that individuals missing outcome data will be missing all subsequent follow-up data at a given time point.

Missing outcome data will only be subject to sensitivity analysis for the primary outcomes (HbA1c and CGM metrics at 24 weeks). These results will be considered as a sensitivity analysis, with the “standard” linear mixed models providing the main results (as described above), which accounts for missing data using a likelihood-based approach under the Missing At Random assumption. The sensitivity analysis will use multiple imputation methods to fill-in any missing outcome data for the outcomes conditional on the nature of the missing data and baseline data (age, sex, site, race/ethnicity, HbA1c, MES, trial arm)

Additional analyses

Per protocol – naïve – exclude any defined non-adherence. This approach provides valid estimates if protocol deviations are noninformative (i.e., completely at random within treatment groups). When protocol deviations are not completely at random within treatment groups, an unstandardized per-protocol effect estimator can be biased.

If we observe < 80 per arm participants who are defined as non-adherent or if there is evidence that the protocol deviations were not completely at random within treatment groups, we will also employ a SAS-macro G-formula RCT and compare results from this to the naïve PP, as this macro and analytic approach employs methods which appropriately allow for time-varying confounder adjustments. This approach is a generalization of standardization that can adjust for both baseline and time-varying confounders and is useful when there are time-varying factors that may be linked to past and future adherence and be associated with outcomes of interest. We define these time-varying adjustments beyond the main analyses, as the MES, covid positive or hospitalization for other health related reasons, baseline diet beverage habits, diabetes quality of life measures from the instruments, food craving inventory scores and weight. For this analysis we will vary the stringency of

the definition of a protocol deviation (70, 80, 90% reported adherence) + inclusion of urinalysis for objective component.

Mediational Analyses: In addition to the analysis evaluating the treatment effects, potential mediating effects of CGM metrics, weight change and diet quality measures + caffeine in the association between treatment and primary (or secondary) outcomes will be evaluated using a SAS macro that can calculate the point and interval estimates for the proportion of treatment effect (PTE) explained by one or more intermediate variables (Mediate SAS).

Sensitivity analyses

-Exclude participants on insulin therapy

-For primary glycemic, weight and CGM outcomes, consider the effect on the outcome measures of further covariate adjustment hypothesized to possibly predict glycemic/weight measures include ASB blend/class, amount of ASB pre-study, history of COVID infection (if tested positive and treated pre-study within 3-months of randomization defined as positive at baseline, if positive and treated in study period post baseline and < week 12 defined as positive at 12, if positive and treated > week 12 and before week 24 visit, defined positive at 24), caffeine levels.

-For secondary clinical measures we will also employ the same Covid infection covariate for all outcomes and include caffeine levels for blood pressure.

-Examine whether the category/type of diet beverage matters for the intervention in respect to primary or secondary outcome measures (e.g., do colas, non-colas/juices, type of sweetener combo matter) Class of habitual intake sweetener includes (aspartame/Ace-K, Sucralose/Ace-K, Aspartame/Sucralose/Ace-K, Sucralose, Stevia, other). We expect to have limited inference beyond aspartame/Ace-K and Aspartame/Sucralose/Ace-K as other formulations are rarer. Thus, this will be exploratory.

-Account for duration of diabetes (does any effect of the intervention depend on the duration of type 2 diabetes?)

-Examine whether habitual artificial sweetened beverage intake habits prior to study influence primary and secondary outcome measures (i.e. do results differ by frequency, amount, and duration of intake?) (use 2 servings a day as threshold, > 16 oz / day).

-We will consider an analysis stratified by sex to inform whether there is any evidence for a differential effect by sex, and thus handling sex as a biological variable

-A modified intent-to-treat analysis will include all participants in the ITT population with at least 1 post-baseline clinical data collection visit.

Statistical software

Statistical analysis will be conducted using SAS 9.4.