



Comprehensive Assessment of Long-Term Effects  
of Reducing Intake of Energy (CALERIE)

**Final Protocol for the Phase 2 Study**

Compiled by:

The CALERIE Research Group

Version 1.5: January 22, 2009

Distributed by the CALERIE Coordinating Center:

Duke Clinical Research Institute  
Duke University  
P.O. Box 17959  
Durham, NC 27715.

Phone: (919) 668-8579  
e-mail: Galan006@DCRI.Duke.edu

# The CALERIE Research Group

**Study Sponsor:** National Institute on Aging (NIA)  
National Institutes of Health

**Project Officials:** Evan Hadley, MD  
Sergei Romashkan, MD  
Judy Hannah, PhD

## **Principal Investigators:**

### **Clinical Centers:**

Eric Ravussin, PhD  
Pennington Biomedical Research Center  
Baton Rouge, LA

Susan Roberts, PhD  
Tufts University  
Boston, MA

John Holloszy, MD  
Washington University  
St. Louis, MO.

### **Coordinating Center:**

James Rochon, PhD  
Kathy Galan, RN  
Duke Clinical Research Institute  
Durham, NC

## **Central Resources:**

### **DLW Laboratory:**

William Wong, PhD  
Baylor College of Medicine  
Houston, TX.

### **Central Biochemistry Laboratory:**

Michael R. Lewis, MD  
University of Vermont  
Burlington, VT.

### **DXA Reading Center:**

Ann Schwartz, PhD  
University of California at San Francisco  
San Francisco, CA

### **Safety Laboratory:**

Linda Evanello, MT(ASCP)  
Esoterix Clinical Trial Services  
Raritan, NJ

### **Nutrition Reading Center:**

Marcia Schmidt, MS, RD, LD  
University of Cincinnati  
Cincinnati, OH.

# Comprehensive Assessment of Long-Term Effects of Reducing Intake of Energy (CALERIE)

## Protocol for the Phase 2 Study

### Table of Contents

	page
SUMMARY OF CHANGES FROM VERSION 1.4 TO 1.5.....	VIII
SUMMARY OF CHANGES FROM VERSION 1.3 TO 1.4.....	IX
SUMMARY OF CHANGES FROM VERSION 1.2 TO 1.3.....	X
SUMMARY OF CHANGES FROM VERSION 1.1 TO 1.2.....	XII
SUMMARY OF CHANGES FROM VERSION 1.0 TO 1.1.....	XIII
1. EXECUTIVE SUMMARY .....	1
2. SPECIFIC AIMS AND OBJECTIVES / HYPOTHESES .....	5
2.1 Primary Specific Aims.....	5
2.2 Secondary Specific Aims.....	5
2.3 Exploratory Aims.....	5
3. BACKGROUND AND SIGNIFICANCE .....	7
3.1 Introduction.....	7
3.2 Energy Metabolism.....	7
3.3 Oxidative Changes and Damage.....	8
3.4 Cardiovascular Risk Factors .....	9
3.5 Insulin Sensitivity and Secretion .....	10
3.6 Immune Function.....	11
3.7 Activation of Neuroendocrine Axes and the Autonomic Nervous System.....	11
3.8 Quality of Life and Cognitive Function.....	13
4. PRELIMINARY STUDIES – CALERIE PHASE 1 STUDIES.....	14
4.1 Introduction to the Phase 1 Studies .....	14
4.2 Description of the Studies .....	14
4.3 Recruitment and Retention.....	14
4.4 Adherence to the Interventions.....	15
4.5 Summary of Adverse Events.....	15
4.6 Laboratory Safety Markers .....	16
5. BASIC STUDY DESIGN FOR THE PHASE 2 STUDY .....	16
6. STUDY POPULATION AND ELIGIBILITY CRITERIA.....	17
6.1 Study Population .....	17
6.2 Inclusion Criteria .....	17
6.3 Exclusion Criteria.....	18
6.3.1 Medical Exclusion Criteria.....	18
6.3.2 Laboratory Exclusion Criteria .....	18
6.3.3 Psychiatric and Behavioral Exclusion Criteria .....	18
6.3.4 Medication Exclusion Criteria .....	18
6.3.5 Other Exclusion Criteria .....	18
7. TREATMENT INTERVENTIONS.....	19
7.1 Treatment Summary.....	19

7.2	Conceptual Framework for CALERIE Phase 2 CR Intervention.....	19
7.3	Treatment Intervention.....	21
7.3.1	Nutritional / Dietary Toolbox .....	22
7.3.2	Behavioral / Environmental Toolbox .....	22
7.3.3	Implementation of the Algorithm .....	23
7.3.4	Initial Intervention Period .....	24
7.3.5	Long-term Intervention .....	24
7.3.6	Group sessions and Individual Counseling.....	25
7.4	Quality Control of the CR Intervention Procedures .....	25
7.5	Control Group.....	26
7.6	Participant Retention .....	26
7.7	Compensation.....	26
8.	RECRUITMENT PROCEDURES .....	26
8.1	Recruitment Strategies .....	26
8.2	Recruitment Procedures at the PBRC.....	27
8.3	Recruitment Procedures at Tufts University .....	27
8.4	Recruitment Procedures at Washington University.....	28
9.	SCREENING PROCEDURES .....	29
10.	RANDOMIZATION AND BLINDING .....	30
10.1	Randomization Procedures .....	30
10.2	Blinding of Study Personnel .....	31
11.	SCHEDULE OF EVALUATIONS AND PROCEDURES .....	31
11.1	Evaluations Performed During the Screening Period.....	31
11.1.1	Phone Screen.....	31
11.1.2	In-person Screening Visit 1 .....	31
11.1.3	In-person Screening Visit 2 .....	31
11.1.4	In-person Screening Visit 3 .....	32
11.1.5	Eligibility Assessment .....	32
11.2	Evaluations and Procedures Performed During the Baseline Period .....	32
11.3	Evaluations Performed During Participant Follow-up .....	33
11.3.1	Month 1 ( $\pm 3$ days) .....	33
11.3.2	Month 3 ( $\pm 7$ days) .....	33
11.3.3	Month 6 ( $\pm 14$ days) .....	33
11.3.4	Month 9 ( $\pm 14$ days) .....	34
11.3.5	Month 12 ( $\pm 14$ days) .....	34
11.3.6	Month 18 ( $\pm 14$ days) .....	35
11.3.7	Month 24 ( $\pm 14$ days) .....	35
11.4	Off-Schedule Evaluations .....	36
11.4.1	Antibody Response to Vaccines.....	36
11.4.2	Additional Weights.....	36
11.4.3	Additional Safety Measures.....	37
11.5	Evaluations for Drop-Outs and Withdrawals .....	37
12.	OUTCOME DETERMINATIONS .....	37
12.1	Energy Metabolism.....	37
12.1.1	Total Energy Expenditure and Energy Intake .....	37
12.1.2	Resting Metabolic Rate .....	37
12.1.3	Core Temperature.....	37
12.2	Cardiovascular Risk Factors .....	37
12.2.1	Blood Pressure .....	38
12.2.2	Serum Lipids and Lipoproteins.....	38
12.2.3	Markers of Inflammation .....	38

12.2.4	Transforming growth factor- $\beta$ 1 .....	38
12.3	Glucose Tolerance and Insulin .....	38
12.4	Immune Function.....	38
12.4.1	Delayed-Type Hypersensitivity .....	38
12.4.2	White Blood Cell Differential .....	38
12.4.3	Antibody Response to Vaccines.....	38
12.5	Endocrine Response.....	39
12.5.1	Norepinephrine .....	39
12.5.2	Dehydroepiandrosterone (DHEA) and Cortisol .....	39
12.5.3	Sex Hormones .....	39
12.5.4	Thyroid Hormones.....	39
12.5.5	Adipokines .....	39
12.5.6	Angiotensin II.....	40
12.5.7	Growth Hormone and Growth Factors .....	40
12.6	Quality of Life, Psychological and Cognitive Functioning .....	40
12.6.1	Rand SF-36.....	40
12.6.2	Profile of Mood States.....	40
12.6.3	Perceived Stress Scale .....	40
12.6.4	Pittsburgh Sleep Quality Index .....	41
12.6.5	Derogatis Interview for Sexual Functioning-Self Report .....	41
12.6.6	Food Craving Questionnaire .....	41
12.6.7	Food Craving Inventory .....	41
12.6.8	Eating Inventory or Three Factor Eating Questionnaire .....	41
12.6.9	Weight Self Efficacy .....	42
12.6.10	Multiaxial Assessment of Eating Disorder Symptoms .....	42
12.6.11	Body Shape Questionnaire.....	42
12.6.12	Tests of Cognitive Impairment .....	42
12.7	Physical Activity Measures .....	43
12.7.1	Cardiorespiratory Fitness using Peak $VO_2$ via Expired Gas Analysis.....	43
12.7.2	Energy Expenditure using the Stanford 7-day Physical Activity Recall (PAR) .....	43
12.7.3	Muscular Strength and Endurance .....	44
12.8	Body Weight and Height .....	44
12.8.1	Body Weight.....	44
12.8.2	Height .....	44
12.9	Body Composition.....	45
12.9.1	Waist Circumference.....	45
12.9.2	Body Composition and Bone Density by DXA .....	45
12.10	Bone Turnover .....	45
12.11	Nutrient Intake .....	45
12.12	Advanced Clinical Endpoints.....	46
12.12.1	Process for Selecting Outcomes and Tests .....	46
12.12.2	Types of Specimens to be Collected and Stored .....	47
12.12.3	Process for Proposing Studies .....	47
13.	DOUBLY LABELED WATER (DLW) METHODS .....	48
14.	ADHERENCE TO THE INTERVENTIONS.....	50
14.1	Adherence Outcome Measures.....	50
14.1.1	Intervals Over Which %CR Will Be Measured.....	51
14.1.2	Long-term Intake / Balance Estimation of %CR .....	51
14.1.3	Short-Term Intake / Balance Estimation of %CR .....	51
14.1.4	Approaches for Estimating Long-Term %CR .....	52
14.2	Adherence Measures to Inform the “Toolbox” Intervention Algorithm .....	52
14.2.1	Degree of Weight Change.....	52
14.2.2	Self-Reported Energy Intake .....	53

14.2.3	Process Measures .....	53
15.	<b>QUALITY CONTROL ACTIVITIES</b> .....	54
15.1	<b>Staff Adherence to Study Procedures and Timetables</b> .....	54
15.1.1	Manual of Procedures .....	54
15.1.2	Training and Certification .....	54
15.1.3	Study Manager .....	54
15.1.4	Multidimensional Monitoring Program.....	55
15.2	Quality Control of Study Measurements .....	55
15.3	Quality Control Documentation and Reports .....	55
15.4	Specific Details on Outcome Domains.....	56
15.4.1	Doubly Labeled Water (DLW) .....	56
15.4.2	DXA.....	56
15.4.3	Resting Metabolic Rate (RMR) .....	56
15.4.4	Core Clinical Chemistry and Biochemistry Laboratory .....	57
15.4.5	Peak Oxygen Uptake Measures (Peak VO <sub>2</sub> ) .....	57
15.4.6	Measures of Physical Activity: 7-day Physical Activity Recall (7-day PAR).....	57
15.5	Reliability Studies.....	57
15.5.1	Test-Retest Reliability Model.....	57
15.5.2	Statistical Analysis.....	58
15.5.3	Sample Size Calculations .....	58
15.5.4	Resampling Rates .....	58
16.	<b>PARTICIPANT SAFETY AND ADVERSE EVENTS</b> .....	59
16.1	Signs, Symptoms and Adverse Events .....	59
16.2	Serious Adverse Events .....	60
16.3	Adverse Events Anticipated in this Study .....	60
16.4	Clinically Significant Laboratory Values .....	61
16.4.1	Clinical Laboratory Tests.....	61
16.5	Potassium Surveillance Protocol.....	61
16.6	Anemia Surveillance Protocol.....	62
16.7	Cholesterol Surveillance Protocol.....	63
16.8	Monitoring for Nutritional Adequacy and Eating Disorders.....	63
16.9	Monitoring for Excessive Weight Loss .....	64
16.10	Depression and Other Mental / Behavioral Health Conditions Surveillance Protocol...	64
16.11	Monitoring Bone Mineral Density .....	64
16.12	Electrocardiogram.....	65
16.13	Medical and Medication History.....	65
16.14	Physical Examination.....	66
16.15	Vital Signs .....	66
16.16	Withdrawal from the Intervention .....	66
16.16.1	Criteria for Temporary Discontinuation of the CR Intervention .....	66
16.16.2	Criteria for Permanent Discontinuation of the CR Intervention.....	66
16.17	Safety Monitoring Procedures .....	67
17.	<b>STATISTICAL CONSIDERATIONS</b> .....	68
17.1	Outcome Measures .....	68
17.2	Power Calculations .....	68
17.3	Analysis Policies .....	69
17.3.1	Type-I Error .....	69
17.3.2	Interim Analyses for Efficacy .....	69
17.3.3	Reporting Format .....	70

17.4	“Intention-to-Treat” Analysis Population.....	70
17.5	Causal Effects of Calorie Restriction .....	70
17.6	Baseline Characteristics.....	71
17.7	Model Building.....	71
17.7.1	Covariates in Linear Models.....	71
17.7.2	Preliminary Investigations to Verify Underlying Assumptions.....	71
17.8	Group Comparisons for Outcome Variables .....	72
17.9	Effect Modifiers and Subgroup Analyses .....	72
17.10	Adherence, Drop-Out, Loss to Follow-up, and Other Mitigating Factors.....	73
18.	DATA MANAGEMENT PROCEDURES .....	73
18.1	Hardware and Software Configuration .....	73
18.2	Centralized Data Management .....	73
18.3	Sources of Data .....	74
18.4	Data Management Activities.....	74
18.5	Data Quality Control Procedures.....	75
18.6	Data Management Reports .....	76
19.	PARTICIPANT RIGHTS AND CONFIDENTIALITY .....	76
19.1	Confidentiality and HIPAA Considerations.....	76
19.2	Institutional Review Boards .....	76
19.3	Informed Consent Procedures.....	77
20.	STUDY ADMINISTRATION.....	78
20.1	NIH Cooperative Agreement Mechanism .....	78
20.2	Steering Committee.....	78
20.3	Subcommittee Structure.....	78
20.4	Data & Safety Monitoring Board .....	79
21.	REFERENCES .....	79
22.	APPENDIX 1: BRIEF SUMMARIES OF THE CALERIE PHASE 1 STUDIES.....	91
22.1	PBRC Study .....	91
22.1.1	Introduction.....	91
22.1.2	Methods.....	91
22.1.3	Results.....	91
22.1.4	Discussion .....	92
22.2	Tufts University Study .....	92
22.2.1	Introduction.....	92
22.2.2	Methods.....	93
22.2.3	Results.....	93
22.2.4	Discussion .....	94
22.3	Washington University Study.....	94
22.3.1	Methods.....	94
22.3.2	Results.....	94
22.3.3	Discussion .....	94
23.	APPENDIX 2: SCHEDULE OF EVALUATIONS.....	96
24.	APPENDIX 3: ORGANIZATIONAL STRUCTURE FOR CALERIE .....	99
25.	APPENDIX 4: INFORMED CONSENT TEMPLATES.....	100

## SUMMARY OF CHANGES FROM VERSION 1.4 TO 1.5

In addition to simple typographical and spelling errors, the following substantive changes were made to Version 1.4 to create Version 1.5.

### **Age Eligibility Criterion (Section 6.2):**

The lower age limit for eligibility has been increased from 20 to 21 for both genders. The lower limit occurs elsewhere in the document, and all such occurrences have been changed accordingly.

### **Calcium Supplement (Section 7.2):**

The amount of the calcium supplement provided to CALERIE participants has been increased from 500 mg to 1,000 mg accordingly. The amount is described elsewhere in the document, and all such occurrences have been changed accordingly.

### **Pneumococcal vaccine (Section 12.4.3):**

The specific vaccine used for this response has been changed to the Pneumovax 23 from Merck.

### **Toolbox Intervention Algorithm (Section 14.2):**

This section has been completely rewritten to describe the approach to opening the toolbox and applying interventions to ensure adherence to the CR intervention. This includes monitoring weight changes over time, self-reported energy intake, and attendance at the individual or group intervention sessions.

### **Monitoring Bone Mineral Density (Section 16.11):**

- The criteria for discontinuing a participant because of a bone mineral density have been changed to the following:

“Bone mineral density will be monitored by measuring BMD by DXA at the total hip and total spine (L1 – L4) at baseline, 6 , 12, and 24 months. Any participant who experiences a decrease in BMD at the total hip or total spine of 10% or greater from the baseline at any time during the study, will have the scan repeated within one month. If the second scan confirms original findings of a decrease in BMD at the total hip or spine of 10% or greater, the participant will have the intervention permanently discontinued. In addition, the BMD *t*-score will also be monitored, and any participant who has BMD *t*-score at the total hip or total spine of less than –2.5 at any time during the study will have the intervention permanently discontinued.

“Additionally, any participant who experiences a decrease in BMD at the total hip or total spine greater than or equal to 5% and less than 10% from baseline to Month 6 or baseline to Month 12 will have BMD measured by DXA at 18 months. If decrease in BMD from baseline to Month 18 is greater than or equal to 10%, the above discontinuation procedure will be followed.”

- Permanent Discontinuation (Section 16.16.2): This section has been revised accordingly.



## SUMMARY OF CHANGES FROM VERSION 1.3 TO 1.4

In addition to simple typographical and spelling errors, the following substantive changes were made to Version 1.3 to create Version 1.4.

### **Sex Hormones in Women:**

- Section 12.5.3: The measurement of sex hormones in women has been deleted. The first paragraph in this section has been revised accordingly.
- Sections 11 and 23 have been revised accordingly.

### **Potassium Monitoring :**

Monitoring potassium abnormalities will be performed at the same frequency as for other laboratory markers, i.e., at baseline and at Months 1, 3, 6, 9, 12 and 24.

- Section 11.4.3: The reference to potassium surveillance has been deleted.
- Section 0: The second paragraph on possible elevations in potassium at the PBRC in Phase 1 has been deleted. The increased potassium testing after the baseline visit has been deleted. The 3rd paragraph now reads, "Only participants with normal potassium levels at screening and baseline will be allowed to enter into the study. After the baseline visit, potassium levels will be measured at Months 1, 3, 6, 9, 12, 18, and 24."
- Section 23 has been revised accordingly.

### **Clinical Laboratory Tests :**

Additional tests will be performed for serum chemistry.

- Section 16.4.1: The following sentence has been added to the serum chemistry paragraph: "Total, HDL- and calculated LDL cholesterol, triglycerides, C-reactive protein, and blood glucose and insulin values will also be performed."

## SUMMARY OF CHANGES FROM VERSION 1.2 TO 1.3

In addition to simple typographical and spelling errors, the following substantive changes were made to Version 1.2 to create Version 1.3.

### **Inclusion / Exclusion Criteria:**

- Section 6.2: The age range has been expanded to between 20 to 50 (inclusive) for men, and 20 to 47 (inclusive) for women.
- Section 6.3.4: Prior documented vaccination with Hepatitis A or pneumococcal vaccine has been deleted as an exclusion criterion.
- Section 12.4.3: The first sentence in the last paragraph has been modified to: “Previous documented vaccination with hepatitis A or pneumococcal vaccine will serve as an exclusion criterion for this component of the study.” However, these individuals will participate in all the other evaluation protocols.
- Section 6.3.4: The restriction to oral contraceptives for women has been removed, so that any form of contraception is allowed including hormone-releasing agents.

The corresponding sections of the Executive Summary have been changed accordingly.

### **Contraception by Female Participants :**

- Section 6.2: An additional inclusion criterion have been added: “Female participants must use acceptable form of contraception (barrier method, oral contraceptive, intrauterine device, or similar) and be willing to continue using such a method while enrolled in the study.”
- Section 16.13: This section has been changed to read: “A complete medical history performed during screening will include a review of all major organ systems and all medications taken during 30 days prior to screening, as well as reproductive status, contraceptive and menstrual history for women, where appropriate ... Use of an appropriate method of contraception by women will be verified at Month 1, 3, 6, 9, 12, 18, and 24 visits.” Acceptable forms are enumerated there.
- Sections 11.1.3, 11.3.1-11.3.7: These sections have been changed to reflect that contraception use will be recorded at Screening and at each of the scheduled follow-up visits, i.e., at Months 1, 3, 6, 9, 12, 18 and 24.

### **Monitoring for Bone Mineral Density:**

There were ambiguities in different sections of the protocol regarding the monitoring for bone mineral density (BMD). As described in Section 12.9.2, DXA is performed at baseline (twice), Month 6 (twice) and at Months 12, 18 and 24 to provide body composition data for the adherence calculations described in Section 14. However, DXA serves a second purpose, i.e., to monitor BMD. This safety monitoring is also assessed measured by DXA, but will only be performed at baseline, 6, 12, and 24 months. The following sections have been changed to make this clear.

- Sections 11.2, 11.3.5 and 11.3.7: BMD and BMC of the hip, spine and forearm using DXA will be performed at baseline, and at Months 12 and 24.
- Section 11.3.3: Body composition using DXA will performed in the CR group (only); BMD and BMC of the hip and spine using DXA in the CR intervention group (only) at Month 6
- Section 11.3.6: Body composition using DXA will performed in the CR group (only) at Month 18.
- Section 12.9.2: This section has been renamed to “Body Composition and Bone Density by DXA”. It has been clarified that whole body scans will be performed to coincide with the DLW measures at Baseline and at Months 6, 12, 18 and 24. The following paragraph has been added:  
“BMD and BMC at the hip, spine and forearm will be measured with DXA. These scans will be obtained at baseline, 6, 12 and 24 months for the hip and spine scans, and at baseline, 12 and 24 months for the forearm. Hip and spine scans will also provide an assessment of participant safety (See Section 16.11). “
- Section 15.4.2: A variety of clarifications were made to the quality control section for DXA.
- Sections 12.10: The timing of the bone turnover measurements has been clarified to: “Bone turnover measurements will be performed to coincide with the DXA measures but will only be performed at baseline, 6, 12 and 24 months.”

- Section 16.11: The first sentence has been changed to read, “BMD will be monitored by measuring BMD by DXA at the hip and spine at baseline, 6 , 12, and 24 months.”
- Section 16.16.2: The BMD criterion for permanent discontinuation has been changed to: “BMD loss at any one site (total hip, femoral neck, or total spine) of  $\geq 5\%$  from the baseline at any time during first 12 months of CR, or 10% or greater at any time during next 12 months of CR.”
- Section 23 (Schedule of Evaluations) has been changed accordingly.

#### **Potassium Monitoring :**

- Section 0: The frequency of potassium monitoring for CR participants during the first 6 months of the intervention has been reduced. This section now reads, “In the 25% CR group, potassium levels will be determined monthly for the first 6 months after intervention start, and at Months 9, 12, 18, and 24 thereafter.”
- Section 0: Creatine phosphokinase and ECG will continue to be measured for CR participants at the same time points as the potassium levels. Their frequency will similarly be reduced to monthly for the first 6 months after intervention start, and at Months 9, 12, 18, and 24 thereafter
- Section 11.4.3 and Section 23 (Schedule of Evaluations) have been updated to reflect this change.

#### **Dosage for the Tetanus Toxoid for the DTH Studies:**

- Section 12.4.1: There was a typographical error in this dosage. It have been changed to read, “DTH will be measured using the Mantoux method in which 0.1ml of a normal saline control and 2 anti-gens, candida and trichophyton, and 0.02 ml (0.20 LF units per dose) for tetanus toxoid (ad-sorbed) will be administered ...”

## SUMMARY OF CHANGES FROM VERSION 1.1 TO 1.2

In addition to simple typographical and spelling errors, the following substantive changes were made to Version 1.1 to create Version 1.2.

### **Bone Mineral Density:**

Clarifications were made to the exclusion criterion and the safety monitoring criterion for bone mineral density (BMD).

1. Exclusion Criteria (Section 6.3): The following change was made to the exclusion criterion:
  - BMD *t*-score at the total hip, femoral neck or total spine (L1-L4) is less than or equal to  $-2.3$  at the first DXA scan during the baseline visit
2. Monitoring BMD (Section 16.11): The following change was made to the discontinuation criterion:
  - ... any participant who has BMD *t*-score at the total hip, femoral neck or total spine (L1-L4) of less than  $-2.5$  at any time during the study will have the intervention permanently discontinued.
3. Criteria for Permanent Discontinuation of the CR Intervention (Section 16.16.2): The following change was made to the discontinuation criterion:
  - BMD *t*-score at the total hip, femoral neck or total spine (L1-L4) of less than  $-2.5$  at any time during the study.

### **Sections 11.3.2 and 11.3.6:**

The safety surveillance section of the protocol indicates that the Multi-axial Assessment of Eating Disorder Symptoms (MAEDS) and the Body Acceptability Morph (BAM) will be performed at months 3, 6, 12, 18 and 24 (Section 16.8). However, the evaluations at Months 3 and 18 were not included in the list of evaluations listed in this section.

- Sections 11.3.2 and 11.3.6 have been changed to show “Monitoring eating disorders” explicitly.

### **Anemia Surveillance Protocol (Section 16.6):**

Clarifications have been made to this section of the protocol.

- Anemia is formally defined as a hemoglobin (Hgb) and/or hematocrit (Hct) level below the lower limit of normal (LLN) for the laboratory.
- One of the criteria for having the hematology panel repeated is a decrease of 5 percentage points from baseline in hematocrit. The previous version indicated “5%” and this was ambiguous.

## SUMMARY OF CHANGES FROM VERSION 1.0 TO 1.1

In addition to simple typographical and spelling errors, the following substantive changes were made to Version 1.0 to create Version 1.1.

1. Exclusion Criteria (Section 6.3): The following additions were made to the exclusion criteria.
  - If the BMD *t*-score at any site (hip or spine) is less than or equal to  $-2.3$  at the first DXA scan during the baseline visit, the volunteer is ineligible.
  - If the LDL-cholesterol level  $\geq 190$  mg/dl at the screening visit, the volunteer is ineligible.

Screening Procedures (Section 9): This section was also changed to reflect this change. The 4th paragraph was changed to read:

“Candidates will also be asked to come in to the center following a 12-hour overnight fast ... Blood for hematology, clinical chemistry (including a lipid panel), a serum pregnancy test, where applicable, and urine for urinalysis will be collected.”

2. Cholesterol Surveillance Protocol: A surveillance protocol has been added for LDL-cholesterol levels of study participants. This required changes to two sections of the protocol.

Section 16.4: The second paragraph was changed to:

“All clinical laboratory tests will be performed by the central laboratory for CALERIE Phase 2 study, at baseline and Months 1, 3, 6, 9, 12, 18 and 24 visits unless otherwise specified.”

“Screening” was deleted because screening tests are already described in Section 6.3.2 and Section 9 and this removes the redundancy.

Section 16.7: A new section has been added to describe on-going LDL-cholesterol monitoring. In summary, for LDL-cholesterol greater than or equal to 160 mg/dl and less than 190 mg/dl prior to randomization, no action will be taken. After randomization, participants assigned to the CR group will follow their CR regimen; control participants will be advised to follow the American Heart Association’s low-cholesterol diet.

The CC will monitor LDL-cholesterol levels provided by the Vermont lab at Baseline and at Months 12 and 24 post-randomization. For a CR participant whose LDL-cholesterol is greater than or equal to 160 and less than 190 mg/dl, the clinical site will not be notified and the participant will continue to follow his/her CR regimen; for a control participant, the CC will notify the clinical site and s/he will be advised to follow the American Heart Association’s low-cholesterol diet. For a participant in either treatment arm with an LDL-cholesterol greater than or equal to 190 mg/dl, the CC will notify the clinical site and the participant will be advised to seek medical help outside of the study. S/He will, however, continue to participate in the study.

## 1. EXECUTIVE SUMMARY

Specific Aims [Section 2]: The overall aim of CALERIE Phase 2 is to test the hypothesis that two years of sustained caloric restriction (CR), involving a reduction in energy intake to 75% of baseline (25% CR), in healthy men aged 21 to 50 and healthy women aged 21 to 47 will result in the same adaptive changes that occur in rodents subjected to CR. Particular emphasis on the adaptive responses thought to be involved in slowing the aging process and protecting against age-related disease processes. Primary outcomes include core body temperature and resting metabolic rate. Secondary outcomes include triiodothyronine and catecholamines (as potential mediators of the predicted metabolic adaptation), and plasma concentrations of TNF- $\alpha$  (because inflammation is one of the adaptive responses suggested as a mediator of the salutary effects of CR on the aging process in rodents). An important secondary aim is to identify potential adverse effects of CR in humans. A number of exploratory aims will be assessed to evaluate the effect of CR on body composition, serum hormones, plasma growth factor concentrations, serum lipid and lipoprotein levels, skeletal muscle, adipose tissue and psychological factors. Full details are provided in Section 2. Consistency between the two sexes and across levels of body composition will be explored. In addition, biological samples will be stored in a biosample repository for future analysis.

Basic Study Design [Section 5]: The study will be conducted as a multi-center, parallel-group, randomized, controlled trial (RCT). A sample of 250 participants will be enrolled, and assigned to either the CR intervention or an *ad libitum* (AL) control group. A 2:1 allocation ratio in favor of the CR intervention will be applied in order to maximize the number of subjects receiving the intervention of greater scientific interest. Participants in both treatment arms will be followed over a period of 24 months. A comprehensive set of evaluations will be performed prior to initiating the intervention, with follow-up evaluations at Months 1, 3, 6, 9, 12, 18 and 24 after randomization. It is expected that 10% of study subjects will drop-out in each of the two follow-up years, so that a sample of approximately 200 subjects is expected to complete the study.

Study Population and Eligibility Criteria [Section 6]: Participants must be between 21 and 50 years of age (inclusive) for men, and between 21 and 47 years of age (inclusive) for women, and body mass index must be greater than or equal to 22.0 and less than 28.0 kg/m<sup>2</sup>. Female participants must use acceptable form of contraception (barrier method, oral contraceptive, intrauterine device, or similar) and be willing to continue using such a method while enrolled in the study. Otherwise, healthy individuals from both genders and all races are eligible to participate. Volunteers will be ineligible if there are significant medical conditions (e.g., history or clinical manifestation of cardiovascular disease, diabetes, cholelithiasis or cancer); abnormal laboratory markers (e.g., elevated potassium levels, hemoglobin or hematocrit below the lower limit of normal); psychiatric or behavioral problems (e.g., eating disorders or a history of drug and alcohol abuse); concomitant medications (e.g., steroids). Never-smokers of tobacco products or ex-smokers who quit completely at least 12 months ago are eligible. Breast-feeding or pregnant women (or those intending to become pregnant before the scheduled end of the intervention) and individuals performing any kind of heavy physical activity will be excluded. Volunteers will be screened out if they are unwilling or unable to adhere to the rigors of the CR intervention or the evaluation schedule over the entire two-year period.

Treatment Interventions [Section 7]: The active intervention will target a sustained 25% restriction in calorie intake *vis-à-vis ad libitum* energy intake measured by doubly labeled water (DLW) at baseline. There will be no gradual ramping of CR, and the 25% energy reduction goal will be maintained for the entire 24 months. Control participants will be advised to continue their current diets on an *ad libitum* basis. The CR intervention will be implemented by a multi-disciplinary team including dietitians, psychologists, and physicians. No specific diet composition will be mandated. Rather, the CR intervention will employ an algorithmic approach that combines specific nutritional and behavioral guidance so that each participant can maintain the prescribed level of calorie restriction. The approach will be tailored to the needs of the individual participant, with specific nutritional and behavioral strategies selected from an intervention “toolbox.” Examples include increasing dietary fiber, modifying recipes to decrease energy density, adding novel foods to relieve boredom, strategies to avoid impulse eating, obtaining desired foods through home delivery or take-home meals, strategies for limiting caloric intake in public settings like restaurants, parties and work, and so on. Selections from the toolbox will be based on the partici-

participant's success in achieving adherence, and on problems arising at that point in time. This algorithmic approach will also provide for continuous feedback and communication between individual participants and interventionists. The behavioral component will include group sessions and individual counseling, and during these sessions, the interventionist will provide specific and individualized dietary information to help the participant adhere to the CR regimen and meet his/her calorie target.

No specific level of physical activity will be required or recommended. However, all participants will be advised of current recommendations from the Surgeon General (Centers for Disease Control) for minimum levels of dedicated physical activity. A complete daily vitamin and mineral supplement will be provided to intervention and control participants to ensure that they meet the current recommendations for these nutrients.

Recruitment and Screening [Sections 8, 9, 11.1 and 23]: Participants will be recruited at the three CALERIE clinical centers using procedures that were successful in the Phase 1 studies. Recruitment will be continuous, and generally include media advertising, direct mail, health promotion events, databases, and referral sources. An effort will be made to recruit an ethnically diverse group based on the demographics of the three clinical sites.

An initial telephone screening will record the volunteer's contact information as well as age, height, weight, and basic eligibility information. Volunteers who are clearly ineligible will be screened out at this point. Then, a staged screening process will be undertaken over a series of 3-4 visits. Exclusion criteria will be evaluated. Volunteers will meet with the study psychologist or a trained member of the behavioral team to assess any barriers to participation. A 14-day food record will be collected to assess the volunteer's ability to adhere and complete a food record continuously over a two-week period.

Randomization and Blinding [Section 10]: CALERIE participants will be assigned to intervention using a random process. A telephone-based, interactive voice-response system (IVRS) will be applied. Randomization will be stratified by sex and BMI within each clinical center, and within each stratum, subjects will be allocated in a 2:1 ratio in favor of the 25% CR intervention. A permuted block randomization technique will be applied so that the desired allocation ratio is maintained at periodic intervals throughout the recruitment process. Given the nature of CR intervention and control conditions, it is not possible to blind study participants or CALERIE staff members to the treatment assignments. Nevertheless, within the resources available to this study, intervention staff will not be engaged in evaluating participants, and evaluation staff will be blinded to the treatment assignments.

Outcome Determinations [Section 12]: A detailed series of evaluations will be performed on participants in both treatment arms at baseline and at periodic intervals during the study. They include the following: measures of energy metabolism, cardiovascular risk factors, glucose tolerance and insulin, immune function, endocrine response, quality of life (QoL), psychological and cognitive functioning, physical activity measures, body height and weight, body composition, bone turnover, and nutrient intake. Biological material including blood, urine, muscle biopsy and abdominal fat biopsy will be stored in a biosample repository for future analyses. A process is described for extending the protocol to incorporate advanced clinical endpoints as the opportunity arises. Complete details are provided in Section 12.

Schedule of Evaluations [Sections 11 and 23]: Evaluations will be performed with participants in both treatment arms at baseline and at periodic intervals during the study. Follow-up visits will be performed at Months 1, 3, 6, 9, 12, 18 and 24 following the start of the assigned intervention. The baseline visits and follow-up visits at Months 12 and 24 are the most elaborate and a complete set of evaluations will be performed. A smaller set of evaluations will be performed at Months 6 and 18, while an abbreviated follow-up will be performed at Months 1, 3, and 9. The schedule of evaluations is summarized in Section 23 and provides complete details.

Adherence Calculations [Section 14]: Adherence measures will be used both as an outcome to determine the degree of CR actually achieved, and to inform decisions about modifications to a participant's intervention program and selecting toolbox options. Adherence will be characterized as the percentage of CR achieved,  $\%CR_P = 100 [1 - (EI_P/EI_{AL(t)})]$ , where  $EI_P$  represents average daily energy intake over the period of interest, and  $EI_{AL(t)}$  represents *ad libitum* long-term average daily energy intake before the start of the intervention. *Ad libitum* EI will be characterized by the average of two consecutive measures of energy expenditure performed at baseline using DLW methods as described in Section 13.

We will use two versions of the intake/balance method, based on the relationship,  $EI = EE + \Delta ES$ , where EE is average daily energy expenditure during the period of interest and  $\Delta ES$  is the change in

body energy stores during the period of interest. The “long-term” version will consider EE over the interval between any two time points by taking the average of the EE estimates across the time points. For intervals spanning more than two DLW measures, the average of the estimates for each interval, weighted by the duration of the interval, will be applied.  $\Delta$ ES will be estimated by calculating the change in energy stores (measured by DEXA) from the beginning to the end of the interval.  $\Delta$ ES will be calculated using standard coefficients for changes in fat mass and in fat-free mass. “Short-term” calculations will be performed at 6 months only. EE will be estimated by 14-day DLW measures;  $\Delta$ ES will be estimated by 14-day changes in FM and FFM, measured by DEXA at the beginning and end of the DLW measurement period.

Quality Control Procedures: [Section 15]: A detailed quality assurance plan will be developed to safeguard the scientific integrity of the study. A comprehensive Manual of Procedures will be developed, initial and ongoing training and certification of each staff member will be conducted, and each clinical site will have a Study Manager with substantial experience in clinical research operations and management of the multi-center clinical trials. S/He will supervise and oversee day-to-day operations of the clinical site, ensure adherence to the Good Clinical Practice guidelines, study protocol, Manual of Procedures, study timeline and budget. Prior to study start, the Coordinating Center (CC), reading centers and laboratories will develop appropriate monitoring plans for their study functional areas with the guidance and oversight by the Quality Control Committee.

Formal reliability studies will be conducted for two of the central facilities, i.e., the DLW lab and the central biochemistry laboratory. Duplicate urine sample sets from the same subject at the same protocol time point will be collected and forwarded to the DLW lab for analysis. The specimens will be labeled in such a way that facility personnel are blinded to which participant (and treatment arm) the specimens correspond, and whether it is the original or duplicate specimen. Similar procedures will be applied for blood samples sent to the biochemistry lab with the exception that the duplicate samples will be drawn from the blood stored in the repository. Statistical analyses will be performed to quantify the reliability of the procedures, and whether there is a significant difference between the test and retest specimens.

Participant Safety and Adverse Events [Section 16]: Protection of subjects from risks related to the study is of paramount concern to investigators and institutions participating in CALERIE. An adverse event (AE) is defined as any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the intervention irrespective of whether it is considered related to the intervention. Participants will be given a diary to record signs, symptoms and adverse events occurring between clinic visits. An expedited reporting protocol will be followed for reporting serious adverse events (SAEs) to the Coordinating Center, the NIA Program Official and the Data and Safety Monitoring Board (DSMB).

Clinical laboratory tests include hematology, serum chemistry and urinalysis will be performed at screening, baseline and at Months 1, 3, 6, 9, 12 and 24. A serum pregnancy test for women at screening only; otherwise, a urine pregnancy test will be performed immediately prior to any DXA evaluation and the Hepatitis vaccine and booster at Months 17 and 23. A heightened surveillance protocol for elevations in potassium, LDL cholesterol and the incidence anemia will be applied. On-going surveillance for mental / behavioral health conditions as well as bone mineral density will be performed. CR participants will be removed from the intervention, either temporarily or permanently, if abnormal values develop. Periodic reports summarizing participant safety will be presented to the CALERIE Steering Committee and the DSMB. Remedial action, including additions and changes to the protocol, will be taken as appropriate.

Statistical Considerations [Section 17]: The study size is limited by the feasibility of finding eligible subjects, enrolling them into the study, and maintaining their commitment to this extensive intervention over a period of 24 months. Enrolling 250 subjects over a period of 20-24 months is thought to be feasible with the resources available. Based on the experience in the Phase 1 studies, a drop-out rate of around 10% per year is expected, so that a sample of approximately 200 subjects is expected to complete the study. Power calculations indicate that based on this sample size, clinically meaningful differences can be detected for the most important outcome variables.

This is the first detailed investigation of effects of CR over an extended period of time in humans. Many of these comparisons are exploratory in nature, and will need to be confirmed by follow-up studies. Type-II error, i.e., failing to detect a significant effect when it exists, is important to this study. Thus,



all tests of significance for between-group comparisons will be performed at the  $\alpha = .05$  level of significance. The primary analysis strategy is under the Intention-to-Treat principles. CALERIE is also interested in mechanistic questions concerning the effect of CR, and to address these issues, the Marginal Structural Model (MSM) of Robins and colleagues will be applied. All major outcomes are observed repeatedly at well-defined time points over participant follow-up, so that statistical methods for longitudinal and repeated measures analysis will be applied. Subgroup analyses will be tested by evaluating the treatment by subgroup interaction. Withdrawal from the intervention, drop-out from the study, crossover to the alternate intervention or death (if any) will be analyzed using the standard techniques for survival data.

Data Management [Section 18]: All data arising from the study will be forwarded to the Coordinating Center. A database will be created on the DCRI computer network specifically for this study. Paper Case Report Forms (CRFs) will be designed to capture all the information required for reports and analyses. Staff at the clinical sites will record the data mandated by the study on the CRFs, and a copy of the CRF will be forwarded to the CC. Double data entry by two different operators at two separate occasions will be performed. As well, a variety of supplementary material and procedures will be conducted with study participants, including blood and urine samples, DXA evaluations and dietary recall. The resulting files and/or materials will be forwarded directly to central laboratories and reading centers for processing and interpretation. At periodic intervals, electronic data files containing the results of these determinations will be forwarded to the CC using a secure FTP server and merged into the master database. A series of validation checks will be conducted on the database. They will search for impossible and implausible values as well as logical inconsistencies across the different data fields. A variety of progress reports will be prepared during the course of a trial reviewed with the appropriate CALERIE working groups.

Participant Rights and Confidentiality [Sections 19 and 25]: All participant data will be kept strictly confidential, and no subject-identifying information will be released to anyone outside the project. Each participant will be assigned an anonymous study ID, which will be used on all study forms. Any study forms and paper records which do contain participant information (e.g., address lists, phone lists) will be kept at the CALERIE clinical site in secured, locked areas. At the Coordinating Center, only authorized personnel will have access to the data files containing study data. Participants will not be identified by name in any reports or publications, nor will the data presented in such a way that the identity of individual participants can be inferred.

Before initiating this study, the protocol, site-specific informed consent form (including HIPAA Authorization) recruitment materials, and other relevant information will be reviewed by a properly constituted Institutional Review Board (IRB) at each participating clinical site. A copy of the IRB approval notification and approved informed consent and HIPAA Authorization Form will be collected by the CC study monitor prior to site initiation and archived at the Coordinating Center. All CALERIE participants will provide written informed consent to participate in the study before any study-related procedures are initiated. The consent will describe the study's aims and objectives, procedures and activities to be undertaken in the study, as well as a summary the potential risks and benefits of participating. Two consents will be undertaken. First, because there is an extended screening phase to determine eligibility and many study procedures are performed during this process, the first consent will occur during the first screening visit. The second informed consent will occur after eligibility has been confirmed and the participant is ready begin the baseline evaluations.

Study Administration [Section 20]: The administrative and funding mechanism for this research is an NIH cooperative agreement (U01). The Steering Committee is the main governing body, and is composed of the Principal Investigators of the clinical centers and the Coordinating Center as well as the NIA Project Scientist. Each member has one vote and all decisions are determined by majority vote. In addition to the Steering Committee, a number of subcommittees will be formed to provide broad direction to the study. Oversight is provided by a Data and Safety Monitoring Board. The DSMB will approve the protocol before the study is initiated; monitor recruitment, retention and adherence; evaluate data completeness and data quality; and, ensure that participant safety is addressed adequately. It reports directly to the Director of NIA, and makes recommendations on all study activities including terminating the study for safety or operational reasons.

