DATA ANALYSIS PLAN

Dietary Composition and Energy Expenditure during Weight-Loss Maintenance

Date – June 14, 2017

Principal Investigators – Cara B. Ebbeling, PhD – David S. Ludwig, MD, PhD

Co-Investigator and Lead Biostatistician – Henry Feldman, PhD

Study Director – Gloria Klein, MS, RD

Data and Quality Manager – Sarah Steltz, MPH

ClinicalTrials.gov – NCT02068885

The purpose of this document is to finalize the Data Analysis Plan prior to breaking the randomization blind. We anticipate that data for the primary outcome (results from doubly-labeled water analyses) will be available in Fall 2017.

TABLE OF CONTENTS

PROJECT SUMMARY

Many overweight and obese people can lose weight for a few months, but most have difficulty maintaining weight loss over the long term. One explanation for the poor long-term outcome of weight loss diets relates to behavior, in that motivation to adhere to restrictive regimens typically diminishes with time. An alternative explanation is that weight loss elicits biological adaptations – specifically a decline in energy expenditure and an increase in hunger – that promote weight regain. We previously examined this question in a cross-over feeding study with 21 overweight or obese young adults and found that total energy expenditure (TEE) during weight-loss maintenance was 325 kcal/d greater on a low-carbohydrate diet compared to a conventional low-fat (high-carbohydrate) diet.[1] A moderate-carbohydrate diet elicited intermediate effects on TEE. We also found potentially important dietary effects on insulin resistance, cortisol excretion, and other chronic disease risk factors. The purpose of the proposed study is to follow-up our initial findings using a parallel design, so that each of three test diets can be examined for 20 weeks, substantially extending duration compared to the 4-week periods in our cross-over study. Following 12±2% weight loss on a standard run-in diet, 150 adults (aged 18 to 65 years) will be randomly assigned to one of three weight-loss maintenance diets controlled for protein content (20% of energy) and varying widely in dietary carbohydrate-to-fat ratio: High-carbohydrate (HI, 60% of energy from carbohydrate, 20% fat), Moderate- carbohydrate (MOD, 40% carbohydrate, 40% fat), Low-carbohydrate (LO, 20% carbohydrate, 60% fat). During the weight-loss maintenance phase, energy intake will be adjusted to prevent changes in body weight. The primary outcome will be change in total energy expenditure (indirect calorimetry using stable isotopes) through 20 weeks. Secondary outcomes will include resting energy expenditure (indirect calorimetry using respiratory gas exchange), physical activity (accelerometry), measures of insulin resistance and skeletal muscle work efficiency, components of the metabolic syndrome, and hormonal and metabolic measures that might inform an understanding of physiological mechanisms. In addition, we will test for effect modification by key baseline covariates, including insulin secretion. We also will assess weight change during a 2-week *ad libitum* feeding phase, as an objective measure of dietary effects on hunger.

TABLES AND FIGURES

Page

Tables

Figures

1. SPECIFIC AIMS, HYPOTHESES, AND STUDY OUTCOMES

1.A. Main Study

Specific Aim #1: To evaluate the effects of 3 diets varying widely in carbohydrate-to-fat ratio (high-carbohydrate, moderate-carbohydrate, low-carbohydrate) on energy expenditure during weight-loss maintenance.

Hypotheses

1a Total energy expenditure (TEE) during weight-loss maintenance will differ among test diets through 20 weeks.

1b Resting energy expenditure (REE) during weight-loss maintenance will differ among test diets through 20 weeks.

Primary outcome: TEE (assessed by indirect calorimetry using stable isotopes, doubly labeled water [DLW]).
Secondary outcomes: REE (assessed by indirect calorimetry using respiratory gas exchange), physical activity (assessed by accelerometry).
Potential effect modification: including insulin secretion (insulin at 30 minutes after standard 75-g oral glucose load).

Specific Aim #2: To evaluate the effects of 3 diets varying widely in carbohydrate-to-fat ratio on chronic disease risk factors during weight-loss maintenance.

Hypothesis

2 Chronic disease risk factors during weight-loss maintenance will differ among test diets through 20 weeks.

Secondary outcomes: insulin sensitivity and insulin secretion (assessed by frequently-sampled oral glucose tolerance test (OGTT), urine C-peptide, glycemic control (HgA1c, 1,5-anhydroglucitol), lipid profiles (total cholesterol, HDL-cholesterol, LDL-cholesterol, non-HDL cholesterol, triglycerides), lipoprotein particle subfraction distribution, coagulopathy (PAI-1, fibrinogen), inflammatory mediators (hsCRP, IL-6), blood pressure, sleep (assessed by accelerometry).

Specific Aim #3: To evaluate physiological mechanisms potentially relating dietary carbohydrate-to-fat ratio to metabolism and risk for chronic disease – including CVD, type 2 diabetes, and cancer.

Hypothesis

3 Differences among diets in measures of skeletal muscle work efficiency, body composition, insulin sensitivity and secretion, anabolic and catabolic hormones, gut microbiome, and metabolomics profiles will provide additional physiological insights into the effects of dietary composition on health outcomes during weight-loss maintenance.

Secondary outcomes: skeletal muscle work efficiency (assessed by cycle ergometry), body composition (assessed by a multi-component model), insulin sensitivity and secretion, urine C-peptide, thyroid functions (thyroxine [T4], free T4, rT3, TSH), growth hormone action (IGF-1, IGF binding proteins), reproductive hormones (LH, FSH, testosterone [total and free], estradiol), stress hormones (24-hour urinary cortisol and catecholamines), leptin, adiponectin (total, high-molecular weight), ghrelin, gut microbiome (saved stool samples), metabolomics profile (saved serum samples).

Specific Aim #4: To evaluate the effects of 3 diets varying widely in carbohydrate-to-fat ratio on body weight during an *ad libitum* feeding phase.

Hypothesis

4 Body weight will differ among diets during the 2-week *ad libitum* feeding phase (implemented after the test phase of weight-loss maintenance).

Secondary outcome: body weight.

1.B. Ancillary Studies

In addition to measurements for the main (parent) study, we will collect data for several ancillary studies. Two of the ancillary studies will require substantial effort (in terms of additional study visits for hospital-based assessments) and have separate registries at ClinicalTrials.gov (Metabolic Fuels/Adipocyte Biology: NCT02235038; Brain Reward Activity: NCT02300857).

