WW Improving Nutrition Study (WINS)

April 21th, 2023
Pre-Analysis Plan

1. **Study Title:**

   WW Improving Nutrition Study (WINS)

2. **Principal Investigators**

   Ana M. Palacios (corresponding)
   Georgia Southern University
   ORCID: 0000-0002-9489-2219

   David B. Allison
   Indiana University Bloomington
   ORCID: 0000-0003-3566-9399

3. **Co-Investigators & Collaborators**

   Alexandra Lee
   WW International, Inc.
   ORCID: 0000-0001-8146-1674

   Chelsie Parker
   Indiana University Bloomington
   ORCID: 0000-0001-5262-9693

   Jing Kersey:
   Georgia Southern University
   ORCID: 0000-0001-7256-3978

   Michelle Cardel
   WW International, Inc.
   University of Florida College of Medicine
   ORCID: 0000-0002-9395-8618

   Stephanie Dickinson
   Indiana University Bloomington
   ORCID: 0000-0002-2998-4467

4. **Geographic location of the study:**

   Contiguous continental states in the US

5. **Key words**

   Behavioral weight management program

Weight Watch
Diet quality (HEI-2015 total score)
Weight loss

6. Abstract

The Weight Watchers Unlimited Workshops and Digital Program (WW) is an evidence-based behavioral weight management program that guides members towards personal weight and wellness goals through a personalized curriculum, complemented with behavioral weekly goals to drive healthy habits. The program includes foods that can be eaten in moderation without the need to tracking, as well as a points system that rates foods. In addition, members have access to food, activity, water, sleep, and weight trackers, meal planning tools, recipes, guided meditations and workouts, peer support, and access to online workshops and WW-trained behavior change coach.

To compare the 6-month changes in diet quality (HEI-2015 total score) in adult participants enrolled in a commercial weight-loss program (WW) vs. control, a total of 376 adults will be recruited via social media and other online platforms in the contiguous 48 States of the US. Participants will be randomly assigned to the WW program, or a control, and will be followed-up for 6 months.

7. Institutional Review Board:

Human Subjects Review: Board Status: Approved
Board Name: Institutional Review Board (IRB)
Board Affiliation: Georgia Southern University
Phone: 912-478-5465
Email: IRB@GeorgiaSouthern.edu
Address: Veazey Hall 3000
PO Box 8005
STATESBORO, GA 30460

8. Research Question:

The main objective of this study is to determine whether a behavior change weight management and wellness program (WW) delivered via an app for 6-months will be effective in improving diet quality (HEI-2015 total score) in U.S. adult participants, relative to a control group through a randomized controlled trial.

9. Study Design:

The two arms and corresponding interventions are as follows:

A. Experimental Arm: WW
- Behavioral WW: Participants will be randomized to enrolled in commercially available weight-loss program (WW) for 6 months.

B. Placebo Comparator Arm: No WW
- Behavioral No WW: Participants will be randomized to receive monthly general health informational emails, including myplate.gov informational handouts, for 6 months.

10. Randomization:

Randomization will be performed using random number generators in SAS or R and will use block randomization with stratification by sex. While enrollment will not be restricted to a certain number of males and females, they will be randomized separately within biological sex at birth to provide balance in treatment assignment.

11. Outcome Measures:

Primary Outcome Measure:

1. Change in Diet Quality (difference between baseline and 6-mo HEI-2015 total score)
   Diet quality will be calculated with the Healthy Eating Index-2015 (HEI-2015). The HEI-2015, is a valid and reliable composite measure that assesses overall diet quality and compliance with the DGA-2015[9]. This tool can help assess diet quality from U.S. populations and racial and ethnic subgroups[8]. The HEI-2015 scores 13 key dietary components to obtain a total score ranging from 0-100. Higher scores reflect greater dietary quality/greater adherence to the Dietary Guidelines (USDA, HEI-2015). Diet quality scores will be calculated by averaging across 3 unannounced 24h dietary recalls collected using the Automated Self-Administered 24-Hour Dietary Assessment Tool (ASA-24[13]). HEI-2015 total scores will be calculated using the "Simple HEI Scoring Algorithm – Per Person" method using the instructions and SAS macro (hei2015.score.macro.sas) provided here: https://epi.grants.cancer.gov/hei/sas-code.html

Secondary Outcome Measures:

2. Percent (%) Body Weight Loss
   Weight (lb); % body weight loss defined as baseline to 6-month weight change divided by baseline weight multiplied by 100.