## 2. SPECIFIC AIMS AND OBJECTIVES / HYPOTHESES

The overall aim of CALERIE Phase 2 is to test the hypothesis that two years of sustained caloric restriction (CR), involving a reduction in energy intake to 75% of baseline (25% CR), in healthy men aged 21 to 50 and healthy women aged 21 to 47, will result in the same adaptive changes that occur in rodents subjected to CR, with particular emphasis on the adaptive responses thought to be involved in a) slowing aging, and b) protecting against age-related disease processes. A second aim is to identify potential adverse effects of CR in humans.

One reason for two years of CR intervention is that the one year in CALERIE Phase 1 was not sufficient to induce the adaptations in humans that are thought to increase longevity in rats. Many of these adaptations are also seen in humans practicing long-term severe CR. A second reason is to isolate the long term effects of CR (i.e., at weight stability) from the acute effects of weight loss.

### 2.1 Primary Specific Aims

Our primary specific aim is to test the hypotheses that CR in humans causes sustained (over two years) metabolic adaptation as defined by:

1. a reduction in core body temperature and
2. reduced resting metabolic rate (RMR) corrected for changes in body composition.

A reduction in metabolic rate has been proposed as one of the mechanisms by which CR slows aging, possibly by reducing oxidative damage. A lower resting metabolic rate, as measured in resting conditions and corrected for body composition, may be mediated by a reduction of serum triiodothyronine, decreased tissue conversion of T4 to T3 and/or a reduction in catecholamines (see Secondary Aim 1).

These aims (core body temperature and RMR) are sufficiently powered to provide a robust test of the hypothesis (see Section 17.2 below), address key unresolved issues in the field of CR, and are feasible as they build upon the experience of the Phase 1 CALERIE team as described in Section 4 below.

### 2.2 Secondary Specific Aims

Our secondary aims are to test the hypotheses that CR in humans:

1. Reduces serum triiodothyronine. Reduced triiodothyronine is a potential mediator of the predicted metabolic adaptation and will provide insight as to the mechanism of this hypothesized primary adaptation to CR. Decreased catecholamines are another potential mediator of the metabolic adaptation and will be explored as part of Exploratory Aim #2
2. Reduces inflammation as reflected in plasma concentrations of Tumor Necrosis Factor- $\alpha$  (TNF- $\alpha$ ). Inflammation is one of the adaptive responses suggested as a mediator of the salutary effects of CR on the aging process in rodents. This measure will be interpreted in the context of additional inflammatory markers enumerated in Exploratory Aim #4.
3. An additional important secondary specific aim is to determine whether CR has adverse effects in humans and to evaluate their seriousness.

These secondary specific aims are sufficiently powered to provide a robust test of the hypotheses (see Section 17.2 below), address key unresolved issues in the field of human CR, and are feasible as they build upon the experience of the Phase 1 CALERIE team as described in Section 4 below.

### 2.3 Exploratory Aims

The rationale for these exploratory aims is to obtain information regarding the mechanisms by which CR mediates its effects at the structural, physiological, cellular and subcellular levels. Analyses and reporting of these tests will take into account the high potential for Type I errors resulting from the large number of comparisons and outcomes.

Exploratory aims are designed to test the hypotheses on the effects of two years of calorie restriction in humans as follows:

1. Changes body composition (fat mass, lean mass), including bone mineral density.

Changes in caloric intake during CR impact body composition. As such, measuring body composition is important as a covariate for the primary aim (core temperature and RMR) in CALERIE phase 2. Intra abdominal fat is an important covariate for the interpretation of the atherosclerosis and Type 2 diabetes measures. Bone mineral density is important to assess the safety of CR.

2. Changes serum hormones, including DHEAS, cortisol, TSH, thyroid binding globulin, growth hormone, leptin, adiponectin, angiotension II, norepinephrine, sex hormones, and hormone-binding protein levels.

Changes in hormone levels could provide clues regarding the mechanisms by which some of the effects of CR are mediated.

3. Lowers plasma growth factor concentrations, including Insulin-Like Growth Factor-1 (IGF-1), Platelet Derived Growth Factor-AB (PDGF-AB) and Transforming Growth Factor- $\beta$  (TGF- $\beta$ ).

A reduction in growth factor levels has been hypothesized to play an important role in mediating the effects of CR on longevity and protection against the development of malignancies.

4. Decreases the intermediate risk factors that are predictive of developing atherosclerosis and Type 2 diabetes as evidenced by improvements in circulating inflammatory cytokines and CRP, serum lipid and lipoprotein levels, lowering serum insulin and glucose levels, reduction in abdominal fat, lowering of blood pressure, and reduction of 10-yr CHD risk using comprehensive population-based risk models.

These adaptations are likely to be among the most important protective effects of CR against secondary aging.

5. Reduces oxidative stress, as reflected in urinary levels of isoprostanes, dinitrotyrosine and 2-deoxyguanosine levels.

CR has been shown to lower oxidative stress in rodents, and this is thought to be one of the mechanisms by which CR slows aging.

6. Modulates immune function as reflected by a change in lymphocyte count, delayed-type hypersensitivity (DTH) and response to vaccine.

Short-term CR has been shown to impact immune function. These studies are important because immune function impacts many other physiological and cellular systems. In addition changes in the immune system may determine responses to environmental pathogens thus affecting mortality.

These measures capture the effects of CR at the structural and physiological levels. Many of the adaptations to CR occur at the cellular and subcellular levels. To explore the effects of CR on cellular structure and function, Phase 2 of CALERIE will collect and process tissue samples as a means of determining the cell structure, signaling, gene expression and pathways that mediate the beneficial effects of CR. This will occur for both candidate pathways (listed below in Exploratory Aims 7 and 8) and across the transcriptome, genome, and proteome.

7. Skeletal muscle. To explore the effects of CR to:

- Increase the expression of SIRT1 and FOXO and decrease the expression of type II deiodinase.
- Lower growth factor expression (IGF-1)
- Reduce protein glycation
- Alter the capacity of metabolic pathways (PDHK)

CR changes the expression of key genes, activation / de-activation of signaling systems. Taken together, these changes result in beneficial downstream changes in structure and function. These effects occur in animals and humans and in diverse tissues including skeletal muscle. The latter tissue is responsible for a large portion of the whole body energy consumption. Understanding the effects of CR in skeletal muscle will provide key contextual information for the interpretation of the primary aim of the study – metabolic adaptation.

8. Adipose tissue biopsies. To explore the effects of CR to:

- Reduce adipocyte size and lower the adipose tissue content of inflammatory cytokines and markers of inflammation as measured by the gene expression (qRT-PCR) of amyloid A, IL-6,

resistin, adiponectin, leptin, MCP1, CD68 and MAC2 and Immunohistochemistry to quantify adipose tissue macrophage infiltration content.

- Change the levels of expression of mitochondrial and lipogenic genes.
- Increase the capacity for free radical scavenging (SOD).

There is evidence that a major beneficial effect of CR is a reduction in fat mass, particularly visceral fat, and adipose tissue is now recognized as a contributor to whole body inflammation. These measures will explore the effects of long term CR on adipose tissue. Adipose tissue is important for the regulation of whole body energy partitioning (vis-à-vis metabolic and hormonal systems) and will provide key contextual information for the interpretation of the primary aim of the study – metabolic adaptation.

9. Psychological factors:

- Quality of life.
- Cognitive function and affective status.

The complex endocrine, physiological, and cellular changes that occur during CR theoretically may impact these two domains in both a positive and negative fashion. As such, they represent important information to determine the safety and benefits of long-term CR.

10. Physical function status.

- VO<sub>2</sub>max
- Strength

Changes in caloric intake during CR impact body composition. How these changes impact function will determine the ultimate desirability of CR in humans and the overall interpretation of the data generated in phase 2 CALERIE.

Given the wealth of basic research in the biology of aging and the mechanisms by which CR enhances lifespan in pre-clinical models, these measures represent only a fraction of the potential scientific value of the phase 2 of CALERIE. As such, an important exploratory aim of phase 2 is:

11. To collect and archive biological specimens (plasma, biopsy samples, circulating cells, urine) and store them such that they can be used for testing of new hypotheses (under the parent protocol) and for ancillary studies (funded from a variety of sources).

A study of the size and scope of CALERIE is unlikely to be replicated in the near future. Many researchers, both basic and clinical, will be interested in testing established and new, novel hypotheses generated in the pre-clinical setting. To meet these demands, additional biological specimens will be collected and archived for this purpose and maintained by the Emerging Science Committee as described in Section 12.12 below.

### **3. BACKGROUND AND SIGNIFICANCE**

#### **3.1 Introduction**

All physiological systems demonstrate age-associated decrements in function that can be considered Primary Aging, and many of these changes are known to be prevented or delayed by long-term caloric restriction (CR) in small animal models [1]. However, the extent to which these benefits of CR may also occur in humans is not known. Moreover, the extent to which sustained human CR might be successful but accompanied by unacceptable side-effects has not been investigated. Comprehensive Assessment of Long-Term Effects of Reducing Intake of Energy (CALERIE) is the first study designed as a randomized clinical trial of sustained caloric restriction in humans. Phase 1 of CALERIE, recently completed, was designed as several pilot protocols of human CR, with the goal of preparing necessary data to inform the planning and conduct of Phase 2 CALERIE, which is described in this protocol and is a 2-year randomized clinical trial of CR vs. ad libitum eating in adult men and women.

#### **3.2 Energy Metabolism**

Several studies have been conducted on the effects of CR on energy expenditure in animal models. Some work has suggested no decrease in metabolic rate with CR throughout the lifespan [2, 3] while

other has suggested [4, 5] a decrease. There is now general recognition [6] that appropriate data normalization is essential for correct interpretation of changes in energy expenditure with experimental treatments that involve weight change but there remains no general consensus on the effects of CR on energy expenditure and metabolic rate in rodents. In monkeys, as in rodents, there are reports of both no change in energy expenditure (measured in this case as oxygen consumption) [7, 8] and decreased expenditure [9, 10]. In summary, studies in monkeys have yielded mixed findings about the effect of CR on energy metabolism.

The first experiments of the effect of energy restriction in humans were in lean men by Keys et al in the 1950s [11]. In these classic experiments lean volunteers received 50% of their habitual intake. There was a decreased basic metabolic rate (BMR) when adjusted for body surface area (-31%), body weight (-20%) and for cell mass (-16%). However, there were indications of malnutrition with deficiencies in many micronutrients. Most of the other studies of the effect of energy restriction on energy metabolism have been performed in obese people. In several studies, a very low calorie diet resulted in a decrease in BMR, which was still significant when expressed per kg of body weight or per kg of fat-free mass [12-14]. A recent meta-analysis of studies in post-obese patients found a lower resting metabolic rate, even after adjustment for body size and body composition [15].

Comparable results for change in metabolic rate and sedentary 24 hour energy expenditure were also obtained in an investigation of Biospherians who were subject to CR for 2 years [16]. Part of the adaptation seen in these different studies may be related to the cost of physical activity as elegantly shown by Weigle and Brunzell [17] but adaptations unaccounted for by changes in body composition or physical activity are also suggested. Therefore, there is evidence that a metabolic adaptation develops in response to CR and loss of weight in humans, in both obese and lean subjects. The reason for the paradox between rodents and humans in regard to an adaptation in energy expenditure in response to CR is not apparent. One possibility is that the human methods to measure energy expenditure are more sensitive and investigators can obtain full cooperation of the subjects during the measurements.

Thus, there is evidence for changes in energy expenditure with CR that include decreases associated with loss of weight and body tissue, and decreases that appear to be independent of body composition change which are indicative of a metabolic adaptation to CR that provides for a greater efficiency of energy utilization and is likely to have both a hormonal basis and be linked to physical manifestations of low energy expenditure such as reduced core temperature. Core temperature is known to vary with metabolic rate [18] and is considered a biomarker of aging in rodents and monkeys [19, 20]. Preliminary data from CALERIE Phase 1 indicate a significant reduction in core body temperature in the caloric restriction and caloric restriction + physical activity groups.

### **3.3 Oxidative Changes and Damage**

Increases with age in oxidative changes in lipids and macromolecules in a variety of tissues in laboratory animals have been reported. However, causal relationships between these changes and rates of physiologic aging changes or life span have not been conclusively demonstrated in vertebrates. Recent studies in transgenic mouse models suggest that certain specific types of oxidative damage (e.g., different types of protein oxidation) in certain subcellular locations (e.g., mitochondria) may play causal roles, while others do not [21, 22]. It is also becoming more apparent that some types of protein oxidation play physiologic roles in antioxidant defenses and regulation of cell functions [23].

There is also evidence that an age-related increase in oxidative stress that leads to lipid and protein oxidation, plays a role in the development of atherosclerosis [24, 25]. More importantly, Protein oxidation by hydroxyl radicals, myeloperoxidase, and reactive nitrogen intermediates (such as nitric oxide) appears to be of key importance because of the fundamental role of proteins as biological catalysts [26, 27]. o-o'-Dityrosine is a marker for damage by hydroxyl and tyrosyl radicals [26]. In mice fed ad libitum, levels of o,o'-dityrosine increase with age in cardiac and skeletal muscle [28]. This increase in dityrosine is prevented by CR [28], and, in aging rats, by antioxidant supplementation [29]. A major methodological breakthrough relative to clinical studies is the finding that levels of the oxidatively modified amino acids in urine mirror those in skeletal muscle and heart, and can serve as a non-invasive measure of oxidative stress in vivo [29, 30]. Some, but not all, laboratory animal studies on the effects of caloric restriction (CR) on the level or accumulation of oxidative changes in tissues have found a relationship. Some laboratory animal studies (but not all) have also found effects on mitochondrial production of oxygen radicals or oxyradical-generating species (e.g., peroxide).

An important consideration in evaluating the laboratory animal literature on CR's effects is the length of time, and percent of life span that had elapsed after initiation of CR at the time oxidative changes were evaluated. In the majority of studies, effects on oxidative status were not measured until after a substantial proportion of the life span had elapsed after the beginning of CR (equivalent to one or more decades in humans). In most cases it is not possible to determine from these studies whether or not differences in oxidative status appeared quickly or whether CR and control animals slowly diverged. There are, however, some shorter-term laboratory animal studies on effects of CR (e.g., six weeks), which have reported effects on oxidative status.

It has been proposed the CR effects on mitochondrial oxyradical production may contribute to effects on oxidative tissue changes. As noted above, this is consistent with some experimental evidence in laboratory animals. If mitochondrial oxyradical production is a primary factor in such changes, it is reasonable to hypothesize that CR's oxidative changes in mitochondrial DNA, proteins, and/or lipids may be especially pronounced than its effects on other cellular components. Effects of CR on oxidative damage to mitochondrial DNA have been reported [31].

### **3.4 Cardiovascular Risk Factors**

The classic risk factors for atherosclerotic diseases (cardiovascular, cerebrovascular and peripheral vascular) include dyslipidemias (high LDL cholesterol, low HDL cholesterol and elevated triglycerides), elevated blood pressure, diabetes mellitus, and smoking. All but smoking are responsive to caloric restriction directly in animal models, or to the weight loss associated with caloric restriction.

More recently, significant cardiovascular risk has been associated with the presence of the metabolic syndrome, characterized by a combination of the elements of atherogenic dyslipidemia (elevated triglycerides in the presence of depressed HDL cholesterol, or more specifically, the predominance of the small dense LDL and HDL particles), pre-hypertension to frank hypertension, elevated waist circumference (reflective of excess in visceral and subcutaneous abdominal adipose tissue) and elevated blood glucose (indicative of the late stages of progression to diabetes) to frank diabetes mellitus itself. Finally, cardiovascular fitness is now recognized as an independent predictor of cardiovascular risk independent of other risk factors.

In epidemiological studies, elevated serum cholesterol, both total and LDL cholesterol, were shown to induce atherosclerosis in the absence of other risk factors. This is evident in human [32] and animal studies [33]. Furthermore, lowering cholesterol levels has been shown to reduce cardiovascular disease mortality [34]. Low HDL cholesterol levels predispose to, and high HDL cholesterol levels protect against, development of atherosclerosis [35]. There is also evidence suggesting that high serum triglyceride levels may be atherogenic. CR improves serum lipid profile, reducing the risk for atherosclerosis [36-39]. To what extent caloric restriction modifies lipoprotein particle size and number is unknown.

Elevated blood pressure is a major risk factor for atherosclerosis, stroke and renal failure. Furthermore, cardiovascular aging, resulting in a progressive increase in arterial stiffness, causes a progressive increase in systolic blood pressure, usually with no change or even a decrease in diastolic blood pressure in the elderly [40]. Data from studies on non-human primates suggest that CR lowers both systolic and diastolic blood pressure [41]. In humans, CR can also favorably effect blood pressure as observed during food restriction for two years in the Biosphere 2 experiment [42].

Chronic inflammation appears to increase with aging [43, 44], The production of proinflammatory cytokines by mononuclear leukocytes in response to stimulation also increases with aging [44]. CR reduces the acute release of proinflammatory cytokines in response to endotoxin and other stimuli in mice and monkeys, thus preventing excessive inflammatory responses and perhaps more importantly, protecting against the aging-related dysregulation of proinflammatory cytokine production [44]. Advanced glycosylation end-products (AGEs) also play an important role in the pathogenesis of aging and of cardiovascular and other diseases [45]. Some studies have shown that calorie restriction attenuates the accumulation of AGEs, including N-(carboxymethyl)lysine (CML) and pentosidine, major products of oxidative modification of glycosylated proteins [46, 47].

Adipose tissue secretes a number of biologically active proteins, including leptin and adiponectin. Blood leptin increases with weight/fat gain and decreases with weight/fat loss [48]. Leptin acutely increases fat utilization and suppresses appetite. However, chronic elevation of leptin levels causes

marked resistance to the appetite suppressive and fat mobilizing effects of leptin [48]. Chronic elevation of leptin has a proinflammatory effect, and promotes smooth muscle cell proliferation and arterial stiffness [48-51]. Adiponectin reverses insulin resistance by increasing muscle and liver glucose uptake, decreasing hepatic glucose production and increasing insulin sensitivity [52-56]. A CR-induced increase in circulating adiponectin concentration could, thus, have an anti-aging effect by decreasing protein glycation and formation of AGE products.

Aminoguanidine markedly diminished age-related arterial stiffening, cardiac hypertrophy, glomerulosclerosis and proteinuria [57, 58]. Several mechanisms may explain the role of AGEs in the pathophysiology of primary and secondary aging. AGEs cause increased matrix production, increased vascular permeability and induction of platelet-derived growth factor (PDGF) and transforming growth factor-beta (TGF- $\beta$ ) [59]. Moreover, stimulation of human monocyte-derived macrophages with AGEs resulted in the induction of TNFalpha, IL-8 and expression of tissue factor, depending on the length of stimulation and different glycation products used [60].

TGF- $\beta$ 1 plays a major role in regulating tissue extracellular matrix protein deposition and degradation [61]. TGF- $\beta$ 1 is involved in tissue repair in response to trauma and inflammatory damage [51]. Sustained production of TGF- $\beta$ 1 induces development of fibrosis in many tissues [62]. TGF- $\beta$ 1 stimulates gene transcription and the production of collagen I, III, V and VI, as well as the production of fibronectin, tenascin, osteonectin, osteopontin, thrombospondin and matrix glycosaminoglycans [62]. TGF- $\beta$  also inhibits collagenase transcription and stimulates the synthesis of metalloproteinase inhibitors [63]. The net result is the accumulation of extracellular matrix proteins. Brooks et al. [64] demonstrated that TGF- $\beta$ 1-deficient mice survived longer, and exhibited less myocardial fibrosis and lower myocardial stiffness compared with age-matched controls. It has also been demonstrated that overexpression of tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ) leads to upregulated expression and release of TGF- $\beta$ 1 which in turn increases myocardial collagen content [65]. Recent studies have shown that renal TGF- $\beta$  mRNA abundance and TGF- $\beta$  immunostaining in the renal interstitium are increased as a function of aging [66]. Although it is quite likely that increased expression of TGF- $\beta$  mediates, in part, the age-related tissue sclerosis, definitive proof is still lacking.

### **3.5 Insulin Sensitivity and Secretion**

With aging, insulin resistance and hyperglycemia develop in parallel with increased abdominal fat and visceral adiposity. Whether beta cell sensitivity to glucose remains intact with aging is unclear. In contrast, CR and consequent weight loss in obese (diabetic and non-diabetic alike) greatly improves glucose metabolism by improving insulin action. In a comprehensive review, Kelley [67] concluded that weight loss in obese patients with type 2 diabetes mellitus not only reduces fasting hyperglycemia (reduction of post-absorptive hepatic a glucose production), but also increases insulin sensitivity (glucose uptake) in peripheral tissues (mostly non-oxidative glucose metabolism, i.e. storage).

These data indicate that CR should prevent insulin resistance associated with aging. However, whether this occurs is not yet known. The most convincing data that long-term CR is an effective means of avoiding the development of insulin resistance occurring with aging are currently from monkey studies [68-70, 41]. In addition, the eight Biospherians who were exposed to a severe CR during most of the 2-year period inside Biosphere 2 exhibited a decrease in fasting blood glucose and fasting insulin [71, 72].

Physiological mechanisms for the improvement may include decreases in circulating free-fatty acid concentration [73], intra-myocellular triglyceride and lipid derivatives [74-76] and secreted cytokines from adipocytes [77-79]. Recently, Shulman [80] has summarized the potential molecular mechanisms involved in the relationship between fat and insulin sensitivity. Importantly, there is now growing evidence that mitochondrial number, size, and location in skeletal muscle may be a major determinant of metabolic flexibility and therefore insulin sensitivity.

Furthermore, there is considerable interest in the possibility that lowering of insulin level by CR may play a role in slowing aging. This possibility was suggested by the finding that loss of function mutations in insulin-like receptors extend the life span of *C. elegans* and *drosophila* [81, 82].

### 3.6 Immune Function

Changes of Immune Function with Age: Considerable evidence indicates that aging is associated with impaired regulation of the immune system [83-87]. This decline in immune function contributes to the increased incidence of infectious, inflammatory and neoplastic diseases observed in elderly subjects as well as their prolonged post-illness recovery periods. Prospective studies indicate a higher incidence of morbidity and mortality in elderly subjects with low delayed-type hypersensitivity (DTH), an *in vivo* measure of cell-mediated immune response [88-92].

Changes of Immune Function with Calorie Restriction: CR has been found to significantly affect many age sensitive immunological responses [93-96] in animal models, but there is little information on the effect of CR on the immune response of humans. The splenocyte response to T cell mitogens, antibody and IL-2 production, response to IL-2, mixed lymphocyte reaction, and T cell cytotoxicity have all been observed to be enhanced by CR [93-96]. Shibolet *et al.* [97] observed that T cell function was considerably enhanced in calorie-restricted mice afflicted with colitis (an immune-related disease) than mice fed *ad libitum*. In addition, an improvement in the number of naïve T cells was found in mice exposed to dietary restriction [98]. It is also interesting to note that CR has been found to reduce production of PGE<sub>2</sub>, a T cell suppressive factor, in both mice and rats [99-100]. Recently, Kim *et al.* [101] showed that energy restriction (30%) reduced the age-associated increase in peripheral blood mononuclear cell LPS- and xanthine and xanthine oxidase-induced IL-6 protein and mRNA level in non-human primates.

*Primate Studies.* In another study, Weindruch *et al.* [102] observed that, like rodents, primates showed lymphopenia following dietary restriction. However, no conclusion regarding the effect of CR on mitogen-induced proliferation could be reached because PWM- and Con A-induced lymphocyte proliferation was lower in primates subjected to CR early in life (up to one year) than in control primates. There also was no significant difference in mitogenic response between primates subjected to CR in young adulthood (three to five years) and controls. Nikolich-Zugich and Messaoudi [103] did not observe a change in circulating peripheral blood mononuclear cells (PBMC) after 24 years of CR in non-human primates. These investigators also reported preservation of naïve CD4 and CD8 T cells as well as a decrease in TNF-alpha and IFN-gamma by CD4 and CD8 cells following CR in rhesus monkey. An increase in proliferative ability of lymphocytes was also reported. The authors attributed the difference between their study and those of Weindruch *et al.* [102], to the longer length of CR in their study.

*Human Studies.* Data on health benefits mediated by changes in immune function of CR in humans are very limited. In a recent study with small number of subjects, we [104] showed that weight loss due to consumption of a low fat diet significantly increased DTH in older moderately hyperlipidemic subjects. Consumption of a low fat diet without weight loss did not improve DTH.

*CALERIE Phase 1.* Preliminary results from the pilot phase of the NIA-supported multi-center clinical trial, CALERIE Phase 1, showed that both 10 and 30% calorie restriction significantly improved DTH, Con A and PHA stimulated lymphocyte proliferation in slightly overweight adults. The 30% CR group tended to have a more consistent improvement, but this difference was not statistically significant (see preliminary results). In addition, LPS stimulated PGE<sub>2</sub> production decreased and anti-CD3-stimulated lymphocyte proliferation increased significantly in the 30% calorie restricted group only. In agreement with the study by Nikolich-Zugich and Messaoudi [103], we observed no difference in peripheral blood percent lymphocytes. Taken together, these data suggest that the immunological effects observed in rodents and non-human primates might be re-producible in humans.

### 3.7 Activation of Neuroendocrine Axes and the Autonomic Nervous System

Caloric restriction (CR) in laboratory rodents has many metabolic and endocrine effects, such as retardation of growth and development in young animals, and a decrease of fertility. Long-term CR reduces gonadal and thyroid function, growth hormone (GH) secretion and insulin-like growth factor-1 (IGF-1) production, but activates the adrenal glucocorticoid system [105]. The molecular mechanisms by which CR affects the neuroendocrine metabolic interactions remain unclear.

Evidence from laboratory animal suggests that effects on the GH-IGF axis may be particularly important in mediating its effects on aging and life span. Other hormonal axes, including the glucocorticoid axis, that have parallel effects to IGF on inflammation and cell proliferation may also be important. Studies on effects of CR in humans on these axes would be valuable.



Mice with impaired pituitary development (Pit-1 and Prop1 mutations) live much longer than normal mice [106]. The hormonal changes induced by Pit-1 or Prop1 mutations are similar to, but more marked than, those induced by CR. They include GH, thyroid stimulating hormone (TSH) and thyroid hormone, luteinizing hormone (LH) and follicle stimulating hormone (FSH) deficiencies and result in IGF-1 deficiency, growth retardation, decreased cell proliferation, infertility and an increase in maximal lifespan. GH/IGF-1 deficiency, in particular, is thought to play a major role in extending longevity of GH deficient mice. This effect on longevity is as great as that induced by CR. This slowing of aging has been demonstrated in both Ames and Snell dwarf mice which have mutations that cause deficits in the embryonic development of the anterior pituitary leading to absence of the GH producing cells [106-108]. These mice are also hypothyroid.

However, lit/lit mice that are defective in the response to growth hormone releasing hormone and only have a GH deficiency also show lifespan extension, as do GH receptor knockout mice, which have increased GH levels but, like the GH deficient animals, have extremely low IGF-1 levels [109,110]. Both IGF-1 deficiency and CR not only extend lifespan but also decelerate aging-dependent changes in multiple cell types and organ systems [111-113]. Moreover, it was recently shown that mutations in single genes can markedly increase lifespan in nematodes, and these and related investigations led to the identification of the IGF/insulin signaling pathways as key players in nematode and rodents lifespan regulation [114-116].

Another parallel between CR effects and effects of low IGF-1 include an apparent lowering of risk for at least some malignancies. In laboratory animals CR is associated with a decrease in the incidence of several types of malignancies [117]. There is extensive evidence that insulin-like growth factors are mitogens that stimulate cell proliferation, inhibit apoptosis and play a major role in carcinogenesis [117-120]. It is well documented that CR results in reductions in plasma and tissue IGF-1 levels [117, 111, 121, 122]. IGF-1 is a powerful stimulator of cell growth and proliferation [117-120]. IGF-1 is also a potent inhibitor of apoptosis. High serum concentrations of IGF-1 are associated with increased risk of breast, prostate, colon, lung and other malignancies [123, 124, 117, 119, 125, 126, 120]. In population studies, individuals in the highest quartile of circulating IGF-1 had a 2.5 to 5-fold higher risk of developing a neoplasm than those in the lowest quartile [123, 119]. These findings suggest that slowing of aging and protection against cancer by both CR and GH deficiency are mediated by a reduction in IGF-1.

Other parallels between CR effects and effects of inhibition of the GH-IGF axis include a diminution of inflammatory responses, and a decrease in oxidative stress and enhancement of antioxidant defense mechanisms [127, 128, 102, 28, 129, 130]. [Effects on oxidative stress may mediate effects on inflammation via the transcription factor NF- $\kappa$ B [131,132]. However the GH/IGF-1 axis may also have protective effects against some types of oxidative damage [133,134].

However, although we hypothesize that a decrease in IGF-1 plays a role in the slowing of aging by CR, other factors may also be involved. This is suggested by the finding that CR further extends longevity in Ames dwarf mice [125]. Furthermore, the effect of CR is as great as, or greater than, that of growth hormone deficiency despite a much less marked reduction in IGF-1 level. In this context, it is our hypothesis that, in addition to the reduction in IGF-1 level, CR results in decreases in a range of other growth factors, hormones and proinflammatory cytokines. We also hypothesize that the simultaneous lowering of the levels of other growth factors [e.g. PDGF, epidermal growth factor (EGF), insulin, TGF $\beta$ 1], hormones [T3, angiotensin II, catecholamines and proinflammatory cytokines (e.g. IL-6, TNF-alpha, IL-1, IL-8, MCP-1)] also plays a role in the extension of maximal life span and reduction in the incidence of malignancies by CR.

There is also evidence that CR produces mild stimulation of the glucocorticoid axis [135]. Several (but not all) effects of CR parallel to those of mild glucocorticoid elevation notably decreased inflammation [44, 102, 136]. Adrenal production of DHEA(S) begins during puberty, and peaks at about 20 years of age. Beginning at age 25 years, plasma DHEA(S) begins to decline markedly and rapidly so that by age 75 years plasma DHEA(S) level is ~80% lower than at 20 years [137-139]. Lane et al. [140] have found that CR markedly slows the decline in serum DHEAS in rhesus monkeys, which, like humans, normally show a steady, age-related decline in DHEAS. This finding raises the possibility that DHEAS may be a marker of primary aging. A 24 mo period of CR is too short for an evaluation of the effect of CR on the rate of decline in DHEAS level. Instead, we will test the hypothesis that CR causes a partial reversal of the age-related decrease in DHEAS.

Plasma norepinephrine (NE) may play an important role in primary aging and may mediate some of CR's effects on aging. NE concentration has been shown to increase with advanced age and may lead to deterioration of cardiovascular and metabolic function [141,142]. The increase in NE concentration is due to both increased secretion and decreased clearance, and it could be a potential marker for primary aging. Moreover, weight loss is associated with a decrease in plasma norepinephrine concentration [143].

The renin–angiotensin system (RAS) contributes to the pathogenesis of several human diseases, including hypertension, congestive heart failure, coronary artery disease and diabetic nephropathy. Several reports have demonstrated that the cardiovascular changes observed during aging are similar to those due to high blood pressure with respect to biochemical, mechanical and electro-physiological properties of the CV system [144,145]. In rats, both hypertension and advanced age [145] produce an increase in absolute and relative left ventricular weight and myocardial fibrosis [146]. Structure and function of the arterial wall also show similar modifications in hypertension and aging. The main vascular structural alteration is related to increased media thickness due to smooth muscle cell hypertrophy/hyperplasia and collagen accumulation [147] resulting in increased stiffness [146].

A common feature of many of the hormonal axes discussed above is regulation of cell proliferation and inflammation. It is our hypothesis that alterations in IGF-1, other growth factors and proinflammatory cytokines, and possibly glucocorticoids, induced by CR, mediate antiproliferative, anti-inflammatory, proapoptotic, and antineoplastic effects, which play a major role in the slowing of aging by CR.

### **3.8 Quality of Life and Cognitive Function**

The extent to which CR affects an individual's perception of their quality of life and psychology is not known. Health-related quality of life has been shown to correlate directly with level of obesity in individuals seeking treatment for weight loss, with the most severely obese reporting the poorest quality of life [148]. Furthermore, physical components of and overall health-related quality of life have been shown to improve both with calorie restriction and weight loss in some studies [149]. On the other hand, long-term severe caloric restriction in the Minnesota Starvation Studies was associated with multiple negative impacts on quality of life including reduced libido, symptoms of depression and obsessions with food [11].

A number of studies have investigated if self-reported dieting or dietary restraint is associated with cognitive impairment [150-152], but there is essentially no work on human CR. The majority of these studies have tested if self-reported dieting or dietary restraint is associated with information processing biases [153,154], or if they affect memory performance, processing speed, or reaction time [150-152]. Information processing biases are demonstrated when certain types of information, usually information that is relevant to a person's "self-schema," are cognitively processed in a distinctive style. For example, certain types of information will be better recalled than other types of information. The two most frequently studied information processing biases associated with dieting, or dietary restraint are memory and attention biases [155,153].

Memory biases occur when certain types of information, such as body shape or food-related information is preferentially recalled. A number of studies have demonstrated that self-reported dieting is associated with poorer working memory or immediate recall, in addition to impairments in vigilance and reaction time. The poorer memory performance of dieters has been attributed to smaller working memory capacity, which appears to be a consequence allocating resources to process task-irrelevant cognitions related to food or body weight/shape [150,151]. Although a number of studies have documented poorer working memory performance among self-reported dieters, other studies have failed to detect memory impairment when CR is documented by weight loss [152,156].

Attention biases occur when certain types of information is preferentially attended to among participants who report dieting or elevated levels of dietary restraint. These biases are most frequently detected using the Stroop Color Naming Task and the modified dot-probe task. A meta-analysis of Stroop task studies showed no evidence of attention biases among dieters [157] while a study in highly restrained females utilizing a Stroop task demonstrated bias for both food and body shape-related material irrespective of whether they were currently dieting [153].

## 4. PRELIMINARY STUDIES – CALERIE PHASE 1 STUDIES

### 4.1 Introduction to the Phase 1 Studies

The original Request for Application (RFA) issued by the NIA described a two-step approach to develop the methodology to be applied in this study. First, “Phase 1” studies were conducted to address methodological issues in conducting studies of calorie restriction in free-living individuals. These issues included recruitment and retention, the racial, ethnic and cultural appropriateness of protocols, the feasibility of delivering a CR intervention, techniques to quantify adherence to the intervention, and identifying adverse effects and assessing tolerability. In Phase 1, each clinical site developed its own protocol and developed its own approach to investigating these issues. The goal was to use these studies to inform the present “Phase 2” study.

A number of criteria for success were developed for the Phase 1 studies. Prominent among them were the ability to recruit and retain sufficient numbers of subjects, maintain a high level of adherence to the CR interventions, and demonstrate that subjects were not exposed to undue or untoward injury or harm. In this section, we focus on these specific goals and provide a brief summary of the main findings from these studies. A detailed summary of the results from all three studies is provided in Section 22 below

### 4.2 Description of the Studies

All three studies were designed as parallel-group, controlled, clinical trials. At the PBRC, subjects were randomized in equal numbers to one of four interventions, i.e., Healthy Diet (AHA Step 1), 25% CR alone, 12.5% CR + 12.5% energy expended in PA, and liquid calorie diet (LCD) until 15% weight loss is achieved, then clamped at the new lower weight. At Tufts, participants were randomized to one of four interventions: 10% calorie restricted modified Pyramid / high fiber diet (90P), 10% calorie restricted low glycemic load / high fiber diet (90G), 30% calorie restricted modified Pyramid / high fiber diet (70P), and 30% calorie restricted low glycemic load / high fiber diet (70G). Subjects were assigned in a 1:1:3:3 allocation scheme across the four interventions. At Washington University, participants were randomized to one of three interventions, i.e., Healthy Lifestyle, 20% CR (16% CR for the first 3 months, and 20% CR thereafter), and 20% increase in caloric expenditure using a physical activity (16% increase in the first 3 months, and 20% thereafter). Subjects were assigned in a 1:2:2 allocation scheme across the three interventions. The Tufts and Washington University protocols called for 12 months of intervention and follow-up for their participants; the PBRC protocol called for a 6-month intervention and follow-up period.

### 4.3 Recruitment and Retention

Recruitment for the CALERIE Phase 1 studies began in the Fall of 2002, and by February, 2004, all three sites had completed recruitment into their respective studies. PBRC subjects completed participant follow-up in the Fall of 2004, while the other two sites completed follow-up in the Winter of 2005. All sites met their recruitment goals of 44 subjects per site in the Phase 1 study. Forty-eight subjects were enrolled at the PBRC, 46 subjects at Tufts, and 48 subjects at Washington University.

Drop-out rates are summarized in Table 4.1. On an annualized basis, the drop-out rates were 8.2% at the PBRC, 15.2% at Tufts and 4.2% at Washington University, for an average annualized rate of 9.2% across the three sites. All sites met their targeted retention rates for the Phase 1 studies. Given the intensity of the intervention and evaluation schedule, drop-out rates were felt to be reasonable for this type of study.

**Table 4.1: Summary of drop-out rates in the Phase 1 Studies**

No. of Subjects	PBRC	Tufts	WashU
<b>Enrolled</b>	48	46	48
<b>Dropped out before Month 1 visit</b>	1	1	0
<b>Dropped out between Months 1 and 3</b>	1	0	0
<b>Dropped out between Months 3 and 6</b>	0	2	1

No. of Subjects	PBRC	Tufts	WashU
Dropped out between Months 6 and 9	—	2	1
Dropped out between Months 9 and 12	—	2	0
Total Drop-outs (pct. of enrolled)	2 (4.2%)	7 (15.2%)	2 (4.2%)
Annualized Drop-out Rate:	8.2%	15.2%	4.2%

#### 4.4 Adherence to the Interventions

An algorithm was developed for quantifying adherence to the Phase 1 interventions. It takes into account the total daily energy expenditure (TDEE) as measured by the doubly-labeled water (DLW) method, the change in body energy stores measured by DXA. From this, one can calculate energy intake (EI), and theoretically derive the percent calorie restriction (%CR) relative to baseline realized by any study subject at any assessment time point. One can then average subject-specific levels to the intervention level and perform statistical comparisons across the interventions and over time.

Applying this algorithm, the following results were observed from the Phase 1 studies.

- There was considerable variability in the %CR achieved among study participants. Trends within the individuals, however, appeared to be relatively stable.
- In the true calorie-restricted treatment arms, a significantly non-null level of calorie restriction was achieved in all studies, both at Month 6 and at Month 12.
- All three studies achieved the nominal level of calorie restriction at Month 6. Significant separation among the treatment arms was evident in the PBRC studies; trends were evident in the other two studies.
- Some erosion in the level of %CR was evident over the period from baseline to Month 12. Although a significantly non-null %CR was evident at Month 12, both studies failed to maintain the nominal level of calorie restriction.
- The WashU study did maintain a significant separation between CR and control condition at Month 12.

On the basis of these analyses, it was concluded that a calorie-restricted intervention with high construct validity can be developed and delivered in the Phase 2 study. The approach is described in detail in Section 7 below.

#### 4.5 Summary of Adverse Events

A detailed review of the safety and tolerability of calorie restriction in the Phase 1 studies was presented to the CALERIE Data & Safety Monitoring Board (DSMB) during a conference call on August 9, 2005. In general, these analyses indicate that the CR interventions were well tolerated by the participants. Specifically, the following observations were made.

- Of the 99 participants who completed the initial screening and started a CR or control intervention (only), 82 participants experienced a total of 517 post-intervention AEs during the study period. Of these, 39 participants experienced a total of 57 AEs that were considered intervention-related.
- No deaths or life-threatening adverse events were observed in the study, and no more than one participant in any CR or control group reported more than one serious adverse event (SAE). An appendicitis experienced by participant 03-1122 (20% CR group, Washington University) and ligament injury experienced by participant 01-A00-00-1601 (25% CR group, PBRC extension study) were reported in two different groups. In the opinion of the investigators, both were not related to the CR interventions.
- There were 6 AEs reported in Phase 1 that led to temporary discontinuation of the intervention. Three of these (moderate dyspnea, moderate oropharyngeal swelling, and fall) were observed in three participants enrolled in the 70P, 70G, and 90G groups at Tufts, respectively. Three other AEs (severe depression, insect bite, and moderate anemia) were reported by three participants enrolled in the Healthy Diet group at PBRC. In the opinion of the investigators, only anemia was related to the intervention.

- Seven participants experienced 7 adverse events in Phase 1 whose severity was characterized as “severe” by the investigator. Five of these events (ligament injury, poison oak exposure, high EDE score, depressed mood, and appendicitis) were reported by five participants, one in the 25% CR group at PBRC, one in the 70P group and one in the 90G group at Tufts, and two in the 20% CR group at Washington University, respectively.
- Two participants in the Healthy Diet group at PBRC reported two severe AEs (depression and toothache) that, in the opinion of the investigator, were not related to the intervention. All of these events, except a high EDE score, were considered not to be related to the intervention. No other adverse events that by nature or severity could significantly limit activities of daily living were reported by the participants enrolled in the CR groups.
- Infections were more often reported by the participants enrolled in the CR group than in the control group. However, comparison of the infections between the CR and control groups is limited to the PBRC site only with a small sample size per group. Infections will be closely monitored in Phase 2 study and the informed consent for CALERIE phase 2 will inform the participants about the possibility of increased incidence of infections.

#### **4.6 Laboratory Safety Markers**

A detailed review of laboratory markers of participant safety was also provided to the DSMB during its conference call of August 9, 2005. The following results were observed.

- No participant discontinued the study because of an abnormal laboratory value.
- At PBRC, potassium levels above normal limits were more often observed in the treatment groups [CR, LCD and CR+PA: 5 (42%), 9 (75%), and 3 (25%) participants, respectively] than in the control group [2 (18%) participants]. Potassium abnormalities were not observed at any Tufts treatment group participants. At Washington University an increased potassium level of 5.3 mmol/L (normal range 3.3 – 4.9 mmol/L) was observed in participant #03-1115 enrolled in the Healthy Lifestyle group at Month 6 visit.
- Based on the available evidence, the CALERIE Steering Committee believes that, although the review of the potassium findings does not conclusively rule out a possible effect of a CR intervention on risk for elevated potassium levels, a Phase 2 study of CR can be conducted safely if it includes an appropriate potassium safety surveillance protocol. This protocol is described in detail in Section 0 below.
- Minor transitory changes in total calcium levels and liver function tests did not reach clinical significance and were about equally distributed between the CR and control groups. No other noticeable abnormalities in the chemistry panel tests were observed in CALERIE Phase 1 study.
- Of the 99 participants who completed the initial screening and started the CR or control interventions, decreases in hemoglobin and hematocrit levels from normal at baseline to abnormal at any point during the study were observed in 25 and 55 participants, respectively. Most of these abnormalities were marginally below the lower limit of normal for a site. At the PBRC, changes in hemoglobin and hematocrit levels were considered clinically significant in three participants enrolled in the Healthy Diet group and in one participant enrolled in the 25% CR group. These abnormalities were reported as AEs. One participant at Washington University was enrolled in the 20% CR group with significantly low hemoglobin, hematocrit and RBC levels that all increased by study completion at Month 12.
- No noticeable changes were observed in any urinalysis parameter at any clinical site.