Aims of Ancillary Studies on Psychological Health, Cognitive Function, and Weight Bias

Specific Aim #5: To evaluate the effects of 3 diets varying widely in carbohydrate-to-fat ratio (high-carbohydrate, moderate-carbohydrate, low-carbohydrate) on psychological health.

Hypothesis
Psychological health will differ among test diets.

Secondary outcomes: depression (Beck Depression Inventory-II [BDI-II]), mood and anxiety (assessed by the Mood and Anxiety Symptom Questionnaire [MASQ]), food addiction (assessed by the Yale Food Addiction Scale [YFAS]), emotional eating (assessed by the Emotional Eating Scale [EES]), binge eating (assessed by the Binge Eating Scale [BES]).

Specific Aim #6: To evaluate the effects of 3 diets varying widely in carbohydrate-to-fat ratio (high-carbohydrate, moderate-carbohydrate, low-carbohydrate) on cognitive function.

Hypothesis
Cognitive function will differ among test diets.

Secondary outcomes: memory (assessed by the California Verbal Learning Test – Second Edition [CVLT-II] and Digit Span Test), processing speed and executive function (assessed by the Trail Making Test Parts A and B [TMT-A, TMT-B]).

Hypothesis-Generating Questions: Weight Bias

Baseline

- Do the study participants express implicit, explicit, and/or internalized weight bias at baseline?
- Is there an association between weight bias and physiological stress, as assessed by 24-hour urinary cortisol at baseline?

End of Run-in Phase

- Does baseline weight bias influence weight loss and attrition during the run-in phase?
- Does weight bias change among participants during the run-in phase?

End of Test Phase

- Does weight bias modify the effect of dietary interventions on 24-hour urinary cortisol, total energy expenditure, and risk factors for chronic disease during weight-loss maintenance?

Aims of Metabolic Fuels/Adipocyte Biology Ancillary Study

Specific Aim #1: To examine metabolic fuel concentration and energy availability during weight-loss maintenance on high-, moderate-, or low-carbohydrate diets.

Hypothesis

Circulating metabolic fuel concentration will be lowest on the high-carbohydrate diet

Primary outcome: Total energy availability (EA) calculated as the sum, in kcal/L, of energy from circulating metabolic fuels (glucose, non-esterified fatty acids, lactate and ketoacids) in the late postprandial period. The late postprandial period is defined as 2.5 to 5 hours after breakfast, lunch, and dinner (consistent with methods in our pilot study[2]).

Secondary outcomes: Fasting and integrated 24-hour EA, insulin, glucose, non-esterified fatty acids, ketoacids, triglycerides, ghrelin, epinephrine, glucagon, and hunger and satiety ratings.

Specific Aim #2: To examine changes in adipocyte biology during weight loss maintenance on high-, moderate- or low-carbohydrate diets.

Hypothesis

A high-carbohydrate diet will induce anabolic changes within the adipocyte, thereby promoting fat storage and obesity.

Main outcome: Differences in adipose tissue gene expression (assessed by mRNA levels of selected candidate genes involved in lipid storage and biosynthesis).

Other outcomes: Histologic examination of adipose tissue samples, adipocyte diameter, measures of angiogenesis and inflammation.

Aims of Macronutrients and Brain Activity by MRI Ancillary Study

Specific Aim #1: To examine the baseline (fasting pre-prandial) effects of long-term adherence to diets varying in carbohydrate-to-fat ratio (high-carbohydrate, moderate-carbohydrate, low-carbohydrate) on blood flow in homeostatic hunger and reward regions.

Hypotheses

- 1.1. Blood flow to the hypothalamus will differ by dietary carbohydrate-to-fat ratio: LO and MOD > HI.
- 1.2. Functional connectivity between the hypothalamus and nucleus accumbens (NAcc) will differ by dietary carbohydrate-to-fat ratio: HI > MOD and LO.

Specific Aim #2: To examine the late postprandial effects of meals varying in carbohydrate-to-fat ratio (high-carbohydrate [HI], moderate-carbohydrate [MOD], low-carbohydrate [LO]) on blood flow in homeostatic hunger and reward regions following long-term adherence to these diets.

Hypotheses

- 2.1. (PRIMARY) Blood flow to the NAcc and connected striatal regions at 4 hours postprandial will differ according to carbohydrate-to-fat ratio: HI > MOD and LO.
- 2.2. Functional connectivity between the hypothalamus and NAcc at 4 hours postprandial will differ by carbohydrate-to-fat ratio: LO and MOD > HI.
- 2.3. Subjective food craving at 4 hours postprandial will be positively associated with NAcc blood flow and hypothalamus-NAcc functional connectivity.

Specific Aim #3: To investigate relationships between long-term weight-loss maintenance, brain activity, and physiological outcomes.

Hypotheses

- 3.1. Blood flow to the hypothalamus at 4 hours postprandial will be significantly related to change in total energy expenditure following adherence to weight-loss maintenance diets.
- 3.2. Blood flow in the NAcc at 4 hours postprandial will be associated with weight change during *ad libitum* feeding.
- 3.3. Functional connectivity between the hypothalamus and NAcc at 4 hours postprandial will be associated with weight change during *ad libitum* feeding.

2. DATA COLLECTION

2.A. Overview. This randomized controlled trial (RCT) will comprise three phases, as shown in **Figure 1**, using a feeding protocol. The purpose of the run-in phase is to obtain baseline measurements, restrict energy intake to achieve a 12±2% decrease in body weight, and then stabilize body weight. The purpose of the test phase is to compare the metabolic effects of high–, moderate–, and low–carbohydrate diets during weight-loss maintenance and the physiological mechanisms underlying these effects. The purpose of the *ad libitum* feeding phase is to evaluate the effects of these diets on body weight. Data will be collected at baseline (BSL, study enrollment), post-weight loss (PWL, end of run-in phase, prior to randomization), and follow-up (post-randomization, midpoint [MID] and end [END] of test phase, 10 and 20 weeks). The primary outcome is total energy expenditure (TEE) through 20 weeks.