3. Achievement of 3, 5, and 10% Weight Loss
   Participants that achieve at least 3/5/10% body weight loss at 6 months or not.

4. Change in Impact of Weight on Quality of Life
   Measured using the Impact of Weight on Quality of Life – Lite (IWQOL) which assesses the perception of how weight affects daily life. IWQOL-Lite includes 31 items rated on a 5-point Likert scale (1 = never true to 5 = always true) with 5 subscales: physical function (11 items), self-esteem (7 items), sexual life (4 items), public distress (5 items) and work (4 items). Subscale scores and total scores range from 0-100, with higher scores reflecting better levels of functioning. Data collected at baseline and 6-months; outcome is change in scores.

5. Change in Feelings of Hunger Over the Past 7 Days
   Measured using a Hunger Visual Analog scale (VAS) which includes a question that asks participants to rate how hungry they felt over the past week on a horizontal line with end points of "Not at all hungry" (0) to "Extremely hungry" (100). VAS are scored by measuring in where the participant

places their tick mark on the horizontal line with endpoints of 0-100. Higher scores indicate greater feelings of hunger. Data collected at baseline and 6-months; outcome is change in scores.

6. Change in Food Cravings
   Measured using the Food-Craving Inventory (FCI-II) which is a validated 33-item self-report measure designed to evaluate the subjective experience of food craving across 33 different foods. The measure assesses the frequency of cravings for a specified food with a five-point Likert scale (1, never; 2, rarely; 3, sometimes; 4, often; 5, always) and consists of 5 factors: high fats, sweets, carbohydrates/starches, fast food fats, & fruits and vegetables, that constitute the total food craving inventory score which averages all 33 items. Scores can range between 1 and 5, with higher scores indicating a greater frequency of cravings. Data collected at baseline and 6-months; outcome is change in scores.

7. Change in Self-Reported Physical Activity Over the Past 7 Days
   Measured using the Global Physical Activity Questionnaire (GPAQ)[1] which collects information on physical activity participation in the following domains: activity at work, travel to-and from- places, recreational activities, and sedentary behavior. From these inputs, the minutes per week spent in moderate activity, vigorous activity, moderate and vigorous activity, and sedentary behavior can be calculated. Data collected at baseline and 6-months; outcome is change in scores.

8. Change in Self-Reported Wellbeing
   Measured using the World Health Organization Well-Being Index (WHO-5)[16] which is a short measure of current, self-reported mental wellbeing. The WHO-5 consists of five statements rated by study participants with a Likert scale: All of the time=5, Most of the time=4, More than half of the time=3, Less than half of the time=2, Some of the time=1, At no time=0. The total raw score, ranging from 0 to 25, is multiplied by 4 to give the final score with 0 representing the worst imaginable well-being and 100 representing the best imaginable well-being. Data collected at baseline and 6-months; outcome is change in scores.

9. Change in Perceived Stress
   Measured using the 10-item Perceived Stress Scale (PSS) [10] which measures the extent to which a participant's perception of life is unpredictable, uncontrollable, and overloading. It was designed for use in older adolescents and adults and is considered to have adequate internal reliability and construct validity. Each question asks about how the participant has felt or thought in the past month and uses a 5-point Likert scale (0=never, 4=very often). Scores are calculated by summing responses, creating a possible score range of 0-40, with higher scores indicating greater perceived stress. Data collected at baseline and 6-months; outcome is change in scores.