Thus, the CALERIE Phase 1 studies suggest that it is indeed feasible to recruit a sufficient number of participants into this study. It should be possible to maintain adequate levels of adherence to the intervention. Moreover, it does not appear that subjects will be exposed to undue and untoward risk from the CR intervention. To be sure, important lessons were learned from the Phase 1 studies. These lessons have been incorporated into the structure of the CR intervention and the safety monitoring protocol described in Sections 7 and 16 respectively.

### **5. BASIC STUDY DESIGN FOR THE PHASE 2 STUDY**

The CALERIE Phase 2 study will be conducted as a multi-center, parallel-group, randomized, controlled trial (RCT). Participants will be recruited from the three CALERIE clinical centers using proce-

dures that were evidently successful in the Phase 1 studies. Recruitment will be continuous, and generally will include media advertising, direct mail, health promotion events, databases, and referral sources. An effort will be made to recruit an ethnically diverse group based on the demographics of their respective regions.

A staged screening process will be undertaken before any volunteer is officially enrolled in the study. At the time of the initial screening visit, men must be between 21 and 50 years of age (inclusive) and women must be between 21 and 47 years of age (inclusive), and body mass index must be greater than or equal to 22.0 and less than 28.0 kg/m<sup>2</sup>. Otherwise, healthy individuals from both genders and all races are eligible to participate. A volunteer will be excluded if there is a history or clinical manifestation of a significant medical condition, suffers from a disease or condition that seriously affects body weight and/or body composition, exhibits behavioral problems such as eating or psychiatric disorders as well as alcohol and drug abuse, or is unable or unwilling to adhere to the rigor of the intervention and the outcome determinations of this study.

A sample of 250 participants will be enrolled with approximately equal representation across the three CALERIE clinical sites. Volunteers will be assigned at random to either the calorie-restricted (CR) intervention or *ad libitum* (AL) control group. A 2:1 allocation ratio in favor of the CR intervention will be applied in order to maximize the number of subjects receiving the intervention of greater scientific interest. It is expected that 10% of study subjects will drop-out in each of the two follow-up years, so that a sample of approximately 200 subjects is expected to complete the study. On the basis of this study size, there is adequate power to detect effects of scientific and clinical interest.

Protocol-specific interventions will be delivered over a period of 24 months. The active intervention will target a 25% restriction in calorie intake vis-à-vis *ad libitum* energy intake measured by doubly labeled water (DLW) at baseline. There will be no gradual ramping of CR, and the 25% CR goal will apply for the entire 24 months. A strict dietary regimen will not be prescribed uniformly across all subjects. Rather, participants will be allowed to make dietary selections on an individualized basis within the range of the Dietary Reference Intakes. A “toolbox” of dietary and behavioral strategies will be implemented to support successful adherence to the CR prescription. Subjects assigned to the *ad libitum* (AL) control group will be advised to continue on their current diets on an *ad libitum* basis with no further dietary advice or recommendation.

The primary outcomes comprise measures of chronic inflammation, metabolic rate, and risk factors for atherosclerosis and type 2 diabetes. A variety of secondary outcomes are also prescribed. A comprehensive set of evaluations will be performed at baseline prior to initiating the intervention, with follow-up evaluations at 1, 3, 6, 9, 12, 18 and 24 months post randomization.

Given the nature of the intervention and control conditions, it is not possible to blind study participants nor the intervention staff to the treatment assignments. Nevertheless, as far as possible, evaluation staff will be blinded to the treatment assignments.

## **6. STUDY POPULATION AND ELIGIBILITY CRITERIA**

### **6.1 Study Population**

CALERIE will be conducted with healthy volunteers from both genders and all races. An effort will be made to recruit an ethnically diverse study population based on the demographics of regions immediately surrounding the three CALERIE clinical sites. Given the intensive interaction with the intervention staff and the number and types of clinical evaluations, participants must live within geographic proximity of one of the CALERIE clinical sites and be willing to make frequent visits there.

### **6.2 Inclusion Criteria**

At the time of the initial screening visit to the CALERIE clinical site, volunteers must satisfy the following Inclusion criteria:

- Age must be between 21 to 50 (inclusive) for men, and 21 to 47 (inclusive) for women;
- Body mass index (BMI) must be greater than or equal to 22.0 kg/m<sup>2</sup> and less than 28.0 kg/m<sup>2</sup>.
- Female participants must use acceptable form of contraception (barrier method, oral contraceptive, intrauterine device, or similar) and be willing to continue using such a method while enrolled in the study.

## 6.3 Exclusion Criteria

Volunteers will be excluded from the study if they meet any of the following criteria:

### 6.3.1 Medical Exclusion Criteria

- History or clinical manifestation of cardiovascular disease or an elevated blood pressure (greater than 140/90 mm Hg)
- Abnormal resting ECG demonstrating: Type II second or third degree heart block; ventricular ischemia; left bundle branch block, cardiac hypertrophy by any criteria, QRS complex > 100 ms in duration; abnormal QTc interval, supraventricular tachycardia of any type but not including APC's, or ventricular arrhythmia of any type (including VPC's more than 60 per minute)
- BMD *t*-score at the total hip, femoral neck or total spine (L1-L4) is less than or equal to -2.3 at the first DXA scan during the baseline visit
- History or clinical manifestation of diabetes
- History or clinical manifestation of cholelithiasis
- History of anaphylaxis, severe allergies, or asthma
- History or clinical manifestation of any other significant metabolic, hematologic, pulmonary, cardiovascular, gastrointestinal, neurologic, immune, hepatic, renal, urologic disorders, or cancer that, in opinion of the investigator, would make the candidate ineligible for the study
- History of stomach or intestinal surgery (except appendectomy) or major abdominal, thoracic or non-peripheral vascular surgery within one year prior to the randomization date
- Any disease or condition that seriously affects body weight and/or body composition

### 6.3.2 Laboratory Exclusion Criteria

- Potassium level above the upper limit of normal at the screening visit confirmed by a test repeated within two weeks
- Hemoglobin, hematocrit, RBC, or iron level below the lower limit of normal at the screening visit confirmed by a test repeated within two weeks
- Evidence of active liver disease or ALT levels above 1.5 times the upper limit of normal
- LDL-cholesterol level  $\geq$  190 mg/dl at the screening visit

### 6.3.3 Psychiatric and Behavioral Exclusion Criteria

- Individuals who practice a vegan dietary lifestyle
- History or clinical manifestation of any eating disorder as determined by IDED-IV when the ratings for each of the diagnostic criteria are rated as "3" or more. Potential participants will also be excluded if they are experiencing a sub threshold eating disorder (defined as IDED-IV ratings of "3" or more on 5 of the 8 combined symptoms for bulimia nervosa and anorexia nervosa)
- Any history of pharmacologic treatment for a psychiatric disorder within one year prior to the randomization date or a history of more than one episode of a pharmacologic treatment for a psychiatric disorder within lifetime
- History of drug or alcohol abuse (up to 14 drinks a week are allowed) within the past two years
- Individuals who present with a BDI score  $\geq$  20 at screening or baseline

### 6.3.4 Medication Exclusion Criteria

- Short-term (less than a month) treatment with steroids within six months prior to the randomization date
- Treatment with steroids for more than a month within five years prior to the randomization date
- Regular use of other medications, except contraceptives

### 6.3.5 Other Exclusion Criteria

- Individuals who participated in the CALERIE Phase 1 studies
- Individuals who have lost or gained  $\geq$ 3 kg over the past six months
- A volunteer must be either be a never-smoker of tobacco products or an ex-smoker who quit completely at least 12 months prior to the screening visit

- Individuals who donated blood within 30 days prior to the randomization date
- Concurrent participation in any other interventional study
- Breast-feeding or pregnant women, or women intending to become pregnant before the scheduled end of the intervention.
- Individuals who were engaged in a regular program of physical fitness involving some kind of heavy physical activity (e.g., jogging, running or riding fast on a bicycle for 30 minutes or more) five or more times per week over the past year
- Unwilling to be assigned at random to the CR or control intervention
- Unwilling or unable to adhere to the rigors of the CR intervention over the entire two-year intervention period
- Individuals who unable or unwilling to discontinue dietary supplements or adhere to the alcohol consumption restrictions during the study
- Unwilling or unable to adhere to the rigors of the data collection and clinical evaluation schedule over the entire two-year period follow-up period

## 7. TREATMENT INTERVENTIONS

### 7.1 Treatment Summary

The goal of the CR intervention is to achieve and maintain a sustained reduction in calorie intake, rather than a specified degree of weight loss. A two year caloric restriction period was selected to provide for a sustained period of weight stability. It is expected that the period of weight loss will last six months and perhaps as long as a year. Outcome measurements obtained during the early weight loss phase will likely parallel results commonly reported in short term weight loss studies and thus are unlikely to provide any new information. However, sustained caloric restriction past the point where weight stability is firmly established will provide data on whether the independent effects of caloric restriction become stable or if the early changes observed are transitory. In addition, a two-year study will permit the examination of long term changes that might be associated with possible negative outcomes such as changes in bone mineral density and/or others.

Thus the intervention section of the protocol emphasizes strategies and techniques that will promote the participant's ability to consistently follow his/her prescribed level of CR. With a major emphasis of the intervention focused on adherence, the protocol includes a proactive and comprehensive plan for providing the participants with an array of supporting services to aid this effort. This plan uses a computerized algorithm for on-going assessment of physiological, dietary and behavioral factors related to study adherence, linking provision of support to these outcomes. Specific dietary strategies, behavioral support and other practical tools necessary to ensure long term adherence to CR are thus provided on an "as needed" basis but in a systematized fashion.

### 7.2 Conceptual Framework for CALERIE Phase 2 CR Intervention

The underlying conceptual framework for the CALERIE Phase 2 intervention strategy is based on:

- the goals of CALERIE in regard to degree and time course of CR,
- the implications of these goals for design of strategies to achieve good adherence to sustained CR.

A crucial purpose of CALERIE is to determine the similarities and differences between CR's effects in humans compared to those in laboratory animal models whose life spans are extended by CR [158-161]. The overwhelming proportion of evidence from laboratory animals has been obtained in animals in which CR is sustained for a long period after the initial phase of weight loss has ended, and weight is relatively stable [162-164]. Thus, to achieve the goals of CALERIE, the period of initial weight loss must be followed by a period of relative weight stability during which adherence to the specified degree of CR is maintained. Based on Phase 1 results, it may take 9 months or 1 year (or perhaps longer) for the initial period of weight loss to end, even if CALERIE participants succeed in maintaining adherence to a steady level of 25% CR. Thus, the CALERIE Phase 2 intervention design optimizes the likelihood that:

- the degree of CR that participants achieve will be substantial,
- the degree of CR that participants achieve will be sustained at a relatively constant level over the course of the intervention.



Achieving a relatively constant degree of CR over the course of the intervention is a particular opportunity and challenge for CALERIE. In almost all of the previous long-term human studies of CR, the period of weight loss was followed by a period in which weight increased faster in the treatment than in the control group [165,166]. In these cases, it is highly likely that, as the study neared completion, the weight-reduced participants were actually decreasing their degree of CR and were in a considerably more positive energy balance state than the control group participants. CALERIE offers an opportunity to avoid this shortcoming and, through its strong emphasis on long-term CR adherence, provide unique information about the effects of sustained CR in human subjects. We will also be able to document the extent to which weight regain (if any) reflects metabolic adaptation or diminished adherence to CR.

An additional unique challenge facing CALERIE stems from the fact that CALERIE participants will not be obese, and some will not even be overweight. Because of this, we expect the following.

- compared to obese persons, CALERIE participants may have weaker motivation for weight loss and perhaps also CR on the basis of the putative health benefits [167];
- CALERIE participants' dietary habits with regard to energy intake are likely to differ less from accepted nutritional guidelines than do those of obese persons.

Thus, the relative degree of reduction in caloric intake achieved by some nutritional and behavioral strategies that have produced significant caloric restriction and weight loss in the obese may be more difficult to achieve in CALERIE. We intend to emphasize the health benefits of CR in recruitment and throughout the intervention to maximize adherence.

However, these considerations pose a challenge for developing intervention strategies that will induce sufficient long-term behavior and dietary change in non-obese participants to accomplish CALERIE's goals for sustained reduction of caloric intake.

To meet this challenge, we have elected not to limit our intervention options to those likely to be economically feasible for widespread public health applicability at this time. A major factor in our decision is the crucial need to obtain data on the physiologic effects of sustained CR after initial weight loss has stopped, including effects on recognized risk factors for chronic diseases of aging. To the extent that CALERIE can obtain these data, the potential health benefits (and possible risks) of sustained CR can be evaluated. This information could in turn be a crucial element in estimating the cost-benefit ratio of various possible public health CR strategies, should CALERIE results indicate the potential value of testing such approaches.

The overall intervention strategy for CALERIE Phase 2 is an intensive behavioral approach coupled with dietary modifications anticipated to enhance adherence to CR, based on strategies that have been found to be effective in long-term weight loss studies [168,169] and based on short-term work supporting dietary composition changes for enhanced satiety and reduced hunger [170]. Each participant will be given their CR goal as well as the techniques necessary to achieve and maintain it. The intervention materials will include both group sessions and individual counseling, and will provide participants with information (e.g. on potential satiating effects of higher fiber), material aids for adherence (e.g., food scales and electronic personal data assistants (PDAs), services (e.g., provision of meals), and incentives for retention [171]. These strategies are adopted to address the long term nature of the CALERIE intervention which distinguishes it from many other weight control studies. Three particularly salient differences between CALERIE and such studies, and our strategic adaptations to them are:

1) CALERIE's goal is to test effects of reducing total caloric intake, rather than effects of changes in dietary composition with regard to specific nutrients. The effects of CR on aging changes in laboratory animal models appear to be largely independent of changes in dietary composition [172,173]. In view of this consideration, CALERIE will not prescribe a strict dietary composition, but will allow each participant to make dietary selections (with the advice of intervention staff) that best help him or her achieve the CR target, and to vary this as needed or desired over the course of the intervention. Macronutrient recommendations for the study, within ranges of intakes from the National Academy of Sciences Dietary Reference Intake (DRI) guidelines [174], will also be provided.

A complete daily vitamin and mineral supplement will be provided to both the intervention and control group to ensure that participants meet the current recommendations for these nutrients. A daily calcium supplement of 1,000 mg. will also be provided.

2) CALERIE participants will not be obese (and some not overweight). Because of this, and because the intervention will require sustaining a high degree of CR, the relative strength of various factors affecting adherence to CR will differ to some extent from those in weight loss studies in obese persons, and will vary among individuals in CALERIE, and are likely to vary within individuals over the course of the intervention. The CALERIE protocol provides for options from a “toolbox” that allows individualizing the intervention to address the major factors affecting each participant’s adherence over the course of the trial [169]. These factors include:

- incentives and disincentives for weight loss (health, body image, etc.)
- incentives and disincentives to avoid weight regain (health, body image, etc.)
- degree of hedonic satisfaction from a lower-calorie diet (taste, volume, etc.)
- degree of novelty and variety in diet over the two-year course of the intervention
- ease and convenience of adhering to a lower-calorie diet
- dietary approaches to suppress hunger and enhance satiety
- social, psychological, and environmental factors affecting eating and drinking (e.g., stresses, alternatives to eating as sources of satisfaction, roles of food in social situations)
- social and psychological effects of participation in the CALERIE intervention (e.g., participation in groups, interacting with nutritional counselors, acquiring knowledge about foods and nutrition).

In particular, one feature of the intervention strategy is to minimize problems related to participant boredom. Boredom has been noted to be a factor that can limit adherence in long-term dietary interventions [168], and will be a factor in decisions about modification of toolbox selections over the course of the intervention. In addition, the intervention will provide information and activities for participants over entire the course of the study to provide novelty and information on alternative ways of maintaining adherence to the intervention and retention in the study.

A general premise of the CALERIE intervention strategy is that the relationship between the participant and the interventionist will benefit from a mutual understanding or implicit contract, in which the participant agrees to adhere as well as possible and, in exchange, the interventionist agrees to make the intervention as interesting and enjoyable as possible, and to help the participant with his/her individual strategy and problems [175].

3) CALERIE is aimed to obtain a specified percentage reduction in caloric intake rather than a specified degree of weight loss. Because weight changes are an imperfect indicator of the degree of CR achieved due to variability among individuals in metabolic responses and changes in spontaneous physical activity induced by CR [176,177], CALERIE faces a particularly strong need for participants and investigators who are able to estimate caloric intake as accurately as possible. The intervention will emphasize strategies and technologies that will enhance the participant’s ability to know or estimate the caloric content of foods and to accurately report amounts of food eaten and, thus, calories consumed. The PDAs (or pen and paper food diaries, if preferred) will be used to report energy intakes to study investigators for the purpose of on-going diet counseling about diet adherence. Changes in body weight and weight maintenance will also be monitored carefully with the caveat that these are general rather than specific indicators of adherence.

### **7.3 Treatment Intervention**

The intervention will target a 25% CR from baseline *ad libitum* energy intake. Total energy expenditure as measured by doubly labeled water (DLW) will be used to assess baseline energy intake [178]. As an additional method of assessment, energy requirements will also be estimated using the DRI formulas for men and women. There will be no gradual ramping of CR and the 25% energy reduction goal will apply for a period of 24 months.

No specific diet composition is mandated; the macronutrient guidelines for the study are the provided DRI guidelines, however, fiber may be increased up to twice the DRI upper guide as a toolbox option if useful. Alcohol consumption will be permitted; however, intake will be limited to not more than two drinks a day and no more than fourteen drinks per week for men and ten drinks per week for women. No specific level of physical activity will be required or recommended. In regard to inquiries from participants about increasing or decreasing their physical activity, staff will provide a standard response indicating that changes in physical activity are acceptable, while reminding participants that the

goal of the intervention is reduced caloric intake, not weight loss *per se*. In addition, all CALERIE participants will receive the advice that current health recommendations from the Surgeon General (Centers for Disease Control) are for a minimum of thirty minutes per day of dedicated physical activity of at least a moderate level on a minimum of five days per week.

All participants in the CR group will keep diet records using a PDA or paper and pencil, and will record home weight at least weekly following a standard method (e.g., re time of day, clothing).

The intervention will employ a computer-based algorithmic approach that combines specific nutritional and behavioral guidance and strategies for each participant to achieve adherence with the ability to optimize individuals' adherence by selecting specific nutritional and behavioral strategies from a "toolbox" for each participant based on individual needs, current and anticipated problems, and preferences. The algorithmic approach will also provide for continuous feedback and communication between individual participants and their interventionist, and modification of "toolbox" selections for each participant over the course of the intervention, based on the individual's success in achieving adherence, and on problems arising or anticipated during the intervention period. A centrally maintained tracking system, containing data needed for decisions about selecting and modifying toolbox selections for each participant (e.g., data on adherence, attendance at group meetings) will be used in implementing the algorithm. The toolbox will be open continuously until a participant completes the study.

### **7.3.1 Nutritional / Dietary Toolbox**

The following are examples of toolbox items that may be used for individual participants. This list is not intended to be inclusive.

- Offering provided food by the research center for a longer period
- Increasing dietary fiber
- Increasing water consumption
- Increasing proportion of protein in diet
- Decreasing consumption of energy-dense, micronutrient-weak foods and also liquid sources of energy
- Increasing consumption of energy-weak foods
- Decreasing portion sizes
- Training in alternative methods for portion size and caloric estimation
- Modifying proportion of calories consumed at different times of the day (meals, snacks)
- Modifying recipes to decrease energy density (e.g., by increasing water and low-energy components or decreasing fat)
- Modifying recipes to provide greater satisfaction with fewer calories
- Use of portion-controlled meals
- Modifying variety in diet (either increase or decrease)
- Identifying novel foods that lessen boredom and provide satisfaction with low caloric intake
- Strategies for limiting caloric intake in specific settings (restaurants, parties, work environment, home).
- Pre-planning detailed, individualized meal plans and recipes
- Improving abilities and equipment for estimating and recording caloric intake in "real time" (e.g., in restaurants or when purchasing food) and retrospectively
- Modification of eating utensils and cooking equipment in home
- Methods for improving adherence to intervention procedures (e.g., weight measures, clinic attendance)

### **7.3.2 Behavioral / Environmental Toolbox**

- Implement a morning daily weighing protocol whereby the participant makes decisions about food selections/intake for that day based upon the pattern of weight changes.
- Strategies to address social and logistical problems related to diet (e.g., different meals for participant vs. rest of family)

- Strategies to address difficulties in study participation (e.g., provision of transportation, solutions to babysitting needs)
- Strategies to relieve symptoms or complaints that may be related to diet (e.g., fatigue, hunger, bloating)
- Strategies to address difficulties in preparing or obtaining desired foods (e.g., home delivery of meals, provision of take-home meals, help in identifying sources of greater variety of foods that meet a particular toolbox need)
- Identifying and engaging in alternatives to eating as sources of satisfaction (e.g., hobbies, reading)
- Increase contacts with interventionists (phone, e-mail, or in-person)
- Improving measurement and recording of body weight.
- Methods for improving adherence to intervention procedures (e.g., weight measures, clinic attendance)
- Cognitive/behavioral strategies to facilitate lower caloric intake, e.g., “mindfulness” techniques
- Attendance at additional group or individual sessions above the “minimal” frequency (see below)
- Strategies to avoid impulse eating
- Techniques to facilitate eating more slowly
- Methods for improving adherence to intervention procedures (e.g., weight measures, clinic attendance)

### **7.3.3 Implementation of the Algorithm**

The intervention will be implemented by a multi-disciplinary team including dietitians, psychologists, and physicians (at a minimum). This team will meet weekly to manage enrollment inclusion/exclusion issues (initially) and individual tailoring of the intervention (throughout the intervention period). The computer-based algorithm will be implemented to provide participants with specific strategies in both of the above toolbox domains (nutritional/dietary and behavioral/environmental) at all times during the intervention. The overall approach is outlined below and described in the following sections.

#### **Step 1. Initial Intervention Period (4 weeks):**

- Provision of meals
- Biweekly or weekly individual meetings with interventionist
- Orientation/overview of toolbox
- Evaluation of participants’ skills and accuracy in estimating their energy intake and weighing themselves, and initial training in techniques to do this.
- Development of individual’s initial plan, including specific nutritional/dietary and behavioral/environmental toolbox components by interventionist and participant.

#### **Step 2: Long-term Intervention Period:**

- Regularly scheduled reviews by participant and interventionist of adherence to current plan, weight change, and degree of CR, and review of problems encountered.
- Regular evaluation of the participant’s skills and accuracy in estimating his/her energy intake, and continuing training for those with inadequate skills/accuracy.
- Evaluation of participant’s responses to his/her estimates of energy intake and weight change, and provision of guidance on appropriate responses when needed.
- Monthly provision to the participant of new information related to possible ways to facilitate adherence and provide novelty.

At regular intervals (to be proactive) or if a review indicates poor adherence, the interventionist and participant will develop a modified plan by adding or dropping toolbox elements, based on decision rules programmed into the computerized toolbox. This will be repeated until adherence is satisfactory and problems are minimized. If review indicates satisfactory adherence and no significant problems, participants will continue with their current tools.

### **7.3.4 Initial Intervention Period**

During the first four weeks of the intervention, participants will be provided meals and detailed day-by-day menus, either directly from a metabolic kitchen, or from other sources, such as pre-prepared foods or by caterers. The intention of this approach is to provide, by direct example, the guidance and training needed to help subjects achieve their 25% CR goal [179,180]. Meals will be varied during this phase to provide participants with direct experience with differing dietary strategies (e.g., higher fiber, higher protein, lower fat, whole fruit substitution for juices, differing types of seasonings), to the extent that this is feasible.

In group classes and the individual sessions, participants will be given an orientation to the toolbox approach, and provided with examples of differing toolbox strategies.

The participant's skills and accuracy in estimating his/her caloric intake will be assessed during this period, and initial training in techniques to do this will be provided.

The participant will team up with an interventionist to develop a mutually agreed upon individual dietary/behavioral plan to achieve the targeted degree of CR. This plan will reflect the participant's input based on his/her preferences, needs and experience with differing types of provided meals, as well as the interventionists' expert opinion on nutritional and behavioral strategies most likely to be effective for them. The plan will include both specific dietary/nutritional and specific behavioral/environmental goals, and activities on which the participant will continue to work with their interventionist for the rest of the study.

### **7.3.5 Long-term Intervention**

During the remainder of the intervention, participants will prepare or purchase their own meals at least part of the time, but may use provided portion-controlled meals or snacks as a toolbox option.

At regularly scheduled intervals, starting in the second month of the intervention, staff working with individual participants will assess how well the participant is adhering to the intervention, and identify problems affecting current or expected future adherence to the intervention. The assessment will be structured to provide active input by both the interventionist and the participant. Changes or additions to the selection of toolbox options will be triggered by three types of criteria: physiological, dietary, and behavioral. Within each category, determination of whether a specific criterion is met will be made on the basis of pre-specified evidence-based decision rules, detailed in the MOP.

The physiological criteria will be triggered when the participant does not meet acceptable ranges for objectively measured %CR (based on long-term intake/balance measures of energy intake), or predicted weight change (i.e., loss, maintenance). These standards will be derived using calculations that account for energy expenditure and changes in body composition of the participant. Details on calculation of %CR and predicted weight change, and their implementation as adherence tools, are described in Section 14 below.

The dietary criteria will reflect the extent to which the participant is successful in following the CR diet protocol, specifically implementation of the individual's selected nutritional/dietary toolbox options, including achieving and maintaining abilities to estimate caloric intake. The intervention team will make this assessment using information collected (using an objective assessment tool developed for CALERIE) from interventionists who interact directly with the subjects. One trigger for considering changes in tool box options will be self-reported %CR (compared to self-reported baseline EI) of less than 10% or greater than 50%.

The behavioral criteria will assess the participant's current and anticipated behavior and problems in regard to issues such as clinic attendance, performance of agreed-on activities and practical difficulties affecting ability to participate in the intervention, social or psychological problems related to the intervention, and other indicators of the study's impact on overall well being. This information will be collected using standardized instruments, and a feedback assessment form developed specifically for CALERIE that solicits information from participants regarding self-evaluation of their ability to follow the protocol and its effects on their quality of life. Specific examples of the sorts of problems expected are missing visits, lack of family support, poor self-reported adherence to the targeted caloric intake, problems in adherence related to particular situations, stressors, or environments, and reports of hunger, fatigue, or boredom. The tool(s) recommended will be determined by the type of challenge the participant

is having but will follow a sequence of program rules. A self-report of needing assistance will also automatically trigger the interventionist to consider other tools.

Assessments will occur no less often than monthly, with regard to these criteria and will be continuous. In addition, the interventionist will attempt to anticipate future problems such as following up with participants if they unexpectedly miss a group session, or planning how to handle eating during the holidays.

Information from the assessment process will be used, by the interventionist and participant working together to add, maintain, or drop toolbox options, following a problem-solving model similar to that used in the LookAhead study, with the following steps: 1) problem identification; 2) brainstorming to identify possible solutions; 3) cost/benefit analysis (in regard to the demands and benefits for the participant of alternative solutions; and 4) selection of a plan (from tool box items). The plan will be mutually agreed upon by the participant and his/her interventionist.

The new plan will be followed for the for a specified period of time, with on-going re-assessments of its effectiveness and modifications when indicated, as described above. A tool can be continued throughout the study if a participant finds it particularly useful.

### **7.3.6 Group sessions and Individual Counseling**

The duration of contact with participants has been found to be associated with adherence to interventions to reduce caloric intake. Additional evidence of this in long-term interventions comes from results of the CALERIE Extension Study at Pennington, in which participants who received two treatment sessions per month did not regain a significant amount of weight from Week 24 to 48, but the participants who received services only when they requested them during this period regained a significant amount of weight.

The CALERIE Phase 2 intervention will include intensive participant contact through both group sessions and individual counseling. During CALERIE Phase 2, participants in the CR group will attend at least 12 group sessions at regular intervals during the first six months of the study. From months 7 to 24, participants in the CR group will attend group sessions at least once a month. The groups will utilize open enrollment, which allows participants to begin attending group sessions as soon as they enroll in the study. This group format can rely on “modules” of group sessions that cover information that is appropriate for the participant, given the length of time that the participant has been enrolled in the study. Weight measurements will be obtained at these sessions and used by the intervention staff.

In addition to the group sessions, a participant will meet individually with his/her interventionist (e.g., registered dietician or psychologist) at least weekly for the first four weeks, and at least monthly thereafter. During these sessions, the interventionist can provide specific and individualized dietary information to help the participant adhere to the CR regimen and meet his/her calorie target. Participants' adherence and problems will be reviewed, and modifications to the selection of toolbox options will be made as described above, following specified decision rules. These options include attendance at additional group or individual sessions if the participant requires further support and guidance, or is having difficulty adhering to the CR regimen.

Based on prior experience, each behavioral group is 1 to 1.5 hours in duration (4 to 6 hours per month). Individual sessions with an RD or behaviorist vary in length, but they are generally 20-30 minutes in duration, though the first sessions may be longer.

## **7.4 Quality Control of the CR Intervention Procedures**

Recognizing the importance of a successful intervention for the study to accomplish its goal, the CALERIE team will develop an intervention manual based on behavioral and nutritional principles and all sites will follow the procedures outlined. Interventionists at the sites will receive initial training on how to implement the intervention and will be certified prior to study start. An annual training session and re-certification will be required for all intervention staff.

Implementation of the intervention at all clinical sites will be monitored by the Chief Interventionist with significant experience in delivering comprehensive behavioral interventions in a multicenter clinical trials. The monitoring will be performed using the central computer-based system developed for CALERIE.

The QC committee and Chief Interventionist will review adherence data quarterly to assure that the intervention goal is achieved by the sites. An overall decline in adherence at any site will trigger a visit by the Chief Interventionist and/or its designees to retrain and re-certify the site's intervention team.

## **7.5 Control Group**

Participants for the control group will be selected from the same population as the CR group participants and will thus represent a comparable population. All participants will undergo identical screening and assessment procedures. However, after randomization, participants randomized to the control group will be advised to continue on their current diets on a completely *ad libitum* basis. These participants will receive the same vitamin and mineral supplements as provided to the intervention group.

Participants randomized to the control group will measure home weights only during DLW periods. Scales will be distributed for these periods but then collected from the participants after the 14-day DLW period ends. Participants in the control group will have periodic contacts with study investigators and will undergo all outcome measurements in accordance with the same schedule as for the CR subjects.

## **7.6 Participant Retention**

Two techniques that were helpful during Phase 1 were motivational interviewing and enhancement, and active problem solving. Both active intervention and control participants will be periodically surveyed for problems that could lead to dropout. Motivational interviewing and enhancement can be administered on an "as needed" basis. When participants indicate that life circumstances or problems are interfering with their participation, active problem solving can help identify if a problem truly exists, and identify methods to resolve the problem while retaining the participant in the study [181]. Additional incentives (e.g., gift certificates, cookware) will be used to enhance retention [171].

Retention of participants in the intervention group will be facilitated by frequent contact with their behavioral coach and study staff. Frequent contact will foster quick problem solving to adherence issues as they arise. In addition, the strong relationship that develops between the participant and their behavioral coach and study staff also facilitates retention [182]. During the group sessions, participants will be coached on ways to develop intrinsic motivation and rewards for adherence to the CR regimen and completion of the study.

## **7.7 Compensation**

Control and CR subjects are eligible to receive equal financial compensation for participation in the study. Compensation will be provided based on participation in scheduled study activities and tests.

# **8. RECRUITMENT PROCEDURES**

## **8.1 Recruitment Strategies**

Recruitment strategies developed for the phase 1 studies were successfully implemented at all 3 sites and thus will be continued and refined for the present study. Each site will build on their past experience to recruit healthy free-living men and women satisfying the age and BMI eligibility criteria. The Phase 2 age range is similar to that previously recruited by Tufts and Pennington while Washington University targeted an older population. A slightly leaner group is being sought in the current study; however, this is not expected to adversely impact recruitment. After six months of recruitment, CALERIE will evaluate its degree of success in attracting participants using this lower BMI range, and may elect to raise the upper limit if deemed necessary for adequate recruitment.

Each site will make an effort to recruit an ethnically diverse group based on the demographics of their region. Elements that should aid in recruiting minority participants, groups generally less represented in clinical trials, are that treatment is largely behavioral and assignment will be known [183,184]. Thus, two common barriers to participation are eliminated by this design however co-morbid conditions may eliminate some potential minority participants [185]. In addition to ethnic diversity, the sites will also increase their efforts to recruit men. The experience in the Phase 1 calorie restriction studies was that it was more difficult to recruit men. To address this issue, CALERIE intends to use the approach Pennington has found to be successful. Letters will be mailed to the heads of households at addresses

obtained from mailing lists. It is expected that this will result in a higher percentage of male participation.

The recruitment documents will make clear that CALERIE is primarily a study of the potential beneficial effects of calorie restriction on health and prolongation of life in humans and not strictly a study of weight loss for overweight individuals. Recruitment will be continuous and generally will include media advertising, direct mail, health promotion events, databases, and referral sources. Each of these sources has been found to be productive in attracting minority participants to trial participation with direct mail being the most effective [186]. Sites will employ approaches designed to deal with issues particular to their location and demographics. Previous experience will be used to enhance recruitment while minimizing costs. Site specific recruiting materials will be reviewed for content and submitted to their IRB for approval.

The Coordinating Center will provide materials to aid sites in recruitment. A website for participants will be maintained and study brochures will be provided. A videotape explaining the study and demonstrating several of the tests is also planned.

## **8.2 Recruitment Procedures at the PBRC**

To recruit participants for Phase 2 of CALERIE, the Pennington Center will use strategies that have proven successful for recruitment of participants with similar characteristics. The Pennington Center will also rely heavily on their recruitment experience from Phase 1 of CALERIE and experiences obtained from recruiting for similar studies.

Direct contact by mailing will be a major recruitment strategy. Information on age and BMI for over 60,000 people interested in research studies is available in the PBRC database. This database will be queried and potential participants identified. Letters will be mailed in batches to avoid overwhelming recruiters. Another approach will be to purchase a mailing list for select zip codes in the Baton Rouge area. Post cards with study information will be mailed to potential participants on this list. In previous studies, post cards have been sent out in batches of a 1000 every 4 days to achieve a manageable load of incoming calls. Additionally, a center-wide email will be sent to Pennington Center employees and information about the CALERIE study will be printed on LSU employee check stubs.

A variety of media will be used to distribute information about the study. Ads will be placed in the large local newspaper (*The Advocate*) as well as in smaller regional papers serving rural areas in the greater Baton Rouge region. Ads will also be placed in the Interfax Daily which is sent to business offices daily. Additionally, radio ads will be used as one of the main media resources since these have been very successful in the past. Finally, TV ads, which have been an effective method for recruitment, will be run. The Pennington Center's Communications Department, led by Glen Duncan, will use local TV media to place study related interviews on local TV station's morning shows and health segments. These venues provide an excellent opportunity to describe the study, and the exposure is free.

Flyers will be placed in the PBRC recruiting areas as well as distributed at libraries, churches, and places of business. Staff who are asked to participate in church health fairs or health awareness days at local businesses will promote the study by distributing information, presenting the study, and talking to potential participants. Lastly, the Pennington Center has found that "word-of-mouth" is a very effective recruitment tool.

## **8.3 Recruitment Procedures at Tufts University**

To recruit participants for Phase 2 of CALERIE, Tufts will use a variety of methods that proved successful for the CALERIE pilot study that recruited participants with similar demographics. Initial recruiting efforts will utilize a database of potential prior volunteers. The Human Nutrition Research Center on Aging at Tufts (HNRCA) recruiting office maintains a database of over 20,000 potential volunteers from which men and women meeting the age criteria can be identified. This will exclude participants from the CALERIE pilot study. Eligible volunteers will be sent a letter written by the PI explaining the study. Postage prepaid return envelopes will be enclosed to facilitate a response.

Simultaneously, advertising in a variety of media will be conducted. This will include but not be limited to advertisements in regional and community newspapers such as the Boston Globe, the Metro, Bay State Banner, Boston Phoenix and City Shopper. Health related articles written by the CALERIE staff will periodically accompany the newspaper section where the study is advertised. Advertisements



will also be placed on local radio and TV. These advertisements will be strategically staggered using a monthly schedule planner to cover both weekdays and weekend days, and at different times during the day in order to reach the most broad and diverse audience possible. The placement of these advertisements will be planned for times and programs geared towards what is projected to be the study demographics.

Since the internet is a popular means of communication for individuals in the target age group, the central website address for the study will be included on all printed materials and flyers. Using the Tufts intranet (via email) announcements describing the study will be sent periodically to all Tufts affiliated institution employees and students, including those at Tufts-New England Medical Center and all of the professional schools. Posted flyers will also be extensively distributed at grocery stores, community gymnasiums, health clubs, corporations, churches and other high traffic gathering areas to improve study visibility in Boston and the surrounding area. These postings and emails will enhance visibility of the study to individuals working or living in close proximity.

Direct mailings are planned for specific zipcodes in the Boston area. Purchased mailing lists will be used to send postcards; however, it is somewhat unclear how successful this approach will be since it has not been used at this site previously. Tufts also has a number of outreach and extension program sites where HNRCA provided speakers often visit to connect with the public. These sites provide an excellent and free opportunity to describe the study, distribute flyers, and recruit participants, since CALERIE staff can speak with and respond to questions from any interested candidates. Thus a variety of local activities such as word of mouth, local community network systems, and all health fairs and community outreach efforts conducted by the HNRCA will be used to reach the target population. A study specific brochure and a letter from the PI will be either mailed or otherwise distributed to all interested candidates.

Recruitment of minorities will be achieved by advertising in local minority publications and using the HNRCA's community board's minority representatives who are connected to specific minority groups. Recruiting efforts will be conducted to ensure that participants reflect the demographic composition of the Boston area.

#### **8.4 Recruitment Procedures at Washington University**

Recruitment for Phase 2 of CALERIE at Washington University will employ several methods to attract individuals in the appropriate age range with diverse racial and ethnic backgrounds. Recruitment materials will include e-mail messages, recruitment letters, flyers, television and radio interviews, and public service announcements delivered using six different approaches.

The first recruitment approach will involve mass e-mail messages distributed to employees of Barnes-Jewish Hospital and Washington University School of Medicine. This has been an effective and simple method of recruitment in several previous studies, resulting in participants who work in close proximity to the clinic. A second recruitment strategy will utilize mass mailings (through the postal service). Mailing labels will be purchased from InfoUSA, Inc., and will be chosen based upon study-specific demographics. Personalized letters describing the study and inviting participation will be mailed to individuals residing in zip codes surrounding Washington University Medical Center. Furthermore, to maximize enrollment of minority subjects, zip codes in the St. Louis metropolitan area known to be heavily populated by minority groups will be selected for mailing. As in previous studies, batches of 50 to 100 letters will be sent out, enabling the CALERIE study team to respond to the phone and e-mail inquiries in a timely manner before sending out subsequent batches.

A third recruitment approach will involve television and radio interviews with TV and radio stations in the St. Louis metropolitan area. Three television networks in St. Louis have health reporters who conduct interviews with the investigators and research participants and air the stories on the news. These interviews include descriptions of the study, perspectives from current or previous participants, eligibility criteria, and contact information so that those who may be interested in participating can call or e-mail the study team for additional information. This has been a very effective recruitment strategy. In addition, radio interviews and advertisements on local radio stations will be utilized as needed.

A fourth approach will involve the Washington University School of Medicine "Volunteer for Health" program. This program is operated by the Washington University Center for Clinical Studies as a recruitment mechanism for clinical trials conducted throughout the medical center. The program provides

information to the public about ongoing research through display cases, flyers, and a website. In addition, this resource maintains a large database that can be searched for potential volunteers who meet the inclusion criteria for a study. A fifth recruitment approach will involve targeting local businesses, churches, colleges, and grocery stores with recruitment flyers. Finally, participants from the phase 1 CALERIE study who expressed interest in promoting our study will be contacted. Historically, a large number of participants enrolled in our studies after being referred by other participants who enjoyed their participation in our studies.

## **9. SCREENING PROCEDURES**

The screening process is designed to be rigorous, thorough, and detailed to ensure that only the most suitable candidates for CALERIE are enrolled in the study. Candidates interested in participating in the CALERIE study will undergo a multi-step screening process including a phone screen and in-person clinic visits consisting of biological and behavioral assessments. A list of the tests and procedures that will be used for the screening process is outlined below.

At each site, trained recruiters familiar with the CALERIE protocol will conduct an initial telephone screen. After obtaining verbal consent, the recruiter will complete a standardized form common to all sites. This form will include questions regarding name, age (year of birth) height, weight, telephone number, address and basic eligibility information on demographics and health. Callers who call in after hours will reach a study-specific voicemail requesting their name and contact telephone number(s). All interested candidates who leave a message will be contacted within one week. Candidates who remain interested and meet the basic eligibility criteria will be scheduled for a clinic visit.

At the first visit, a screening informed consent will be sought by the PI or his/her representative and documented on forms approved by the site's Institutional Review Board (IRB). Body mass index (BMI, kg/m<sup>2</sup>) will be verified using measured height and weight to ensure that the candidate is within the eligible BMI range of 22.0 to 27.99 kg/m<sup>2</sup>. After establishing eligibility, the candidate will be assigned a CALERIE ID which they will retain until completion of the study. In the event that a candidate is determined to be ineligible during the screening process, his/her ID number will be discontinued. Candidates will watch a video presentation early in the screening process so they understand the time and study commitment that would be required for participation, and the study coordinator will be available to answer any questions. Candidates may choose to withdraw from screening at any time during the process.

Candidates will also be asked to come in to the center following a 12-hour overnight fast. A fasted clinic weight will be measured at this visit. A study physician will obtain a detailed medical and medications history and perform a physical examination to determine medical eligibility of the candidate. A standard 12-lead electrocardiogram will be recorded and reviewed by a physician and, if necessary, by a cardiologist. Blood for hematology, clinical chemistry (including a lipid panel), a serum pregnancy test, where applicable, and urine for urinalysis will be collected. A repeat blood draw may be required if any abnormal laboratory values are found in the first sample, thereby initiating another clinic visit.

The behavioral screening will target factors expected to affect adherence to the intervention and study retention (sense of commitment, interest in personal health, behavioral and mental health factors, daily life events including social situations and support, and stage of change). Candidates will fill out questionnaires to assess eating patterns, weight history, eating disorders, lifestyle and psychopathology. The tests and measures that will be performed include the following questionnaires: the Eating Inventory [187], the Multiaxial Assessment of Eating Disorder Symptoms (MAEDS) [188], the Structured Clinical Interview for Diagnosis of DSM-IV Personality Disorders (SCID II) Personality Questionnaire [189] and the Beck Depression Inventory-II (BDI-II) [190]. A General Health Questionnaire (GHQ) will also be used as a screening measure. The study psychologist or a trained member of the behavioral team will score these questionnaires. These questionnaires have been selected based upon their ability to predict weight loss during CALERIE Phase 1 and their ability to quantify the degree to which potential participants exhibit eating disorder attitudes and behaviors.