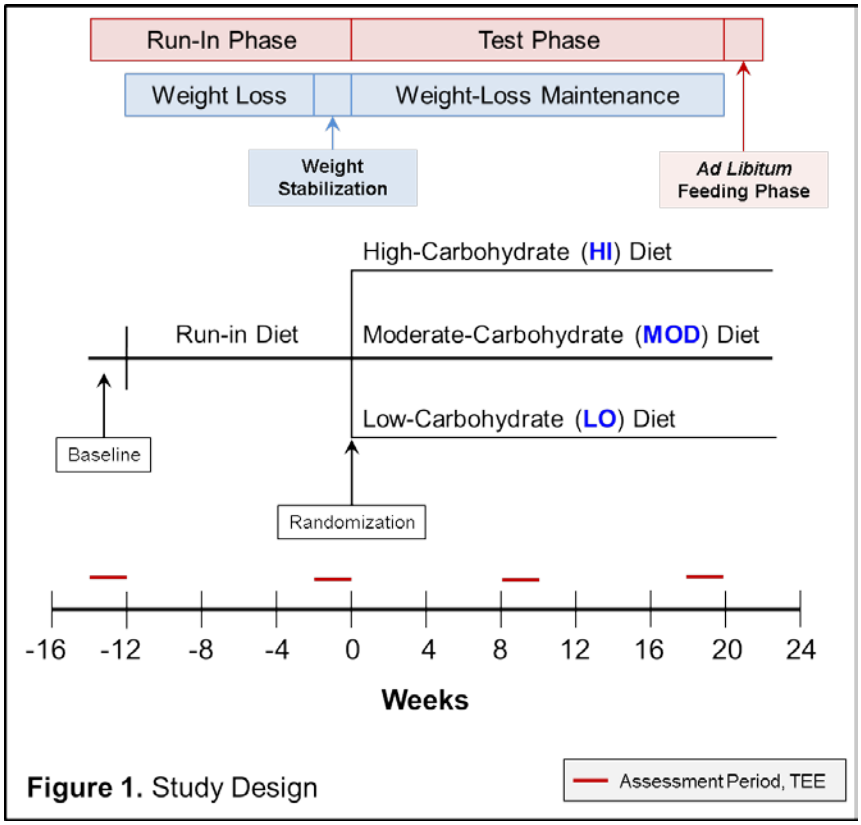

Figure 1. Study Design

2.B. Measurements. The measurement schedule is outlined in **Table 1**, and brief descriptions of outcomes are presented in **Table 2**.

Table 1. Schedule of Measurements.

Measurements	Time Point					
	Screening *	**Baseline (BSL)** †	**Post-Weight Loss (PWL)** ‡	**Follow-up (post-randomization)** §		***Ad libitum*** ‖
	Study Eligibility	Study Enrollment	Prior to Randomization	Midpoint of Test Phase (MID) 10 Weeks	End of Test Phase (END) 20 Weeks	Free Feeding
Medical history	X	X				
Screening labs	X					
Demographics		X				
Height	X	X				
Weight, BMI	X	X	X	X	X	X
Blood pressure (auscultation)	X	X	X	X	X	
TEE (DLW)		X	X	X	X	
REE (respiratory gas exchange)		X	X	X	X	
Physical activity		X	X	X	X	
Sleep		X	X	X	X	
Fasting labs		X	X	X	X	
Insulin sensitivity and secretion		X	X	X	X	
24-hour urine collection		X	X		X	
Body composition		X	X		X	
Skeletal muscle work efficiency		X	X		X	
Stool collection for gut microbiome ¶		X	X		X	
Blood sample for genetics ¶		X				
Psychological health and cognition ¶		X	X		X	
Weight bias ¶		X	X		X	
Postprandial metabolic fuels ¶				10-15 weeks		
Adipocyte biology ¶			X	10-15 weeks		
Brain activity ¶					14-20 weeks	
Dietary energy			X	X	X	
Palatability of test diets					X	

* *Screening Visit (Study Eligibility).* Data will be collected to assess eligibility. Particular emphasis will be placed on identifying participants likely to comply with the rigors of a 9-month feeding protocol requiring frequent study visits.

† *Baseline Data Collection (Study Enrollment).* At baseline, we will collect data on demographic variables, any changes in medical history since screening, body weight, and height. These data will be used to determine each participant's final eligibility status. Upon confirmation of study eligibility, participants will move through assessments.

‡ *Post-Weight Loss Data Collection (Prior to Randomization).* Participants achieving weight-loss corresponding to 10% to 14% of baseline body weight will undergo post-weight loss assessments.

§ *Follow-up Data Collection (Post-Randomization)*. Assessments will be repeated at 10 and 20 weeks post-randomization.

‖ *Ad libitum. Instructions to participants: "Have your meals according to your usual schedule. Eat as much or as little of each meal as you like until you are satisfied. If you finish your meal and are hungry before the next meal, eat something of your own choosing until you are satisfied. If you do not finish your meal, do not eat anything before the next meal. We ask that you eat until satisfied, but avoid overeating to the point of feeling too full."*

¶ *For participants who opt-in for measurement and provide informed consent.*

2.C. Description of Outcomes. Participant visits for data collection will be conducted at dedicated research facilities located on the campus of Framingham State University, except where noted otherwise.

Table 2. Description of Outcomes.