10. Change in Habit Strength
    Measured using the Self-Reported Behavioral Automaticity Index (SRBAI), which captures habitual patterns of behavior. Each behavior of interest is assessed by 4 items rated on a Likert scale 1-strongly disagree to 7-strongly agree. Scores are calculated for each behavior by taking an average of the response, creating a possible score range between 1 and 7. Higher scores indicate greater habit strength for the behavior being measured[5]. Data collected at baseline and 6-months; outcome is change in scores.

11. Change in Alternative Mediterranean Diet Score

The alternative Mediterranean diet score (AMED) score was modified and adapted to the Mediterranean diet scale designed by Trichopoulou et al.[15] The AMED score is made up of 9 components: seven "healthy" components: a. fruits, b. vegetables, c. fish, d. legumes, e. nuts, f. whole grains, and g. ratio of monounsaturated fat to saturated fat, and two additional components: h. red and processed meat, and i. alcohol consumption as described by Dr. Zhilei Shan et al.[12]. Each component, except alcohol, will be categorized into quintiles (Q). Positive scores to the seven healthy components will be assigned as follows: (Q1=1, Q2=2, Q3=3, Q4=4, Q5=5). Reverse scores to red and processed meat will be assigned as follows: (Q5=1, Q4=2, Q3=3, Q2=4, Q1=5). For alcohol consumption (g/d), points will be assigned as follows: 5-15=5, 0-5 or 15-25=4, 0 or 25-30=3, 30-35=2, and ≥35=1 for females and 10-30=5, 0-10 or 30-40=4, 0 or 40-45=3, 45-50=2, and ≥50=1 for males. The analysis will use the total AMED composite score, ranging from 9 to 45, with a higher score representing closer resemblance to a healthy Mediterranean diet. Data collected at baseline and 6-months; outcome is change in scores.

Other Pre-Specified Outcome Measures:

12. Dietary Intake

Macro- and micro-nutrient intakes measured with the validated Automated Self-Administered 24-hour (ASA24®) Dietary Assessment Tool.

13. Self-Reported Sleep Quality

Measured with the sleep assessment module from the validated Automated Self-Administered 24-hour (ASA24®) Dietary Assessment Tool.

14. HEI-2015 Component-Scores

Diet quality sub scores from HEI-2015 will be measured with the Automated Self-Administered Dietary Assessment Tool 24-hour dietary recalls (ASA24®).

12. Statistical Analysis:

To test the effect of WW, the main outcome is change in diet quality (HEI-2015 score). Participants are randomized to one of two groups (WW or Control) with n=188 per group shown in table 1.

Table 1. Sample sizes expected for completion, considering a 30% attrition rate.

	WW	Control	Total
N to Enroll	188	188	376
Expected Completers	132	132	264

All data will be inspected for invalid data including entry errors. Continuous variables will be examined for approximate normality and transformed for analysis as needed. Descriptive statistics will be analyzed at all time points.

For primary analysis of HEI-2015 score:

Analysis of covariance (ANCOVA) will be used to test baseline to 6-month changes in HEI-2015 total scores for Diet Quality between the WW group and control. Covariates will be included for baseline HEI-2015 total score, participant's biological sex at birth, age, race/ethnicity, and education.

For secondary analysis:

For percent weight loss, ANCOVA will be used to test the WW group vs control. Covariates will be included for participant's biological sex at birth, age, race/ethnicity, education, and baseline weight.

Achievement of 3/5/10 % weight loss is only observed at the 6-month follow-up and is a binary variable. Logistic regression will be used, and covariates will be included for participant's biological sex at birth, age, race/ethnicity, education, and baseline weight.

The remaining secondary outcomes are measured at baseline and the 6-month post-intervention as continuous scores. ANCOVA will be used to test the baseline to 6-month changes in scores between WW and control groups. Covariates will be included for baseline outcome measures, participant's biological sex at birth, age, race/ethnicity, and education. The other pre-specified outcome variables may be analyzed in a similar manner.