Barriers to participation in the study will be assessed using a standardized interview which will also be administered by trained behavioral staff. This assessment will ascertain participants' willingness to commit to the study, support from household members to participate in the study, motivation and challenges facing the participant and other similar study specific issues of importance. A computer based

body morph assessment test [191] will be performed to identify candidates with both current and potential body image issues.

Candidates will meet with the study psychologist or a trained member of the behavioral team for a SCID-II interview [189] if indicated (i.e., if person meets criteria for personality disorders based on the SCID-II Questionnaire), and the Interview for Diagnosis of Eating Disorder (IDED-IV) [192] if indicated (i.e., if the cutoff scores from the eating inventory or MAEDS are exceeded). A template of the study schedule will be provided and reviewed and the candidate will be asked to complete a calendar of their work schedule and travel obligations.

Candidates will also meet with study personnel to review work or personal schedule conflicts as well as regarding any test scores from questionnaires that may require a more detailed discussion. Detailed demographic information pertinent to the study such as physician information, education level, race, ethnicity, occupation, marital status, emergency contacts etc. will be obtained using a demographic questionnaire. A Stanford brief physical activity questionnaire [193] will be used to assess the activity level of each candidate to screen out those who are very physically active or are in training. A study dietitian will use a study-specific diet questionnaire to obtain detailed information about the candidate's diet. Information will be collected on major food allergies or special conditions (gastrointestinal conditions) which may affect participation in the study, current supplement use and feelings about discontinuing supplements for the study, alcohol consumption patterns and other diet related issues. The dietitian will also provide detailed verbal and written instructions on how to complete a 14-day food record. This process is largely designed to serve as a behavioral run-in and to assess the ability of a participant to adhere and complete a food record continuously over a 2 week period. Candidates will then bring back their completed 14-day food records and meet with the dietitian to discuss the record. The completeness, accuracy, and contents of the diet record will be discussed. Issues such as underreporting, not reporting accurate portion size, not reporting recipes, and any missing information will be discussed and, if necessary, a repeat food record will be recommended. The candidate's willingness to reduce his/her dietary intake in the event that s/he is randomized to the CR group will be discussed.

Eligibility will be assessed using predefined criteria presented in Section 6 above. Results from the various assessments will be summarized on screening checklists and these documents will be used to provide an overall eligibility assessment for each candidate. A multidisciplinary team consisting of behavioral experts, dietitians, clinical staff and CALERIE study staff will meet to jointly discuss and approve a candidate for official admission to the study. The study coordinator at each site will inform candidates of their acceptance to the study and prepare for the participants' enrollment in the study. Ineligible candidates will be informed as to why they were not eligible for the study.

## **10. RANDOMIZATION AND BLINDING**

### **10.1 Randomization Procedures**

To protect against selection biases, CALERIE participants will be assigned to intervention using a random process. Randomization will be stratified by sex and BMI within each clinical center, with BMI dichotomized into two categories, namely normal weight, i.e., 22.0 to 24.9 kg/m<sup>2</sup>, and overweight, i.e., 25.0 to 27.9 kg/m<sup>2</sup>. Within each stratum, subjects will be allocated in a 2:1 ratio in favor of the 25% CR intervention.

Within each stratum, randomization sequences will be generated *a priori* using a permuted block randomization technique [194-196]. Blocks of size 3 or 6 will be chosen, with the block size varied in a random manner. Within each block, allocations will be assigned in the 2:1 ratio in favor of the 25% CR intervention, and as a result, the desired allocation ratio will be maintained at periodic intervals throughout the recruitment process.

A CALERIE ID number will be assigned to each volunteer at the beginning of the screening process. The ID number uniquely identifies a specific participant, and will be used for that participant throughout the study including the screening, baseline and follow-up phases. If the participant is found to be ineligible during the screening process, or drops out or is withdrawn from the intervention at any point, the ID number is retired and not re-assigned to any future volunteer.

A study subject will be randomized and assigned to treatment only when it has been determined that s/he has satisfied all the eligibility requirements for the study, has completed all the baseline as-

assessments, and informed consent for the intervention phase has been obtained. Treatment assignment will be performed using an interactive voice-response system (IVRS). This is an automated telephone system maintained at the Coordinating Center, available 24 hours per day, 7 days per week. When a subject is ready to be randomized, the CALERIE staff member phones a toll-free telephone number to access the IVRS system. The CALERIE staff member logs on to the system with an ID number and password. S/He then keys in the participant's CALERIE ID number, identifies the appropriate stratum, and verifies that the participant fulfills all the eligibility requirements. The software then accesses the randomization sequence for that stratum, determines the next treatment assignment, and communicates the treatment assignment to the CALERIE staff member. The software records the CALERIE ID number and the time and date of that treatment assignment so that the Coordinating Center can verify that subjects were randomized in the correct order.

## **10.2 Blinding of Study Personnel**

Given the nature of the CR intervention and control conditions, it is not possible to blind study participants or CALERIE staff members to the treatment assignments. Nevertheless, an effort will be made to distinguish CALERIE staff delivering the interventions versus those performing the outcome evaluations. That is, within the resources available to this study, intervention staff will not be engaged in evaluating participants, and evaluation staff will be blinded to the treatment assignments. Moreover, an effort will be made to minimize communication between the two groups of staff members. Moreover, reliability studies will be conducted with the central resources and reading centers. These facilities will be blinded to the identities of the samples selected for the reliability testing. This is described in more detail in Section 15.5 below.

# **11. SCHEDULE OF EVALUATIONS AND PROCEDURES**

Whenever possible, each participant will be evaluated by the same site personnel throughout the study. Section 23 below shows the procedure schedule for the study.

## **11.1 Evaluations Performed During the Screening Period**

### **11.1.1 Phone Screen**

- Brief study overview
- Phone screening questionnaire
- Schedule in-person Screening Visit #1

### **11.1.2 In-person Screening Visit 1**

- Obtain screening informed consent
- Height and weight, BMI
- Assign CALERIE ID
- Review information video
- Demographic information
- General Health Questionnaire
- 7-day Stanford PAR
- Eating Inventory
- SCID II
- BDI
- MAEDS
- Dietitian interview
- Assessments calendar
- Schedule in-person Screening Visit #2

### **11.1.3 In-person Screening Visit 2**

- Blood chemistry, hematology and urinalysis
- Serum pregnancy test (for women)

- Physical examination
- Medical and Medications History
- Contraception use
- ECG
- Barriers to Participation interview
- Body Morph Assessment
- 14-day food record
- Schedule in-person Screening Visit #3

#### **11.1.4 In-person Screening Visit 3**

- Blood chemistry, hematology and urinalysis review
- Repeat blood chemistry, hematology and urinalysis, if necessary
- 14-day food record review

#### **11.1.5 Eligibility Assessment**

- Screening tests review by the screening team
- Repeat 14-day food record, if necessary
- Schedule the Baseline visit.

### **11.2 Evaluations and Procedures Performed During the Baseline Period**

- Obtain informed consent and HIPAA Authorization form
- Abbreviated medical and medication history
- Physical examination
- Vital signs
- Clinic weight and height
- Home weight
- ECG
- Blood chemistry, hematology and urinalysis
- Urine pregnancy test
- BDI
- EE by DLW (two 14-days periods)
- RMR (two measurements)
- Core body temperature
- Cardiovascular risk factors
- OGTT, insulin and C-peptide
- Immune function measurements
- Endocrine response and growth factors
- Sex hormones (for men)
- Quality of Life assessments
- Psychological assessments
- Cognitive function measurements
- Six day diet record (two 6-day periods)
- Stanford 7-day PAR
- Muscular strength and endurance
- Maximal oxygen uptake ( $VO_{2max}$ )
- Body composition using DXA
- BMD and BMC of the hip, spine and forearm using DXA
- Markers of bone turnover
- Blood and urine for archive

- Muscle biopsy
- Abdominal fat biopsy
- Concomitant medications
- Randomization

### **11.3 Evaluations Performed During Participant Follow-up**

Most of the outcomes and safety measurements will be performed at the following scheduled visits. Some of the outcomes and safety measurements will follow a different schedule and are described in Section 11.4 below. Month 6, 12, 18, and 24 examinations cannot be completed in one day and will be performed over the period of 14 days. The length of these visits is driven by the length of the DLW study. Thus, a visit window refers to the start (first day) of the visit period that must be within a window rather than to a single day within this window.

Month 1 and 9 visits do not involve substantial number of measurements and can be completed in one day. Visit window for Month 1 and 9 visits refers to a single day that must be within the indicated window. Randomization date will be used as a reference to calculate a scheduled visit date.

#### **11.3.1 Month 1 ( $\pm 3$ days)**

- Clinic weight
- Resting BP
- Vital signs
- Waist circumference
- ECG
- Blood chemistry, hematology and urinalysis
- Concomitant medications
- Contraception use
- Adverse events
- BMI (CR intervention group only)
- BDI

#### **11.3.2 Month 3 ( $\pm 7$ days)**

- Clinic weight
- Resting BP
- Vital signs
- Waist circumference
- ECG
- Blood chemistry, hematology and urinalysis
- Blood for archive
- Concomitant medications
- Contraception use
- Adverse events
- Monitoring eating disorders
- BMI (CR intervention group only)
- BDI

#### **11.3.3 Month 6 ( $\pm 14$ days)**

- EE by DLW (CR intervention group only)
- RMR (CR intervention group only)
- Core body temperature
- Clinic and home weight
- Resting BP

- Vital signs
- ECG
- Blood chemistry, hematology and urinalysis
- Urine pregnancy test (CR intervention group only)
- Quality of Life assessments
- Psychological assessments
- Cognitive function measurements
- Six day diet record (CR intervention group only)
- Stanford 7-day PAR (CR intervention group only)
- Body composition using DXA (CR intervention group only)
- BMD and BMC of the hip and spine using DXA (CR intervention group only)
- Waist circumference
- Markers of bone turnover
- Blood for archive
- Concomitant medications
- Contraception use
- Adverse events
- Monitoring eating disorders
- BMI (CR intervention group only)
- BDI

#### **11.3.4 Month 9 ( $\pm 14$ days)**

- Clinic weight
- Resting BP
- Vital signs
- Waist circumference
- ECG
- Blood chemistry, hematology and urinalysis
- Concomitant medications
- Contraception use
- Adverse events
- BMI (CR intervention group only)
- BDI

#### **11.3.5 Month 12 ( $\pm 14$ days)**

- EE by DLW
- RMR
- Core body temperature
- Clinic and home weight
- Resting BP
- Vital signs
- Physical examination
- ECG
- Blood chemistry, hematology and urinalysis
- Urine pregnancy test
- Cardiovascular risk factors
- OGTT, insulin and C-peptide
- Immune function measurements

- Endocrine response and growth factors
- Sex hormones (for men)
- Quality of Life assessments
- Psychological assessments
- Cognitive function measurements
- Six day diet record
- Stanford 7-day PAR
- Muscular strength and endurance
- Maximal oxygen uptake ( $VO_{2max}$ )
- Body composition using DXA
- BMD and BMC of the hip, spine and forearm using DXA
- Waist circumference
- Blood and urine for archive
- Muscle biopsy
- Abdominal fat biopsy
- Adherence to the intervention
- Concomitant medications
- Contraception use
- Adverse events
- Monitoring eating disorders
- BMI (CR intervention group only)
- BDI

#### **11.3.6 Month 18 ( $\pm 14$ days)**

- EE by DLW (CR intervention group only)
- RMR (CR intervention group only)
- Clinic and home weight
- Resting and 24-h BP
- Vital signs
- Waist circumference
- ECG
- Blood chemistry, hematology and urinalysis
- Urine pregnancy test (CR intervention group only)
- Six day diet record (CR intervention group only)
- Stanford 7-day PAR (CR intervention group only)
- Body composition using DXA (CR intervention group only)
- Blood for archive
- Adherence to the intervention
- Concomitant medications
- Contraception use
- Adverse events
- Monitoring eating disorders
- BMI (CR intervention group only)
- BDI

#### **11.3.7 Month 24 ( $\pm 14$ days)**

- EE by DLW
- RMR



- Core body temperature
- Clinic and home weight
- Resting BP
- Vital signs
- Physical examination
- ECG
- Blood chemistry, hematology and urinalysis
- Urine pregnancy test
- Cardiovascular risk factors
- OGTT, insulin and C-peptide
- Immune function measurements
- Antibody response to vaccines
- Endocrine response and growth factors
- Sex hormones (for men)
- Quality of Life assessments
- Psychological assessments
- Cognitive function measurements
- Six day diet record
- Stanford 7-day PAR
- Muscular strength and endurance
- Maximal oxygen uptake ( $VO_{2max}$ )
- Body composition using DXA
- BMD and BMC of the hip, spine and forearm using DXA
- Waist circumference
- Markers of bone turnover
- Blood and urine for archive
- Muscle biopsy
- Abdominal fat biopsy
- Adherence to the intervention
- Concomitant medications
- Contraception use
- Adverse events
- Monitoring eating disorders
- BMI (CR intervention group only)
- BDI

#### **11.4 Off-Schedule Evaluations**

The following evaluations will be performed at times different than the scheduled times outlines above (i.e., off-schedule):

##### **11.4.1 Antibody Response to Vaccines**

- Pneumococcal, hepatitis A and tetanus/diphtheria vaccines administration – Month 17
- Booster doses of hepatitis A vaccine – Months 23
- Antibody response to vaccines - Months 17, 18, 23, and 24

##### **11.4.2 Additional Weights**

- Clinic weight: at every clinic visit including those at Months 17, 18, 23 and 24 for the antibody response to vaccines.
- Home weight: Control group – during each DLW period for 14 days; CR group – daily

### **11.4.3 Additional Safety Measures**

- Adverse events: at every clinic visit including those at Months 17 and 23 for the antibody response to vaccines.
- Concomitant medications: at every clinic visit including those at Months 17 and 23 for the antibody response to vaccines.
- Women who develop amenorrhea for two months - FSH, LH and estradiol at any time during the study.

### **11.5 Evaluations for Drop-Outs and Withdrawals**

Every effort will be made to retain participants in the study. If withdrawal occurs for any reason, as many procedures as possible for the closest scheduled visit (Months 1, 3, 6, 9, 12, 18, or 24) will be performed.

## **12. OUTCOME DETERMINATIONS**

CALERIE Phase 2 study is designed to test some of current hypotheses on the mechanisms underlying the biological effect of CR on aging and extended longevity, and to assess the feasibility and safety of prolonged 25% reduction in caloric intake in non-obese humans. The rationale for selecting the outcomes described below is presented in Section 2 above. It is important to note at this point that in order to examine the longitudinal effects of CR on aging from a mechanistic perspective, complicated, sophisticated and expensive techniques are selected.

### **12.1 Energy Metabolism**

Energy intake (EI), total energy expenditure (TEE), resting metabolic rate (RMR), and core body temperature will be measured at baseline and during the study at the intervals indicated below.

#### **12.1.1 Total Energy Expenditure and Energy Intake**

TEE will be measured by the doubly-labeled water (DLW) method as described in Section 13 below. In the CR group, DLW studies will be performed during 2-week periods at baseline (two periods), and at Months 6, 12, 18, and 24. In the control group, TEE will be measured during 2-week periods at baseline (two periods), and at Months 12 and 24 only. All DLW measurements will be performed by USDA/ARS Children's Nutrition Research Center Stable Isotope Laboratory, Houston, TX, under the direction of Dr. William Wong.

Total energy expenditure (TEE) will be calculated as,

$$TEE = VCO_2 (1.231 + 3.815 / RQ),$$

where  $VCO_2$  (L/d) is rate of carbon dioxide production and RQ is the respiratory quotient that is provisionally estimated as 0.86 but will be calculated for individual participant using indirect calorimetry and/or reported dietary intake.

#### **12.1.2 Resting Metabolic Rate**

RMR and respiratory quotient (RQ) will be measured via indirect calorimetry using Vista-MX metabolic measurement system (Vacumed, Ventura, CA). Baseline RMR will be determined from the average of two trials conducted on consecutive days and compared with follow-up measurements at Months 6, 12, 18, and 24 visits in the CR group and at Months 12 and 24 in the control group. Changes over time in RMR will be analyzed in relation to body composition change.

#### **12.1.3 Core Temperature**

Core temperature will be measured and recorded every minute during the 24 hour period spent in the metabolic ward at baseline, Months 6, 12, and 24 visits. Volunteers will swallow a 8.7 x 23-mm Jonah™ radio-capsule (Mini Mitter Co., Inc., Bend, OR) and will be fitted with The VitalSense Monitor: that records core temp. Average 24 hour, day (8:00 a.m. to 10:30 p.m.) and night (2:00 a.m. to 5:00 a.m.) temperatures will be computed.

### **12.2 Cardiovascular Risk Factors**

Resting blood pressure (BP), serum lipids and lipoproteins, inflammation markers, advanced glycosylation end-products, and transforming growth factor-beta will be measured at baseline, Months 12 and 24

unless otherwise is indicated. The assays listed in the following sections will be performed in the Clinical laboratories of the University of Vermont.

### **12.2.1 Blood Pressure**

Resting BP will be measured at baseline and at Months 1, 3, 6, 9, 12, 18, and 24. Resting BP will be assessed in the morning. Two readings will be taken on each of two separate days during the clinic visit. The exception is during the minimal clinic visits at Months 1 and 9. For those visits, two readings will be taken on the single clinic day only. Pulse pressure will be calculated and used as an index of arterial stiffness. Since pulse pressure is influenced by stroke volume and heart rate, it will be normalized for these 2 factors in analyses.

**Unless otherwise stated, the assays listed in the following sections will be performed in the Clinical laboratories of the University of Vermont. See Section 23 below for timing of the assays.**

### **12.2.2 Serum Lipids and Lipoproteins**

Total cholesterol, HDL-cholesterol, LDL-cholesterol (calculated), and triglycerides concentrations will be determined in samples obtained after an overnight fast by the Laboratory for Clinical Biochemistry (LCBR) at the University of Vermont.

### **12.2.3 Markers of Inflammation**

Commercial ELISA kits will be used to measure MCP-1, IL-6 and ICAM-1 (Quantakine High Sensitive – R&D Systems, Minneapolis, MN) and nephelometry (Dade Behring BNII) - to measure plasma C-reactive protein. IL-1, IL-8 and TNF $\alpha$  will be measured using Multiplex Biomarker Immunoassay (Linco Research, St. Charles, MO).

### **12.2.4 Transforming growth factor- $\beta$ 1**

Plasma TGF- $\beta$ 1 will be measured using an ELISA kit (R&D Systems, Minneapolis, MN).

## **12.3 Glucose Tolerance and Insulin**

Serum glucose, insulin and C-peptide levels in the fasting state and during oral 75 G glucose tolerance test will be measured at baseline, Months 12, and 24. Multiplex Biomarker Immunoassay (Linco Research, St. Charles, MO) will be used to determine insulin levels and Automate Chemuliminoetric Immunoassay for measuring C-peptide.

## **12.4 Immune Function**

Immune status will be assessed from measurements of delayed type hypersensitivity response (DTH), white blood cell (WBC) differential and antibody response to 3 different vaccines. The proposed measurements are justified as standard indicators of CR and were chosen based on the following: (a) they are sensitive to CR as shown in animal models and in CALERIE Phase 1; (b) age-associated changes in these measurements have been reported; and (c) the ability of immune cells to produce antibodies is important in determining resistance and/or pathogenesis of infectious and inflammatory diseases.

### **12.4.1 Delayed-Type Hypersensitivity**

DTH will be measured using the Mantoux method in which 0.1ml of a normal saline control and 2 antigens, candida and trichophyton, and 0.02 ml (0.20 LF units per dose) for tetanus toxoid (adsorbed) will be administered intradermally on the volar surface of the arm at baseline, 12 and 24 months. The diameter of indurations will be measured 24 and 48 hours following administration of the test.

### **12.4.2 White Blood Cell Differential**

Whole blood will be examined with hematology analyzer for red blood cell and WBC counts and a blood smear will be used to estimate WBC differentials. Assays will be performed at baseline, Months 6, 12, and 24 visits by Esoterix laboratory.

### **12.4.3 Antibody Response to Vaccines**

Three different vaccines representing different arms of the immune response will be used. Hepatitis A represents a primary T cell dependent antibody response, tetanus/diphtheria represents a secondary T cell dependent response and pneumococcal vaccine represents a T cell independent response.

Standard doses of pneumococcal vaccine, polyvalent (Pneumovax 23, Merck), hepatitis A (HAVRIX®, GlaxoSmithKline, or VAQTA®, Merck and Co. Inc), and tetanus/diphtheria will be administered at month 17 to avoid the confounding effect of the repeated administration of the vaccines which might interfere with the ability to determine CR-induced changes. A booster dose of hepatitis A vaccine will be administered at month 23.

Blood for measurement of antibodies will be collected at baseline (i.e., prior to start of caloric restriction), Months 17 (pre-vaccination antibody titer for all 3 vaccines), Month 18 (after vaccination antibody titer for all 3 vaccines), Month 23 (Hepatitis A antibody titer before Hepatitis A booster), and Months 24 (Hepatitis A antibody titer after Hepatitis A booster). At Month 24, antibody titers for pneumococcal and tetanus/diphtheria vaccines will also be measured to determine if caloric restriction has an effect on maintaining antibody titers for a longer period of time, i.e., prevents drop in titer levels.

Previous documented vaccination with hepatitis A or pneumococcal vaccine will serve as an exclusionary criterion for this component of the study. That is, if a participant has already been vaccinated with the hepatitis A vaccine, s/he should not be administered hepatitis A vaccine; similarly, if a participant has received pneumococcal vaccine, s/he would not be administered that vaccine. Moreover, approximately 30% of the population may have been exposed to Hepatitis A naturally at baseline. Rather than excluding these individuals, any volunteer whose baseline antibody level is above 10 will be excluded from the analysis. Fold change in antibody titer will be determined by dividing after vaccination antibody levels to that of pre-vaccination antibody levels. The baseline measurements are performed to determine if calorie restriction will affect total (non-specific) immunoglobulin levels.

## **12.5 Endocrine Response**

DHEA(S), cortisol, TSH, triiodothyronine (T3), growth hormone, leptin, angiotensin II, norepinephrine (NE), total adiponectin, IGF-1, IGFβP-1, IGFβP-3, and PDGF-AB will be measured at baseline, Months 12 and 24 as indicated below. Because of diurnal variation, all blood samples will be obtained in the morning (at about 8:00 a.m.).

### **12.5.1 Norepinephrine**

NE concentrations will be measured in arterialized venous blood (hot box or similar method) by GC/HPLC by LCBR at the University of Vermont. Blood samples will be obtained after the subject has been recumbent for 15-minutes. Two NE draws, 5-minutes apart, will be performed at baseline and at Months 12 and 24.

### **12.5.2 Dehydroepiandrosterone (DHEA) and Cortisol**

Chemiluminescent immunoassay kits will be used to measure DHEA (Immulite 1000, Diagnostics Products Corp.) and cortisol (ADVIA Centaur, Bayer Health Care) by LCBR at the University of Vermont.

### **12.5.3 Sex Hormones**

In men, the effect of CR on reproductive function will be assessed at baseline, 12 and 24 months by measuring fasting LH, FSH, total testosterone, sex-hormone binding globulin (SHBG) and free testosterone. Normal values will be based upon laboratory standards. Automated analysis procedures are set up for all these assays, which will be performed using the most precise commercially available kits where available (ADVIA Centaur, Bayer Health Care and Immulite 1000, Diagnostics Products Corp.). Additionally, the following hormonal determinations will be made in the fasting state for any woman who develops amenorrhea over a period of two months: LH, FSH, and estradiol levels.

### **12.5.4 Thyroid Hormones**

TSH and T3 will be measured using chemiluminescent immunoassays (ADVIA Centaur, Bayer Health Care) by LCBR at the University of Vermont.

### **12.5.5 Adipokines**

Immunoassay kits will be used to measure Leptin and total adiponectin will be measured using Multiplex Biomarker Immunoassay (Linco Research, St. Charles, MO).

### **12.5.6 Angiotensin II**

Angiotensin II will be measured using an immunoassay kit (Linco Research).

### **12.5.7 Growth Hormone and Growth Factors**

GH will be measured by the Mayo Clinic Laboratory using two-site immunoenzymatic assay. IGF-1, IGF $\beta$ P-1, IGF $\beta$ P-3, and PDGF-AB will be measured using commercially available ELISA kits from R&D Systems (Minneapolis, MN).

## **12.6 Quality of Life, Psychological and Cognitive Functioning**

Three types of measures, quality of life, psychological, and cognitive functioning, will be assessed at four measurement points: baseline, Months 6, 12, and 24. Administration of these measures at other assessment times (e.g., 3 months and 18 months) was considered, but rejected based upon the findings from studies during CALERIE phase 1 and based upon concerns about time burdens upon the participants. This assessment battery should be scheduled at the same time of day and in the same sequence, relative to other assessment methods, to control for temporal factors such as time of day, hunger, effects of intrusive tests, etc. All of the measures, except some of the cognitive functioning measures, require English-speaking participants. A review of the literature found that there were no other good alternative measures of the relevant constructs that had been validated for other languages. Therefore, for non-English-speaking participants, these measures should not be administered.

The measures of quality of life and psychological measures involve the use of self-report (self-administered) questionnaires. Cognitive bias measures are administered to individual participants by a trained evaluator. Measures of cognitive impairment are computer-administered. The measures were selected based upon the following features: 1) validation studies, 2) reliability studies, 3) sensitivity to change, 4) coverage of a broad range of constructs that should be measured in the phase 2 study, and 5) relatively low subject burden. The total time for administration of these measures is estimated to be 2.5 to 3 hours for each assessment period.

### **Quality of Life Measures:**

Changes in quality of life will be assessed using the Rand SF-36, the Profile of Mood States, and the Perceived Stress Scale. These measures were selected to measure both positive and negative changes in quality of life indicators.

#### **12.6.1 Rand SF-36**

The Rand Short Form with 36 items is frequently used in research studies as a measure of quality of life. The SF-36 was used at all three CALERIE sites during phase 1. The findings from these studies indicated that caloric restriction caused neither improved quality of life nor deteriorated quality of life. The SF-36 was selected for use in CALERIE phase 2 to assess worsening quality of life indicators associated with long-term caloric restriction. The SF-36 has eight subscales: physical functioning, role limitations due to physical problems, role limitations due to emotional problems, vitality, bodily pain, social functioning, mental health, and general health perceptions.

#### **12.6.2 Profile of Mood States**

Profile of Mood States (POMS) is an adjective checklist of 65 items rated on a five-point scale that ranges from (1) "not at all" to (5) "extremely". Six subscales are derived from the POMS: anxiety-tension, depression-dejection, anger-hostility, fatigue-inertia, vigor-activity and confusion-bewilderment. The subscales are combined to calculate a total mood disturbance score. The POMS is a reliable and valid self-report inventory that is especially sensitive to changes in energy, fatigue, and mood [197]. The POMS was selected for use in CALERIE phase 2 because it is sensitive to both positive and negative changes in quality of life indicators.

#### **12.6.3 Perceived Stress Scale**

The Perceived Stress Scale (PSS), a four-item questionnaire, will be administered to assess perceived changes in stress during a period of prolonged caloric restriction. This brief inventory should require a few minutes to complete. The PSS has been found to be sensitive to changes in stress in other studies [198].

## **Psychological Assessments:**

The following psychological assessments were selected to measure sleep habits, sexual functioning, eating habits, mood, eating disorder symptoms, and body size concerns: Pittsburgh Sleep Quality Index (PSQI), Derogatis Interview for Sexual Functioning-Self Report (DISF-SR), Food Craving Questionnaire (FCQ), Food Craving Inventory (FCI), Three Factor Eating Questionnaire (TFEQ)—also known as the Eating Inventory, Weight Self-Efficacy, Multiaxial Assessment of Eating Disorder Symptoms (MAEDS), Body Shape Questionnaire (BSQ).

### ***12.6.4 Pittsburgh Sleep Quality Index***

PSQI is a reliable and valid measure of sleep quality and disturbance over a one-month interval. Nineteen items yield seven subscale scores (subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medications, and daytime dysfunction) and one total score.

Changes in the amount and quality of sleep may occur with caloric restriction. Amount and quality of sleep are variables strongly affected by age [199], and are thought to be determined in large part by the balance of different factors, in particular growth-hormone releasing hormone (sleep promoting) and corticotrophin-releasing hormone (sleep impairing) [200]. Since CR is hypothesized to delay biological aging and have numerous effects on different hormonal systems, it is possible that CR influences sleep and that this change underlies some other key variables in CALERIE 2 such as quality of life variables and hormonal variables.

### ***12.6.5 Derogatis Interview for Sexual Functioning-Self Report***

DISF-SR is a reliable and valid measure of sexual functioning for men and women; there are different parallel forms for men and women. The DISF-SR has five subscales (sexual cognition and fantasy, sexual arousal, sexual behavior and experience, orgasm, sexual drive and relationship) and a total score is derived. The DISF-SR was selected to measure beneficial and adverse effects of long-term caloric restriction.

### ***12.6.6 Food Craving Questionnaire***

FCQ was developed to measure trait (generalized) and state (at the moment) craving for food. The reliability and validity of the state and trait versions of the FCQ have been established. The FCQ was selected for use in CALERIE phase 2 as a measure of general food craving. The Trait FCQ measures a general susceptibility to craving foods. The 39-item Trait FCQ will be administered only at baseline. The 15-item State FCQ measures the strength of food cravings at the moment of administration and will be administered at baseline and at each six-month assessment period during the study to assess changes in food craving as a function of long-term caloric restriction.

### ***12.6.7 Food Craving Inventory***

FCI is a 28-item self-administered measure designed to assess the subjective experience of specific food craving across 28 different foods. The measure consists of 4 scales: high fats, sweets, carbohydrates/starches, and fast food fats. Also, a total score is derived. The FCI is scaled in a frequency format assessing the frequency of an individual experiencing a craving for a particular food. All items are scored as: Never = 1, Rarely = 2, Sometimes = 3, Often = 4, Always = 5. The four scales of the FCI and the FCI total score have been found to have good internal consistency and test-retest reliability. The strong association of the FCI with other measures of food cravings supports the concurrent validity of this measure, and discriminant validity is supported by its lack of association with measures that assess other aspects of eating behavior.

### ***12.6.8 Eating Inventory or Three Factor Eating Questionnaire***

TFEQ (more recently called the Eating Inventory) assesses dietary restraint, disinhibition, and perceived hunger and has well-established reliability and validity. The questionnaire produces three scores for the three subscales: Dietary Restraint, Disinhibition, and Perceived Hunger. A total score is not computed. For all three subscales, the higher number indicates the greater levels of restraint, disinhibition, and/or perceived hunger. The TFEQ has been validated in a variety of studies and has been found to be very sensitive to change in a variety of outcome studies, including CALERIE Phase 1 at PBRC.

### **12.6.9 Weight Self Efficacy**

Self efficacy theory is an integrative cognitive-social learning framework which posits that an individual has a measurable level of confidence as to their perceived ability to perform a certain task such as weight loss [201]. Weight self-efficacy is measured by the 20-item Weight Efficacy Lifestyle (WEL) Questionnaire which includes 5 subscales (food availability, negative emotions, social pressure, physical discomfort and positive activities) and a total WEL score [202]. Both the total WEL score, and the “food availability” subscale in particular, predict weight loss in Phase 1 CALERIE at Tufts. This finding is consistent with previous work suggesting that higher levels of weight self-efficacy are associated with greater weight loss [203,204].

### **12.6.10 Multiaxial Assessment of Eating Disorder Symptoms**

The MAEDS is a self-report inventory that measures the six symptom domains related to eating disorders: binge eating, restrictive eating, purgative behavior, fear of fatness, avoidance of forbidden foods, and depression. The MAEDS has been validated in a variety of studies and has been found to be very sensitive to change in a variety of outcome studies, including CALERIE Phase 1 at PBRC.

### **12.6.11 Body Shape Questionnaire**

The BSQ is a self-report instrument that assesses concerns about body size and shape, specifically the experience of “feeling fat”. In this study, an 8-item shortened form of the BSQ will be used. The shortened form measures global appearance satisfaction, as well as the experience of “feeling fat.” The BSQ has been validated in a variety of studies and has been found to be very sensitive to change in a variety of outcome studies, including CALERIE Phase 1 at PBRC.

### **Tests for Cognitive Biases**

Attention bias will be measured using the Stroop color-naming task. The number of Stimuli will be limited to 40 words per test, since this test is prone to habituation effects.

Selective interpretation (judgment) bias will be measured using ambiguous stimuli (polysemous words and homophones, such as “waste” vs. “waist”) that must be interpreted in a self-referent context (participants will be asked to write a sentence that includes themselves and use of the word, which will be presented orally).

Memory bias will be tested using free recall of the word stimuli used in the selective interpretation task. Alternatively, another set of stimuli can be utilized, where a self-referent encoding paradigm will be used followed by an implicit memory task (word stem completion) and an explicit task (free recall).

### **12.6.12 Tests of Cognitive Impairment**

The Cambridge Neuropsychological Test Automated Battery (CANTAB) is a computerized cognitive assessment battery used to measure reaction time, processing speed, attention, concentration, executive function, short-term memory, and delayed memory. The CANTAB has been extensively evaluated and validated in healthy adults. The battery of tests included in the CANTAB is normed for a wide range of ages (4-90 years) and is suitable for longitudinal studies and repeated testing. The tests are predominantly non-verbal and non-cultural, making them highly suitable for use in diverse populations. Since the tests are computerized, administration and scoring are simple and reliable. The CANTAB uses a touch screen to make it easier for the subject to follow instructions, and the tasks have a game-like quality and objective feedback, which encourage motivation and high subject compliance. The program scores the tests and exports the data into statistical analysis packages or spreadsheet packages. For the present study, we have chosen the following CANTAB modules to assess for attention, vigilance, executive function, and memory.

The verbal recognition memory module will be used to measure immediate free recall and delayed recognition memory.

The reaction time module has three purposes, (a) it trains the subject in holding down the press-pad and touching the screen, (b) it provides a screen for the ability to acquire and perform this motor skill and (c) it acts as a simple single and multiple choice reaction time tasks. Its five stages require increasingly complex chains of responses. In each case, the subject must react as soon as a yellow dot appears. In some stages it may appear in one of five locations, and the subject sometimes responds by using the press-pad, sometimes by touching the screen, and sometimes both.

The rapid visual information processing module is a test of sustained attention with a small working memory component. A white box appears in the center of the computer screen, inside which digits, from 2 to 9, appear in a pseudo-random order, at the rate of 100 digits per minute. Subjects are requested to detect consecutive odd or even sequences of digits (e.g. 2-4-6, 3-5-7, 4-6-8, 5-7-9, etc.) and to register responses using the press-pad. Initially, the computer prompts the subject when sequences appear and gives feedback when the pad is pressed. As the practice part of the test progresses, these cues are gradually phased out, and in the assessment part, no cues or feedback are given.

The spatial working memory module is a test of spatial working memory and strategy performance. The aim of the test is that the subject should find a blue 'token' in each of the boxes displayed and use them to fill up an empty column on the right hand side of the screen, while not returning to boxes where a blue token has previously been found. The color and position of the boxes used are changed from trial to trial to discourage the use of stereotyped search strategies.

The ID/ED shift module is a test of the subject's ability to attend to the specific attributes of compound stimuli, and to shift that attention when required. Two artificial dimensions are used, color-filled shapes and white lines. Two stimuli (one correct, one incorrect) are displayed, initially each of only one dimension, then each of both dimensions (first adjacent, then overlapping as illustrated). Feedback teaches the subject which stimulus is correct, and after six correct responses, the stimuli and/or rules are changed. These shifts are initially intra-dimensional (e.g., color filled shapes remain the only relevant dimension), then later extra-dimensional (white lines become the only relevant dimension).

The delayed matching to sample module is a memory test that presents the subject with a complex visual pattern (the sample) and then, after a brief delay, four patterns between which she or he must choose. Each pattern is made up of four sub-elements, each of a different color. One of the choice patterns is identical to the sample, one is a novel distracter pattern, one has the shape of the sample and the colors of the distracter, and the fourth has the colors of the sample and the shape of the distracter. To discourage strategies based on encoding single quadrants, all four choice patterns have a quadrant in common with the sample.

## **12.7 Physical Activity Measures**

Physical activity measures will include peak  $VO_2$ , Stanford 7-day Physical Activity Recall (PAR), and muscular strength and endurance.

### ***12.7.1 Cardiorespiratory Fitness using Peak $VO_2$ via Expired Gas Analysis***

Maximal oxygen uptake, commonly referred to as peak  $VO_2$ , is the most widely recognized measure of cardiorespiratory fitness. It is physiologically reflective of the capacity of the cardiovascular and pulmonary systems to transport oxygen to working skeletal muscle, as well as of the capacity of the working muscles to utilize oxygen during exercise. Peak  $VO_2$  is therefore a reflection of both central and peripheral cardiovascular capacity. It has been shown to be an independent risk marker for several diseases, most notably cardiovascular disease [205-208]. In Phase 2, peak  $VO_2$  will be used as an index of physical fitness at baseline and to determine changes in physical fitness over the 24-month intervention, at baseline, 12 and 24 months. Peak  $VO_2$  is measured as an absolute value (L/min) and is commonly expressed relative to body mass (ml/kg/min). Peak  $VO_2$  will be expressed as both an absolute and relative value for CALERIE and will be measured using common testing protocols and testing procedures standardized as described in Section 15.4.5 below.

### ***12.7.2 Energy Expenditure using the Stanford 7-day Physical Activity Recall (PAR)***

There exist many standardized survey methods for the measurement of physical activity. For CALERIE, the investigative sites are interested in using survey instrument for the measurement of physical activity for several reasons. First, a valid physical activity survey tool can be used as an independent, perhaps validating assessment of energy expenditure to be compared with the values obtained using other, simultaneously performed measures (e.g., doubly labeled water). Second, a valid physical activity survey and analysis has the potential of providing energy expenditure information faster than any other method, and therefore valuable for the calculation of caloric restriction prescription in CALERIE subjects. Third, such a survey tool can provide information about physical activity energy expenditure not available by global measures such as DLW.



Specifically, the PAR will provide information about weekly physical activity in each of five intensity levels, the time spent in physical activity in each of these levels and an estimate of total weekly EE via physical activity and changes in physical activity in response to the intervention.

Administration coincide with the doubly labeled water measures and will occur at baseline and at Months 6, 12, 18, and 24 in the CR group, and at baseline and Months 12 and 24 in the control group. We will administer the PAR twice during the 14-day DLW periods. In order to ensure uniformity of collection and scoring of the 7-day PAR in Phase 2, study personnel will be trained in a common training session by a qualified 7-day PAR trainer prior to initiation of Phase 2 studies as described in Section 15.4.6 below.

### **12.7.3 Muscular Strength and Endurance**

After 10-min of low- to moderate-effort warm-up exercise on a stationary bicycle, isometric and concentric isokinetic strength of the knee flexors and knee extensors (dominant leg) will be assessed on a Biodex System 3 dynamometer (Biodex Medical Systems, Shirley, New York). The test sequence will always be performed as follows: (a) isometric knee extension at 45° of knee flexion, (b) isometric knee flexion at 60° of knee flexion (c) 60°/second isokinetic knee extension and flexion, and (d) 180°/second isokinetic knee extension and knee flexion. Each test will only be performed once since repeated isokinetic measures have been shown to be highly reproducible and there appears to be no learning effect during repeated measures.

Peak torque from the isometric tests and power, work and peak torque data from the isokinetic tests will be interpreted as muscular strength data. The results will be expressed in absolute terms and relative to body weight. Muscle fatigue index will be calculated from the 180°/second isokinetic test as the percent decrease in peak torque during the 30 repetition set as follows:

Percent decrease =  $100 - [(\text{mean of the last 5 repetitions}) / \text{mean of the highest consecutive 5 repetitions}] \times 100$ .

Grip strength that is predictive of whole body muscle mass and future risk of frailty will be assessed using a Jamar dynamometer (Asimow Engineering Company, Los Angeles, CA). The right and left hands will be tested alternately with each being tested three times. The peak force from each trial will be recorded and means for the right hand, left hand, and all tests combined will be calculated for comparison with body weight independent and body weight dependent normative data.

This will be performed at baseline and at Months 12 and 24.

## **12.8 Body Weight and Height**

### **12.8.1 Body Weight**

Clinic weights will be utilized as outcome measurement. Daily home weights will be used in the short-term intake / balance estimation of %CR, as well as for enhancing adherence to the CR and *ad libitum* interventions tool.

Clinic Weights: Clinic weight will be measured at each clinic visit (including the off-schedule evaluations during Months 17, 18, 23 and 24) using a standardized electronic standing scale. Supplemental measurements will be taken during the DLW periods. All measurements will be performed after an overnight fast at approximately the same time and under similar conditions in pre-weighed gowns and the weight of the gown will be subtracted to obtain a true naked weight.

Daily Weights: For details, please refer to Section 7 above

### **12.8.2 Height**

Height will be measured at baseline only using a wall-mounted stadiometer in conjunction with a right-angled head board. Shoes and head dressing will be removed and participant's head will be positioned in the Frankfort plane (the orbitale (lower edge of the eye socket) is in the same horizontal plane as the tragion (the top notch of the ear)). Height measurements will be repeated to obtain 2 values that are identical or close to one another (within 0.1cm) and the average of the two measurements will be used.

## **12.9 Body Composition**

### **12.9.1 Waist Circumference**

Waist circumference (WC) will be measured at the minimal and umbilical points. WC is the best surrogate marker for visceral adiposity, and a rapid and inexpensive measure with clinical applicability. It will be measured at all clinic visits, i.e., at baseline and at Months 1, 3, 6, 9, 12, 18 and 24.

### **12.9.2 Body Composition and Bone Density by DXA**

Body fat, lean, bone; bone mineral density (BMD) bone mineral content (BMC), and bone areas will be measured by DXA. This method is extensively employed for the assessment of total and regional body fat and bone [209,210]. Measurements using DXA have replaced underwater weighing as the reference standard for body composition studies, reflecting the ease of use, the low radiation dose, and the fact that it provides a 3-compartment model of body density, i.e., lean body mass, bone mass, and fat mass, compared to the 2-compartment model of underwater weighing.

Whole body DXA measurements will be performed to coincide with the doubly labeled water measures and will occur at baseline and at Months 6, 12, 18, and 24 in the CR group, and at baseline and Months 12 and 24 in the control group. At baseline, two consecutive DLWs will be performed. The first whole body DXA will be performed at the beginning of the first DLW period, while the second will be performed at the end of the second DLW period. At Month 6, one whole body scan will be performed at the beginning of the 14-day DLW period while a second will occur at the end. Otherwise, one whole body DXA will be performed at each of the other follow-up time points. All scans will be performed using the Hologic whole body scanners.