Measurements	Pre-specified Purpose	Outcome Variables and Assessment Methods
Medical history	Screening	
Screening labs	Screening	• We will measure hematocrit; BUN, blood urea nitrogen; creatinine; ALT, alanine transaminase; HgA1c, hemoglobin A1c; TSH, thyroid stimulating hormone. • Plasma samples will be sent to the Core Lab at Boston Children's Hospital for analysis, using standard procedures.
Demographics	Descriptive data Covariates	• We will collect descriptive data regarding age, gender, ethnicity, race and socioeconomic status by self-report. • The following variables will be used as covariates in analyses: age, gender, ethnicity, race.
Height	Descriptive data	• Data will be collected using a calibrated stadiometer.
Weight, BMI	Covariate Secondary outcome	• We will measure weight during assessment visits at FSU using a calibrated electronic scale. Weight and BMI will be covariates for SA#1, SA#2, and SA#3. • Participants will weigh themselves daily using Wi-Fi scales linked to the SetPoint Health Website. These data will be used to monitor weight loss during the run-in phase, to monitor weight stability and inform adjustments in energy intake during the test phase, and as a secondary outcome during the *ad libitum* phase (SA#4).
Blood pressure	Secondary outcome	• Systolic and diastolic blood pressure will be measured by auscultation at the right arm using a sphygmomanometer (System 5, American Diagnostic Corporation, Hauppauge, New York), 3 times at each assessment time point.
Total energy expenditure	** **Primary outcome**	• Data will be expressed per kg body weight (TEE/kg) for primary analysis, with adjustment for baseline covariates. • We will measure TEE by indirect calorimetry using stable isotope methodology (doubly labeled water). • TEE will be calculated from rCO2 using the food quotient (FQ) as an estimate of respiratory quotient (RQ).[3] • Two spot urine samples will be collected on separate days prior to dosing. Additional urine samples will be collected at 7 time points over 2 weeks subsequent to dosing. • Urine samples (collected at FSU or AV) will be sent to Baylor College of Medicine for determination of isotopic enrichment using gas-isotope-ratio mass spectrometry.
Resting energy expenditure	Secondary outcome	• We will measure REE by indirect calorimetry using respiratory gas exchange methodology (TrueOne 2400 System, Parvo Medics, Sandy, UT).
Physical activity	Secondary outcomes	• We will measure total physical activity and minutes of moderate- to vigorous-intensity physical activity under free-living conditions over 7 days using an ActiGraph monitor (Pensacola, FL) worn on the right hip.
Sleep	Secondary outcomes	• We will measure total sleep time, sleep onset latency, wake after sleep onset, and sleep efficiency using an ActiGraph monitor (Pensacola, FL) worn on the non-dominant wrist for 7 nights. We also will ask subjects to complete a daily sleep questionnaire to evaluate timing, amount, and quality of sleep.
Fasting labs	Secondary outcomes	• We will assess the following variables: *glycemic control* (HgA1c, 1,5-anhydroglucitol), *lipid profile* (total cholesterol, HDL-cholesterol, triglycerides, LDL-cholesterol [direct enzymatic/spectrophotometric methodology], non-HDL-cholesterol [calculated], lipoprotein particle subfraction distribution (nuclear magnetic resonance [NMR] spectroscopy); *coagulopathy* (PAI-1, fibrinogen); *chronic Inflammation* (high-sensitivity C-reactive protein, IL-6); *adipokines* (leptin, total and high-molecular weight adiponectin); *thyroid function* (T4, free T4, rT3, TSH); *growth hormone action* (IGF-1, IGF-Binding Proteins); *reproductive hormones* (LH, FSH, testosterone, estradiol); *other*: ghrelin, metabolomics profiling. • Most of the samples will be sent to the Clinical and Epidemiological Research Laboratory (CERLab) at Boston Children's Hospital for analysis, using standard procedures. NMR spectroscopy will be done at a commercial laboratory, and metabolomics profiling will be done at the Broad Institute.

Table 2 (CONTINUED). Description of Outcomes.

Measurements	Pre-specified Purpose	Outcome Variables and Assessment Methods
Insulin sensitivity and secretion	Secondary outcomes Covariates	• We will measure hepatic and peripheral insulin sensitivity and insulin secretion by frequently-sampled OGTT. We will use a standard 75-gram dose of dextrose (Trutol). Blood for determination of plasma glucose and insulin will be obtained by indwelling venous catheter at −10, −5, 0, 10, 20, 30, 60, 90, and 120 minutes relative to the start time of dextrose consumption. • Insulin sensitivity will be calculated using the method of Abdul-Ghani et al.[4] Insulin secretion will be determined based on insulin level at 30 minutes (insulin-30). • Fasting plasma glucose and insulin (averaged across the −10, −5, 0 time points), insulin sensitivity, and insulin secretion will be analyzed as secondary outcomes and also used as covariates. We will test for effect modification by baseline insulin-30 and other variables related to glucose homeostasis. • We will also measure urinary C-peptide. • Samples will be sent to the Clinical and Epidemiological Research Laboratory (CERLab) at Boston Children's Hospital for analysis, using standard procedures.
24-hour urine sample	Secondary outcomes	• We will collect 24-hour urine samples for analysis of cortisol and catecholamines. We will also measure creatinine. • Aliquots will be sent to the Clinical and Epidemiological Research Laboratory (CERLab) at Boston Children's Hospital for analysis, using standard procedures.
Body composition	Secondary outcomes Covariates	• We will assess body composition using a multi-component model to estimate percentage body fat using measures of total body volume from air displacement plethysmography (ADP, BodPod, Cosmed USA Inc., Concord, CA), total body water by isotope dilution, and bone mineral content by dual-energy x-ray absorptiometry (DXA, Horizon A, Hologic Inc., Bedford, MA).[5]
Skeletal muscle work efficiency	Secondary outcome	• We will measure work efficiency during graded cycle ergometry, according to published methods.[6,7] Following a 10-minute warm-up period, participants will pedal at 60 rpm against graded resistance to generate power corresponding to grades of 10W, 25W, and 50W in 4-minute stages. • We will measure oxygen uptake and carbon dioxide production by indirect calorimetry and convert oxygen consumption to energy expenditure based on respiratory exchange ratio (RER). • We will define skeletal muscle work efficiency at each grade as power generated per increase in energy expenditure above resting.
Gut microbiome	Secondary outcomes – Opt-in	• We will collect stool samples to assess changes in gut flora previously associated with metabolic changes and obesity (e.g., relative abundance of Firmicutes and Bacteroidetes species), using published methods.[8-10]
Obesity-related genes	Covariates Opt-in	• We will isolate and save buffy coat from blood samples for extracting DNA. Genetic studies may include candidate gene analysis and whole genome/exome sequencing. • We have particular interest in amylase gene copy number as an effect modifier.[11]
Psychological health and cognition	Secondary outcomes – Ancillary study – Opt-in	• We will assess the following variables pertaining to psychological health: *depression* (Beck Depression Inventory-II [BDI-II]), *mood and anxiety* (Mood and Anxiety Symptom Questionnaire [MASQ]), *food addiction* (Yale Food Addiction Scale [YFAS]), *emotional eating* (Emotional Eating Scale [EES]), *binge eating* (Binge Eating Scale [BES]). • We will assess the following variables pertaining to cognition: *memory* (California Verbal Learning Test – Second Edition [CVLT-II], Digit Span Test), *processing speed and executive function* (Trail Making Test Parts A and B [TMT-A, TMT-B]).
Weight bias	Secondary outcomes – Ancillary study – Opt-in	• We will explore weight bias using the following instruments: Implicit Associations Test, Obese Persons Trait Survey, Weight Bias Internalization Scale, and Beliefs About Obese Persons Scale.
Postprandial metabolic fuels	Primary and secondary outcomes – Ancillary study – Opt-in	• We will calculate total energy availability (EA, primary outcome) as the sum, in kcal/L, of energy from circulating metabolic fuels (glucose, non-esterified fatty acids, lactate, ketoacids) in the late postprandial period (2.5 to 5 hours after breakfast, lunch, and dinner). • We will assess fasting and integrated 24-hour EA, insulin, glucose, non-esterified fatty acids, ketoacids, and hunger and satiety ratings. • Data will be collected in a research unit at Brigham and Women's Hospital.
Adipocyte biology	Secondary outcomes – Ancillary study – Opt-in	• We will evaluate differences in adipose tissue gene expression, assessed by mRNA levels of selected candidate genes involved in lipid storage and biosynthesis. • We will examine the histology of adipose tissue samples, adipocyte diameter, and measures of angiogenesis and inflammation. • Fat biopsies will be done at Boston Medical Center.

