# STATISTICAL ANALYSIS PLAN

1. Original Statistical Plan: The funding provided by UL1 TR001419 was administered through the UAB Center for Clinical and Translational Science (CCTS), which required a preassigned statistician to draft the statistical analysis plan, as a stipulation of applying for funding. The original statistical analysis plan drafted by the required statistician was:

> Analyses will be performed blinded and using two-sided tests with a Type I error rate of $\alpha$=0.05. All data will initially be checked for missing values, cleaned, and inspected to determine the ranges, identify outliers, and check for concordance with parametric statistical assumptions (continuous variables only), including normality and homoscedasticity. Data will be analyzed using linear (or generalized linear) mixed models, including ANCOVA models. Data for endpoints that involve repeated measures will also control for baseline values as covariates, following procedures outlined by Allison [42]. These endpoints will also be tested for associations with other potential covariates, such as biological sex, race, BMI, age, and chronotype, and significant covariates will be included in our models. The fat loss and lean mass retention analyses will be performed per-protocol, while all analyses with repeated measures will be conducted using intention-to-treat analysis and will be followed by per-protocol sensitivity analyses. Categorical variables (from the surveys/interviews) will be analyzed using the chi-square test, or Fisher's exact test if the assumptions for the chi-square test do not hold. We will try to recollect any missing data, but that which remains may be treated using multiple imputation as appropriate.

2. Amendments: After the grant was awarded, the PI was allowed to switch the statistician and did so. Since the original statistical plan was unclear or ambiguous in several respects, we made the following modifications and interpretations after the trial commenced and under the guidance of the new statistician:

*Modifications:* Instead of adjusting each individual endpoint for its unique set of statistically significant covariates, we adjusted all analyses for a single set of covariates: sex, age, and race. Adjustment for these covariates did not change the statistical significance of any endpoint when compared to unadjusted analyses. We followed principles in Allison (1995), and we adjusted for baseline values only when the baseline value correlated with the change score and there was a baseline imbalance. Second, to be consistent with research norms, questionnaire data were analyzed in completers-only without adjustment and were not imputed.

*Clarifications and Interpretations:* In the 6th sentence, we clarify that "per-protocol" meant "completers-only." We interpreted the 3rd sentence to mean that linear (or generalized linear) mixed models will be applied to repeated measures data, while ANCOVA or the equivalent will be applied to non-repeated measures data. Since it was ambiguous which outcomes should be treated as repeated measures, we consulted with three prominent journal editors in the realm of obesity. Based on their consensus, we treated the body weight data (which was a combination of fasting and non-fasting weights) and the adherence data as repeated measures. This was also by far the most common approach

followed by manuscripts published in *JAMA* journals. We tested whether there was a between-group difference in weight loss using least-square means at the week 14 time point, while differences in mean adherence were tested via the main effect of group. Lastly, it was ambiguous which data would be multiply imputed. Since linear mixed models handle missing data well, but ANCOVA does not, we imputed all non-repeated measures data (with the exception of questionnaire data).

3. Previous Presentation: Results from preliminary analyses, which did not use linear mixed modeling, were presented at *ObesityWeek* 2020 and a handful of invited seminars. Full analyses, which included linear mixed models for weight loss and adherence, were conducted later.