Intention-to-treat (ITT) will be used for analyses of all outcomes to include all randomized participants for primary analysis. To account for missing data, multiple imputation will be used. Model assumptions of normality and equal variance of residuals will be evaluated for each outcome[3,6,7]. Normality of residuals will be considered satisfied with a skewness <|2| given large sample sizes and; and equal variance between groups will be considered satisfied with (0.5<(Var1/Var2)<2) with approximately balanced group sizes[2,7]. Significance tests for baseline differences will not be conducted per CONSORT guidelines[11]. A sensitivity analysis excluding observations with standardized residuals >|3| will be conducted to assess the impact of outliers.

Attendance/adherence to intervention will also be explored, and secondary analysis for per-protocol data will only include participants who complete the study and maintained acceptable adherence to intervention.

13. Power Analysis and Sample Size

With 264 estimated completers (30% attrition of n=376 enrolled), the study will show 80.9% power for detecting significant differences between groups (two-sided alpha=0.05) for an effect size of Cohen's d=0.35 and 89.9% power for d=0.40. This corresponds to a difference in HEI score of 4.9 (SD=14) or 5.6, respectively.

Figure 2a:

			Power, assuming two-sided 0.05 alpha, for comparison of means:								
	30%		Effect size (d)								
N to enroll	**N after attrition**	**N per group**	0.2	0.25	0.3	**0.35**	**0.4**	0.45	0.5	0.55	0.6
200	140	70	21.6%	31.2%	42.2%	53.8%	65.2%	75.3%	83.6%	89.8%	94.1%
226	158	79	23.9%	34.5%	46.6%	58.9%	70.5%	80.3%	87.8%	93.0%	96.3%
251	176	88	26.1%	37.8%	50.8%	63.6%	75.1%	84.3%	91.0%	95.2%	97.7%
277	194	97	28.3%	41.0%	54.7%	67.9%	79.2%	87.7%	93.4%	96.8%	98.6%
300	210	105	30.2%	43.8%	58.1%	71.4%	82.2%	90.1%	95.0%	97.8%	99.1%
326	228	114	32.4%	46.8%	61.6%	74.9%	85.2%	92.3%	96.4%	98.5%	99.5%
351	246	123	34.5%	49.7%	64.9%	78.1%	87.8%	94.0%	97.4%	99.0%	99.7%
376	264	132	36.7%	52.5%	68.0%	**80.9%**	**89.9%**	95.4%	98.2%	99.4%	99.8%
400	280	140	38.5%	55.0%	70.6%	83.1%	91.5%	96.3%	98.6%	99.6%	99.9%
		SD:					**Mean diff:**				
	HEI	14	2.80	3.50	4.20	**4.90**	5.60	6.30	7.00	7.70	8.40
	Weight loss (kg) at 6 mo	**5.04**	1.01	1.26	1.51	**1.76**	2.01	2.27	2.52	2.77	3.02
	Weight loss % at 6mo	**5.10**	1.02	1.28	1.53	**1.79**	2.04	2.30	2.55	2.81	3.06

Figure 2a shows the power obtained with various sample sizes depending on Cohen's d effect size. The standardized mean differences from Cohen's d were then converted to mean differences between the two groups on the original scales of HEI diet quality scores and weight loss using standard deviations taken from the literature review displayed in Figure 2b.

Figure 2b:

	Ventura-Marra HEI at 3mo post		Weight loss (kg) at 6 mo		Weight loss, % 6 mo	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
Trt	71.3	13.9	4.00	5.17	4.31	5.2
Ctrl	63.9	14.8	1.77	4.90	1.85	5
Mean diff	7.4	14.36	2.23	5.04	2.46	5.10
Effect size d=	0.52		0.44		0.48	

Figure 2b displays effects sizes from previous literature for studies on weight loss and HEI diet quality. Cohen's d is calculated from means and SD for differences between groups, where d=(Mean1-Mean2)/SD. Ventura Marra, M, *et al.* 2019[17] found effect sizes of 0.52 (Mean difference=7.4, SD=14.4) for differences between groups in HEI-2015 scores after a weight loss intervention at 3-months. This is consistent with a recent systematic review which showed a range of effects from 4 to 7 on HEI[4]. For weight loss, 6-month statistics were computed based on weight loss data collected at 3 and 12-month timepoints from Tate et. Al. 2022[14]

References

1. Armstrong T, Bull F. Development of the World Health Organization Global Physical Activity Questionnaire (GPAQ). *Journal of Public Health.* 2006;14(2):66-70.