BMD and BMC at the hip, spine and forearm will be measured with DXA. These scans will be obtained at baseline, 6, 12 and 24 months for the hip and spine scans, and at baseline, 12 and 24 months for the forearm. Hip and spine scans will also provide an assessment of participant safety (See Section 16.11).

Measurements will be made using a standardized protocol for subject positioning, scan mode and analysis. All DXA measures will be transmitted and assessed by a centralized reading center to maintain consistency and reduce variability. The measurements will take approximately twenty minutes to perform. The radiation doses for the scans are as follows: whole body =1 mREM; hip = 20 mREM; spine = 20 mREM; and forearm = 5 mREM.

Because the precision of the body composition measurements is of particular importance, the primary DXA operator at each CALERIE clinic will perform a precision assessment for whole body scans obtained by DXA by repeating the whole body scan on 30 participants. After the first whole body scan is acquired, these participants will be asked to stand up, and then be re-positioned on the scanner table for a second whole body scan. All scans will be analyzed by the same person at the UCSF Reading Center. The results of the repeated scans will be used to calculate appropriate measures of precision, including the coefficient of variation (with 90% confidence interval), for the outcomes of interest, including fat mass, lean soft tissue mass, %fat, and BMC for the whole body and for subregions.

## **12.10 Bone Turnover**

Caloric restriction and weight loss can lead to increased bone resorption and therefore a potentially increased risk for osteoporosis, although the mechanisms associated with this response remain unclear [211]. Measurement of serum markers of bone turnover may detect more immediate alterations in bone turnover than direct measures of bone density by DXA.

Bone turnover measurements will be performed to coincide with the DXA measures but will only be performed at baseline, 6, 12 and 24 months. As marker of bone resorption, we will measure serum CTX (serum c-terminal telopeptide of type 1 collagen). As marker for bone formation, we will measure serum PINP. Both markers will be measured using ELISA. As we are only measuring one marker of bone resorption and bone formation in this study, we need to use those with the highest clinical sensitivity (i.e. highest signal to noise ratio) as indicated by the serum CTX and serum PINP.

## **12.11 Nutrient Intake**

Six-day diet diaries will be collected using paper logs, and coincide with the doubly labeled water measures i.e., at baseline and at Months 6, 12, 18, and 24 in the CR group and Months 12 and 24 in

the control group. They will be analyzed for calories, macronutrient composition (percent of energy), fiber (total, fermentable and non-fermentable) and variety (where variety within a day and variety over the measurement period will be calculated as number of unique food items in different food groups). Micronutrients (vitamins and minerals) will also be analyzed for archive information on the subjects and specific analyses as described below. Food records will be entered into the University of Minnesota Nutrient Data System (NDS) analysis program by the Central Coding Site at the University of Cincinnati and reported to the Coordinating Center.

Macronutrients, fat, fiber and variety will be examined to determine whether these variables predict adherence to CR. Calcium intake during the intervention will be correlated with BMD changes, to examine whether the anticipated loss of BMD with CR is related to Ca intake after body weight change is taken into account.

## **12.12 Advanced Clinical Endpoints**

CALERIE will conduct mechanistic studies on endpoints which go beyond typical clinical trial endpoints (i.e., those which are measured routinely in clinical trial settings like clinical chemistry, body composition, or blood pressure).

Two basic approaches to the selection of these outcomes will be applied. First, specific candidate hypotheses, based on pre-clinical animal and cell culture data regarding the effects of CR, and on other data on mechanisms affecting aging, will be tested. Six major candidate pathways/domains will be investigated as delineated below. This does not preclude the addition of other endpoints pending on the discovery of novel mechanisms related to CR or aging and the development of additional hypotheses. (These domains are obviously interrelated and overlapping.)

- Oxidative stress
- Growth factors
- Mitochondrial function
- Energy and substrate metabolism
- Cell proliferation, differentiation, and turnover
- Cell signaling and nutrient sensing.

Secondly, tools will be applied which capture a global, integrated picture of the effects of CR. Specifically, these measures will allow for a global view of the effects of CR on multiple systems simultaneously. The approaches may include screening in domains such as:

- Proteome
- Metabolome
- Transcriptome
- Cytokine and adipokine surveys.

### **12.12.1 Process for Selecting Outcomes and Tests**

Optimal selection of the most appropriate outcomes and tests requires expertise beyond that found in CALERIE. CALERIE has obtained expert advice in some of the relevant domains, but additional advice will be needed. However, our protocol provides a method (described below) for incorporating this advice into the final selection of outcome measures without requiring that this advice be obtained before startup of Phase 2. This approach will also allow CALERIE to obtain expert advice in regard to new opportunities for outcome measures that arise over the course of CALERIE, based on findings from other studies.

In addition, we recognize that not all of these outcome determinations will require measurement of the outcome in all CALERIE participants. Measurements of many of these advanced clinical endpoints require one or more of the following: specialized training for data collection, specialized equipment, extraordinary processing during sample collection, or complicated techniques for measurements. Also, because of their exploratory nature, cost, or subject burden, some desirable measures may not need to be done on all participants in CALERIE. Some of these endpoints will obviously rely upon local technical expertise or equipment (e.g. magnetic resonance spectroscopy [MRS] measurement of ATP turnover rates) or simply are not feasible for reasons of cost [MRS] or subject burden across all sites and all subjects. A thoughtful, ordered selection process for study of some outcomes in subsets of the

CALERIE population can allow study of a wider variety of outcomes, given constraints in cost and distribution of equipment, etc. The process described below will allow well-informed decisions regarding this issue for all tests.

Rather than specify before the start of enrollment all the tests that will be done, we specify: a) the types of specimens that will be collected and stored, and b) the process we will use in selecting specific tests using these specimens, and for selecting tests that require significant participant involvement (e.g., magnetic resonance spectroscopy studies) besides specimen provision, or which require assays on freshly-collected tissue.

### 12.12.2 *Types of Specimens to be Collected and Stored*

The types of specimens to be collected and stored are outlined in Table 12.1.

**Table 12.1: Summary of the Types of Specimens to be Collected**

Specimen	Volume, and time collected	Storage Conditions	Study Time Point					
			Baseline	3M	6M	12M	18M	24M
<b>Blood</b>	70 ml (fasting, a. m)		X	X	X	X	X	X
<b>Blood components (from above specimen):</b>								
<b>White cells</b>	Buffy coat	Cryopreserved (liquid N <sub>2</sub> ), 10 aliquots	X	X	X	X	X	X
<b>Plasma</b>	5 ml	Liquid N <sub>2</sub> , 10 aliquots	X	X	X	X	X	X
<b>Serum</b>	30 ml	Liquid N <sub>2</sub> 30 aliquots	X	X	X	X	X	X
<b>Urine</b>	30 ml (24 hour)	- 80 C 6 aliquots)	X			X		X
<b>Muscle (Quadri-ceps)</b>	200-750 mg	Liquid N <sub>2</sub> Divided into 2-5 samples	X			X		X
<b>Fat (Abdominal)</b>	1 g	Liquid N <sub>2</sub> Divided into 5 samples	X			X		X

These specimens will be stored in a central repository at the University of Vermont which will maintain appropriate quality control and data management.

### 12.12.3 *Process for Proposing Studies*

CALERIE has established an “Advanced Clinical Endpoints” (ACE) Committee, which (with the assistance of independent reviewers) will consider each proposed test. For each proposed test, a detailed description will be provided to the ACE Committee, which will include, at a minimum, the following information:

- Hypotheses to be tested (if hypothesis-based study) or descriptive information to be gathered (e.g., for array studies) and rationale for the importance of the hypothesis or information in regard to CALERIE goals.
- Methods, including pilot data (if needed) on feasibility and reliability of the test in humans or human specimens
- Materials and equipment needed
- Requirements (if any) for performance in “real-time”, e.g., imaging studies or assays on fresh tissue

- Participant time and travel burden, if any
- Type and amount of stored samples needed
- Sample shipping and handling requirements
- Selection of study time points at which outcome will be measured
- Number of participants from which samples are needed and rationale for this selection
- Number of sites from which samples are to be used and rationale for this number
- Analytic plan
- Personnel, including qualifications and time commitment
- Time line
- Budget (by CALERIE budget year).

The ACE Committee will review proposed tests and determine whether they merit further consideration. For tests meriting further consideration, the ACE Committee will suggest at least two independent reviewers (non-CALERIE investigators) and Committee's Chair will present their names to the Steering Committee for approval. After the selection is approved, Chair of the ACE Committee will forward the proposed protocol to reviewers. Based on this review and CALERIE's scientific priorities, the ACE Committee will make a recommendation to the CALERIE Steering Committee on each proposed test. Tests approved by the Steering Committee will be presented to the DSMB for final review and approval. CALERIE will hold funds in reserve in each remaining CALERIE budget period for approved tests. NIA will release these funds for DSMB-approved tests as needed.

The timetable for decisions on specific tests will differ depending on whether the test requires "real time" performance (e.g., imaging studies or assays on fresh tissue). Such proposed tests will be reviewed and submitted to the DSMB before the start of the intervention phase in order to allow collection of baseline data. (In some cases, such studies may be proposed later, if they do not need to be performed on all CALERIE participants.) Informed consent for any tests not described in the initial consent form will be requested.

For tests that do not require real time performance and can be done exclusively or stored specimens, decisions on selection of tests may in many cases be deferred until relatively late in the intervention period. This will allow CALERIE to accommodate proposals for tests based on new research findings from other studies over the course of CALERIE. This will allow selection of the combination of tests that optimizes the scientific information that CALERIE can obtain.

In addition to storage of samples for use by CALERIE investigators, CALERIE will reserve a proportion of its stored samples for potential use in studies by other investigators. (It is expected that funding for performing such studies, including shipping of samples, will come from sources other than CALERIE.) Requests for samples by other investigators must describe the proposed test, providing the same information (as appropriate) as outlined above for tests proposed for CALERIE. The procedures for review and approval of these requests will be the same as described above for internal proposals for tests using CALERIE funds.

### **13. DOUBLY LABELED WATER (DLW) METHODS**

Energy is required for muscular activity, growth, reproduction and synthesis of metabolites such as proteins, fatty acids, nucleic acids and steroids, which are essential to maintain basal metabolic functions as well as optimal growth and development. Numerous methods such as the food record, <sup>13</sup>C-bicarbonate infusion, and indirect calorimetry have been used to estimate energy expenditure (EE) in humans. However, dietary recalls seldom reflect the true caloric content of ethnic foods, and this approach does not work well with children. It is also well documented that overweight individuals often underreport their food intake [212-219]. The <sup>13</sup>C-bicarbonate infusion method is invasive and of short duration (< 24 h). Activity of the subject is restricted during the infusion. Therefore, EE measured by this method is not representative of the true daily EE of the free-living subject. Whole room calorimetry is considered the gold standard for measuring EE in humans. Although the subject is free to move around in the calorimetric chamber, spontaneous physical activity is greatly reduced. Furthermore, the measurement is carried out under strictly controlled, artificial environmental conditions and often is of short duration (<24 h).

The doubly labeled water (DLW) method yields an average EE for a period of 5-14 d. The procedure is noninvasive, nonrestrictive and reflective of actual EE under free-living conditions [220-227]. Briefly, the DLW method involves enrichment of body water with the stable (natural and nonradioactive) isotopes, deuterium ( $^2\text{H}$ ) and oxygen-18 ( $^{18}\text{O}$ ), and then determination of their monoexponential wash-out kinetics in a physiological fluid such as urine. The DLW method is based on the principle that the disappearance rate of  $^2\text{H}$  reflects water turnover rate whereas the disappearance rate of  $^{18}\text{O}$  reflects both water and  $\text{CO}_2$  turnover rates. Therefore, with time, the difference between the disappearance rates of  $^2\text{H}$  and  $^{18}\text{O}$  represent the rate of  $\text{CO}_2$  production. Knowing the respiratory quotient (RQ) or food quotient (FQ), EE can be calculated from the  $\text{CO}_2$  production rates.

Currently the DLW method is considered to be the reference method for the estimation of EE or caloric requirements in free-living subjects. Six validation studies using calorimetry show that DLW method provides an accurate assessment of the  $\text{CO}_2$  production rate and hence EE; mean $\pm$ SD errors in  $\text{CO}_2$  production were 1.9 $\pm$ 2.0% [228], 1.8% [229], 1.5 $\pm$ 7.6% [230], 1.0 $\pm$ 6.3% [231], 1.0 $\pm$ 8.2% [232] as well as -2.5 $\pm$ 5.8% [233].

Therefore, the DLW method will be used as a criterion method for the estimation of EE in the CALERIE Phase 2 clinical trial. The baseline EE information collected on each study participant will be used to define the caloric restriction prescription. Follow-up EE information measured by the DLW method will be used (1) to determine the effects of the treatment on energy expenditure and (2) together with body composition measurements by DXA, to estimate energy intake.

The DLW tracers will be ordered from stable isotope manufacturers. The DLW mixtures will be prepared by the manufacturers based on the dosage design of the CALERIE Phase 2 clinical trial. The mixtures will be tested for accuracy prior to dispensing into 1-liter leak-proof plastic bottles. The DLW mixtures will be shipped to the Central DLW Laboratory for distribution to the study sites.

For each DLW period, two baseline urine samples will be collected from each study participant. The participant will then receive by mouth, a mixed cocktail containing 0.086 g of  $^2\text{H}_2\text{O}$  at 99.98 atom %  $^2\text{H}$  and 0.138 g of 100%  $^{18}\text{O}$  per kg body weight. Based on body weight, the staff at each study site will weigh out an appropriate amount of the DLW mixture for each study participant. The DLW dosage is designed to minimize potential errors introduced by the anticipated fluctuation in natural abundances of the two isotopes during treatment, to reduce the effect of analytical errors on the precision of the DLW method, and to ensure there are sufficient isotopes at the end of the 14-day DLW study period for accurate and precise isotope ratio measurements [224,234,235]. The cup containing the DLW dose will be rinsed thoroughly three times with approximately 5-10 mL of drinking water and the participant will be instructed to finish drinking all the rinses. Six postdose urine samples will be collected: two on day 0, two on day 7, and two on day 14. Study participants will be instructed to void at home in the morning on days 7 and 14 before the postdose urine samples are collected. The exact time of dosing and sample collection times will be recorded. Urine samples will be transferred to o-ring cryovials in preparation for shipment. Procedures will be carefully followed at each site so the central DLW laboratory remains blinded to treatment assignment and participant ID for CALERIE Phase 2. Encrypted ID labels for each site, created and printed by the Coordinating Center, will be affixed by site personnel to the follow-up urine cryovials, and then stored at  $-70^\circ\text{C}$ . One set of cryovials will be shipped on dry ice to the central DLW laboratory, where they will be stored at  $-70^\circ\text{C}$  until analysis. Samples collected during two baseline DLW studies must be shipped to the central DLW laboratory immediately upon completion so that the samples can be analyzed to meet the 2-week turnaround requirement. Samples collected in the follow-up DLW studies can be shipped in batches. DLW sample analyses for the follow-up studies will be completed as soon as possible, preferably within 6 weeks of receipt of the samples at the central DLW laboratory. The study sites will provide their own dose administration and sample collection kits. The Central DLW Laboratory will provide the sample boxes, shipment boxes, and prepaid Federal Express labels for the sample shipments. The urine samples will be prepared for hydrogen and oxygen isotope ratio measurements by gas-isotope-ratio mass spectrometry according to the procedures developed and validated by Dr. William W. Wong [236,237]. Briefly, for  $^2\text{H}$  assays, 10  $\mu\text{l}$  of urine without further treatment is converted to  $\text{H}_2$  using the zinc reduction method [236,237]. The  $\text{H}_2$  is introduced via the automated sample inlet system directly into a Finnigan instrument for hydrogen isotope ratio measurement. For  $^{18}\text{O}$  assays, the ISOPREP-18  $\text{H}_2\text{O}$ - $\text{CO}_2$  equilibration chambers are used in which 100  $\mu\text{l}$  of urine is allowed to equilibrate with 300 mbar of  $\text{CO}_2$  of known  $^{18}\text{O}$  content for 10 h prior to admission to the ion source of a VG instrument for oxygen isotope ratio measurement [236]. The precision (standard

deviation) for  $^2\text{H}$  assay is 1.0 ‰ for samples with natural abundance of  $^2\text{H}$  and 1.8 ‰ for samples with enriched levels of  $^2\text{H}$  [237]. For  $^{18}\text{O}$  assays, the precision is 0.21 ‰ for samples with natural abundance of  $^{18}\text{O}$  and 0.97 ‰ for samples with enriched levels of  $^{18}\text{O}$  [236].

The Gas-Isotope-Ratio Mass Spectrometry Laboratory at the USDA/ARS Children's Nutrition Research Center has established quality control procedures. The quality control procedures include regular preventive maintenance and/or calibration of the mass spectrometers, the electronic scales, the heating blocks, the  $-70^\circ\text{C}$  freezers, the central water cooler, and the central compressed air supply. Prior to any sample analysis, the operating conditions such as vacuum readings, viscous leak flow rates, compressed air pressure, cooling water pressure, peak shapes, sensitivity, and  $^1\text{H}_3^+$  correction for  $^2\text{H}$  analysis of each mass spectrometer will be evaluated. Dose dilutions also will be analyzed prior to any actual sample analysis. The dose dilution values also will be converted to theoretical isotope fractional turnover rates and EE values in order to monitor the variability of the isotope ratio measurements over time as described in the DLW Quality Control Manual. Reference materials will be analyzed, at least semi-annually, when a filament is replaced or a major tune up of the ion source of the mass spectrometer is required. Isotopic results as well as the chart recorder tracings of the mass spectrometers will be reviewed daily in order to identify nonconformity. If nonconformity is not the result of analytical errors, the study site will be notified immediately for confirmation of the DLW data. If necessary, the samples will be reanalyzed and the justification will be recorded. As part of the quality control procedures, some of the samples collected in the CALERIE Phase 2 clinical trial will be repeated. However, the Central DLW Laboratory will be blinded to the identities of these samples.

The isotopic results will be normalized against two international water standards: Vienna-Standard Mean Ocean Water and Standard Light Antarctic Precipitation [238]. The isotope dilution spaces for  $^2\text{H}$  ( $N_H$ ) and  $^{18}\text{O}$  ( $N_O$ ) will be calculated as follows:

$$N_H \text{ or } N_O \text{ (mol)} = \frac{d \times A \times E_\alpha}{\alpha \times E_d \times 18.02}$$

where "d" is the dose of  $^2\text{H}_2\text{O}$  or  $\text{H}_2^{18}\text{O}$  in grams, "A" is the amount of laboratory water in grams used in the dose dilution, " $\alpha$ " is the amount of  $^2\text{H}_2\text{O}$  or  $\text{H}_2^{18}\text{O}$  in grams added to the laboratory water in the dose dilution, " $E_\alpha$ " is the rise in  $^2\text{H}$  or  $^{18}\text{O}$  abundance in the laboratory water after the addition of the isotopic water, and " $E_d$ " is the rise in  $^2\text{H}$  or  $^{18}\text{O}$  abundance in the urine samples at time zero obtained from the zero-time intercepts of the  $^2\text{H}$  and  $^{18}\text{O}$  decay curves in the urine samples. The use of dose dilution in the calculation of isotope dilution spaces was recommended by the International Dietary Energy Consultancy Groups to assure accuracy of the isotope dilution calculations [239]. Carbon dioxide production rate ( $\dot{V}\text{CO}_2$ ) will be calculated from the fractional turnover rates of  $^2\text{H}$  ( $k_H$ ) and  $^{18}\text{O}$  ( $k_O$ ) as follows [240]:

$$\dot{V}\text{CO}_2 \text{ (mol/d)} = 0.4812 \times [(k_O \times N_O) - (k_H \times N_H)] - 0.0246 \times r_g$$

where  $r_g$  is the fractionated water loss which is calculated as  $1.05 \times (N_O \times k_O - N_H \times k_H)$ . The  $\dot{V}\text{CO}_2$  is converted to TEE based on an energy equivalent of a liter of  $\text{CO}_2$  to be  $3.815/\text{RQ} + 1.2321$  where RQ is the respiratory quotient provisionally estimated as 0.86 [241].

## 14. ADHERENCE TO THE INTERVENTIONS

Adherence measures will be used for the following purposes in CALERIE:

- As a process measure to determine the degree of CR actually achieved by the intervention, and to account for its effect in the analysis of other outcome measures.
- As criteria for decisions during the course of the intervention about modifications of an individual's intervention programs, using the "toolbox" algorithm described in Section 7.3.

Methods for characterizing adherence for these two purposes are described separately below.

### 14.1 Adherence Outcome Measures

The degree of adherence will be characterized as the percentage of CR achieved. This may be expressed as follows,

$$\%CR_P = 100 [1 - (EI_P/EI_{AL(t)})],$$

where  $\%CR_P$  represents an individual's or group's estimated percent reduction in average daily energy intake over the period of interest compared to their *ad libitum* long-term average daily energy intake,  $EI_P$  represents average daily energy intake over the period of interest, and  $EI_{AL(t)}$  represents *ad libitum* long-term average daily energy intake before the start of the intervention. This is the "habitual" EI that is to be maintained, or reduced by 25%, depending on the participant's assignment in CALERIE. For the purposes of comparing the  $\%CR$  achieved to the targeted 25% CR, percent adherence can be also readily characterized as  $(\%CR \text{ achieved} / 25) \times 100$ . However, in the following discussion of CR as an outcome measure, "adherence" simply refers to  $\%CR$  achieved.

*Ad libitum* EI will be characterized by the average of two consecutive measures of energy expenditure by DLW, using methods described in Section 13 above. This approach is based on the assumption that, for short intervals, weight on *ad libitum* EI is approximately stable, and EE thus approximates EI. A weight change of greater than 5% during the interval spanning the DLW measurement periods will be a basis for exclusion.

#### **14.1.1 Intervals Over Which $\%CR$ Will Be Measured**

For  $\%CR$  as process measure, differing time intervals will be of interest. For example, we will wish to know how the intervention group's short-term  $\%CR$  changes from one time point to another over the course of the intervention, and we will also wish to characterize average  $\%CR$  over the entire intervention. For different biologic outcomes, different intervals of CR may be relevant: For example, some biologic outcomes may be strongly affected by the  $\%CR$  occurring shortly before the time of measurement of the outcome. Other biologic outcomes may be more related to the long-term  $\%CR$ , and may change in response to the duration of time that a given  $\%CR$  is maintained.

To analyze  $\%CR$  over different intervals, we will use two versions of the intake/balance method, based on the relationship,

$$EI = EE + \Delta ES,$$

where EE is average daily energy expenditure during the period of interest, and  $\Delta ES$  is the change in body energy stores during the period of interest. These two "intake / balance" (I/B) approaches are described below.

#### **14.1.2 Long-term Intake / Balance Estimation of $\%CR$**

The average  $\%CR$  occurring over intervals of months (e.g., from baseline to month 6, or from month 6 to month 12) will be estimated by the following approach to estimate EI over these intervals. EE over the interval between any two time points at which DLW estimates of EE are made will be estimated by taking the average of the EE estimates from each of the time points. EE over intervals spanning more than two consecutive DLW EE estimates will be based on the average of the estimates for each interval, weighted by the duration of the interval.  $\Delta ES$  will be estimated by calculating the change in energy stores (measured by DEXA) from the beginning to the end of the interval, divided by the duration of the interval in days.  $\Delta ES$  will be calculated using standard coefficients for changes in fat mass and in fat-free mass. Coefficients for calculating energy content of changes in FFM will be reviewed after analyses examining changes in the hydration of fat free mass by subject group and duration of CR. Changes in body water over the interval will be estimated by deuterium dilution measures conducted during DLW determinations.

#### **14.1.3 Short-Term Intake / Balance Estimation of $\%CR$**

This approach will be used to make "point estimates" of  $\%CR$  being achieved at specific time points in the course of the intervention, e.g., at 6 months or at 12 months. The potential limitation of this approach is the inherent noise in assessment of change in body energy stores; thus individual estimates will be subject to error and with this limitation in mind the individual data will be used cautiously. For these calculations, EI at these time points will be estimated using the following approach.

EE will be estimated by 14-day DLW measures at these time points. These will be the same DLW determinations used for the long-term I/B method described above. Short-term  $\Delta ES$  will be estimated by 14-day changes in FM and FFM, measured by DEXA at the beginning and end of the DLW measurement period, or by daily home-scale weight changes estimated from the regression of all home-scale weight measures from the beginning to the end of the DLW measurement period. This value for



short-term change in energy stores will be used to calculate short-term EI using the standard intake/balance equation  $EI = EE + \Delta ES$ .

The importance of home weight measures will be emphasized to participants. However, if fewer than 6 home weight measures over the 14-day doubly-labeled water measurement period are recorded, or if the home weight measure for either day 1 or day 14 is not available, then up to 3 fasting clinic weight data will be substituted for daily home weights.

#### **14.1.4 Approaches for Estimating Long-Term %CR**

For estimates of long-term adherence over the course of the intervention, both long-term and short-term I/B balance data can be used. In this regard, the following considerations are pertinent.

The long-term intake / balance method provides good characterization of long-term  $\Delta ES$ , but has potential problems with estimation of long-term average EE (and therefore of long-term EI) over the interval of interest, since EE may not change linearly, or may fluctuate, between the two DLW measurement periods. We note parenthetically that increasing the frequency of EE determinations will tend to lessen errors due to this problem. The short-term intake/balance methods provides a greater degree of certainty of EE over the period over which  $\Delta ES$  is measured. However, it has problems due to low values of  $\Delta ES$  relative to the level of measurement error. Also, because it is a point estimate, its ability to estimate long-term EI is also limited by the variability of point estimates of EI over a longer interval than the 14-day measurement period. This limitation can also be addressed by increasing the frequency of short-term EI estimates.

In view of the strengths and weaknesses of the two approaches, we will estimate long-term adherence over the course of the intervention by both methods as follows.

- Average of short-term I/B estimates (at 6 months, 12 months, etc.) weighted by the duration of time from the previous estimate.
- Long-term I/B estimates as described in Section 14.1.2 above. (For any interval, e.g., Month 12-24),  $\Delta ES$  will be estimated by changes in FFM and FM (measured by DEXA) from the beginning to the end of the interval, divided by the number of days in the interval.

We will analyze results from the two approaches in regard to concordance, systemic bias, and differences in variability.

## **14.2 Adherence Measures to Inform the “Toolbox” Intervention Algorithm**

Over the course of the intervention, estimates of the degree of an individual’s adherence to the dietary intervention will be used to make decisions about changing the selection or number of “toolbox” interventions. These measures include:

- degree of weight change
- self-reported energy intake (for active intervention group only)
- process measures

### **14.2.1 Degree of Weight Change**

Weight loss in response to a given percentage reduction in EI will vary considerably among individuals, being affected by factors such as body composition, relative ability to mobilize different types of energy stores (e.g., fat or muscle), and changes in BMR and physical activity in response to the change in EI. Thus the ability to estimate an individual’s adherence to a prescribed %CR based on weight change alone is limited. Nonetheless, certain degrees of weight change are outside reasonable confidence limits for expected weight change, and will trigger consideration of modifications of the individual’s intervention program, using the “toolbox” algorithm.

The goal of this analysis was to provide adherence boundaries (similar to normograms in psychological experiments) for CR interventions by employing body weights and %CR as measured in CALERIE Phase 1. Across strata and sites, the prescribed %CR and individual adherence were variable. For the development of the normogram, a set of estimated curves was developed and compared. To accomplish this, each individual’s percentage weight change during the intervention period was calculated, and the functional form of the curve was estimated across individuals. One objective of these calculations was to estimate a curve which fit the observed data over the course of 12 months. For parsimony, a set of simple models was estimated for the first 12 months of CR to be used in Phase 2.

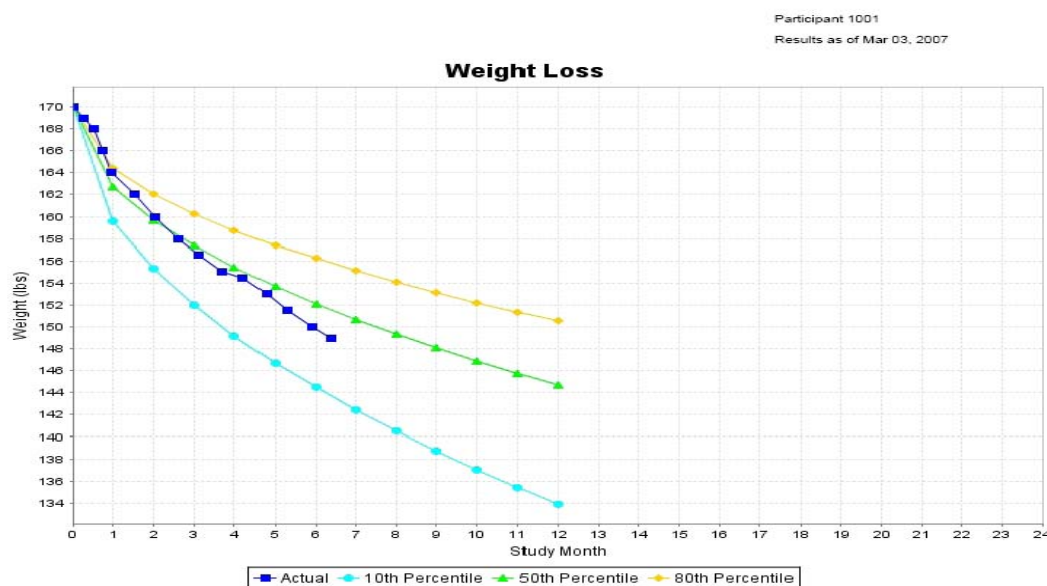
Using the relationship between the change in weight over time and calculated %CR at 6 months in CALERIE phase I, the estimated weight was adjusted to an expected level at 25% CR, the level of CR prescribed in CALERIE Phase 2. As a second step, the distribution of these expected changes was assessed. Finally, using pre-determined percentile values from this expected change, a normogram for boundaries of expected change was developed for use in CALERIE Phase 2.

Suboptimal adherence with caloric restriction to inform the “toolbox” intervention algorithm is defined by referring to the participant’s individual weight graph described above as provided by the Computer Tracking System. “Closed Toolbox” strategies should be applied immediately when a participant’s weight is outside the adherence curve. A participant shall not be considered out of adherence for this toolbox options until s/he enters the 5<sup>th</sup> week of the study. “Closed Toolbox” strategies shall be discontinued when the participant no longer meets criteria necessary to open the “Closed Toolbox.”

In order to determine whether criteria for using the “Closed Toolbox” are met, the following definition was developed:

- the upper boundary of adherence (lowest level of adherence before opening the “Closed Toolbox”) should be the 80<sup>th</sup> percentile; and
- the lower boundary (most rapid rate of allowable weight loss) should be the 10<sup>th</sup> percentile

Below is a graphic representation of a participant’s weight change in relation to the 10<sup>th</sup>, 50<sup>th</sup>, and 80<sup>th</sup> percentiles lines in order to determine adherence.



If participant’s rate of weight change is within expected parameters, but the participant reports difficulty following his/her CR prescription, then “Open Toolbox” strategies should be used.

### 14.2.2 Self-Reported Energy Intake

Sub-optimal adherence with self-monitoring is defined by the failure to self-monitor eating behavior and/or dietary intake. The “Open Toolbox” options shall be used first. However, if criteria are met for opening the “Closed Toolbox,” the Counselor should open the “Closed Toolbox” strategies. “Closed Toolbox” strategies shall be applied when a participant self-monitors for less than 70% of days since the previous session for month one through six, less than 50% for month seven through month twelve, and less than 30% of days for month thirteen through month twenty-four. A day is considered incomplete when the Counselor ascertains that the reported calorie intake is less than 80% of the individual’s actual intake. “Closed Toolbox” strategies may be discontinued when the participant no longer meets criteria necessary to open the “Closed Toolbox.”

### 14.2.3 Process Measures

Consideration of changes in use of toolbox options to improve adherence will be triggered if there is:

- failure to attend two consecutive individual or group sessions or missing three out of the last four sessions for unjustified reasons;
- failure to implement behavioral/nutritional strategies agreed on at previous visit(s);
- inability to achieve goals for ability to estimate energy content of diet, or other self-monitoring tasks.

## **15. QUALITY CONTROL ACTIVITIES**

Prior to study start, the Quality Control (QC) Committee will develop a detailed quality assurance plan and will lead and oversee development of the quality control procedures in accordance with the plan. The quality assurance plan will address the areas presented below. During the study, the QC Committee will ensure implementation of the plan and strict adherence to the QC procedures by all CALERIE Phase 2 structural components.

### **15.1 Staff Adherence to Study Procedures and Timetables**

Several approaches will be used to ensure adequate adherence to study procedures and timetables by the study staff. These approaches include development of a comprehensive Manual of Procedures, initial and ongoing training and certification of each staff member, oversight and management of the daily clinical site operations by a Study Manager, and multidimensional monitoring program.

#### **15.1.1 Manual of Procedures**

Study Manual of Procedures (MOP) will be developed to facilitate consistency in protocol implementation and data collection across participants and clinical sites. The MOP will detail study conduct and operations and provide reassurance that scientific integrity of the study and participants' safety are closely monitored. This will increase the likelihood that the results of the study will be scientifically valid. The Manual will transform the study protocol into a set of procedures describing the study organization, operational definitions of the outcomes, participant recruitment, screening, enrollment and randomization as well as follow-up procedures, data collection methods, data flow, case report forms (CRFs) completion and quality control procedures.

The MOP development process will be led by the QC Committee and will involve investigators and study staff from the CC, clinical sites, central and DLW laboratories, DXA and other reading centers, the Core Clinical Laboratories, and NIA representatives. The MOP will be updated throughout the conduct of the study to reflect any protocol or consent amendments as well as possible refinement of the CRFs and study procedures. Sufficient number of copies of the MOP will be provided to each study structural component, and electronic copy will be available to all study staff on the CALERIE website.

#### **15.1.2 Training and Certification**

Under the guidance from and supervision by the QC Committee, training materials on appropriate study procedures will be developed by the CC, laboratories and reading centers. Training will be a dynamic process. Prior to study start, staff members from all clinical sites and other study structural components will attend the Investigators' Meeting where they will receive initial overall training on the study protocol and procedures. The general training session will be followed by the group training sessions where staff members will be trained and certified in their functional areas (e.g., vital signs, height and weight, questionnaires and instruments, RMR). Training sessions for some specialized study procedures (e.g., DXA) may be held at the appropriate study structural components (e.g., a reading center).

The first training session after the Investigators' Meeting will be held by the CC's monitor during site initiation visit, and then subsequent training sessions and recertification will be required annually. These training sessions and recertification for individual staff members may also be held on an "as needed" basis. Any new staff member who joins the research team after study start will be subject to all training and certification requirements in his/her functional area. Training materials will be available to all study staff on the CALERIE website.

#### **15.1.3 Study Manager**

Each clinical site will have a Study Manager with substantial experience in clinical research operations and management of the multi-center clinical trials. This person will work in collaboration with the clinical site's Principal Investigator (PI) to coordinate and implement the CALERIE Phase 2 protocol. The Study Manager will supervise and oversee day-to-day operations of the clinical site, ensure adherence to the

Good Clinical Practice guidelines, study protocol, Manual of Procedures, study timeline and budget. He/she will identify operational problems, assess staff performance and provide feedback to the PI. When necessary, the Study Manager will direct remedial action and determine the training needs at their site. This person will serve as liaison between the clinical site, NIA, Coordinating Center, laboratories, and other study structural components.

#### **15.1.4 Multidimensional Monitoring Program**

Prior to study start, the CC, reading centers and laboratories will be required to develop appropriate monitoring plans for their study functional areas with the guidance and oversight by the QC Committee.

The CC's monitoring plan will cover participant recruitment, consent process, screening, performance of general procedures (e.g., BP measurement), equipment calibration and maintenance, data collection and transmission, AE collection and reporting, and compliance with the IRB, OHRP, and NIH regulations. To implement this plan, the CC's monitor will conduct regular semi-annual visits to the clinical sites. S/HE will review the documentation and verify that procedures are being followed. The reading centers and laboratories monitoring plans will include regular site visits by experienced staff from the reading centers and laboratories to observe performance of the procedures by the clinic staff.

All deficiencies and deviations will be brought to the attention of the site coordinator, Study Manager and study leadership and documented in monitoring reports and follow-up letters to the site PIs. If necessary, monitors will provide on-site training to the study staff.

### **15.2 Quality Control of Study Measurements**

Study measurements will be performed by clinical sites, laboratories and reading centers. Several approaches are being and will be used to ensure quality of the measurements. Quality assurance starts with selection of laboratories and reading centers. Any central facility will be evaluated according their ability to perform required measurements, level of staffing and previous experience with multi-center trials, capacity in conducting their evaluations in a timely manner, their standard operating procedures and quality control procedures.

Quality control procedures ensuring consistency and uniformity of measurements by all structural components during the study will be developed by the QC Committee in collaboration with the appropriate structural components as part of the overall quality assurance plan. Clinical sites, laboratories and reading centers will follow the study MOP and their SOPs, and maintain appropriate records documenting their operations and measurements. The CC will require the central lab(s) provide annual updated documentation of their SOPs and certifications.

### **15.3 Quality Control Documentation and Reports**

Adequate documentation and reporting of study progress, performance of structural components and evaluation of adherence to the protocol and MOP requirements are important part of the quality assurance plan that allows the study leadership to effectively manage the study.

Each structural study component will keep detailed records to document and monitor performance, quality of operations and progress in their functional area. Some of this documentation will be developed by the CC (e.g., weight measurement log), and some will be mandated by the study structural component Standard Operating Procedures (e.g., samples tracking reports for the laboratories, etc.). This documentation will be subject to review by the study monitors and will be used to compile a number of comprehensive progress, operational and quality control reports that will be generated by the CC for review by the study leadership.

A number of reports will assess performance in all study functional areas. These reports will describe:

- Recruitment, follow-up, drop-out and withdrawal from the intervention
- Adherence to intervention
- Adherence to study procedures
- Protocol deviations
- DLW laboratory progress
- Central laboratory progress

- Reading centers progress
- Laboratories test-retest reliability reports

Reports will be regularly reviewed by the Quality Control Committee to evaluate performance of the individual structural study components and initiate appropriate corrective action, if necessary.

## **15.4 Specific Details on Outcome Domains**

### ***15.4.1 Doubly Labeled Water (DLW)***

All samples from a doubly labeled water study in a patient will be run together, and at least three internal standards that span the range of enrichments anticipated in CALERIE Phase 2 will be run with each batch of samples. To present variability in standard data over time to the QC Committee, theoretical time points will be created (separate time points for  $^2\text{H}$  and  $^{18}\text{O}$ ) for each batch of internal standards so that the standards on average in the first 10 runs can be used to calculate a rate constant for  $^2\text{H}$  and a rate constant for  $^{18}\text{O}$  disappearance, respectively, of 0.1 and 0.13. For each set of internal standards run, the standard data will be used with the theoretical time points to calculate theoretical rate constants for isotope disappearance for that standard run. The difference between each set of theoretical rate constants from the internal standards will be expressed as a percentage difference from 0.03 (the defined difference between the two rate constants for the average of the first 10 runs). These transformations of standard data will be presented graphically on the y axis of a graph in which x axis is date of analysis. By this means, the QC Committee will be presented the changes in DLW analytical error over time. Although this presentation of standard data does not accurately reflect error in patient data (since there are fewer standards than the number of urine specimens measured for each patient) changes over time can be used to evaluate machine accuracy and precision.

Reports will be reviewed by the QC Committee of the transfer and receipt of samples at the DLW Core Laboratory, the generation of results and the transfer of these results to the field sites and CC.

### ***15.4.2 DXA***

Quality Control Procedures for the Instrument: All clinical sites and the central reading center will use an Hologic DXA machine. The instrument at each site will be tested regularly with an Hologic Spine Phantom. The spine phantom is measured each day that a study participant is scanned, or at least 3 days per week for each DXA scanner. There will be a weekly scan at each site of the local Whole Body Phantom. These scans will be saved and sent to the Core Reading Laboratory for analysis.

Cross Calibration of the study DXA Scanner will be done at least once during Phase 2 and will be supervised by the Coordinating Center and Core Reading Laboratory. Additional cross calibrations will be performed pending input from the Core Reading Laboratory. This will be a three point calibration with ten scans at each configuration. The scan data will be saved and transferred to the Core Reading Laboratory for analysis and calculation of the study wide CV's for the various analyses. These results will be presented to the QC Committee for review.

Quality Control Procedures for Technologists and Technicians: Technicians will be required to be familiar with key components of the Hologic Manual, the relevant sections of the CALERIE Protocol and MOP. We will require certification of the chief operator at each site by the International Society for Clinical Densitometry (ISCD) and training with the Hologic video. They will also be required to provide a set of scans to the Core Reading Laboratory, including whole-body, hip, spine and forearm scans on several individuals, that demonstrate proper scan acquisition technique in order to be certified to obtain DXA scans in CALERIE.

### ***15.4.3 Resting Metabolic Rate (RMR)***

RMR measurements will be an important endpoint for CALERIE Phase 2. The same Vacumed metabolic cart will be used at all three field sites, with a common protocol and a training of the operators. A training session for RMR data collection will be held before the initiation of Phase 2. A QC frequent alcohol burn test will be performed regularly (a least bimonthly) and those QC tests will be reported to the QC Committee for review. Measurements will be done fasting in the morning after spending the night at the research facility and fed a standard dinner the night before. Data will be collected between 7 and 9 AM at least after lying for ten minutes on a bed plus another ten minutes with the hood system over the head. Data will then be collected for 30 minutes with stable data.

#### **15.4.4 Core Clinical Chemistry and Biochemistry Laboratory**

The Central Laboratory will standardize procedures for collecting, processing, storing and shipping blood, tissues, and urine samples by providing a detailed manual with explicit instructions for the clinical sites and laboratory staff. Centralized training for all staff involved in sample handling and processing will be provided as well as annual re-certification training to promote standardization across the study. Mechanisms will also be put in place to train and certify individuals who join field center staffs in the interim periods between annual re-certifications.

Quality control for analytes will be assessed internally through the use of pooled control materials for most assays and externally through proficiency testing via participation in standardization programs such as those conducted by the College of American Pathology, the American Association for Clinical Chemistry, and others. All assays where the results are returned to the participants are done under CLIA certification. Reproducibility will be determined by assaying a percent of randomly selected samples in duplicate and by calculating technical error values. Instrument calibration and maintenance will be routinely conducted and documented. Data whenever possible will be entered directly from instruments into databases; however, data entered manually will be double entered and reviewed before release.

Sample labels will be provided to the clinical sites along with shipping supplies. A database at the laboratory will link the participant ID number to its sample ID number at the laboratory. Laboratory results will be reported to the Coordinating Center on a weekly basis or as requested. Critical levels will be predetermined for all assays, and any results considered to be life threatening will trigger a call to the clinical site followed by faxing of the results.