Table 2 (CONTINUED). Description of Outcomes.		
Brain activity	Primary and secondary outcomes – Ancillary study	• We will measure blood flow to the nucleus accumbens and connected striatal regions (primary outcome). Regional cerebral blood flow (rCBF) will be measured using arterial spin labeling (ASL) at rest, in neuroanatomically-defined a priori regions (nucleus accumbens, caudate, putamen, pallidum). • We will measure blood flow to the hypothalamus. Regional cerebral blood flow (rCBF) will be measured using arterial spin labeling (ASL) at rest, in neuroanatomically-defined a priori regions (hypothalamus). • We will measure functional connectivity between the hypothalamus and nucleus accumbens. Blood-oxygen-level-dependent (BOLD) fMRI connectivity between neuroanatomically-defined regions of interest (hypothalamus, nucleus accumbens) will be measured during rest. • Scanning will take place on a Siemens 3T scanner (Siemens, Erlangen, Germany) at the Brigham and Women's Hospital Imaging Center.
Palatability of test diets	Covariate	• We will measure perceived palatability (tastiness) of test diets using a 10-cm Visual Analog Scale (VAS).

2.D. Timeline. The 4-year study will consist of an initial start-up period; recruitment, preparation, and implementation periods for each of 3 cohorts; and a data management and analysis period (**Table 3**). During the start-up period, we will develop the diets, establish protocols, prepare a manual of operations, hire and train staff, and begin recruitment. The study will be conducted at Framingham State University (FSU, Framingham, MA), with a satellite feeding site at Assabet Valley Regional Technical High School (AV, Marlborough, MA). For each cohort, recruitment will occur during the spring semester prior to study participation. We will prepare for the cohort during the summer so that feeding can begin in the fall semester and continue through the next spring semester. The data management and analysis period will be devoted to preparation of data sets and programs for statistical analysis.

Table 3. Study Timeline.

	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun
Year 1 *(2013-2014)*												
Study Start-up												
Recruitment, Cohort 1								n=25				
Year 2 *(2014-2015)*												
Preparation, Cohort 1												
Implementation, Cohort 1												
Recruitment, Cohort 2								n=65				
Year 3 *(2015-2016)*												
Preparation, Cohort 2												
Implementation, Cohort 2												
Recruitment, Cohort 3								n=60				
Year 4 *(2016-2017)*												
Preparation, Cohort 3												
Implementation, Cohort 3												
Data Management and Analysis												

2.E. Windows for Study Visits. Data will be collected within a 3-week window at each time point (BSL, PWL, MID, END).

3. GROUP ASSIGNMENT

3.A. Stratification and Randomization. Subjects who successfully complete the run-in phase will be eligible for randomization. A blocked randomization design will be employed to ensure close balance among the three diet arms at any point in the study. The randomization will be stratified by feeding site (FSU, AV), sex (male, female), ethnicity-race (non-Hispanic white, other), age (18–39.9 years and 40.0–65.9 years), and BMI (overweight: 25.0–29.9 kg/m^2, obese: ≥30.0 kg/m^2) to ensure balance at the completion of enrollment within every subcategory, regardless of size. Enrollment logs, one for each stratum, will be prepared with a numerical sequence of identifiers. Diet assignment lists, identical to the enrollment logs except for the addition of a randomly chosen diet, will be prepared under supervision of the Lead Biostatistician, using specialized software developed for that purpose. The diet assignments will be randomly permuted within blocks of 3, 6, and 9, and the blocks themselves will be randomly permuted. Each upcoming assignment will thus be unpredictable, preventing any deliberate or inadvertent bias on the part of those conducting enrollment.

Master randomization lists, one for each of 32 strata, will be prepared with a numerical sequence of identifiers. For example, AV-MWhJOv-001 is an identifier corresponding to Assabet Valley, Male, non-Hispanic White, Junior Age (18-39.9 years), Overweight, 001 (first subject assigned to this particular stratum). The master randomization list will be prepared by the Data and Quality Manager (DQM), with supervision from the Lead Biostatistician, using specialized software developed for that purpose. The software takes into account the three treatment arms, two feeding sites, four stratification factors (sex, ethnicity-race, age, BMI) each with two levels. Each identifier on the master randomization list will be assigned to one of three treatment arms (Low-, Moderate-, or High-Carbohydrate diet).

The assignment list will be kept in the private custody of the DQM and maintained in a secure shared drive folder accessible only to the DQM and an assigned back-up staff member trained in randomization. The DQM, after confirming eligibility criteria (including adequate weight loss during the run-in phase) with the study director, will assign the next available randomization ID according to the subject's stratum. Randomization lists will be maintained in an Excel spreadsheet, and assignments will be made electronically. Randomization will occur in waves (i.e., groups of subjects randomized at a time) for each cohort. The DQM will relay the diet assignment of each subject to intervention staff by email. Randomization will be done at Boston Children's Hospital (1 Autumn Street).