2. Blanca MJ, Alarcón R, Arnau J, Bono R, Bendayan R. Effect of variance ratio on ANOVA robustness: Might 1.5 be the limit? Behav Res Methods 2018;50(3):937-62. doi: 10.3758/s13428-017-0918-2.

3. Casella G, Berger R. Statistical Inference. Second ed. USA, 2021.

4. Cheng, J., Liang, H.-W., Klem, M. L., Costacou, T., & Burke, L. E. Healthy eating index diet quality in randomized weight loss trials: A systematic review. Journal of the Academy of Nutrition and Dietetics 2022; 2212-2672. doi: 10.1016/j.jand.2022.08.114

5. Gardner B, Abraham C, Lally P, de Bruijn GJ. Towards parsimony in habit measurement: testing the convergent and predictive validity of an automaticity subscale of the Self-Report Habit Index. *Int J Behav Nutr Phys Act.* 2012;9:102.

6. Glass G, Hopkins KD. Statistical methods in psychology and education. Third ed. Needham Heights, MA: Allyn & Bacon, 1996.

7. Glass GV, Peckham PD, Sanders JR. Consequences of Failure to Meet Assumptions Underlying the Fixed Effects Analyses of Variance and Covariance. Review of Educational Research 1972;42(3):237-88. doi: 10.3102/00346543042003237.

8. Kirkpatrick SI, Reedy J, Krebs-Smith SM, et al. Applications of the Healthy Eating Index for Surveillance, Epidemiology, and Intervention Research: Considerations and Caveats. *J Acad Nutr Diet.* 2018;118(9):1603-1621.

9. Reedy J, Lerman JL, Krebs-Smith SM, et al. Evaluation of the Healthy Eating Index-2015. *J Acad Nutr Diet.* 2018;118(9):1622-1633.

10. Roberti JW, Harrington LN, Storch EA. Further Psychometric Support for the 10-Item Version of the Perceived Stress Scale. *Journal of College Counseling.* 2006;9(2):135-147.

11. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Ann Intern Med 2010;152(11):726-32. doi: 10.7326/0003-4819-152-11-201006010-00232.

12. Shan Z, Wang F, Li Y, et al. Healthy Eating Patterns and Risk of Total and Cause-Specific Mortality. JAMA Intern Med. Published online January 09, 2023. doi:10.1001/jamainternmed.2022.6117

13. Subar AF, Kirkpatrick SI, Mittl B, et al. The Automated Self-Administered 24-hour dietary recall (ASA24): a resource for researchers, clinicians, and educators from the National Cancer Institute. *J Acad Nutr Diet.* 2012;112(8):1134-1137.

14. Tate DF, Lutes LD, Bryant M, et al. Efficacy of a Commercial Weight Management Program Compared With a Do-It-Yourself Approach: A Randomized Clinical Trial. *JAMA Netw Open.* 2022;5(8):e2226561.

15. Teresa T Fung, Marjorie L McCullough, PK Newby, JoAnn E Manson, James B Meigs, Nader Rifai, Walter C Willett, Frank B Hu, Diet-quality scores and plasma concentrations of markers of inflammation and endothelial dysfunction, *The American Journal of Clinical Nutrition*, Volume 82, Issue 1, July 2005, Pages 163–173, https://doi.org/10.1093/ajcn/82.1.163

16. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. *Psychother Psychosom.* 2015;84(3):167-176.

17. Ventura Marra M, Lilly CL, Nelson KR, Woofter DR, Malone J. A Pilot Randomized Controlled Trial of a Telenutrition Weight Loss Intervention in Middle-Aged and Older Men with Multiple Risk Factors for Cardiovascular Disease. *Nutrients.* 2019;11(2).