The Central Laboratory will also provide long term storage for unanalyzed samples. Frozen samples will be maintained at either  $-80^{\circ}\text{C}$  or  $-145^{\circ}\text{C}$ , depending on the sample, in locked freezers with an emergency backup system. These alarmed freezers are monitored 24 hours/day, 7 days/week, and routine testing is performed monthly. Samples will be inventoried and tracked for location by rack, storage, and box position.

#### **15.4.5 Peak Oxygen Uptake Measures (Peak $\text{VO}_2$ )**

Maximum oxygen uptake (peak  $\text{VO}_2$ ) is being assessed in CALERIE via incremental treadmill exercise tests to determine if any of the interventions results in changes in cardiovascular fitness. Generally  $\text{VO}_2$  data are vulnerable to measurement error from differences in subject motivation among tests, due to lack of calibration of the metabolic cart during preparation for testing a subject or due to instrument error. All study personnel that will be performing peak  $\text{VO}_2$  testing will be trained at a common training session on the trouble-shooting of instrument errors, on the proper calibration of the instrument, and on performing a true maximum test using physiologic criteria ( $\text{RER} \geq 1.05$ ;  $\text{RPE} \geq 17$ , plateau in the  $\text{VO}_2$  curve).

#### **15.4.6 Measures of Physical Activity: 7-day Physical Activity Recall (7-day PAR)**

The 7-day Physical Activity Recall (7-day PAR) is being used at all three of the field sites to measure spontaneous and planned physical activity of study subjects. The instruments will be scored by study personnel at each site. The scored data will then be transferred to the Coordinating Center for the generation of final reports using a common algorithm. In Phase 2 CALERIE, in order to ensure uniformity of collection and scoring of the 7-day PAR study personnel will be trained in a common training session by a qualified 7-day PAR training (e.g., Dr. William Haskell or Dr. Steven Blair) prior to initiation of the Phase 2 studies.

### **15.5 Reliability Studies**

#### **15.5.1 Test-Retest Reliability Model**

The statistical analyses and scientific inferences drawn from this study depend critically on having valid results from the central laboratories and reading centers. Although it would be desirable to conduct reliability studies on all the central facilities, this is not feasible. There are costs associated with this approach including the financial cost of conducting the analysis on the retest samples, the extra time and effort by lab personnel to analyze these samples, the capacity taken away from the central function of the facility, and the extra complication of taking duplicate samples, blinding their identities and forwarding them to the central facility. For the DLW lab, in particular, there is a requirement to turn the analyses

around quickly to derive the correct energy prescription at baseline and to monitor adherence during follow-up. Thus, reliability studies will only be conducted for two of the central facilities, i.e., the DLW lab and the central biochemistry laboratory.

The two studies will be conducted independently of each other, and in both cases a test-retest reliability study will be performed. Duplicate urine sample sets from the same subject at the same protocol time point will be collected and forwarded to the DLW lab for analysis. Only a sample of all specimens collected from the participants will be re-sampled (see below). The specimens will be labeled in such a way that facility personnel are blinded to which participant (and treatment arm) the specimens correspond, and whether it is the original or duplicate specimen. A “key” will be created at the Coordinating Center to link the identify of the test and retest specimens to the true CALERIE ID number so that statistical analyses can be performed. Similar procedures will be applied for blood samples sent to the biochemistry lab with the exception that the duplicate samples will be drawn from the blood stored in the repository.

### **15.5.2 Statistical Analysis**

Statistical analyses will be performed to address two questions: (1) what is the “reliability” of the results in the facility (defined below), and (2) is there a significant difference between the test and retest specimens. Simple descriptive statistics will be generated on the “test” and “retest” specimens and the difference between them. A plot of the retest results against the test results will be made with a reference line at 45° to observe whether there are systematic differences. A Bland-Altman plot will be made to determine whether differences are consistent across the entire range of the outcome measure.

Formal statistical methods for analyzing a reliability studies are described in Shrout and Fleiss [242] and Eliasziw *et al.* [243]. The underlying statistical model is a one-way ANOVA model with a random subject effects term, and from this the intraclass correlation  $\rho$  can be derived. Heuristically, this is the proportion of total variance accounted for by subject-to-subject variability, and a large value indicates that there is relatively little within-subject variability. The intraclass correlation is also known as the “reliability coefficient,” and it will be used as the primary reliability measure. Additionally, a paired  $t$ -test will be conducted to determine if there are systematic differences between the specimen pairs. These analyses are complementary, i.e., the first addresses variability while the second evaluates differences in means.

### **15.5.3 Sample Size Calculations**

Sample size calculations for reliability studies were performed following the procedures in Donner and Eliasziw [244]. Let  $\rho_0$  represent the minimum reliability level that is acceptable for the central facility, and consider the hypotheses,  $H_0: \rho = \rho_0$  vs.  $H_A: \rho > \rho_0$ . Donner and Eliasziw suggested that the minimum acceptable reliability for a test-retest study is  $\rho_0 = 0.80$ , and it was therefore adopted for these calculations. The type-I error was set to 0.05 and type-II error to 0.20. Applying this methodology and assuming that the true reliability in the facility is 0.90, a minimum sample of 46 duplicated specimens is required. Rounding up, a study consisting of 50 test-retest specimens will be performed.

### **15.5.4 Resampling Rates**

For DLW laboratory, two DLW determinations will be conducted on study participants at baseline. However, baseline DLW studies will not be included in the test-retest study. This is to avoid delaying the lab in deriving the baseline TEE values so that the correct energy prescription can be derived. Otherwise, follow-up DLW studies will be performed at Months 6, 12, 18 and 24 in the CR intervention and at Months 12 and 24 (only) in the control group. Assuming that subjects will be enrolled over a period of 20 months with an annualized drop-out rate of 10%, it follows that approximately 700 DLW studies across both treatment arms will be eligible for the reliability study. Thus, a sample of  $50 / 700 = 7.1\%$  of the eligible DLW studies will be selected at random and included in this study. To protect against incorrect assumptions, this is rounded up to 8% of eligible samples.

Thus of the 700 DLW studies performed post randomization and sent to the lab for analysis, a subset of 8% of them would be selected at random and included in this study. A duplicate urine sample set will be created at the clinical site, and forwarded in a blinded manner to the lab for analysis. Thus, 50 duplicate TEE (and associated parameters) would be derived by the lab and made available for statistical analysis.

For the central biochemistry lab, the main evaluations will be performed at baseline and at Months 12 and 24. Under the same assumptions as above, approximately 650 blood samples will be drawn and sent to the laboratory for analysis. Thus, a random sample of  $50 / 650 = 7.6\%$  of the eligible blood samples will be selected and included in the biochemistry lab reliability study. This is rounded up to 8% of eligible samples. Thus, of the 650 blood samples to be collected, a subset of 8% of them would be selected at random and included in this study. Duplicate blood samples will be drawn from the blood stored in the repository and forwarded to the lab for analysis. Thus, 50 duplicate sets of blood markers would be derived by the lab and made available for statistical analysis.

## 16. PARTICIPANT SAFETY AND ADVERSE EVENTS

### 16.1 Signs, Symptoms and Adverse Events

Protection of subjects from risks related to the study is of paramount concern to investigators and institutions participating in CALERIE. Due diligence will be applied to ensure that study subjects are not exposed to undue or untoward risk. Periodic reports summarizing participant safety will be presented to the CALERIE Steering Committee and to the Data & Safety Monitoring Board (DSMB). Remedial action, including additions and changes to the protocol, will be taken as appropriate.

According to ICH Guideline E2a, an adverse event (AE) is, “Any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and which does not necessarily have to have a causal relationship with this treatment.” Of course, CALERIE is not being performed in a regulatory environment and no pharmaceutical product is being investigated. Nevertheless, this is a reasonable definition to apply for this intervention with this study population, and it will therefore be adopted. Thus, for the purposes of this study, an AE is defined as any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the intervention irrespective of whether it is considered related to the intervention. ICH guidelines also distinguish between “expected” and “unexpected” adverse events. An unexpected AE is one which by its nature or severity is inconsistent with pre-existing knowledge concerning the nature of the intervention. Given, however, that is the first detailed investigation into the medium-term effects of calorie restriction in humans, we do not draw this distinction. All AEs, expected or unexpected, will be reported.

Participants will be given a diary to record signs, symptoms and adverse events occurring during participant follow-up. This includes anecdotal reports of medical problems such as headaches, dizziness, fatigue, pain, nausea and infections as well as menstrual irregularities for women. In addition to the regularly scheduled clinic visits, monthly contact will be made with participants in both treatment arms, either by telephone or in person. Diary-reported AEs will be reviewed by CALERIE staff, and newly emerging AEs, as well as the status of any AEs that have not yet resolved, will be recorded in the case report forms. During these contacts, the CALERIE staff member will proactively ask whether any sign, symptom or AEs has occurred since the previous contact so that a complete accounting is made.

Medical conditions and adverse events will also be captured at the baseline visit in order to establish the prevalence in each treatment arm prior to starting the assigned intervention. Treatment-emergent AEs, i.e., those newly observed or those pre-existing at baseline that worsen after starting the intervention, will be therefore be identified.

For each sign, symptom or adverse event, the following information will be recorded:

- a brief descriptor of the adverse event
- start and stop dates
- intensity (mild / moderate / severe)
- whether the AE was “serious” or not (as defined below)
- Causal association with the intervention assigned (none / doubtful / possibly / probably / very likely)
- outcome (resolved / resolved with sequelae / improving / still present and unchanged / death)
- action taken with respect to the intervention (none / intervention temporarily discontinued / medical therapy required / intervention permanently discontinued / other).



## 16.2 Serious Adverse Events

ICH Guideline E2a provides a separate definition for a “serious” adverse event (SAE), and it is adopted in this study. Specifically, a SAE is one that satisfies one or more of the following criteria:

- results in death;
- is life-threatening;
- requires in-patient hospitalization, or prolongation of an existing hospitalization;
- results in persistent or significant disability or incapacity;
- results in a congenital anomaly / birth defect; or
- is an otherwise important medical event.

“Life-threatening” means that, in the view of the Investigator, the participant was placed at immediate risk of death from the event as it occurred. A “disability” is a substantial disruption of a person’s ability to conduct normal life functions. An “important medical event” is any other event that jeopardized the health of the participant and required medical or surgical intervention to prevent one of the other outcomes listed above from occurring.

A “Serious Adverse Event Form” will be specifically used to capture the information for each distinct occurrence. In addition to the information described above, the following data will be collected:

- the defining SAE criterion (from the list above)
- relationship to assigned treatment intervention (none / doubtful / possibly / probably / very likely)
- relevant participant history
- concomitant medications
- a narrative providing greater detail of the events surrounding the SAE
- source of the information.

An expedited reporting protocol will be followed for reporting SAEs. An initial SAE report will be completed within 24 hours of first knowledge of the event and forwarded to the Coordinating Center; the CC in turn will inform the NIA Project Official and the Chairman of the DSMB. A follow-up to the original SAE report will be completed within 5 additional calendar days and forwarded to the CC. As soon as feasible, the CC will forward a copy of the SAE Report to the other members of the CALERIE Steering Committee. Each CALERIE Principal Investigator is responsible for informing his/her own IRB according to local procedures. The Chairman of the DSMB has the option of informing the other members of the Board, convening a full meeting of the DSMB, or adopting any other action he deems appropriate for the event. The site will also forward a follow-up report describing the SAE resolution and outcome to the CC for distribution to the NIA Project Official, the Chairman of the DSMB and the CALERIE Steering Committee.

SAEs not resolved by the end of the study or that have not resolved upon discontinuation of the subject’s participation in the study, will be followed by the CALERIE clinical site until the event resolves, stabilizes or returns to baseline.

## 16.3 Adverse Events Anticipated in this Study

The CALERIE Phase 2 is being conducted with healthy volunteers, and the interventions are thought to be relatively benign, so that the signs, symptoms and AEs anticipated in this study are expected to be low in quantity and severity. Moreover, analyses of the Phase 1 data indicated that in general, the CR interventions were well tolerated by the participants. No deaths or life-threatening SAEs were observed. Only 6 AEs were reported in Phase 1 that led to a temporary discontinuation of the intervention. They included moderate dyspnea, moderate oropharyngeal swelling, a fall, severe depression, insect bite, and moderate anemia. However, in the opinion of the investigators, only anemia was considered “associated” with the intervention. Most AEs were characterized as “mild” or “moderate” by the CALERIE investigators. Seven adverse events were characterized as “severe,” and they included ligament injury, poison oak exposure, a high Eating Disorder Examination (EDE) score, depressed mood, appendicitis, depression and toothache. All of these events, except the high EDE score, were considered “not associated” with the intervention by the study investigators. Infections were more often reported by the participants enrolled in CR groups than in control groups, and they will be closely monitored in Phase 2.

Otherwise, no other adverse events, which by their nature or severity could significantly limit activities of daily living, were reported by the participants in the CR groups in Phase 1.

The following AEs were reported by more volunteers in the Phase 1 studies:

- headache
- stress symptoms
- sleep disorders
- insomnia
- back pain
- diarrhea
- nasopharyngitis
- sinus congestion
- irregular heart rate
- arthralgia
- asthenia
- irritability
- dizziness
- depressed mood
- anemia
- fatigue
- pain in extremity
- multiple allergies.
- constipation
- nausea

They will be given particular attention in this study.

## 16.4 Clinically Significant Laboratory Values

Analysis of CALERIE Phase 1 laboratory test results showed that particular attention should be paid to the possibility of elevation in potassium levels and an increased incidence of anemia in participants exposed to the CR regimen.

All clinical laboratory tests will be performed by the central laboratory for CALERIE Phase 2 study, at baseline and Months 1, 3, 6, 9, 12, 18 and 24 visits unless otherwise specified.

The central laboratory will provide the centers with all materials necessary for specimen collection, temporary storage, and shipping of samples. The clinical centers will collect the specimens and store them in accordance with applicable storage conditions. All sites will ship samples to the central laboratory for clinical laboratory evaluation and issuance of a laboratory report.

### 16.4.1 Clinical Laboratory Tests

For all fasting laboratory tests, participants must fast for a minimum of eight hours. This includes no coffee and juices. Water is allowed. No study procedure will be administered until fasting laboratory samples are obtained. These evaluations will be performed at baseline and at Months 1, 3, 6, 9, 12, 18 and 24.

Hematology: Hematological tests will include: white cell count (WBC) with differential, red cell count (RBC), RBC morphology, hemoglobin (Hgb), hematocrit (Hct), and platelet count.

Serum Chemistry: Serum chemistry tests will include sodium, potassium, calcium, magnesium, iron, albumin, total protein, total bilirubin, alanine transaminase (ALT), aspartate aminotransferase (AST), gamma-glutamyl transpeptidase (GGT), alkaline phosphatase, creatine phosphokinase, creatinine, blood urea nitrogen (BUN), glucose and uric acid. Total, HDL- and calculated LDL cholesterol, triglycerides, C-reactive protein, and blood glucose and insulin values will also be performed.

Urinalysis: Urinalysis will include: pH, protein, glucose, ketones, occult blood, specific gravity, and a microscopic examination.

Pregnancy Test (serum human chorionic gonadotrophin – qualitative and urine: A serum pregnancy test will be performed at screening for all female participants. A urine pregnancy test will be performed at baseline and then at Months 6, 12, 18 and 24 for CR participants, and at Months 12 and 24 only for controls. It will also be performed prior to the Hepatitis A vaccine (and booster) at Months 17 and 23.

## 16.5 Potassium Surveillance Protocol

Hyperkalemia is defined as a potassium level greater than 5.5 mEq/L. Levels between 5.5 mEq/L and 6.0 mEq/L indicate a mild condition, between 6.1 mEq/L and 7.0 mEq/L – a moderate condition, and above 7.0 mEq/L – a severe hyperkalemia [245].

Only participants with normal potassium levels at screening and baseline will be allowed to enter into the study. After the baseline visit, potassium levels will be measured at Months 1, 3, 6, 9, 12, 18, and 24.

Any participant enrolled in the CR group who has potassium level between 5.5 mEq/L and 6.0 mEq/L at any point during the study will have the test repeated in one week. If a repeated test shows elevated above 5.5 mEq/L potassium level, the CR intervention will be temporarily discontinued and a

participant will be advised to seek medical help outside of the study. The CR intervention will be restarted after the potassium level decreases to 5.0 mEq/L or below.

Any participant enrolled in the CR group who has potassium level of 6.1 mEq/L and above at any point during the study will have the test repeated within 48 hours. If a repeated test shows elevated above 5.5 mEq/L potassium level, the CR intervention will be temporarily discontinued and a participant will be advised to seek medical help outside of the study. The CR intervention will be restarted after the potassium level decreases to 5.0 mEq/L or below.

If potassium level is still elevated above 5.0 mEq/L after one month of treatment, the CR intervention will be permanently discontinued and a participant will follow all other study procedures to the study end. The CR intervention will also be permanently discontinued if any participant in the CR group shows increase in potassium level of 5.5 mEq/L and above after the CR was restarted. A participant will follow all other study procedures to the study end.

Participants enrolled in the control group who have potassium level of 5.5 mEq/L and above at any point during the study will have the test repeated in accordance with the criteria presented above and then advised to seek medical help outside of the study. Because no intervention is administered in this group, participants will continue in the study.

Creatine phosphokinase levels will be measured and electrocardiogram (ECG) will be recorded at the same time points when potassium measurements are performed.

ECG is essential and may be instrumental in diagnosing hyperkalemia in the appropriate clinical setting. ECG changes have a sequential progression of effects, which, however roughly, correlate with the potassium level. The ECG will be watched for the early signs of hyperkalemia that include peaked T waves, shortened QT interval, and ST segment depression. All ECGs for the participants who have potassium levels elevated above 5.5 mEq/L will be reviewed by the cardiologist at each clinical site.

## **16.6 Anemia Surveillance Protocol**

In CALERIE Phase 1 study a decrease in hemoglobin and hematocrit levels from baseline to any point during the study was observed more often in the participants enrolled in the CR group than in those enrolled in the control group. Thus, to ensure participants' safety, a comprehensive anemia surveillance protocol will be followed in CALERIE Phase 2 study. In addition, participants will be advised not to donate blood during the study period.

For the purpose of this study anemia is defined as hemoglobin (Hgb) and/or hematocrit (Hct) level below the lower limit of normal (LLN) for the laboratory.

Only participants with normal as per clinical laboratory reference range Hgb, Hct, and RBC levels at screening and baseline will be allowed to enter into the study. During the study, hematology tests will be performed at Months 1, 3, 6, 9, 12, 18, and 24 on all participants.

Any participant enrolled in the CR group who shows decrease in Hgb, Hct, or RBC level below the lower limit of normal for the laboratory or demonstrates a decrease of 5 percentage points in Hct level (even if it stays within the normal range) at any point during the study, will have the hematology panel repeated in two weeks. Iron level will also be repeated.

If repeated tests confirm previous findings, the participant will be advised to seek medical help outside of the study. Iron, folic acid and/or vitamin B<sub>12</sub> are allowed. A participant will continue in the study and the hematology panel and iron level will be repeated one month after the treatment was initiated. If anemia is not improving or worsens, the CR intervention will be temporarily discontinued and will be restarted after the test results return to normal levels. If anemia is not improving or is worsening after the CR intervention was temporarily discontinued for one month, the intervention will be permanently discontinued and a participant will be asked to follow all other study procedures to the study end.

Participants enrolled in the control group who develop anemia at any point during the study will have the tests repeated in accordance with the criteria presented above and advised to seek medical help outside of the study. Because no intervention is administered in this group, participants will continue in the study.

## 16.7 Cholesterol Surveillance Protocol

At screening, total cholesterol, HDL-cholesterol, LDL-cholesterol (calculated), and triglycerides concentrations will be measured by the Esotereix/LabCorp and results will be provided by the laboratory to the clinical sites as part of the safety chemistry panel. Only candidates with LDL-cholesterol level < 190 mg/dl at screening will be allowed to enter into the study. Any volunteer who has LDL-cholesterol  $\geq$  190 mg/dl will be advised to seek medical help outside of the study.

No action will be taken for any candidate whose LDL-cholesterol level is greater than or equal to 160 mg/dl and less than 190 mg/dl prior to randomization. After randomization, participants assigned to the CR group will follow their CR regimen. Participants randomized to the control group will be advised to follow the American Heart Association's low-cholesterol diet.

During the study, total cholesterol, HDL-cholesterol, LDL-cholesterol (calculated), and triglycerides concentrations will be measured at Baseline, and Month 12 and 24 by the University of Vermont laboratory as part of the outcome measures. These results will be provided to the CC, but will not be provided to the clinical sites except as indicated below.

*LDL-cholesterol level greater than or equal to 160 and less than 190 mg/dl:*

CR group: the clinical site will not be notified and the participant will continue to follow his/her CR regimen.

Control group: the CC will notify the clinical site and the participant will be advised to follow the American Heart Association's low-cholesterol diet.

*LDL-cholesterol level greater than or equal to than 190 mg/dl:*

Both groups: the CC will notify the clinical site and the participant will be advised to seek medical help outside of the study. S/He, however, will continue to participate in the study.

## 16.8 Monitoring for Nutritional Adequacy and Eating Disorders

Nutritional Adequacy: The selection of a nutritionally adequate CR diet will be assured in the early weeks of the intervention by close nutritional supervision. Meal plans developed by research dietitians to meet the recommended dietary intakes of all essential nutrients for gender and age will be followed during the first four weeks of the study to establish a pattern of nutritional adequacy. Nutritionists will continue to review the weekly food diaries (being kept for behavioral purposes) for the entire study duration and discuss any concerns directly with the subject. In addition, subjects will receive a daily vitamin/mineral supplement providing 100% of the recommended adult intakes of all known essential micronutrients, including a separate supplement of 1,000 mg elemental calcium as calcium citrate. Regular intake of vitamin/mineral and calcium supplements will be confirmed. Careful monitoring of body weights and blood chemistries will also provide assurance of nutritional balance and hydration. In addition, bone mineral density and body composition will be monitored at 6, 12 and 24 months.

Eating Disorders: The study investigators realize that a small subgroup of the population could be susceptible to developing disordered eating behaviors if enrolled in a CR protocol. A screening protocol has been developed that will exclude participants from the study who display eating disorder pathology. During the study, participants will be administered the Multi-axial Assessment of Eating Disorder Symptoms (MAEDS) and the Body Acceptability Morph (BAM) at months 3, 6, 12, 18 and 24 as safety measures. Participants who score 2 standard deviations (2 SDs is equal to a *t*-score of 70) above the mean on any subscale of the MAEDS will be administered the Interview the Diagnosis of Eating Disorders-Fourth Version (IDED-IV) to determine if eating disorder pathology is present. Each participant who has signs of an eating disorder will be interviewed by a clinical specialist with relevant experience.

The IDED-IV will also be administered if participants report body image disturbance on the BAM, defined as: 1) scoring 2 standard deviations or more above the mean on the Current Body Size scale of the BAM (i.e., 2 SDs is equal to a *t*-score of 70, and suggests that the participant views his/her body as larger than it actually is); 2) scoring 2 standard deviations or more below the mean on the Ideal Body Size scale (i.e., 2 SDs is equal to a *t*-score < 30, and suggests that the participant's perception of ideal body size is excessively small); and/or 3) confirming acceptability of the extreme ideal body size shown to the participant in the acceptability phase (phase II) of the measure.

If a participant meets the diagnostic criteria for anorexia nervosa, bulimia nervosa, or binge eating disorder (IDED-IV ratings for each of the diagnostic criteria are "3" or more) or experiencing a sub

threshold eating disorder (defined as IDED-IV ratings of “3” or more on 5 of the 8 combined symptoms for bulimia nervosa and anorexia nervosa), the CR intervention will be permanently discontinued and a participant will be advised to seek medical help outside of the study.

These objective safety measures will help the study team identify participants who are developing eating disorder pathology and provide an appropriate intervention or referral. Additionally, participants will be in close contact with study staff through attendance of the behavioral groups and individual sessions with the RD. These visits serve as a good tool to assure the safety of the participants, since abnormal behavior and excessive weight loss or gain can be monitored closely by study staff, who are well acquainted with the participants. Should a participant exhibit abnormal behavior or a dramatic weight change, the study staff will consult with the site’s physician, behavioral expert, and other study team members to determine the most appropriate action. These actions could range from minor dietary changes to referral for treatment for more serious behavioral concerns.

### **16.9 Monitoring for Excessive Weight Loss**

BMI will be calculated at screening and in the CR group at Months 1, 3, 6, 9, 12, 18, and 24. Any volunteer who has a BMI < 22 kg/m<sup>2</sup> at screening will not be eligible for the study, and the screening process will be terminated at that point.

Any participant enrolled in the CR group who has BMI <18.5 kg/m<sup>2</sup> at any point during the study will be advised about the risks of excessive weight loss and will be prescribed a diet plan with increased number of calories up to the baseline level for up to one month. If BMI is still <18.5 after one month of increased calorie intake, CR intervention will be permanently discontinued, and a participant will follow all other study procedures to the study end. The CR intervention will also be permanently discontinued if any participant in the CR group shows decrease in BMI <18.5 kg/m<sup>2</sup> after the CR was restarted. A participant will follow all other study procedures to the study end.

### **16.10 Depression and Other Mental / Behavioral Health Conditions Surveillance Protocol**

Beck Depression Inventory (BDI) questionnaire will be administered at screening, baseline, month 1, 3, 6, 9, 12, 18, and 24 visits. Any candidate who presents with a score of ≥20 at screening or baseline will not be eligible for the study and the screening process or baseline testing will be terminated at that point.

If a participant enrolled in the CR group has BDI score ≥ 30 suggestive of a severe depression at any point during the study, an interview will be conducted with the participant to determine whether the CR intervention needs to be permanently discontinued or whether the protocol for moderate depressions should be followed. Any participant enrolled in the CR group who has BDI score between ≥ 20 and < 30 suggestive of a moderate depression at any point during the study will have the questionnaire re-administered in one week. If a repeated test score is still suggestive of a moderate depression, the CR intervention will be temporarily discontinued and a participant will be advised to seek medical help outside of the study. The CR intervention will be restarted after the BDI score decreases to < 20, or if a qualified mental health professional treating participant’s depression indicates in writing that it is safe to restart the CR intervention.

If BDI score is still ≥ 20 after one month of treatment, or a qualified mental health professional indicates that it is unsafe to restart the CR, intervention will be permanently discontinued and a participant will follow all other study procedures to the study end. The CR intervention will also be permanently discontinued if any participant in the CR group shows a BDI score ≥ 20 after the CR was restarted. A participant will follow all other study procedures to the study end.

Other adverse mental/behavioral health conditions may also develop in the study. The study psychologist and/or the study physician will advise any participant who develops any of such conditions to seek medical help outside of the study.

### **16.11 Monitoring Bone Mineral Density**

Bone mineral density will be monitored by measuring BMD by DXA at the total hip and total spine (L1 – L4) at baseline, 6, 12, and 24 months. Any participant who experiences a decrease in BMD at the total hip or total spine of 10% or greater from the baseline at any time during the study, will have the scan

repeated within one month. If the second scan confirms original findings of a decrease in BMD at the total hip or spine of 10% or greater, the participant will have the intervention permanently discontinued. In addition, the BMD *t*-score will also be monitored, and any participant who has BMD *t*-score at the total hip or total spine of less than -2.5 at any time during the study will have the intervention permanently discontinued. A participant will be advised to seek medical help outside of the study and will follow all other study procedures to the study end.

Additionally, any participant who experiences a decrease in BMD at the total hip or total spine greater than or equal to 5% and less than 10% from baseline to Month 6 or baseline to Month 12 will have BMD measured by DXA at 18 months. If decrease in BMD from baseline to Month 18 is greater than or equal to 10%, the above discontinuation procedure will be followed.

### **16.12 Electrocardiogram**

A standard resting 12-lead ECG will be recorded at screening, baseline, month 1, 3, 6, 9, 12, 18, and 24 visits. The ECGs will be reviewed by the clinical site physician and cardiologist. At screening or baseline, any candidate who has any of the following ECG abnormalities will not be eligible for the study, and the screening process or baseline testing will be terminated at that point:

- Signs of hyperkalemia (for details please refer to Section 0 above)
- Type II second or third degree heart block
- Ventricular ischemia
- Left bundle branch block, cardiac hypertrophy by any criteria, QRS complex > 100 ms in duration
- Abnormal QTc interval, supraventricular tachycardia of any type not including APC's, or ventricular arrhythmia of any type (including VPC's more than 60 per minute)
- Exercise ECG recorded during the  $VO_{2max}$  test demonstrating any of the above mentioned abnormalities occurring with exercise (if ventricular ischemia is observed, a stress imaging study to exclude a false positive result will be recommended)

At months 1, 3, 6, 9, 12, 18, and 24 any of these abnormalities, except ventricular ischemia occurring with exercise, observed in any participant enrolled in the CR group will lead to permanent discontinuation of the CR intervention, and a participant will follow all other study procedures to the study end.

If ventricular ischemia is observed with the exercise, the CR intervention will be temporarily discontinued and a stress imaging study performed within two weeks. If a stress imaging study confirms presence of ventricular ischemia, the CR intervention will be permanently discontinued and a participant will follow all other study procedures to the study end.

### **16.13 Medical and Medication History**

A complete medical history performed during screening will include a review of all major organ systems and all medications taken during 30 days prior to screening, as well as reproductive status, contraceptive and menstrual history for women, where appropriate.

An abbreviated medical and medication history will be reviewed at baseline to ensure that the candidates did not develop any disease during the screening period. Use of an appropriate method of contraception by women will be verified at Month 1, 3, 6, 9, 12, 18, and 24 visits. Acceptable forms of contraception include:

- tubal ligation
- partial or complete hysterectomy
- oral contraceptive pills
- implanted progesterone "Implanon™"
- contraceptive vaginal ring "NuvaRing™"
- barrier method ± spermicide (condoms, contraceptive sponge, diaphragm)
- intrauterine device (IUD)
- spousal vasectomy
- abstinence

- natural family planning (NFP) when other contraceptive methods are prohibited due to religious reasons. Participant reason and competence / training and basal body temperature records must be documented.

### **16.14 Physical Examination**

A complete physical examination will be performed for each participant at screening, baseline, Months 12, and 24 visits.

### **16.15 Vital Signs**

In addition to blood pressure as described in Section 12.2.1 above, vital signs will include oral temperature, respiratory rate and pulse rate. They will be captured at every clinic visit, i.e., at baseline and at Months 1, 3, 6, 9, 12, 18 and 24.

### **16.16 Withdrawal from the Intervention**

This subsection presents criteria for temporary and permanent discontinuation of the CR intervention in a participant due to the safety concerns. In the event that participants develop short- or long-term illnesses, or undergo surgical procedures, that may cause their prescribed caloric intake to be inappropriate for some period of time, they will be advised to discontinue caloric restriction until the condition is resolved, but to continue study visits if possible. For example, these conditions may include severe infections and recovery from trauma or surgery.

Women who develop menstrual irregularities or acyclicity will be followed for one year. If, after one year the menstrual cycle is not returning to the baseline status, CR intervention will be permanently discontinued and a participant will be advised to seek medical help outside of the study.

Any participant who develops a condition that would have excluded him or her from participation at baseline, but for which the prescribed energy intake is not inappropriate, will continue on caloric restriction.

All participants who require permanent discontinuation of the CR intervention will follow all other study procedures to the study end, unless a participant refuses to continue in the study by withdrawing his/her consent.

#### **16.16.1 Criteria for Temporary Discontinuation of the CR Intervention**

Temporary discontinuation of the CR intervention is defined as cessation of the 25% CR regimen for up to 30 days. The CR intervention will be temporarily discontinued if any of the following is observed:

- Increase in potassium level > 5.5 mEq/L and <6.0 mEq/L at any point during the study confirmed by a second test repeated in one week
- Increase in potassium level of 6.1 mEq/L and above at any point during the study confirmed by a second test repeated within 48 hours if at this second measurement potassium level is elevated above 5.5 mEq/L)
- Treatment-resistant anemia (anemia that is not improving or is worsening after one month of treatment)
- Decrease in BMI <18.5 at any point during the study
- Moderate depression (BDI score  $\geq$ 20)
- Any disease or condition that requires temporary discontinuation of the CR intervention, including but are not limited to severe infections, recovery from trauma or surgery.

#### **16.16.2 Criteria for Permanent Discontinuation of the CR Intervention**

The CR intervention will be permanently discontinued and a participant advised to seek medical help outside of the study if any of the following is observed:

- Increase in potassium level > 5.5 mEq/L resistant to one month of treatment
- Rechallenge hyperkalemia (increase in potassium level of 5.5 mEq/L and above after the CR was temporary discontinued and then restarted)
- Persistent anemia (anemia that is not improving or worsening after temporary discontinuation of the CR intervention)

- Cancer
- Cardiovascular MACE (Major Adverse Clinical Event, including myocardial infarction, stroke, transient ischemic attack) confirmed by appropriate medical professional
- Eating or psychiatric disorder including severe depression
- Further decrease in BMI after temporary discontinuation of the CR intervention or repeated decrease in BMI <18.5 after the CR intervention was restarted
- Rechallenge moderate depression (reoccurrence of moderate depression after the CR intervention was restarted) or moderate depression that is not improving or is worsening after temporary discontinuation of the CR intervention
- Trauma requiring prolonged hospitalization or bed rest for more than one month
- Pregnancy
- Menstrual irregularities or acyclicity for more than one year
- A decrease in BMD at the total hip or total spine (L1 – L4) of 10% or greater from baseline to Month 6 or Month 12 confirmed by a second scan performed within one month of the initial scan.
- A decrease in BMD at the total hip or total spine from baseline to Month 6 or baseline to Month 12 greater than or equal to 5% and less than 10% together with a decrease of 10% or greater from baseline to Month 18.
- BMD *t*-score at the total hip, femoral neck or total spine (L1-L4) of less than –2.5 at any time during the study.

### **16.17 Safety Monitoring Procedures**

Safety of the participants during CALERIE Phase 2 study will be ensured by the clinical sites' physician-investigators, the Safety Committee and the DSMB. The Safety Committee will include physician-investigators from each clinical site, the NIA and the CC. For details on the DSMB role please refer to Section 20.4 below

At screening and baseline, the physician-investigators will perform medical and medications history, physical examination, review the results of ECGs, laboratory, DXA and other tests and confirm that the candidates meet medical eligibility criteria. Any candidate who meets any exclusion criterion due to a newly diagnosed disease or condition, ECG, DXA, laboratory or any other safety test abnormality will be informed about the findings and the screening process will be terminated at that point. Physician-investigator will not perform any extensive testing to confirm a diagnosis, but instead, will advise a candidate to seek medical help outside of the study.

Exclusion criteria provide a reasonable room for applying the best medical judgment by a physician-investigator when assessing eligibility of the participants. Where the participant eligibility is questionable either due to some disease/condition or an abnormal test result, and in order to ensure uniformity in applying exclusion criteria across the three clinical sites, the CALERIE Safety Committee will review medical information for such candidate provided by the site's physician-investigator and confirm candidate's eligibility.

To ensure that participants' safety is not compromised during the study, the physician-investigators will perform physical examination and review the results of ECGs, laboratory, DXA and other safety tests. Physician-investigator will identify any abnormalities, assess their clinical significance and determine what action is necessary. The actions may range from repeating a test or examination, to discontinuing the CR intervention, and, if indicated by the participant's condition or required by the study protocol, advising a participant to seek medical help outside of the study.

Physician-investigator will inform the Safety Committee about all serious and severe adverse events, all cases of elevated potassium levels, anemia, depression, newly diagnosed diseases that require treatment and other significant safety issues observed at their clinical sites. The Safety Committee will review this information on its periodic conference calls and determine whether and what corrective action is necessary.

The CC will prepare semi-annual summary safety reports presenting adverse events, laboratory and other safety test abnormalities. These reports will be reviewed by the Safety Committee that will



determine whether there are concerns about the participants' safety and whether the study-wide corrective action is necessary.

## 17. STATISTICAL CONSIDERATIONS

### 17.1 Outcome Measures

As described in detail in Section 12 above, a number of outcome measures have been identified of interest to CALERIE. In general, they relate to the following broad categories.

- Energy metabolism (e.g., core temperature, RMR)
- Cardiovascular risk factors (e.g., blood pressure, serum lipids and lipoproteins)
- Markers of inflammation (e.g., CRP, TNF $\alpha$ , ICAM-1)
- Transforming growth factors (e.g., TGF- $\beta$ )
- Glucose tolerance and insulin (e.g., serum glucose, insulin and C-peptide levels)
- Immune function (e.g., DTH, WBC differential, responses to vaccines)
- Endocrine response (e.g., sex hormones for men, IGF-1, PDGF-AB)
- Quality of Life measures, Psychological and cognitive function (e.g., Rand SF-36, Food Craving Inventory, Body Shape Questionnaire, Cambridge Neuropsychological Test Automated Battery)
- Physical activity (e.g., peak VO<sub>2</sub>, Physical Activity Recall, muscle strength and endurance)
- Muscle strength and endurance (e.g., knee flexors and knee extensors, grip strength)
- Clinic weight
- Body composition (e.g., subcutaneous and visceral abdominal adipose tissue)
- Bone mineral density and markers of bone turnover
- Nutrient intake (e.g., micro- and macro-nutrient composition of food intake)
- Oxidative stress
- Mitochondrial function
- Intracellular signaling

These outcomes are measured on a ratio-scale, so that techniques for continuous random variables will be applied.

### 17.2 Power Calculations

The overall aim of CALERIE Phase 2 is to test the effects of CR, maintained at a 25% reduction from baseline energy intake, over a period of 24 months. The study is designed as a multi-center, parallel-group, randomized controlled trial. A 2:1 allocation ratio in favor of the CR intervention will be applied in order to maximize the number of subjects receiving the intervention of greater scientific interest, and to make the study more attractive to this target population.

Given the nature of the intervention and duration of follow-up, recruitment is expected to present a significant challenge in this study. Moreover, as described in Section 7 above, an intensive intervention, with frequent participant contact, is envisaged for subjects assigned to the CR intervention. The study size is therefore limited by the feasibility of finding eligible subjects, enrolling them into the study, and maintaining their commitment for the two-year interval. On the basis of these considerations, enrolling 250 subjects over a period of 20-24 months is thought to be feasible with the resources available. Based on the experience in the Phase 1 studies, a drop-out rate of around 10% per year is expected, so that a sample of approximately 200 subjects is expected to complete the study.

Power calculations were therefore performed to determine the magnitude of the treatment effect that can be detected with this sample size and this allocation ratio. The type-I error rate was set to  $\alpha = 0.05$ , and to ensure that definitive results are forthcoming, the type-II error was set to  $\beta = 0.10$ , representing power at 90% to detect meaningful differences. Applying standard sample size procedures for Gaussian outcomes [246,247], the standardized effect (i.e., the between-group difference in means divided by the standard deviation) that can be detected is 0.4862.

Table 17.1 applies this effect to some of the primary and secondary outcomes identified in Section 2. One set of analyses was performed to derive the between-group difference that can be detected at a

single time point, e.g., at the end of the study. Another set was performed for the between-group difference in the change score, e.g., from baseline to the final time point. It can be shown that if the correlation between the baseline and follow-up time point is greater than 0.5, and the standard deviation at the two time points is the same, the change score analysis will have greater power. Standard deviations in all cases were derived from the Phase 1 studies.

**Table 17.1: Treatment differences that can be detected in the Phase 2 study for the primary and secondary outcomes for two types of analyses**

Outcome Variable	Single Time Point		Change Score Analysis	
	Std. Dev.	Difference	Std. Dev.	Difference
Core Temperature (°C)	0.32	0.16	0.36	0.17
RMR (kcal/day)	242.29	117.81	123.14	59.87
Serum Triiodothyronine (ng/dL)	22.02	10.71	20.68	10.06
TNF- $\alpha$ (pg/mL)	1.15	0.56	0.64	0.31

In all cases, the detectable differences are as small or smaller than the average differences that were actually observed in the phase 1 studies. Thus, if the same differences are observed in the Phase 2 study, there will be adequate power to declare these differences statistically significant. Inspection of Table 17.1 also reveals that the standard deviations of the changes scores are generally smaller than those at a single time point. This implies that the change scores analysis is expected to have greater power. By extension, including the baseline value of the outcome as a covariate in the regression models (as described in Section 17.7.1 below) is expected to result in slightly greater power again.

Power calculations were also performed for a number of the exploratory outcome variables (not shown). They indicate that there will be adequate power for many (but not all) of the exploratory variables. They include glucose, insulin, insulin sensitivity, HDL cholesterol, BMD lumbar spine and BMD pelvis, fat mass and fat-free mass, DTH and IGF-1.

Finally, as described 17.8 below, the main analytic strategy for this study is repeated measures analysis. This analysis includes data at all intermediate time points as well as partial information provided by subjects before they drop out. All things being equal, there should be greater power from this analysis, so that the detectable differences described above are somewhat conservative.

### 17.3 Analysis Policies

#### 17.3.1 Type-I Error

There is no *a priori* reason to assume that the effects of the CR intervention will be better than that in the control arm, and departures on both sides of the null hypothesis are of interest. Thus, two-sided tests of significance will be performed for all comparisons.

As described above, the effects of the CR intervention will be evaluated using a variety of outcome measures. The value of a calorie restricted diet will not be decided on the results from a single outcome variable. Rather, the combined weight of evidence across the different collections of outcomes variables will be considered. Thus, as discussed in Cook and Farewell [248] and Proschan and Waclawiw [249], there is no requirement to control the family-wise error rate, i.e., the error rate across the set of outcome measures. The global null hypothesis of no between-group difference across the set the outcome measures is not meaningful here. Moreover, this is the first detailed investigation of effects of CR over an extended period of time in humans. Many of these comparisons are exploratory in nature, and will need to be confirmed by follow-up studies. Type-II error, i.e., failing to detect a significant effect when it exists, is important to this research. Thus, all tests of significance for between-group comparisons will be performed at the  $\alpha = .05$  level of significance. Type-I error will be controlled, however, for any subgroup comparisons, and this is described in more detail below.

#### 17.3.2 Interim Analyses for Efficacy

The immediate health benefit of the CR intervention during the course of the intervention is expected to be relatively modest, and the health risk to control participants from not obtaining the CR intervention is thought to be minimal. Rather than terminating the study early for an apparent treatment effect, it is

judged to be more important to continue observation to evaluate the longer-term physiological effects of the CR intervention. Thus, no interim analyses for efficacy will be performed during this study.

This notwithstanding, the safety of all study participants will be monitored closely during the study, and a decision to terminate the study, or one of its treatment arms, on basis of safety considerations may be made as described in Section 16 above.

### **17.3.3 Reporting Format**

The CONSORT statement [250,251] provides a structured format for reporting results from randomized clinical trials. It has been endorsed by a number of medical journals including *JAMA*, *BMJ*, and the *Lancet*, and will be applied for reporting the results from these studies.