3.B. Masking. All urine samples collected for calculating the primary outcome, TEE by DLW, will be sent to Baylor College of Medicine (BCM) for analysis by isotope-ratio mass spectrometry (IRMS). All personnel in the IRMS Core Laboratory at BCM will be masked to diet group assignments. Secondary outcomes will be assessed by research assistants and laboratory technicians masked to assignments. The medical (safety) monitor for the trial also will be masked to assignments.

The principal investigators, study physicians, intervention team, and food service production team will be aware of diet group assignments. No clinical investigator will review trial outcomes, by arm or in aggregate, until the end of the trial.

4. STATISTICAL METHODS – Specific Aims #1–#6

4.A. Analysis Plan. We will follow the *a priori* analysis plan described below, randomizing a sample size of N=150 adults to ensure robust power. The primary outcome measure of the trial is total energy expenditure (TEE) per kg body weight, measured at four time points: baseline (pre-weight loss), week 0 (post-weight loss, pre-randomization), week 10 (midway through the test phase), and week 20 (end of test phase). The primary null hypothesis is that the time course of TEE between week 0, week 10, and week 20 will be the same for all three diets: high (HI), moderate (MOD), or low (LO) carbohydrate-to-fat ratio. An alternative hypothesis is detailed below.

The analytic framework for addressing both primary and secondary hypotheses will be repeated-measures analysis of variance (ANOVA), with the outcomes of interest as dependent variables and study arm (diet: HI, MOD, LO) as a three-level independent variable. Although covariates are theoretically balanced by the randomization and thus should have little influence, we will adjust the ANOVA for a number of baseline and time-varying covariates in order to reduce residual variance and improve our power to detect diet differences. These include the outcome of interest at baseline (pre-weight loss); change in body weight over the test phase (weeks 0−20); demographic characteristics (sex, race, ethnicity, age); baseline anthropometric measures (body-mass index, percentage lean mass, percentage weight lost pre-randomization); and design variables (study site, cohort, enrollment wave). We will employ an autoregressive covariance structure to account for potentially diminishing within-subject correlation over time. To minimize the influence of extreme values on the

fitted model, we will employ an outlier-deletion algorithm equivalent to robust regression with iterative reweighting.[12] A single subject who developed a disqualifying medical condition (hypothyroidism, as documented by 2 elevated TSH values) post-randomization will be excluded from the primary analysis.

To test the primary hypothesis, we will construct appropriate contrasts from parameters of the fitted repeated-measures model (namely, adjusted mean TEE at week 10 and week 20 – adjusted mean TEE at week 0, HI vs. MOD vs. LO), and test their significance with a 2-df F-test and critical p-value 0.05. If the overall null hypothesis is rejected, pairwise contrasts between study arms will be constructed and compared to zero. The principle of closed testing[13] dictates that in this special situation of three groups, compared pairwise only if the overall null hypothesis is rejected, we may make each pairwise comparison with a critical p-value of 0.05 and still preserve the Type I error rate for the family of four comparisons at 5%.

4.B. Sample Size, Power, and Detectable Effects. For comparability with pilot data from prior studies, our sample-size calculation is formulated in terms of changes in TEE uncorrected for body weight between week 0 and week 20. The primary null hypothesis is that the mean value of ΔTEE, defined as change in total energy expenditure at week 20 of the test phase compared to week 0 (post-weight loss), will be the same in all three diet arms. To estimate our power to detect deviations from this null hypothesis, we use the power characteristics of a simple one-factor, three-level ANOVA using covariate-adjusted residual variance estimates taken from our pilot data. As a conservative, minimal-impact alternative, pictured in the inset panel of **Figure 2**, we hypothesize that one of the three diets (high carbohydrate-to-fat ratio, or HI) will differ from the other two. The power of the ANOVA test to reject the null hypothesis increases with the magnitude of difference, as detailed in **Figure 2** for three illustrative sample sizes. These curves were derived from the non-central F distribution under a range of alternatives from 0 to 400 kcal/d, taking account of the parallel-group design (as contrasted with our pilot study, a 3-arm crossover design) and assuming a standard deviation of 412 kcal/d for ΔTEE among subjects as observed in the pilot data. Our proposed sample size, 45 completers per diet (allowing for 10% attrition from recruited sample of 50), provides 80% power to detect a difference of 237 kcal/d, as indicated by the point on the middle curve. This is a smaller effect than was discerned in the pilot, where mean ΔTEE under a high-carbohydrate (low-fat) diet differed from the average under moderate- and low-carbohydrate diets by 263 kcal/d. The smaller sample size, 25 per arm, would provide only 50% power for that magnitude of effect, while the larger sample size, 75 per arm, would be much more costly and would reduce the effect detectable with 80% power only to 175 kcal/d. We thus believe our design choice to be optimal in terms of feasibility and statistical power.

Detectable-effect figures for a variety of secondary outcomes are tabulated in **Table 4**. Based on the same design, alternative hypothesis, and analytic strategy, each figure represents 0.575 standard deviations of the outcome, a moderate effect size by conventional benchmarks.

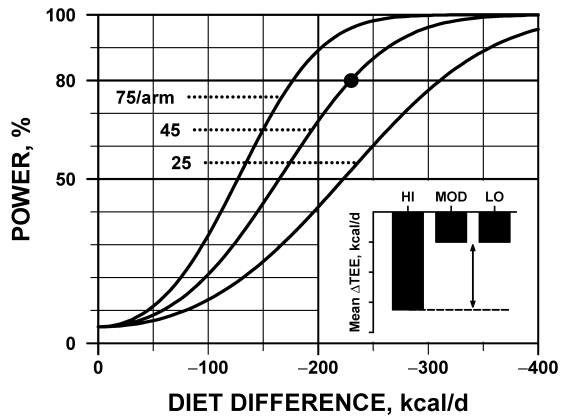

Figure 2. Power of one-way ANOVA to detect deviation of mean ΔTEE under one diet from the other two (inset). Each curve represents a fixed sample size. Dot indicates detectable effect, 237 kcal/d at 80% power, with proposed 45/arm. Curves based on non-central F distribution, 5% Type I error, between-subject standard deviation 412 kcal/d as in preliminary data.

4.C. Interim Analysis and Early Stopping Rule. No interim analyses are planned for the trial; therefore, no early stopping rule will be employed.