## **17.4 “Intention-to-Treat” Analysis Population**

Following ICH Guideline E9, an Intention-to-Treat (ITT) analysis [252] will be applied to address the specific hypotheses identified in this protocol. ITT principles are summarized as follows:

1. All participants randomized to an intervention are included in all analyses irrespective of protocol violations and intercurrent events arising post randomization (e.g., poor adherence, receiving off-study interventions, the occurrence of side effects and adverse reactions).
2. Subjects are analyzed according to the treatment arm to which they were originally assigned, irrespective of cross-over and drop-in to the alternate treatment arm. We note, however, that an excessive cross-over rate erodes the credibility of the study and attenuates statistical power to detect treatment differences under ITT procedures [253,254].
3. All subjects continue to be followed beyond withdrawal from the intervention, cross-over and any other intercurrent events until the scheduled end of participant follow-up [255].

Because randomization carries the expectation of creating treatment groups equivalent with respect to known and unknown prognostic factors, removing randomized subjects from the analysis, even for the best of intentions, runs the risk of tampering with this balance and introducing differential selection bias into the treatment comparisons.

Some interpretations of the ITT approach extend these principles to include protocol violations occurring at or before the point of randomization. Tsiatis [256], however, indicated that this is unnecessary provided that the following conditions are met: detection of the relevant violation can be made objectively, the probability of the protocol violation is the same across the treatment arms, and all subjects receive equal scrutiny for the protocol violation, i.e., no ascertainment bias. We will follow this approach in broad measure in this study. Specifically, the following considerations will be made.

- A randomized subject who is discovered to have failed one or more eligibility criteria will be removed from the study population and no further follow-up will be performed.
- A randomized subject who is assigned to one intervention but immediately starts the alternate intervention due to an administrative error will be analyzed according to the treatment actually received. We note, however, that this is different than the issue of cross-over to the alternate intervention after the intervention has begun.
- A randomized subject who fails to start his/her assigned intervention, however, will be subject to the ITT principles above. Because CALERIE is unblinded, the probability of this protocol violation may well be different across the two treatment arms.

It is recognized that a large number of “simple” administrative errors will erode the integrity of the study. Thus, the disposition of these protocol violations will be considered on a case-by-case basis by the CALERIE Steering Committee. A large number of such errors in CALERIE as a whole, or one site in particular, will prompt retraining on the process for randomizing subjects and initiating the intervention.

## **17.5 Causal Effects of Calorie Restriction**

The ITT analysis evaluates a treatment “regimen” that is centered around the nominal intervention but includes factors that are inherently linked to the intervention such as adherence, side-effects, withdrawal from the intervention and participant drop-out. The value of an ITT analysis in CALERIE is that it provides a conservative estimate, i.e., a “lower bound,” for the effects of calorie restriction. This notwithstanding, CALERIE is also interested in mechanistic questions concerning the effect of CR. This

analysis asks the following question: if we control for, or otherwise adjust for, these confounding influences, what is the direct physiological effect of calorie restriction?

To address these mechanistic questions, causal analysis [257-259] will also be applied – specifically, the Marginal Structural Mean (MSM) of Robins and colleagues [260,261]. Under this approach, a repeated measures model [262,263] similar to that described in Section 17.8 below is developed to relate the outcome to treatment effect, the time effect, stratification variables and baseline covariates, as well as the confounding variable. However, the analysis is a weighted analysis, with weights inversely proportional to the probability of observing the confounder profile. This is similar to the familiar Horvitz-Thompson estimator in sample survey methodology. Robins demonstrated that a weighted analysis yields a consistent estimator of the causal effects of the regression parameters provided that the model for deriving the weights is correctly specified. Matsuyami [264] and Yamaguchi and Ohashi [265,266] provide case studies on applying this technique in randomized clinical trials, and this approach will be followed in broad measure in the analysis.

Thus, the causal model will be used to address the following types of questions:

- What is the estimated treatment difference adjusting for imperfect participant adherence to the protocol?
- What is the estimated treatment difference adjusting for participant drop-out during the study?
- What is the estimated treatment difference adjusting for protocol violations during the study?

The specific details of this analysis will be laid out in the Statistical Analysis Plan.

## **17.6 Baseline Characteristics**

Descriptive statistics will be generated to describe the characteristics of the study population. For categorical variables, this includes the numbers and percentages falling into the various categories, together with bar charts and histograms to provide a visual depiction. Summary statistics for continuous variables will include the mean and standard deviation as well as the various percentiles of the distribution. This will be supplemented with box-and-whisker plots for graphical display.

Statistical comparisons will be performed to compare treatment arms with respect to demographic characteristics and important prognostic factors at baseline. For categorical variables, e.g., sex and race, standard techniques for categorical data [267] will be applied, including Fisher's exact test, Pearson  $\chi^2$  procedures, and Mantel-Haenszel procedures for ordinal data, e.g., income levels. For continuous variables, e.g., age and baseline BMI, non-parametric procedures, e.g., the Wilcoxon-Mann-Whitney test or the Kruskal-Wallis test, will be applied as appropriate. The baseline values of the outcome variables will be analyzed in a similar manner to determine whether any systematic differences between the treatment arms remained after randomization.

## **17.7 Model Building**

### ***17.7.1 Covariates in Linear Models***

The following strategy will be applied to determine which baseline covariates to include as covariates in the various linear models [268]. Because randomization is stratified by clinical site, and sex and BMI within clinical site, these variables will be included as fixed-effects in all regression models. Secondly, for a continuous outcome measure, the baseline value of that variable, centered about its sample mean, will be entered to adjust for any baseline imbalances and to increase precision. Finally, any variable for which there is a meaningful imbalance between the treatment arms at baseline may also be entered to adjust for the imbalance. The final set of covariates will be described in the Statistical Analysis Plan. This plan will be reviewed by the CALERIE Steering Committee before the blind is broken and the statistical analyses are begun.

### ***17.7.2 Preliminary Investigations to Verify Underlying Assumptions***

The outcome measures are continuous random variables, and as described below, will be analyzed using techniques appropriate for Gaussian outcomes. Standard techniques will be performed to ensure that the data are consistent with the underlying assumptions. In Phase 1 of CALERIE, simple transformations were generally successful in reducing skewness and stabilizing the variance. If necessary, these techniques will be explored in this study.

## 17.8 Group Comparisons for Outcome Variables

CALERIE investigators have articulated three important goals in analyzing the outcomes from this study. They include the following:

- characterize trends in the outcome measures over time;
- compare the two treatment arms with respect to their experience over follow-up;
- evaluate changes within the CR intervention arm, especially between the weight-loss phase and the weight-stabilization phase.

All the major outcomes are observed repeatedly at well-defined time points over participant follow-up. These observations are expected to be correlated, so that statistical methods for longitudinal and repeated measures analysis [269] will be applied. Assuming that the set of repeated measures follows (or, can be transformed to follow) the multivariate normal distribution, techniques for Gaussian data [270] will be followed using, for example, PROC MIXED in SAS. Preliminary analyses will be performed to identify a parsimonious covariance structure among the repeated measures. A variety of models will be explored including the AR(1) and ARMA(1,1) time series models, as well as the mixed effects model of Laird and Ware [271]. The likelihood ratio test and the Akaike Information criterion will be applied to identify this structure.

The model for the expected value will consist of fixed-effect terms for the overall mean, the treatment effect, the time effect, the treatment  $\times$  time interaction, plus the stratification variables and covariates described in Section 17.7.1 above. Initially, time will be treated as a categorical variable to avoid arbitrary assumptions on the nature of the trend over time; a test for linearity (or, other trends) will be performed if it is of specific interest. Specific hypotheses will be tested by deriving contrasts among the model parameters. The overall treatment effect will be assessed by testing the null hypothesis that averaged over time, there is no treatment difference. Supporting analyses will investigate between-group differences at specific time points, e.g., at Months 6, 12 and 24. If the overall treatment effect is significant, comparisons at each time point will be performed at the  $\alpha = .05$  level of significance. If it is not significant, between-group comparisons at each time point may still be performed; however, a Bonferroni correction will be applied to maintain control of type-I error rate for this family of comparisons. Similarly, contrasts will be defined among the model parameters to characterize changes within the CR intervention arm over time, especially between the weigh-loss and weight-stabilization phases.

RMR and TEE will be analyzed in relation to changes in body composition measures (e.g., FFM) over time, and a different model will be applied. Body composition is known to be correlated with RMR, and differences between the two treatment arms are expected. Thus, body composition must be considered as a time-dependent confounder. Robins [272] demonstrated that including a time-dependent confounder as a covariate in a regression model will give rise to inconsistent estimators of the parameters. The approach of Rochon [273,274] will therefore be applied for this analysis instead.

## 17.9 Effect Modifiers and Subgroup Analyses

There is strong scientific interest in observing whether the overall treatment effect is consistent across subpopulations of interest [275]. Power calculations described above were performed to detect an overall treatment difference, so that in terms of power, it will be difficult to observe a significant deviation in the treatment effect in a particular subpopulation [276]. Subgroup analyses have also been severely abused over the years [277,278], particularly when there is no overall treatment difference. Yusuf *et al.* [279] warned that dramatic subgroup effects should be greeted with “considerable skepticism.” In terms of bias, therefore, subgroup analyses will be considered exploratory in nature and will be used to generate hypotheses for future studies.

Subgroup effects will be evaluated using formal statistical procedures. The subgrouping variable and its interactions with the other terms will be added to the statistical models described above. Contrasts among the model parameters will be derived to assess whether the treatment effect is consistent across the levels of the subgrouping variable. If it is statistically significant, treatment comparisons will be performed at each level of the subgrouping variable at the  $\alpha = .05$  level of significance. If it is not significant, between-group comparisons within each level of the subgrouping variable may still be performed; however, a Bonferroni correction will be applied to maintain control of type-I error rate for this family of comparisons.

In CALERIE Phase 2, the primary subgrouping variables of interest are the demographic characteristics of the participants (e.g., age, sex and race) and the baseline body composition measures (e.g., BMI). Other potential effect modifiers may also be considered in secondary analyses.

### **17.10 Adherence, Drop-Out, Loss to Follow-up, and Other Mitigating Factors**

Time-to-event outcomes such as permanent withdrawal from the intervention, drop-out from the study, cross-over to the alternate intervention or death (if any) will be analyzed using the standard techniques for survival data [280]. Survival distributions will be estimated using the Kaplan-Meier method. The Cox proportional hazard model will be used to compare the survival curves across the treatments, controlling for baseline covariates. A check will be made for proportionality in the hazard function over follow-up time, and remedial action will be taken if required. If the exact time of the event is not observed, but is only known to have occurred at some point since the previous evaluation, survival techniques for interval-censored data [281] will be applied instead.

Adherence will be measured as the percent calorie restriction (%CR) actually achieved by the study participant at any time point. Methods for deriving the adherence score are described in detail in Section 14 above. %CR will be treated as a continuous random variable and analyzed using the same techniques are described in Section 17.8 above. The Phase 1 data suggested that %CR was skewed towards the lower values. In this case, appropriate transformation will be applied to make it consistent with Gaussian assumptions.

## **18. DATA MANAGEMENT PROCEDURES**

### **18.1 Hardware and Software Configuration**

All data forthcoming from the study will be stored in an Oracle database system. The application and database will be hosted on Solaris Unix servers at the CC. Data entry into Oracle will be performed using Clintrial. It provides a validated system and complete audit trail in compliance with ICH Guideline E6. SAS will be applied for the management of analysis data files and for performing statistical computations. S-Plus will be used to provide supplementary statistical computations as required.

Database and web servers will be secured by a firewall and through controlled physical access. Oracle has many security features to ensure that any staff member accessing the database has the proper authority to perform the functions s/he requests of the system. Within the secondary SAS databases, Unix group-access control maintains similar security. The Sun workstation login is secured by extensive user-password facilities under Unix. Access to databases will be controlled centrally by the CC through user passwords linked to appropriate privileges. This protects the data from unauthorized changes and inadvertent loss or damage.

Database back-up will be performed automatically every day, and standard DCRI policies and procedures will be applied to dictate tape rotation and retention practices. All disk drives that provide network services, and all user computers, will be protected using a virus scanning software. Standard DCRI policies will be applied to update these protection systems periodically through the study.

### **18.2 Centralized Data Management**

All data arising from the study will be forwarded to the Coordinating Center and standard data management activities for multi-center clinical trials [282] will be applied. No data management on collaborative data will be conducted at the clinical sites. This would duplicate the efforts at the CC and create ambiguity concerning the “official” data arising from the study.

In general, the study database will be designed to address the following issues:

- provide on-going, administrative reports on progress in conducting the study including recruitment, screening and enrollment summaries, and the progress of study participants through the different milestones of the protocol;
- provide an on-going summary of progress in compiling the study database and the overall quality of the accumulating data;
- provide on-going reports on quality assurance issues including protocol deviations at baseline and follow-up, as well as reliability studies from the central laboratories and reading centers as described in Section 15;

- provide on-going reports on participant adherence to the CR intervention and study procedures as described in Section 14;
- provide on-going reports on participant safety and adverse events as described in Section 16;
- provide on-going, descriptive information on the baseline characteristics of participants assigned to intervention as described in Section 17.6;
- provide formal statistical analyses at the end of the study on all the outcomes arising from this study as summarized in Section 12.

### **18.3 Sources of Data**

Data for the central database will come from a variety of sources and include the following.

- Basic clinical information and data obtained from outcome assessments performed at each site, i.e., screening and eligibility criteria, demographic information, vital signs and physical measures, metabolic measures, food intake, food pattern and physical activity assessments, the results of the physical examinations, adverse events, quality of life and cognitive function questionnaires, etc., will be recorded on paper case report forms (CRFs).
- Laboratory tests, i.e., serum lipids and lipoproteins, inflammatory markers, insulin and glucose, immune function, endocrine response, etc., will be performed by a central biochemistry laboratory. Electronic files will be transferred to the CC using a secure FTP server and merged into the main CALERIE database.
- Safety laboratory tests, i.e., potassium, hemoglobin, hematocrit, etc., will be performed by a second central biochemistry laboratory. Electronic files will be transferred to the CC using a secure FTP server and merged into the main CALERIE database.
- Urine samples from study participants will be forwarded to the central DLW reading center at Baylor University. Electronic files with the resulting DLW measures will be forwarded to the CC using a secure FTP server and merged into the main CALERIE database.
- DXA files will be forwarded to the central DXA reading center for reading and interpretation. Electronic files with the resulting body composition measures will be forwarded to the CC using a secure FTP server and merged into the main CALERIE database.
- Six-day food record data will be recorded on paper data forms and forwarded to a dietary coding center. There, they will be entered into a database using the Nutritional Data System (NDS). Electronic files with the nutrition analysis will be forwarded to the CC using a secure FTP server and merged into the main CALERIE database.

In all cases, data for a particular participant will be identified by the CALERIE ID number, the protocol time point and the date of the corresponding procedure so that they can be merged successfully in the CALERIE database.

### **18.4 Data Management Activities**

All data management activities and data quality at the DCRI will be guided by a comprehensive set of data management Standard Operating Procedures (SOPs). At the time of writing, there are 22 written procedures covering all areas starting with CRF receipt and entry, through database lock. The SOPs provide a detailed description of every staff member's roles and responsibilities in each step of the data management process, regulate data base development, describe internal audit procedures, data audit trail, data status reports, version control measures, and training requirements. In general, the following data management procedures will be applied.

1. Paper Case Report Forms (CRFs) will be designed specifically for this study. They will be designed to capture all the information required to address the reports and analyses described above.
2. Appropriate conventions for specific data fields will be defined according to the according to the DCRI's standard operating procedures (SOPs). They include, for example, conventions for text fields, numeric fields, Yes/No fields, date fields, checklists, and missing data.
3. Key fields, e.g., the participant's ID number, the protocol time point and date of the evaluation, will be recorded for each major component of the CRF.
4. Personnel at clinical sites will record the data mandated by the study on the CRFs. Every CRF page will be accompanied by detailed instructions to promote consistency and reliability across the

sites and over time. All CRFs will be completed according to the current Good Clinical Practices (GCP) guidelines [283,284]. Training on completing (and correcting) the CRFs as well as forwarding them to the CC will be included in the initial training session and any follow-up training sessions.

5. At periodic intervals, a copy of the CRF will be forwarded to the CC by parcel delivery service for data entry and processing.
6. As described above, a variety of supplementary material and procedures will be conducted with study participants, including blood and urine samples, DXA evaluations and the dietary recall. The resulting files and/or materials will be forwarded directly to central laboratories and reading centers for processing and interpretation.
7. A database will be created on the DCRI computer network specifically for this study. As described above, the database will be managed with Oracle using Clintrial for data entry.
8. Each record in the database is identified by the participant's ID, the protocol time point and a unique record identifier.
9. CRFs forwarded from the clinical sites will be entered into the study database. Double data entry, by two different operators at two separate occasions, will be performed [285-287] to ensure a high level of confidence in the data entered.
10. Similarly, a data dictionary will include information for each of the electronic data records. Each record type will be designed to integrate with the rest of the database. Key fields will specifically be identified for these records.
11. At periodic intervals, these electronic data files will be forwarded to the CC using a secure FTP server. From there, they will be merged into the master database according to the DCRI's SOPs.
12. A series of validation checks will be developed for the database. They will search for impossible and implausible values as well as logical inconsistencies across the different data fields; longitudinal checks will evaluate consistency in variables over time; other checks will search for digit preferences or peculiar or fabricated values. For any exception uncovered, a data clarification form (DCF) will be generated and forwarded to the clinical site for investigation and resolution. Corrections will be made on the DCF and returned to the CC for data entry.
13. If a correction is required to the original CRF prior to submission to the CC, it will be made according to GCP guidelines. That is, a single line will be drawn through the old value so that the original entry is still visible. The correct value will be written close to the field, and the correction initiated and dated by the CALERIE staff member at the site making the change.

## **18.5 Data Quality Control Procedures**

Several levels of database quality control on the accumulating database will be performed. The first two levels are the double data-entry process and the programmatic consistency checks and/or range checks as described above. The third level is a record or panel level of control. Programs will be written to identify suspected duplicate and blank or missing records and records not double-entered within and across database tables. Next, an independent auditing group will perform the fourth level of database quality control. These internal data quality and process compliance audits will be conducted to document the frequency of random errors and identify systematic deviations so that they can be corrected. Other periodic quality control checks will document the frequency of random entry errors and identify systematic and process errors.

All electronic transmissions from the laboratories and central reading centers will be tested and validated by the CC according to the DCRI's SOPs prior to incorporating data into the study data base. All data transmissions will be documented and files will be stored according to the DCRI's SOPs and work instructions. Routine quality control checks of the file headings, participant IDs, inspection of dates, etc., for all transferred files will be performed by DCRI's data management team upon receipt of the files. The DCRI will also perform regular database checks to ensure completeness of data

In addition, at periodic intervals during the study, the Coordinating Center will conduct on-site, monitoring visits, and provide source document verification. A 10% random sample of eligible participants will be selected, and their CRFs will be verified against source documents by the monitor. Error rates will be quantified and compared against established norms for multi-center clinical trials, e.g., 3



errors per 1,000 data fields or 0.3%. In all cases, remedial action will be taken as appropriate. Training and recertification will be made available to redress deficiencies and misunderstandings.

## **18.6 Data Management Reports**

A variety of progress reports [288] will be prepared during the course of a trial. Specifically, one report will document progress with the central database and include the following:

- numbers of CRFs expected and received at the various protocol milestones;
- descriptive statistics on the lag time in sending CRFs to the CC for data entry;
- numbers of overdue and delinquent CRFs;
- numbers of queries and descriptive statistics on the lag time in resolving data queries;
- numbers of overdue and delinquent queries;
- numbers and percents of permanently missing data fields;
- numbers of CRFs with no outstanding queries.

A similar report will document progress with transferring electronic data files from each of the laboratories and central reading centers to the CC, and include the following summaries:

- numbers of samples sent from the sites to the central facility;
- numbers of records received from the laboratory at the CC;
- descriptive statistics on the lag time in sending samples from the site to the central lab, and the lag time in sending the corresponding electronic record from the lab to the CC;
- numbers of overdue and delinquent electronic records according to acceptable windows.

## **19. PARTICIPANT RIGHTS AND CONFIDENTIALITY**

### **19.1 Confidentiality and HIPAA Considerations**

Participant confidentiality will be protected throughout the study. All participant data will be kept strictly confidential, and no subject-identifying information will be released to anyone outside the project. Confidentiality will be through several mechanisms. First, each participant will be assigned an anonymous study ID, which will then be used on all study forms. Secondly, any study forms and paper records that do contain participant information (e.g., address lists, phone lists) will be kept at the CALERIE clinical site in secured, locked areas. No participant identifiers will be placed on biological samples and other CALERIE materials forwarded to central labs and facilities. Only the study ID number and the date of the evaluation will be provided. Third, access to all participant data and information, including biological samples, will be restricted to authorized personnel. In the case of computerized data, this restricted access will be assured through user logon IDs and password protection.

At the Coordinating Center, only authorized personnel will have access to the data files containing study data. Security will be assured through user logon IDs, passwords and appropriate access privileges. Study participants will be identified only by their CALERIE ID number, and no personal identifying information, such as name, address, social security number, etc., will be entered into the Coordinating Center database. Any participant-specific data reported to the Steering Committee or will be identified only by the CALERIE ID number.

Finally, participants will not be identified by name in any reports or publications, nor will the data presented in such a way that the identity of individual participants can be inferred. Analysis files created for further study by the scientific community will have no participant identifiers. These data files will be created in accordance with the Ancillary Studies policy and Publication policy of CALERIE.

### **19.2 Institutional Review Boards**

Before initiating this study, the protocol, site-specific informed consent forms, HIPAA Authorization Form, recruitment materials, and other relevant information will be reviewed by a properly constituted Institutional Review Board (IRB) at each participating clinical site. A copy of the IRB approval notification and approved informed consent and HIPAA Authorization Form will be collected by the study monitor prior to site initiation and archived at the Coordinating Center. Any amendments to the protocol, other than simple administrative and typographical changes, will be approved by each IRB before they

are implemented. The sites will seek annual renewals of their IRB approvals in accordance with local procedures. A copy of the approved informed consent and approved renewal will be reviewed and collected by the study monitor during site visits and archived at the Coordinating Center.

### **19.3 Informed Consent Procedures**

All CALERIE participants will provide written informed consent to participate in the study before any study-related procedures are initiated. The consent will describe the study's aims and objectives, procedures and activities to be undertaken in the study, as well as a summary the potential risks and benefits of participating. The participant will also be informed that s/he has the option of declining participation in the study without penalty or prejudice. If participation is declined, no further contact with the volunteer will be made. Otherwise, informed consent will be undertaken in person with the participant. Informed consent (including the HIPAA authorization) will be documented on the IRB-approved consent form, and signed by the participant.

Two consents will be undertaken. First, because there is an extended screening phase to determine eligibility and many study procedures are performed during this process, the first consent will occur during the first screening visit as described in Section 9 above. The procedures to be undertaken, and the risks and benefits associated with these procedures will be explained to the volunteer. The second informed consent will occur after eligibility has been verified and the participant is ready begin the baseline evaluations. This consent will describe the interventions, the associated risk and benefits, and describe study procedures.

Following the International Conference on Harmonisation (ICH) Guideline for Good Clinical Practice (E6), each informed consent document will at a minimum contain the following elements as appropriate for that consenting process.

1. A statement that the study involves research, an explanation of the purposes of the research and the expected duration of the subject's participation, a description of the procedures to be followed, and identification of any procedures which are experimental.
2. A description of any reasonably foreseeable risks or discomforts to the subject.
3. A description of any benefits to the subject or to others which may reasonably be expected from the research.
4. A disclosure of appropriate alternative procedures or courses of treatment, if any, that might be advantageous to the subject.
5. A statement describing the extent, if any, to which confidentiality of records identifying the subject will be maintained and that notes the possibility that the records may be inspected by outside authorities.
6. For research involving more than minimal risk, an explanation as to whether any compensation and an explanation as to whether any medical treatments are available if injury occurs and, if so, what they consist of, or where further information may be obtained.
7. An explanation of whom to contact for answers to pertinent questions about the research and research subjects' rights, and whom to contact in the event of a research-related injury to the subject.
8. A statement that participation is voluntary, that refusal to participate will involve no penalty or loss of benefits to which the subject is otherwise entitled, and that the subject may discontinue participation at any time without penalty or loss of benefits to which the subject is otherwise entitled.

A copy of the each signed informed consent document will be given to the participant and the original copy will be retained in the investigator's file at the clinical site. All informed consent documents will be audited by the Coordinating Center during its periodic monitoring visits to the clinical sites.

A sample informed consent template is provided in Section 25 below. It may be modified according to the specific needs of the IRB at each participating clinical site.

## 20. STUDY ADMINISTRATION

### 20.1 NIH Cooperative Agreement Mechanism

The administrative and funding instrument used for this program is an NIH cooperative agreement (U01). This is an assistance mechanism (rather than an acquisition mechanism) in which substantial scientific and/or programmatic involvement by the National Institute on Aging (NIA) is anticipated during performance of the activity. Under this agreement, the NIA's purpose is to support and/or stimulate study activity by working jointly with the Principal Investigators in a partner role. However, the NIA does not assume direction, prime responsibility or a dominant role in the activity. As described below, this responsibility resides with the Steering Committee. Specific tasks and activities in carrying out the collaborative aspects will be shared among the Principal Investigators and the designated NIA Program Administrator as members of the Steering Committee.

The protocols and governance policies call for the continual submission of data centrally to the Coordinating Center for the collaborative database as described in Section 18 above. Upon study completion, Principal Investigators retain custody of and have primary rights to the site-specific and collaborative data, subject to Government rights of access consistent with current DHHS, PHS, and NIH policies. The NIA Project Scientist, on behalf of the NIH, will have the same access, privileges and responsibilities regarding the collaborative data as the other members of the Steering Committee.

Upon completion of the project, CALERIE will put its intervention materials and procedure manuals into the public domain and/or make them available to other investigators, according to the approved plan for making data and materials available to the scientific community and the NIA, for the conduct of research at no charge other than the costs of reproduction and distribution.

The NIA reserves the right to terminate or curtail the study (or an individual award) in the event of substantial shortfall in participant recruitment, follow-up, data reporting, quality control, or other major breach of the protocol; if human subject safety or ethical issues dictate a premature termination; or if there is failure to develop or implement mutually agreeable collaborative protocols.

Any disagreement that arise concerning scientific/programmatic matters (within the scope of the U01 award), between U01 awardees and the NIA may be brought to arbitration. An arbitration panel will be composed of three members: one selected by the Steering Committee (without NIA representatives voting) or by the individual U01 awardee in the event of an individual disagreement; a second member selected by the NIA; and, the third member selected by the two prior selected members. For U01 awardees, this special arbitration procedure will in no way affect the investigator's right to appeal an adverse action in accordance with PHS regulations at 42 CFR Part 50, Subpart D, and HHS regulations at 45 CFR Part 16.

### 20.2 Steering Committee

The Steering Committee is the main governing body of the project. It is composed of the Principal Investigators of the clinical centers, the Principal Investigator of the Coordinating Center, and the NIA Project Scientist. The clinical centers, the Coordinating Center and the NIA each have one vote on the Steering Committee. All decisions are determined by majority vote.

All major scientific decisions are determined by the Steering Committee. It assumes overall responsibility for the design and conduct of the trial. It appoints (and disbands) subcommittees as the need arises; designs, approves and implements the study protocol; oversees the development of the Manual of Procedures; monitors participant recruitment and treatment delivery; evaluates data collection and management; oversees quality assurance procedures; monitors participant safety; and, implements changes and enhancements to the study as required. It also has primary responsibility for facilitating the conduct of the study and reporting results.

### 20.3 Subcommittee Structure

In addition to the Steering Committee, a number of subcommittees will be formed to provide broad oversight to the study. Here, we provide a brief description of each subcommittee. An organizational chart is provided in Section 24 below.

Executive Subcommittee: Although the Steering Committee is the decision and policy-making body of the study, an Executive Subcommittee will consist of the chair of the Steering Committee, the NIA Pro-

ject Official and the Principal Investigator of the Coordinating Center. This group will address the day-to-day activities of the trial and provide overall direction for the Steering Committee.

Safety Monitoring Subcommittee: This committee will oversee safety concerns in study, and review periodic safety reports as described in Section 16 above

Quality Control Subcommittee: This committee will monitor the quality of the measures and evaluations performed in the study as described in Section 15 above. It will also provide oversight to the activities in the central labs and reading centers, including the DLW lab described in Section 13 above.

Intervention Subcommittee: This committee will monitor fidelity in delivering the interventions as described in Section 7 above, and to make recommendations for revising the protocol as appropriate.

Operation Managers/Clinical Coordinators Subcommittee: This committee will consist of the Operations Managers and clinical coordinators at the clinical sites, the study monitor and the Project Leader at the CC. These individuals will discuss issues and procedures for implementing the study, discuss problems and solutions, monitor recruitment; and makes recommendation on any aspect of the study to the Steering Committee.

Analysis & Publications Subcommittee: This committee will consider, suggest and approve analysis requests (e.g., for abstracts) during the study, and develop policies regarding publication issues.

Advance Clinical Endpoints (ACE) Committee: This committee is an external advisory committee whose purpose is to advise the CALERIE Steering Committee on opportunities for performing state-of-the-art mechanistic studies. Examples includes metabolomic analyses on the biological samples collected.

The Steering Committee may appoint other committees and working groups as required by the study.

## 20.4 Data & Safety Monitoring Board

NIH policy issued in June of 1998 requires a DSMB for "... multi-site clinical trials involving interventions that entail potential risk to the participants." The DSMB is a multidisciplinary group which serves in an advisory capacity to CALERIE [289-293]. Members are appointed by the Director of NIA, and include experts in human physiology, nutritional studies, clinical trial design, biostatistics, and research ethics. The chair of the Steering Committee, the PI of the CC, and the NIA Project Official serve as ex officio members.

The DSMB will approve the protocol before the study is initiated; monitor recruitment, retention and adherence; evaluate data completeness and data quality; and, ensure that participant safety is addressed adequately. It reports directly to the Director of NIA, and makes recommendations on all study activities including terminating the study for safety or operational reasons.

## 21. REFERENCES

1. Weindruch R, Walford RL. The retardation of aging and disease by dietary restriction. Springfield, IL: CC Thomas; 1988.
2. McCarter R, Masoro EJ, Yu BP. Does food restriction retard aging by reducing the metabolic rate? *Am J Physiol* 1985;248:E488-E90.
3. McCarter RJ, Palmer J. Energy metabolism and aging: a lifelong study of Fischer 344 rats. *Am J Physiol* 1992;263(3 Pt 1):E448-52.
4. Ballor DL. Effect of dietary restriction and/or exercise on 23-h metabolic rate and body composition in female rats. *J Appl Physiol* 1991;713:801-6.
5. Gonzales-Pacheco DM, Buss WC, Koehler KM, Woodside WF, Alpert SS. Energy restriction reduces metabolic rate in adult male fisher-344 rats. *J Nutr* 1993;123:90-7.
6. Lynn WS, Wallwork JC. Does food restriction retard aging by reducing metabolic rate? *J Nutr* 1992;1229:1917-8.
7. Kemnitz JW, Weindruch R, Roecker EB, Crawford K, Kaufman PL, Ershler WB. Dietary restriction of adult male rhesus monkeys: design, methodology, and preliminary findings from the first year of study. *J Gerontol* 1993;48:B17-B26.
8. Lane MA, Baer DJ, Tilmont EM, et al. Energy balance in rhesus monkeys (*Macaca mulatta*) subjected to long-term dietary restriction. *J Gerontol A Biol Sci Med Sci* 1995;505:B295-B302.

9. DeLany JP, Hansen BC, Bodkin NL, Hannah J, Bray GA. Long-term calorie restriction reduces energy expenditure in aging monkeys. *J Gerontol A Biol Sci Med Sci* 1999;541:B5-11; discussion B2-3.
10. Ramsey JJ, Roecker EB, Weindruch R, Kemnitz JW. Energy expenditure of adult male rhesus monkeys during the first 30 mo of dietary restriction. *Am J Physiol* 1997;2725 pt 1:E901-E7.
11. Keys A, Brozek J, Henschel A, Mickelson O, Taylor H. The biology of human starvation. Minneapolis: University of Minnesota Press; 1950.
12. Fricker J, Rozen R, Melchior JC, Apfelbaum M. Energy-metabolism adaptation in obese adults on a very-low-calorie diet. *Am J Clin Nutr* 1991;534:826-30.
13. Keesey RE. Physiological regulation of body weight and the issue of obesity. *Med Clin North Am* 1989;731:15-27.
14. Vansant G, Van Gaal L, Van Acker K, De Leeuw I. Short and long term effects of a very low calorie diet on resting metabolic rate and body composition. *Internatl J Obesity* 1989;13Suppl 2:87-9.
15. Astrup A, Gotzsche PC, van de Werken K, et al. Meta-analysis of resting metabolic rate in formerly obese subjects. *Am J Clin Nutr* 1999;696:1117-22.
16. Weyer C, Walford RL, Harper IT, et al. Energy metabolism after 2 y of energy restriction: the biosphere 2 experiment. *Am J Clin Nutr* 2000;724:946-53.
17. Weigle DS, Brunzell JD. Assessment of energy expenditure in ambulatory reduced-obese subjects by the techniques of weight stabilization and exogenous weight replacement. *Int J Obes* 1990;14(suppl 1):69-81.
18. Rising R, Keys A, Ravussin E, Bogardus C. Concomitant interindividual variation in body temperature and metabolic rate. *Am J Phys* 1992;263(4 Pt 1):E730-40.
19. Lane MA, Baer BJ, Rumpler WV, et al. Calorie restriction lowers body temperature in rhesus monkeys, consistent with a postulated anti-aging mechanism in rodents. *Proc Nat Acad Sci* 1996;93:4159-64.
20. Roth GS, Lane MA, Ingram DK, et al. Biomarkers of caloric restriction may predict longevity in humans. *Science* 2002;811.
21. Schriener SE, Linford NJ, Martin GM, et al. Extension of Murine Life Span by Overexpression of Catalase Targeted to Mitochondria. *Science* 2005;3085730:1909-11.
22. Van Remmen H, Ikeno Y, Hamilton M, et al. Life-long reduction in MnSOD activity results in increased DNA damage and higher incidence of cancer but does not accelerate aging. *Physiol Genom* 2003;161:29-37.
23. Stadtman ER. Cyclic oxidation and reduction of methionine residues of proteins in antioxidant defense and cellular regulation. *Arch Biochem Biophys* 2004;4231:2-5.
24. Asif M, Egan J, Vasani S, et al. An advanced glycation endproduct cross-link breaker can reverse age-related increases in myocardial stiffness. *Proc Natl Acad Sci USA* 2000;97:2809-13.
25. Morgan CR, Lazarow A. Immunoassay of insulin: two antibody system. *Diabetes* 1963;12:115-26.
26. Leeuwenburgh C, Rasmussen JE, Hsu FF, Mueller DM, Pennathur S, Heinecke JW. Mass spectrometric quantification of markers for protein oxidation by tyrosyl radical, copper, and hydroxyl radical in low density lipoprotein isolated from human atherosclerotic plaques. *J Biol Chem* 1997;272:3520-6.
27. Stadtman ER. Protein oxidation and aging. *Science* 1992;257:1220-4.
28. Leeuwenburgh C, Wagner P, Holloszy JO, Sohal RS, Heinecke JW. Caloric restriction attenuates dityrosine cross-linking of cardiac and skeletal muscle proteins in aging mice. *Arch Biochem Biophys* 1997;3461:74-80.
29. Leeuwenburgh C, Hansen PA, Holloszy JO, Heinecke JW. Oxidized amino acids in the urine of aging rats: potential markers for assessing oxidative stress in vivo. *Am J Physiol* 1999;276:R128-35.
30. Leeuwenburgh C, Hansen PA, Holloszy JO, Heinecke JW. Hydroxyl radical generation during exercise increases mitochondrial protein oxidation and levels of urinary dityrosine. *Free Radic Biol Med* 1999;27:186-92.
31. Drew B, Phaneuf S, Dirks A, et al. Effects of aging and caloric restriction on mitochondrial energy production in gastrocnemius muscle and heart. *Am J Physiol Regul Integr Comp Physiol* 2003;2842:R474-80.
32. Goldstein JL, Hobbs HH, Brown MS. Familial hypercholesterolemia. In: Scriver CR BA, Sly WS and Valle D, ed. *The Metabolic & Molecular Bases of Inherited Disease*. New York: McGraw-Hill; 2001:2863-913.
33. Reardon CA, Getz GS. Mouse models of atherosclerosis. *Curr Opin Lipidol* 2001;122:167-73.
34. Gotto AMJ, Grundy SM. Lowering LDL cholesterol: questions from recent meta-analyses and subset analyses of clinical trial Data Issues from the Interdisciplinary Council on Reducing the Risk for Coronary Heart Disease, ninth Council meeting. *Circulation* 1999;998:E1-7.
35. Assmann G, Nofer JR. Atheroprotective effects of high-density lipoproteins. *Annu Rev Med* 2003;54:321-41.
36. Katznel LI, Bleecker ER, Colman EG, Rogus EM, Sorkin JD, Goldberg AP. Effects of weight loss vs aerobic exercise training on risk factors for coronary disease in healthy, obese, middle-aged and older men. A randomized controlled trial. *JAMA* 1995;27424:1915-21.

37. Ross R, Dagnone D, Jones PJ, et al. Reduction in obesity and related comorbid conditions after diet-induced weight loss or exercise-induced weight loss in men. A randomized, controlled trial. *Ann Intern Med* 2000;133:92-103.
38. Verdery RB, Walford RL. Changes in plasma lipids and lipoproteins in humans during a 2-year period of dietary restriction in Biosphere 2. *Arch Intern Med* 1998;158:900-6.
39. Walford RL, Harris SB, Gunion MW. The calorically restricted low-fat nutrient-dense diet in Biosphere 2 significantly lowers blood glucose, total leukocyte count, cholesterol, and blood pressure in humans. *Proc Nat Acad Sci* 1992;89:11533-7.
40. Burt VL, Whelton P, Roccella EJ, et al. Prevalence of hypertension in the US adult population: results from the Third National Health and Nutrition Examination Survey, 1988-1991. *Hypertension* 1995;25:305-13.
41. Lane MA, Ingram DK, Roth GS. Calorie restriction in nonhuman primates: effects on diabetes and cardiovascular disease risk. *Toxicol Sci* 1999;522 Suppl:41-8.
42. Walford RL, Mock D, Verdery R, MacCallum T. Calorie restriction in biosphere 2: alterations in physiologic, hematologic, hormonal, and biochemical parameters in humans restricted for a 2-year period. *J Gerontol A Biol Sci Med Sci* 2002;57:6:B211-24.
43. Cohen HJ, Pieper CF, Harris T, Rao KM, Currie MS. The association of plasma IL-6 levels with functional disability in community-dwelling elderly. *J Gerontol A Biol Sci Med Sci* 1997;52:4:M201-8.
44. Ershler WB, Sun WH, Binkley N, et al. Interleukin-6 and aging: blood levels and mononuclear cell production increase with advancing age and in vitro production is modifiable by dietary restriction. *Lymphokine Cytokine Res* 1993;12:4:225-30.
45. Basta G, Schmidt AM, De Caterina R. Advanced glycation end products and vascular inflammation: implications for accelerated atherosclerosis in diabetes. *Cardiovasc Res* 2004;63:4:582-92.
46. Cefalu WT, Bell-Farrow AD, Wang ZQ, et al. Caloric restriction decreases age-dependent accumulation of the glycoxidation products, N epsilon-(carboxymethyl)lysine and pentosidine, in rat skin collagen. *J Gerontol A Biol Sci Med Sci* 1995;50:6:B337-41.
47. Novelli M, Masiello P, Bombara M, Bergamini E. Protein glycation in the aging male Sprague-Dawley rat: effects of antiaging diet restrictions. *J Gerontol A Biol Sci Med Sci* 1998;53:2:B94-101.
48. Assayag P, Charlemagne D, de Leiris J, et al. Senescent heart compared with pressure overload-induced hypertrophy. *Hypertension* 1997;29:1:15-21.
49. Benetos A, Laurent S, Hoeks AP, Boutouyrie PH, Safar ME. Arterial alterations with aging and high blood pressure. A noninvasive study of carotid and femoral arteries. *Arterioscler Thromb* 1993;13:90-7.
50. Gaballa MA, Jacob CT, Raya TE, Liu J, Simon B, Goldman S. Large artery remodeling during aging: biaxial passive and active stiffness. *Hypertension* 1998;32:4:437-43.
51. Michel JB, Salzmann JL, Cerol ML, et al. Myocardial effect of converting enzyme inhibition in hypertensive and normotensive rats. *Am J Med* 1988;84:1:12-21.
52. Basso N, Paglia N, Stella I, et al. Protective effect of the inhibition of the renin-angiotensin system on aging. *Regul Pept* 2005;128:2:247-52.
53. Brilla CG, Pick R, Tan L, Janicki JS, Weber KT. Remodeling of the rat right and left ventricles in experimental hypertension. *Circ Res* 1990;67:6:1355-64.
54. Brilla CG, Weber KT. Reactive and reparative myocardial fibrosis in arterial hypertension in the rat. *Cardiovasc Res* 1992;26:6:671-7.
55. Loffreda S, Yang SQ, Lin HZ, et al. Leptin regulates proinflammatory immune responses. *FASEB J* 1998;12:5:57-65.
56. Sader S, Nian M, Liu P. Leptin: a novel link between obesity, diabetes, cardiovascular risk, and ventricular hypertrophy. *Circulation* 2003;108:6:644-6.
57. Corman B, Duriez M, Poitevin P, et al. Aminoguanidine prevents age-related arterial stiffening and cardiac hypertrophy. *Proc Natl Acad Sci U S A* 1998;95:13:1301-6.
58. Li YM, Steffes M, Donnelly T, et al. Prevention of cardiovascular and renal pathology of aging by the advanced glycation inhibitor aminoguanidine. *Proc Natl Acad Sci U S A* 1996;93:13:3902-7.
59. Vlassara H. Protein glycation in the kidney: role in diabetes and aging. *Kidney Int* 1999;49:6:1795-804.
60. Pertynska-Marczewska M, Kiriakidis S, Wait R, Beech J, Feldmann M, Paleolog EM. Advanced glycation end products upregulate angiogenic and pro-inflammatory cytokine production in human monocyte/macrophages. *Cytokine* 2004;28:1:35-47.
61. Border WA, Ruoslahti E. Transforming growth factor-beta in disease: the dark side of tissue repair. *J Clin Invest* 1992;90:1:1-7.
62. Ignatz RA, Massague J. Transforming growth factor-beta stimulates the expression of fibronectin and collagen and their incorporation into the extracellular matrix. *J Biol Chem* 1986;261:10:4337-45.