5. STATISTICAL METHODS – Postprandial Metabolic Fuels/Adipocyte Biology Ancillary Study

5.A. Analysis Plan. Energy Availability (EA) will be modeled as a function of diet and time, with an interaction term to allow the time course to vary by diet, using repeated-measures analysis of variance (ANOVA). For each period (using parameters of the model), we will estimate the mean EA for each diet and test the hypothesis that EA was equal in all 3 diets. Assuming this hypothesis is rejected at $p \leq 0.05$, the principle of closed testing permits additional pairwise comparisons to be performed with a critical p-value of 0.05 while preserving the familywise Type I error rate below 5%.[14] Secondary outcomes (metabolic fuel concentrations, hunger and satiety ratings, % change in adipose tissue gene expression) will be analyzed similarly.

5.B. Statistical Power and Sample Size Considerations. We utilized data and findings from the preliminary study by Walsh et al.[2] to estimate statistical power. The residual standard deviation for postprandial EA in this preliminary study was 0.32 kcal/L. Assuming that two diet groups are equal with the third diet differing by a certain amount, the detectable differences in postprandial energy availability (kcal/L) based on subject number per group are indicated in the center column of table 2. Pairwise detectable differences in postprandial energy availability are indicated in the rightmost column.

Table 5. Detectable difference for postprandial energy availability resulting from various subject numbers per group, under the hypothesis that only one diet will differ from the other two diets (center column), or between two given diets (rightmost column), with 80% power and assuming 5% type 1 error.

N per group	Detectable Difference (kcal/L)	
	A vs. B=C	A vs. B
5	0.62	0.64
8	0.46	0.48
10	0.41	0.42
12	0.37	0.38
15	0.33	0.34

The above calculations assume equal subject numbers per test diet group. However, subjects will be enrolled prior to randomization procedures for the main study, and therefore we will have no control over the final group distribution for this ancillary study. Based on statistical calculations, the expected loss of power from imbalance in group sizes will be small. For example, group sizes of 7-8-9 instead of 8-8-8 would inflate the detectable difference between two given diets by only 3.6% (i.e. 0.5 kcal/L instead of 0.48 kcal/L). Even with a more extreme group imbalance, such as 6-7-11 (likelihood 1.5%), the detectable difference between the two smaller groups would be inflated by only 11%, while the detectable difference between the two larger groups would in fact be reduced by 3%.

Based on these results, a tentative sample size of 8 subjects per group appears to be tenable, as this would reveal any differences as large as 0.5 kcal/L, which is within the range of the difference between low fat and the other two diets in the late postprandial period as previously reported by Walsh et al.[2] We have designed our enrollment goal of 30 to account for subject drop out and to provide increased power to buffer against the possibility of unequal subject numbers among diet groups.

6. STATISTICAL METHODS – Macronutrients and Brain Activity by MRI Ancillary Study

6.A. Arterial Spin Labeling (ASL) Perfusion MRI Data Analysis. ASL perfusion images will be analyzed using Advanced Normalization Tools (ANTs), FMRIB Software Library (FSL), and Freesurfer. Images will be transferred from the imager through DICOM transfer to a Linux analysis computer. DICOM images will be converted to Nifti format and anonymized for manipulation within additional programs. Briefly, a brain mask will be generated from the individual subject T1 image using ANTs and the IXI brain template. The pixel value range of the ASL CBF image will be adjusted using FSL and the image matrix size scaled to match the T1 using ANTs. The T2-weighted image will be skull-stripped and registered along with the matched axial T1 and T2 and CBF images to the T1 image using ANTs. T1 images will be segmented and labeled in native space using Freesurfer, and scaled to match the T1 nifti format. The SPM canonical T1 will be mapped to the subject and subject's T1 skull-stripped, followed by registration between the skull-stripped canonical BET image to the skull-stripped subject T1 image, and the in-house hypothalamus ROI mask mapped to the subject's native space. Finally, individual subject mean perfusion image will be calculated for each subject.

Regional of interest data analysis will be performed as previously described. Mean perfusion signal at each scan within regions of interest adapted from Freesurfer segmentation masks (or, for the hypothalamus, an in-house mask based on parcellation of the MNI template brain) will be calculated. Signal in our primary hypothesized areas of NAcc (R/L), hypothalamus, globus pallidus (R/L), caudate (R/L), and putamen (R/L) will be calculated and used for subsequent analysis.

6.B. Resting State fMRI Data Analysis. Resting state data will be analyzed in SPM8 and using the Conn toolbox to examine functional connectivity between the NAcc and hypothalamus. For each subject, time series will be extracted and temporal band pass filtered, outliers from ART included as nuisance regressors, and Pearson's correlation coefficients calculated between the mean time series of each ROI and every voxel in the brain. For a priori ROIs (NAcc, hypothalamus), anatomic borders have been defined using a manually segmented MNI-152 brain, from which ROIs have been segmented and parcellated as individual structures. All structures have been segmented using a contour line and manual editing, producing core files for subcortical grey matter and cortical parcellation units that can be implemented as seed ROI regions within the Conn toolbox. These will be converted to z-scores and average z-maps from seed ROI regions calculated and submitted to second-level ANOVA analyses to examine differences in functional connectivity between diets [height threshold: $p<0.001$ (voxel level), extent threshold: $p<0.05$ (cluster level)]. As a tertiary aim, we will also explore functional connectivity in regions of the default mode and salience networks.

6.C. Data Analysis Corresponding to Specific Aims. For each of the planned analyses below, we will use age, medication history, ethnicity, and menstrual cycle day (for cycling females) as covariates. All analyses will be held to a significance level of $p<0.05$, unless otherwise noted. For all analyses using ASL/BOLD data from the 4 hours postprandial time point, the stabilizing effect of using the difference between the preprandial to postprandial change in blood flow will be assessed using similar models, but may not improve power in this within subject design.

Specific Aim 1

- Hypothesis 1.1. Whole brain normalized blood flow within the hypothalamus at (fasting, pre-prandial) baseline will be analyzed using a 1-way ANCOVA design [between-subjects factor: diet group (LO/MOD, HI)], within SPSS, with post-hoc comparisons to test the hypothesized differences between LO/MOD and HI.
- Hypotheses 1.2. We will use a 1-way ANCOVA design [between-subjects factor: diet group (LO/MOD, HI)], implemented through the factorial design modules in SPM8 and Conn, to examine the main effects of diet group on BOLD functional connectivity between NAcc and hypothalamic regions, with post-hoc comparisons to test the hypothesized differences between HI and LO/MOD, with post-hoc comparisons to test the hypothesized differences between LO/MOD and HI.

Specific Aim 2

- Hypothesis 2.1. Whole brain normalized blood flow within the right NAcc at 4 hours postprandial will be analyzed using a 1-way ANCOVA design [between-subjects factor: diet group (LO/MOD, HI)], within SPSS, with post-hoc comparisons to test the hypothesized differences between LO/MOD and HI.
- Hypotheses 2.2. We will use a 1-way ANCOVA design [between-subjects factor: diet group (LO/MOD, HI)], implemented through the factorial design modules in SPM8 and Conn, to examine the main effects of diet group on BOLD functional connectivity between NAcc and hypothalamic regions at 4 hours postprandial, with post-hoc comparisons to test the hypothesized differences between LO/MOD and HI.
- Hypothesis 2.3. In SPSS, we will use Pearson correlations to examine relationships across groups between food cravings (area under the curve from 9 ratings from preprandial to 4 hours postprandial) and blood flow to the NAcc and functional connectivity between the hypothalamus and NAcc.

Specific Aim 3

- Hypothesis 3.1. We will use regression models with blood flow to the hypothalamus at 4 hours postprandial as the independent variable and the outcome variable of change in TEE (to parallel the Parent RCT, from post-weight-loss to week 20 in the Test Phase) as the dependent variable.
- Hypothesis 3.2. We will use regression models with right NAcc blood flow at 4 hours postprandial as the independent variable and weight change during the two-week *ad libitum* feeding period as the dependent variable.
- Hypothesis 3.3. We will use multiple regression models with functional connectivity between the NAcc and hypothalamus at 4 hours postprandial as the independent variable and weight change during two-week *ad libitum* feeding period as the dependent variable.

6.D. Statistical Power and Sample Size Considerations. ASL power analysis (for the primary outcome: blood flow to the NAcc at 4 hours postprandial) draws on established data sets on test-retest reproducibility of similar ASL methods[15,16] and on our pilot study comparing perfusion images acquired 4 hours postprandial for two different glycemic index meals.[17] Regional retest reliability of regional perfusion has been reported at 10% using similar methods to ours, though this reliability does not include the further stabilization provided by correction for global flow variations. A conservative estimate of 7.1% regional SD (10% test-retest SD) in perfusion and a diet group effect of 7% requires an n=20 for a 2-tailed paired t-test with $\alpha=0.05$ and a power of 86%. Note that we readily detected a greater than 8% change in the NAcc with 11 subjects in our pilot study. For the longitudinal study, a slightly more conservative n=25/diet group is proposed.

7. REFERENCES CITED

1. Ebbeling CB, Swain JF, Feldman HA, et al. Effects of dietary composition on energy expenditure during weight-loss maintenance. *JAMA.* 2012;307:2627-2634. PMID: 22735432.
2. Walsh CO, Ebbeling CB, Swain JF, Markowitz RL, Feldman HA, Ludwig DS. Effects of diet composition on postprandial energy availability during weight loss maintenance. *PLoS One.* 2013;8:e58172. PMID: 23483989.
3. Ravussin E, Harper IT, Rising R, Bogardus C. Energy expenditure by doubly labeled water: validation in lean and obese subjects. *Am J Physiol.* 1991;261:E402-409. PMID: 1909495.
4. Abdul-Ghani MA, Matsuda M, Balas B, DeFronzo RA. Muscle and liver insulin resistance indexes derived from the oral glucose tolerance test. *Diabetes Care.* 2007;30:89-94. PMID: 17192339.
5. Heymsfield SB, Ebbeling CB, Zheng J, et al. Multi-component molecular-level body composition reference methods: evolving concepts and future directions. *Obes Rev.* 2015;16:282-294. PMID: 25645009.
6. Goldsmith R, Joanisse DR, Gallagher D, et al. Effects of experimental weight perturbation on skeletal muscle work efficiency, fuel utilization, and biochemistry in human subjects. *Am J Physiol Regul Integr Comp Physiol.* 2010;298:R79-88. PMID: 19889869.
7. Baldwin KM, Joanisse DR, Haddad F, et al. Effects of weight loss and leptin on skeletal muscle in human subjects. *Am J Physiol Regul Integr Comp Physiol.* 2011;301:R1259-1266. PMID: 21917907.
8. Ley RE, Turnbaugh PJ, Klein S, Gordon JI. Microbial ecology: human gut microbes associated with obesity. *Nature.* 2006;444:1022-1023. PMID: 17183309.
9. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature.* 2006;444:1027-1031. PMID: 17183312.
10. Ley RE, Backhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. Obesity alters gut microbial ecology. *Proc Natl Acad Sci U S A.* 2005;102:11070-11075. PMID: 16033867.
11. Rukh G, Ericson U, Andersson-Assarsson J, Orho-Melander M, Sonestedt E. Dietary starch intake modifies the relation between copy number variation in the salivary amylase gene and BMI. *Am J Clin Nutr.* 2017. PMID: 28539377.
12. Rousseuw PJ, Leroy AM. *Robust Regression and Outlier Detection*. New York, NY: John Wiley & Sons; 1987.
13. Bauer P. Multiple testing in clinical trials. *Stat Med.* 1991;10:871-889; discussion 889-890. PMID: 1831562.
14. Bender R, Lange S. Adjusting for multiple testing--when and how? *J Clin Epidemiol.* 2001;54:343-349. PMID: 11297884.
15. Pfefferbaum A, Chanraud S, Pitel AL, et al. Volumetric cerebral perfusion imaging in healthy adults: regional distribution, laterality, and repeatability of pulsed continuous arterial spin labeling (PCASL). *Psychiatry Res.* 2010;182:266-273. PMID: 20488671.
16. Xu G, Rowley HA, Wu G, et al. Reliability and precision of pseudo-continuous arterial spin labeling perfusion MRI on 3.0 T and comparison with 15O-water PET in elderly subjects at risk for Alzheimer's disease. *NMR Biomed.* 2010;23:286-293. PMID: 19953503.
17. Lennerz BS, Alsop DC, Holsen LM, et al. Effects of dietary glycemic index on brain regions related to reward and craving in men. *Am J Clin Nutr.* 2013;98:641-647. PMID: 23803881.