63. Edwards DR, Murphy G, Reynolds JJ, et al. Transforming growth factor beta modulates the expression of collagenase and metalloproteinase inhibitor. *Embo J* 1987;6:1899-904.
64. Brooks WW, Conrad CH. Myocardial fibrosis in transforming growth factor beta(1)heterozygous mice. *J Mol Cell Cardiol* 2000;32:187-95.
65. Sivasubramanian N, Coker ML, Kurrelmeyer KM, et al. Left ventricular remodeling in transgenic mice with cardiac restricted overexpression of tumor necrosis factor. *Circulation* 2001;104:826-31.
66. Ruiz-Torres MP, Bosch RJ, O'Valle F, et al. Age-related increase in expression of TGF-beta1 in the rat kidney: relationship to morphologic changes. *J Am Soc Nephrol* 1998;9:782-91.
67. Kelley D. Effects of weight loss on glucose homeostasis in NIDDM. *Diabetes Review* 1995;3:366-77.
68. Bodkin NL, Ortmeier HK, Hansen BC. Long-term dietary restriction in older-aged rhesus monkeys: effects on insulin resistance. *J Gerontol A Biol Sci Med Sci* 1995;50B:142-7.
69. Cefalu WT, Wagner JD, Bell-Farrow AD, et al. Influence of caloric restriction on the development of atherosclerosis in nonhuman primates: progress to date. *Toxicol Sci* 1999;52(2 Suppl):49-55.
70. Gazdag AC, Sullivan S, Kemnitz JW, Cartee GD. Effect of long-term caloric restriction on GLUT4, phosphatidylinositol-3 kinase p85 subunit, and insulin receptor substrate-1 protein levels in rhesus monkey skeletal muscle. *Ann Intern Med* 2000;132:44-6; discussion B7-8.
71. Walford RL, Mock D, MacCallum T, Laseter JL. Physiologic changes in humans subjected to severe, selective calorie restriction for two years in biosphere 2: health, aging, and toxicological perspectives. *Toxicol Sci* 1999;52(2 Suppl):61-5.
72. Walford RL, Weber L, Panov S. Caloric restriction and aging as viewed from Biosphere 2. *Receptor* 1995;51:29-33.
73. Boden G. Free fatty acids-the link between obesity and insulin resistance. *Endocr Pract* 2001;7:44-51.
74. Forouhi NG, Jenkinson G, Thomas EL, et al. Relation of triglyceride stores in skeletal muscle cells to central obesity and insulin sensitivity in European and South Asian men. *Diabetologia* 1999;42:932-5.
75. Jacob S, Machann J, Rett K, et al. Association of increased intramyocellular lipid content with insulin resistance in lean nondiabetic offspring of type 2 diabetic subjects. *Diabetes* 1999;48:1113-9.
76. Perseghin G, Scifo P, De Cobelli F, et al. Intramyocellular triglyceride content is a determinant of in vivo insulin resistance in humans: a 1H-13C nuclear magnetic resonance spectroscopy assessment in offspring of type 2 diabetic parents. *Diabetes* 1999;48:1600-6.
77. Flier JS. Diabetes. The missing link with obesity? *Nature* 2001;409:292-3.
78. Lofgren P, van Harmelen V, Reynisdottir S, et al. Secretion of tumor necrosis factor-alpha shows a strong relationship to insulin-stimulated glucose transport in human adipose tissue. *Diabetes* 2000;49:688-92.
79. Simpson HL, Umpleby AM, Russell-Jones DL. Insulin-like growth factor-I and diabetes. A review. *Growth Horm IGF Res* 1998;8:83-95.
80. Shulman GI. Cellular mechanisms of insulin resistance. *J Clin Invest* 2000;106:171-6.
81. Buse MG, Robinson KA, Marshall BA, Hresko RC, Mueckler MM. Enhanced O-GlcNAc protein modification is associated with insulin resistance in GLUT1-overexpressing muscles. *Am J Physiol Endocrinol Metab* 2002;283:E241-50.
82. Tissenbaum HA, Guarente L. Model organisms as a guide to mammalian aging. *Dev Cell* 2002;2:9-19.
83. Siskind GW. Immunological aspects of aging: an overview. In: Schimke RT, ed. *Biological Mechanism in Aging: USDA, NIH*, 1980:455-467.
84. Miller RA. The aging immune system: primer and prospectus. *Science* 1996;273:70-4.
85. Miller RA. Aging and immune function In: *Fundamental Immunology*. In: Paul WE, ed: Lippincott-Raven, Philadelphia., 1999.
86. McElhaney JE. The unmet need in the elderly: designing new influenza vaccines for older adults. *Vaccine* 2005;23 Suppl 1:S10-25.
87. Haynes L. The effect of aging on cognate function and development of immune memory. *Curr Opin Immunol* 2005;17:476-9.
88. Christou NV, Meakins JL, Gordon J, et al. The delayed hypersensitivity response and host resistance in surgical patients. 20 years later. *Ann Surg* 1995;222:534-46; discussion 546-8.
89. Rodysill KJ, Hansen L, O'Leary JJ. Cutaneous-delayed hypersensitivity in nursing home and geriatric clinic patients. Implications for the tuberculin test. *JAGS* 1989;37:435-443.
90. Wayne SJ, Rhyne RL, Garry PJ, Goodwin JS. Cell-mediated immunity as a predictor of morbidity and mortality in subjects over 60. *J Gerontol* 1990;45:M45-8.
91. Roberts-Thomson IC, Whittingham S, Youngchaiyud U, Mackay IR. Ageing, immune response and mortality. *Lancet* 1974;2:368-70.

92. Cohn JR, Hohl CA, Buckley CE. The relationship between cutaneous cellular immune responsiveness and mortality in a nursing home population. *J. Am. Geriatr. Soc.* 1983;31:808-9.
93. Weindruch R. Immunogerontologic outcomes of dietary restriction started in adulthood. *Nutr Rev* 1995;53:66-71.
94. Fernandes G, Venkatraman JT, Turturro A, Attwood VG, Hart RW. Effect of food restriction on life span and immune functions in long-lived Fischer-344 x Brown Norway F1 rats. *J Clin Immunol* 1997;17:85-95.
95. Esquifino AI, Cano P, Jimenez V, Cutrera RA, Cardinali DP. Experimental allergic encephalomyelitis in male Lewis rats subjected to calorie restriction. *J Physiol Biochem* 2004;60:245-52.
96. Spaulding CC, Walford RL, Effros RB. Calorie restriction inhibits the age-related dysregulation of the cytokines TNF-alpha and IL-6 in C3B10RF1 mice. *Mech Ageing Dev* 1997;93:87-94.
97. Shibolet O, Alper R, Avraham Y, Berry EM, Ilan Y. Immunomodulation of experimental colitis via caloric restriction: role of Nk1.1+ T cells. *Clin Immunol* 2002;105:48-56.
98. venkatraman JT, Attwood, V>G, Turturro, A., Hart, R.W., Fernandes, G. Maintenance of virgin T cells and immune functions by food restriction during aging in long-lived B6D2F1 female mice. *Aging: Immunol. Infect. Dis.* 1994;5:13-26.
99. Venkatraman J, Fernandes G. Modulation of age-related alterations in membrane composition and receptor-associated immune functions by food restriction in Fischer 344 rats. *Mech. Ageing Dev.* 1992;63:27-44.
100. Meydani SN, Lipman R, Blumberg JB, Taylor A. Dietary energy restriction decreases ex vivo spleen prostaglandin E<sub>2</sub> synthesis in Emory mice. *J. Nutr.* 1990;120:112-115.
101. Kim MJ, Aiken JM, Havighurst T, Hollander J, Ripple MO, Weindruch R. Adult-onset energy restriction of rhesus monkeys attenuates oxidative stress-induced cytokine expression by peripheral blood mononuclear cells. *J Nutr* 1997;127:2293-301.
102. Weindruch R, Sohal RS. Seminars in medicine of the Beth Israel Deaconess Medical Center. Caloric intake and aging. *N Engl J Med* 1997;337:986-94.
103. Nikolich-Zugich J, Messaoudi I. Mice and flies and monkeys too: caloric restriction rejuvenates the aging immune system of non-human primates. *Exp Gerontol* 2005;40:884-93.
104. Santos MS, Lichtenstein AH, Leka LS, Goldin B, Schaefer EJ, Meydani SN. Immunological effects of low-fat diets with and without weight loss. *J Am Coll Nutr* 2003;22:174-82.
105. Nelson JF. Neuroendocrine involvement in the retardation of aging by dietary restriction: A hypothesis. In: Modulation of aging processes by dietary restriction. Boca Raton: CRC Press; 1994.
106. Brown-Borg HM, Borg KE, Meliska CJ, Bartke A. Dwarf mice and the ageing process. *Nature* 1996;384:33.
107. Bartke A, Wright JC, Mattison JA, Ingram DK, Miller RA, Roth GS. Extending the lifespan of long-lived mice. *Nature* 2001;4146862:412.
108. Flurkey K, Papaconstantinou J, Miller RA, Harrison DE. Lifespan extension and delayed immune and collagen aging in mutant mice with defects in growth hormone production. *PNAS* 2001;986736-6741.
109. Coschigano KT, Holland AN, Riders ME, List EO, Flyvbjerg A, Kopchick JJ. Deletion, but not antagonism, of the mouse growth hormone receptor results in severely decreased body weights, insulin, and insulin-like growth factor I levels and increased life span. *Endocrinology* 2003;1449:3199-810.
110. Zhou Y, Xu BC, Maheshwari HG, et al. A mammalian model for Laron syndrome produced by targeted disruption of the mouse growth hormone receptor/binding protein gene (the Laron mouse). *Proceedings of the National Academy of Sciences of the United States of America* 1997;9424:13215-20.
111. Longo VD, Finch CE. Evolutionary medicine: from dwarf model systems to healthy centenarians? *Science* 2003;2995611:1342-6.
112. Miller RA, Chang Y, Galecki AT, Al-Regaiey K, Kopchick JJ, Bartke A. Gene expression patterns in calorically restricted mice: partial overlap with long-lived mutant mice. *Mol Endocrinol* 2002;1611:2657-66.
113. Tatar M, Kopelman A, Epstein D, Tu MP, Yin CM, Garofalo RS. A mutant Drosophila insulin receptor homolog that extends life-span and impairs neuroendocrine function. *Science* 2001;2925514:107-10.
114. Holzenberger M, Dupont J, Ducos B, et al. IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice. *Nature* 2003;4216919:182-7.
115. Kenyon C. A conserved regulatory system for aging. *Cell* 2001;1052:165-8.
116. Kenyon C, Chang J, Gensch E, Rudner A, Tabtiang R. A C. elegans mutant that lives twice as long as wild type. *Nature* 1993;366:461-4.
117. Hursting SD, Lavigne JA, Berrigan D, Perkins SN, Barrett JC. Calorie restriction, aging, and cancer prevention: mechanisms of action and applicability to humans. *Annu Rev Med* 2003;54:131-52.
118. Grimberg A, Cohen P. Role of insulin-like growth factors and their binding proteins in growth control and carcinogenesis. *J Cell Physiol* 2000;1831:1-9.



119. Khandwala HM, McCutcheon IE, Flyvbjerg A, Friend KE. The effects of insulin-like growth factors on tumorigenesis and neoplastic growth. *Endocr Rev* 2000;213:215-44.
120. Yu H, Rohan T. Role of the insulin-like growth factor family in cancer development and progression. *J Natl Cancer Inst* 2000;9218:1472-89.
121. Straus DS. Nutritional regulation of hormones and growth factors that control mammalian growth. *FASEB J* 1994;8:6-12.
122. Thissen J-P, Ketelslegers J-M, Underwood LE. Nutritional regulation of the insulin-like growth factors. *Endocr Rev* 1994;15:80-101.
123. Furstenberger G, Senn HJ. Insulin-like growth factors and cancer. *Lancet Oncol* 2002;35:298-302.
124. Giovannucci E. Insulin, insulin-like growth factors and colon cancer: a review of the evidence. *J Nutr* 2001;131(11 Suppl):3109S-20S.
125. O'Brien MF, Watson RW, Fitzpatrick JM. Insulin-like growth factor I and prostate cancer. *Urology* 2001;581:1-7.
126. Shi R, Berkel HJ, Yu H. Insulin-like growth factor-I and prostate cancer: a meta-analysis. *Br J Cancer* 2001;857:991-6.
127. Brown-Borg HM, Johnson WT, Roakoczy S, Romanick M. Mitochondrial oxidant generation and oxidative damage in ames dwarf and GH transgenic mice. *J Amer Aging Assoc* 2001;24:85-100.
128. Forster MJ, Sohal BH, Sohal RS. Reversible effects of long-term caloric restriction on protein oxidative damage. *J Gerontol Med Sci* 2000;5511:B522-9.
129. Sohal RS, Weindruch R. Oxidative stress, caloric restriction, and aging. *Science* 1996;2735271:59-63.
130. Weindruch R, Sohal RS. Seminars in medicine of the Beth Israel Deaconess Medical Center. Caloric intake and aging. *N Engl J Med* 1997;33714:986-94.
131. Dixit V, Mak TW. NF-kappaB signaling. Many roads lead to madrid. *Cell* 2002;1115:615-9.
132. Li N, Karin M. Is NF-kappaB the sensor of oxidative stress? *FASEB J* 1999;1310:1137-43.
133. Sanz A, Gredilla R, Pamplona R, et al. Effect of insulin and growth hormone on rat heart and liver oxidative stress in control and caloric restricted animals. *Biogerontology* 2005;61:15-26.
134. Yang S, Chintapalli J, Sodagum L, et al. Activated IGF-1R Inhibits Hyperglycemia-Induced DNA Damage and Promotes DNA Repair by Homologous Recombination. *Am J Physiol Renal Physiol* 2005.
135. Han ES, Evans TR, Shu JH, Lee S, James F, Nelson JF. Food restriction enhances endogenous and corticotropin-induced plasma elevations of free but not total corticosterone throughout life in rats. *J Gerontol A Biol Sci Med Sci* 2001;569:B391-7.
136. Matsuzaki J, Kuwamura M, Yamaji R, Inui H, Nakano Y. Inflammatory responses to lipopolysaccharide are suppressed in 40% energy-restricted mice. *J Nutr* 2001;1318:2139-44.
137. Belanger A, Candas B, Dupont A, et al. Changes in serum concentrations of conjugated and unconjugated steroids in 40- to 80-year-old men. *J Clin Endocrinol Metab* 1994;79:1086-90.
138. Orentreich N, Brind JL, Rizer RL, Vogelmann JH. Age changes and sex differences in serum dehydroepiandrosterone sulfate concentrations throughout adulthood. *J Clin Endocrinol Metab* 1984;59:551-5.
139. Orentreich N, Brind JL, Vogelmann JH, Andres R, Baldwin H. Long-term longitudinal measurements of plasma dehydroepiandrosterone sulfate in normal men. *J Clin Endocrinol Metab* 1992; 75:1002-4. *J Clin Endocrinol Metab* 1992;75:1002-4.
140. Lane MA, Ingram DK, Ball SS, Roth GS. Dehydroepiandrosterone sulfate: a biomarker of primate aging slowed by calorie restriction. *J Clin Endocrinol Metab* 1997;2093-6.
141. Morrow LA, Linares OA, Hill TJ, et al. Age differences in the plasma clearance mechanisms for epinephrine and norepinephrine in humans. *J Clin Endocrinol Metab* 1987;653:508-11.
142. Veith RC, Featherstone JA, Linares OA, Halter JB. Age differences in plasma norepinephrine kinetics in humans. *J Gerontol* 1986;413:319-24.
143. Emdin M, Gastaldelli A, Muscelli E, et al. Hyperinsulinemia and autonomic nervous system dysfunction in obesity: effects of weight loss. *Circulation* 2001;1034:513-9.
144. Shamsuzzaman AS, Winnicki M, Wolk R, et al. Independent association between plasma leptin and C-reactive protein in healthy humans. *Circulation* 2004;109:2181-5.
145. Singhal A, Farooqi IS, Cole TJ, et al. Influence of leptin on arterial distensibility: a novel link between obesity and cardiovascular disease? *Circulation* 2002;106:1919-24.
146. Yamauchi T, Kamon J, Waki H, et al. The fat-derived hormone adiponectin reverses insulin resistance associated with both lipoatrophy and obesity. *Nat Med* 2001;7:941-6.
147. Berg AH, Combs TP, Du X, Brownlee M, Scherer PE. The adipocyte-secreted protein Acrp30 enhances hepatic insulin action. *Nat Med* 2001;7:947-53.
148. Fontaine KR, Cheskin LJ, Barofsky I. Health-related quality of life in obese persons seeking treatment. *J Fam Pract* 1996;433:265-70.

149. Fine JT, Colditz GA, Coakley EH, et al. A prospective study of weight change and health-related quality of life in women. *JAMA* 1999;282:2136-42.
150. Green MW, Elliman NA, Rogers PJ. Impaired cognitive processing in dieters: Failure of attention focus or resource capacity limitations? *Br J Health Psychol* 1997;2:259-67.
151. Green MW, Rogers PJ. Impaired cognitive functioning during spontaneous dieting. *Psychol Med* 1995;25:1003-10.
152. Kretsch MJ, Green MW, Fong AKH, Elliman NA, Johnson HL. Cognitive effects of a long-term weight reducing diet. *Int J Obesity* 1997;21:14-21.
153. Green MW, Rogers PJ. Selective attention to food and body shape words in dieters and restrained nondieters. *Int J Eat Disord* 1993;14:515-7.
154. Placanica JL, Faunce GJ, Soames Job RF. The effect of fasting on attentional biases for food and body shape/weight words in high and low Eating Disorder Inventory scorers. *Int J Eat Disord* 2002;32:79-90.
155. Berry SL, Beatty WW, Klesges RC. Sensory and social influences on ice cream consumption by males and females in a laboratory setting. *Appetite* 1985;6:41-5.
156. Landers DM, Arent SM, Lutz RS. Affect and cognitive performance in high school wrestlers undergoing rapid weight loss. *Journal of Sport & Exercise Psychology* 2001;23:307-16.
157. Dobson KS, Dozois DJ. Attentional biases in eating disorders: a meta-analytic review of Stroop performance. *Clin Psychol Rev* 2004;23:1001-22.
158. Spindler SR. Calorie restriction enhances the expression of key metabolic enzymes associated with protein renewal during aging. *Ann N Y Acad Sci*. 2001;928:296-304.
159. Mattison JA, Lane MA, Roth GS, Ingram DK. Calorie restriction in rhesus monkeys. *Exp Gerontol*. 2003;38:35-46.
160. Bordone L, Guarente L. Calorie restriction, SIRT1 and metabolism: understanding longevity. *Nat Rev Mol Cell Biol*. 2005;6:298-305.
161. Lee CK, Klopp RG, Weindruch R, Prolla TA. Gene expression profile of aging and its retardation by caloric restriction. *Science* 1999;285:1390-3.
162. Masoro EJ, Yu BP, Bertrand HA. Action of food restriction in delaying the aging process. *Proc Natl Acad Sci* 1982;79:4239-4241.
163. Yu BP, Masoro EJ, Mc Mahan CA. Nutritional influences on aging of Fischer 344 rats: physical, metabolic and longevity characteristics. *J Gerontol* 1985;40:657-670.
164. Lambert J, Lamonthe JM, Zernicke RF, Auer RN, Reimer RA. Dietary restriction does not adversely affect bone geometry and mechanics in rapidly growing male Wistar rats. *Pediatr Res*. 2005; 57:227-31.
165. Knowler WC, Barrett-Connor E, Fowler SE, et al. for the Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med*. 2002;346:393-403.
166. Garner DM, Wooley SC. Confronting the failure of behavioral and dietary treatments for obesity. *Clin Psychol Rev* 1991;11:729-80.
167. Bacon L, Stern JS, Van Loan MD, Keim NL. Size acceptance and intuitive eating improve health for obese, female chronic dieters. *J Am Diet Assoc*. 2005;105:929-36
168. Wadden TA, Butryn ML, Byrne KJ. Efficacy of lifestyle modification for long-term weight control. *Obes Res*. 2004;12 Suppl:151S-62S.
169. Williamson DA, Stewart TM. Behavior and lifestyle: approaches to treatment of obesity. *J La State Med Soc*. 2005;157:Spec No 1:S50-5.
170. Howarth, NC, Saltzman, E, Roberts, SB. Dietary Fiber and Weight Regulation. *Nutr Revs* 2001;59:129-139.
171. Aitken L, Gallagher R, Madronio C. Principles of recruitment and retention in clinical trials. *Int J Nurs Pract*. 2003;9:338-46.
172. Shimokawa I, Higami Y, Yu BP, Masoro EJ, Ikeda T. Influence of dietary components on occurrence of and mortality due to neoplasms in male F344 rats. *Aging (Milano)* 1996;8:254-62.
173. Vasselli JR, Weindruch R, Heymsfield SB, Pi-Sunyer FX, Boozer CN, Yi N, Wang C, Pietrobelli A, Allison DB. Intentional weight loss reduces mortality rate in a rodent model of dietary obesity. *Obes Res*. 2005;13:693-702.
174. *Dietary Reference Intakes: Applications in Dietary Planning* (2003). Food and Nutrition Board (FNB), Institute of Medicine (IOM). The National Academies Press.
175. Jackson M, Berman N, Huber M, Snetselaar L, Granek I, Boe K, Milas C, Spivak J, Chlebowski RT. Research staff turnover and participant adherence in the Women's Health Initiative. *Controlled Clin Trials*. 2003;24:422-35.
176. DeLany JP, Hansen BC, Bodkin NL, Hannah J, Bray GA. Long-term calorie restriction reduces energy expenditure in aging monkeys. *J Gerontol Biol Sci*. 1999;54A:B5-B11.

177. Ramsey JJ, Rocker EB, Weindruch R, Kemnitz JW. Energy expenditure of adult male rhesus monkeys during the first 30 months of dietary restriction. *Am J Physiol.* 1997;272:E901–E907.
178. Schoeller DA. Measurement of energy expenditure in freelifing human beings by using doubly labeled water. *J Nutr* 1988;118:1278–89.
179. Jeffery RW, Wing RR, Thorson C, et al. Strengthening behavioral interventions for weight loss: a randomized trial of food provision and monetary incentives. *J Consult Clin Psychol* 1993;61:1038–45.
180. Wing RR, Jeffery RW, Burton LR, Thorson C, Nissinoff KS, Baxter JE. Food provision vs. structured meal plans in the behavioral treatment of obesity. *Int J Obes Relat Metab Disord* 1996;20:56–62.
181. Perri MG, Corsica JA. Improving the maintenance of weight lost in behavioral treatment of obesity. In: Wadden TA, Stunkard AJ, eds. *Handbook of Obesity Treatment*. New York: Guilford Press; 2002, pp. 357–79.
182. Perri MG, Nezu AM, McKelvey WF, Shermer RL, Renjilian DA, Viegner BJ. Relapse prevention training and problem-solving therapy in the long-term management of obesity. *J Consult Clin Psychol.* 2001;69:722-6.
183. Moreno-John G, Gachie A, Fleming CM, Napoles-Springer A, Mutran E, Manson SM, Perez-Stable EJ. Ethnic minority older adults participating in clinical research: developing trust. *J Aging Health* 2004;16(5 Suppl):93S-123S.
184. Oddone EZ, Olsen MK, Lindquist JH, Orr M, Horner R, Reda D, Lavori P, Johnson G, Collins J, Feussner JR. Enrollment in clinical trials according to patients race: experience from the VA Cooperative Studies Program (1975-2000). *Controlled Clin Trials* 2004 ;25(4):378-87.
185. Adams-Campbell LL, Ahaghotu C, Gaskins M, Dawkins FW, Smoot D, Polk OD, Gooding R, DeWitty RL. Enrollment of African Americans onto clinical treatment trials: study design barriers. *J Clin Oncol* 2004 15;22:730-4.
186. Fouad MN, Corbie-Smith G, Curb D, Howard BV, Mouton C, Simon M, Talavera G, Thompson J, Wang CY, White C, Young R. Special populations recruitment for the Women's Health Initiative: successes and limitations. *Controlled Clin Trials* 2004;25(4):335-52.
187. Stunkard, A. J., Messick, S. (1985). The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. *J Psychosomatic Res* 29:71-83.
188. Anderson D.A, Williamson D.A, Duchmann E.G, Gleaves D.H, & Barbin J.M. (1999). Development and validation of a multifactorial treatment outcome measure for eating disorders. *Assessment* 6:7-20.
189. First, MB, Spitzer, RL, Gibbon M, Williams, JBW: *Structured Clinical Interview for DSM-IV Personality Disorders, (SCID-II)*. Washington, D.C.: American Psychiatric Press, Inc., 1997.
190. Beck AT, Beamesderfer A. Assessment of depression: the depression inventory. *Mod Prob Pharmacopsych* 1974;7:151-169.
191. Stewart, T.M., Williamson, D.A., Smeets, M.A.M., and Greenway, F.L. A Computerized Assessment of Body Image: A Psychometric Study. *Obesity Research* 2000; 9:43-50.
192. Williamson, D.A., (1990) *Assessment of Eating Disorders*. New York: Pergamon.
193. Taylor-Piliae RE, Norton LC, Haskell WL, et al. Validity of a new brief physical activity survey among older adults. *Circulation* 2005;111.
194. Lachin JM. Properties of randomization in clinical trials. *Controlled Clin Trials* 1988; 9: 289-311.
195. Matts MP and Lachin JM. Properties of permuted-block randomization in clinical trials. *Controlled Clin Trials* 1988; 9: 327-44.
196. Lachin JM, Matts JP and Wei LJ. Randomization in clinical trials: Summary and Recommendations *Controlled Clin Trials* 1988; 9: 365-74.
197. O'Connor PJ. Evaluation of four highly cited energy and fatigue mood measures. *J of Psychosomatic Res* 2004;57:435-41.
198. Klem ML, Wing RR, Simkin-Silverman L, Kuller LH. The psychological consequences of weight gain prevention in healthy, premenopausal women. *Int J Eat Disord* 1997;21:167-74.
199. Bundlie SR. Sleep in aging. [Review] *Geriatrics* 1998;53 Suppl 1:S41-3.
200. Steiger A. Sleep and endocrinology. [Review] *J Intern Med* 2003;254:13-22.
201. Bandura A. Self-efficacy: toward a unifying theory of behavior change. *Psychological Review.* 1977;84:191-215.
202. Clark MM, Abrams DB, Niaura RS, Eaton CA, Rossi JS. Self-efficacy in weight management. *Journal of Consulting and Clinical Psychology.* 1991;59:739-744.
203. Bernier M, Avard J. Self-efficacy, outcome, and attrition in a weight-reduction program. *Cognitive Therapy and Research.* 1986;10:319-338.
204. Martin PD, Dutton GR, Brantley PJ. Self-efficacy as a predictor of weight change in African-American women. *Obesity Research.* 2004;12:646-651.
205. Blair SN, Kohl HW III, Paffenbarger RS Jr, Clark DG, Cooper KH, Gibbons LW. Physical fitness and all-cause mortality: a prospective study of healthy men and women. *JAMA* 1989;262:2395-401.

206. Gulati M, Black HR, Shaw LJ, et al. The prognostic value of a nomogram for exercise capacity in women. *N Engl J Med* 2005;353: 468-75.
207. Blair SN, Kampert JB, Kohl HW III, et al. Influences of cardiorespiratory fitness and other precursors on cardiovascular disease and all-cause mortality in men and women. *JAMA* 1996;276:205-10.
208. Kraus WE, Douglas PS. Where does fitness fit in [editorial]? *N Engl J Med* 2005; 353:517-9.
209. Glickman SG, Marn CS, Supiano MA, Dengel DR. Validity and reliability of dual-energy X-ray absorptiometry for the assessment of abdominal adiposity. *J Appl Physiol* 2004; 97:509-14.
210. Brownbill RA, Ilich JZ. Measuring body composition in overweight individuals by dual energy x-ray absorptiometry. *BMC Med Imaging* 2005; 5:1.
211. Ricci, TA, Heymsfield, SB, Pierson, RN, jr., Stahl, T, Chowdhury, HA, Shapses, SA. Moderate energy restriction increases bone resorption in obese post-menopausal women. *Am. J. Clin. Nutr.* 73: 187-197, 2001.
212. Bandini LG, Schoeller DA, Cyr HN, Dietz WH. Validity of reported energy intake in obese and nonobese adolescents. *Am J Clin Nutr* 1990;52:421-425.
213. Schoeller DA, Bandini LG, Dietz WH. Inaccuracies in self-reported intake identified by comparison with the doubly labelled water method. *Can J Physiol Pharmacol* 1990;68:941-949.
214. Livingstone MBE, Prentice AM, Coward WA, Strain JJ, Black AE, Davies PSW, Stewart CM, McKenna PG, Whitehouse RH. Validation of estimates of energy intake by weighed dietary record and diet history in children and adolescents. *Am J Clin Nutr* 1992;56:29-35.
215. Bandini LG, Must A, Cyr H, Anderson SE, Spadano JL, Dietz WH. Longitudinal changes in the accuracy of reported energy intake in girls 10-15 y of age. *Am J Clin Nutr* 2003;78:480-484.
216. Mennen LI, Jackson M, Cade J, Mbanya JC, Lafay L, Sharma S, Walker S, Chungong S, Wilks R, Balkau B, Forrester T, Cruickshank JK. Underreporting of energy intake in four populations of African origin. *Int J Obes Relat Metab Disord* 2000;24:882-887.
217. Black AE, Goldberg GR, Jebb SA, Livingstone MBE, Cole TJ, Prentice AM. Critical evaluation of energy intake data using fundamental principles of energy physiology: 2. Evaluating the results of published surveys. *Eur J Clin Nutrition* 1991;45:583-599.
218. Martin LJ, Su W, Jones PJ, Lockwood GA, Tritchler DL, Boyd NF. Comparison of energy intakes determined by food records and doubly labeled water in women participating in a dietary-intervention trial. *Am J Clin Nutr* 1996;63:483-490.
219. Schoeller DA. How accurate is self-reported dietary energy intake? *Nutr Rev* 1990;48:373-379.
220. Lifson N, Gordon GB, McClintock R. Measurement of Total Carbon Dioxide Production by Means of D<sub>2</sub>O<sup>18</sup>. *J Appl Physiol* 1955;7:704-710.
221. Prentice AM. Applications of the doubly-labelled water (<sup>2</sup>H<sub>2</sub>-<sup>18</sup>O) method in free-living adults. *Proc Nutr Soc* 1988;47:259-268.
222. Schoeller DA, Leitch CA, Brown C. Doubly labeled water method: in vivo oxygen and hydrogen isotope fractionation. *Am J Physiol* 1986;251:R1137-R1143.
223. Schoeller DA, Ravussin E, Schutz Y, Acheson KJ, Baertschi P, Jequier E. Energy expenditure by doubly labeled water: validation in humans and proposed calculation. *Am J Physiol* 1986;250:R823-R830.
224. Schoeller DA. Energy expenditure from doubly labeled water: some fundamental considerations in humans. *Am J Clin Nutr* 1983;38:999-1005.
225. Schoeller DA, van Santen E. Measurement of energy expenditure in humans by doubly labeled water method. *J Appl Physiol* 1982;53:955-959.
226. Speakman JR. The history and theory of the doubly labeled water technique. *Am J Clin Nutr* 1998;68:932S-938S.
227. Wong WW, Butte NF, Ellis KJ, Hergenroeder AC, Hill RB, Stuff JE, Smith EO. Pubertal African-American girls expended less energy at rest and during physical activity than Caucasian girls. *J Clin Endocrinol Metab* 1999;84:906-911.
228. Coward WA, Prentice AM, Murgatroyd PR, Davies HL, Cole TJ, Sawyer M, Goldberg GR, Halliday D, Macnamara JP. Measurement of CO<sub>2</sub> and water production rates in man using <sup>2</sup>H, <sup>18</sup>O-labelled H<sub>2</sub>O: comparisons between calorimeter and isotope values. In: van Es AJ, ed. *Clinical Nutrition and Metabolic Research*. Munich: Basel:Karger, 1984. 169-177.
229. Klein PD, James WP, Wong WW, Irving CS, Murgatroyd PR, Cabrera M, Dallosso HM, Klein ER, Nichols BL. Calorimetric validation of the doubly-labelled water method for determination of energy expenditure in man. *Hum Nutr Clin Nutr* 1984;38:95-106.
230. Schoeller DA, Webb P. Five-day comparison of the doubly labeled water method with respiratory gas exchange. *Am J Clin Nutr* 1984;40:153-158.
231. Seale JL, Rumpler WV. Comparison of energy expenditure measurements by diet records, energy intake balance, doubly labeled water and room calorimetry. *Eur J Clin Nutr* 1997;51:856-863.

232. Seale JL, Conway JM, Canary JJ. Seven-day validation of doubly labeled water method using indirect room calorimetry. *J Appl Physiol* 1993;74:402-409.
233. 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-E409.
234. Horvitz MA, Schoeller DA. Natural abundance deuterium and 18-oxygen effects on the precision of the doubly labeled water method. *Am J Physiol Endocrinol Metab* 2001;280:E965-E972.
235. Trabulsi J, Troiano RP, Subar AF, Sharbaugh C, Kipnis V, Schatzkin A, Schoeller DA. Precision of the doubly labeled water method in a large-scale application: evaluation of a streamlined-dosing protocol in the Observing Protein and Energy Nutrition (OPEN) study. *Eur J Clin Nutr* 2003;57:1370-1377.
236. Wong WW, Lee LS, Klein PD. Deuterium and oxygen-18 measurements on microliter samples of urine, plasma, saliva, and human milk. *Am J Clin Nutr* 1987;45:905-913.
237. Wong WW, Clarke LL, Llauroador M, Klein PD. A new zinc product for the reduction of water in physiological fluids to hydrogen gas for  $^2\text{H}/^1\text{H}$  isotope ratio measurements. *Eur J Clin Nutr* 1992;46:69-71.
238. Gonfintini R. Standards for stable isotope measurements in natural compounds. *Nature* 1978;271:534-536.
239. The Doubly-Labeled Water Method for Measuring Energy Expenditure: *Technical Recommendations for Use in Humans*. In: *A Consensus Report by the IDECG Working Group*. Editor: Prentice AM. International Atomic Energy Agency, P.O. Box 100, A-1400, Vienna, Austria, 1990.
240. Schoeller DA. Measurement of energy expenditure in free-living humans by using doubly labeled water. *J Nutr* 1988;118:1278-1289.
241. 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:E402-9.
242. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psych Bull* 1979; 86:420-28.
243. Eliasziw M, Young SL, Woodbury MG, Fryday-Field K. Statistical methodology for the concurrent assessment of interrater and intrarater reliability: using goniometric measurements as an example. *Phys Ther* 1994; 74:777-88.
244. Donner A, Eliasziw M. Sample size requirements for reliability studies. *Stat Med* 1987; 6: 441-8.
245. David Garth. Hyperkalemia. eMedicine 2005; January 5.
246. Lachin JM. Introduction to sample size determination and power analysis for clinical trials. *Control Clin Trials* 1981; 2:93-114.
247. Machin D, Campbell MJ, Fayers PM, Pinol APY. *Statistical Tables for Clinical Studies* 2nd edn.. Oxford: Blackwell Scientific Ltd., 1997.
248. Cook RJ, Farewell VT. Multiplicity considerations in the design and analysis of clinical trials. *J Roy Stat Soc, Series A* 1996; 159: 93-110.
249. Proschan MA, Waclawiw MA. Practical guidelines for multiplicity adjustment in clinical trials. *Controlled Clin Trials* 2000; 21: 527-39.
250. Moher D, Schulz KF, Altman DG, Lepage L. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001; 357:1191-4.
251. Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001; 285:1987-91.
252. Lewis JA. Statistical principles for clinical trials (ICH E9): an introductory note on an international guideline. *Stat Med* 1999; 18: 1903-42.
253. Lee YJ, Ellenberg JH, Hirtz DG, Nelson KB. Analysis of clinical trials by treatment actually received: Is it really an option? *Statist Med* 1991; 10:1595-1605.
254. Weinstein GS, Levin B. Effect of crossover on the statistical power of randomized studies. *Ann Thorac Surg* 1989; 48:490-5.
255. Lavori PW. Clinical trials in psychiatry: Should protocol deviation censor patient data? (with discussion) *Neuropsychopharmacology* 1992; 6:39-63.
256. Tsiatis A. Analysis and interpretation of trial results. Intent-to-treat analysis. *J Acquir Immune Defic Syndr* 1990; 3 (Suppl 2): S120-S123.
257. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol* 66; 688-701.
258. Holland PW. Statistics and causal inference (with discussion). *J Am Stat Assoc* 1986; 81:945-70.
259. Little RJ, Rubin DB. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytic approaches. *Ann Rev Public Health* 2000; 21-121-45.
260. Hernán MA, Brumback B, Robins JM (2000). Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 11, 561-70.

261. Robins JM, Hernán MA, Brumback B (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology* 11, 550-60.
262. Liang K-Y, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986; 73:13-22.
263. Zeger SL, Liang K-Y. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986; 42:121-130.
264. Matsuyama Y (2003). Sensitivity analysis for the estimation of rates of change with non-ignorable drop-out: an application to a randomized clinical trial of vitamin D<sub>3</sub>. *Stat Med* 22, 811-27.
265. Yamaguchi T, Ohashi Y. Adjusting for differential proportions of second line treatment in cancer clinical trials. Part I: Structural nested models and marginal structural models to test and estimate treatment arm effects. *Stat Med* 23, 1901-2003.
266. Yamaguchi T, Ohashi Y. Adjusting for differential proportions of second line treatment in cancer clinical trials. Part II: An application in a clinical trial of unresectable non-small-cell lung cancer. *Stat Med* 23, 2005-22.
267. Agresti A. *Categorical Data Analysis*. New York: John Wiley & Sons; 1990
268. Grouin JM, Lewis J. Committee for Proprietary Medicinal Products (CPMP): points to consider on adjustment for baseline covariates. *Stat Med* 2004; 23: 701-9.
269. Diggle PJ, Heagerty PJ, Liang K-Y, Zeger SL. *Analysis of Longitudinal Data*. 2nd edn. New York: Oxford University Press; 2002.
270. Jennrich RI, Schluchter MD. Unbalanced repeated-measures models with structured covariance matrices. *Biometrics* 1986; 42:805-820.
271. Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics* 1982;38:963-974.
272. Robins JM. Marginal structural models versus structural nested models as tools for causal inference. In *Statistical Methods in Epidemiology*, Halloran ME and Berry D (eds). New York: Springer-Verlag, pp95-134.
273. Rochon J. Supplementing the intent-to-treat analysis: Accounting for covariates observed post randomization in clinical trials. *J Am Stat Assoc* 1995; 90: 292-300.
274. Rochon J. Accounting for covariates observed post randomization for discrete and continuous repeated measures data. *J Royal Stat Soc, Series B* 1996; 58: 205-219.
275. Rothwell PM. Treating individuals 2. Subgroup analysis in randomized controlled trials: importance, indications, and interpretation. *Lancet* 2005; 265: 176-86.
276. Brookes ST, *et al*. Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. *J Clin Epidemiol* 2004; 57: 229-36.
277. Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet* 2000; 355: 1064-9.
278. Parker AB, Yusuf S, Naylor CD. The relevance of subgroup-specific treatment effects: the Studies Of Left Ventricular Dysfunction (SOLVD) revisited. *Am Heart J* 2002; 144: 941-7.
279. Yusuf S, Wittes J, Probstfield J, Tyroler HA. Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. *JAMA* 1991;266:93-98.
280. Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data* 2nd edn. New York: John Wiley & Sons, 2002.
281. Efron B. Logistic regression, survival analysis, and the Kaplan-Meier curve. *J Amer Stat Assoc* 1988; 83:414-425.
282. McFadden E. *Management of Data in Clinical Trials*. New York: John Wiley & Sons, 1998.
283. Allen ME. Good Clinical Practice: current state of the art for clinical trial documentation. *Med Leg J* 1998; 66:61-4.
284. Bohaychuk W, Ball G. *Conducting GCP-Compliant Clinical Research*. New York: Wiley, 1999.
285. Reynolds-Haertle RA, McBride R. Single vs. double data entry in CAST. *Controlled Clin Trials* 1992; 13:487-494.
286. Day S, Fayers P, Harvey D. Double data entry: What value, what price? *Controlled Clin Trials* 1998; 19:15-24.
287. Gassman JJ, Owen WW, Kuntz TE, Martin JP, Amoroso WP. Data quality assurance, monitoring and reporting. *Controlled Clin Trials* 1995; 16:104S-136S.
288. McBride R, Singer SW. Interim reports, participant closeout, and study archives. *Controlled Clin Trials* 1995; 16:137S-167S.
289. Smith MA, Ungerleider RS, Korn EL, Rubinstein L, Simon R. Role of independent data-monitoring committees in randomized clinical trials sponsored by the National Cancer Institute. *J Clin Oncol* 1997; 15:2736-43.
290. Armitage P. Data and safety monitoring in the Concorde and Alpha trials *Controlled Clin Trials* 1999; 20:207-28.
291. Freidlin B, Korn EL, George SL. Data monitoring committees and interim monitoring guidelines. *Controlled Clin Trials* 1999; 20:395-407.

292. Pocock S, Furberg CD. Procedures of data and safety monitoring committees. *Am Heart J* 2001; 141:289-94.
293. Ellenberg SS. Independent data monitoring committees: rational, operations and controversies. *Stat Med* 2001; 20:2573-84.

## 22. APPENDIX 1: BRIEF SUMMARIES OF THE CALERIE PHASE 1 STUDIES

### 22.1 PBRC Study

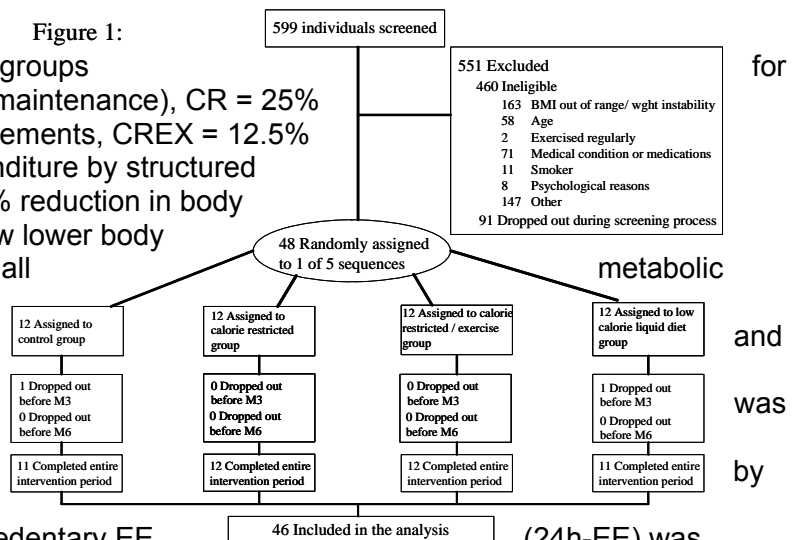
#### 22.1.1 Introduction

Caloric restriction (CR) extends life span and retards age-related chronic diseases in a variety of animal species including rats, mice, fish, flies, worms and yeast. The mechanism/s through which this occurs is unclear. CR reduces metabolic rate and oxidative stress, improves insulin sensitivity and alters neuro-endocrine and sympathetic nervous system function in animals. Whether prolonged CR increases life span (or improves biomarkers of aging) in humans is unknown. In experiments of nature, humans have been subjected to periods of non-volitional partial starvation. The Pennington CALERIE phase 1 study was designed to test the feasibility and safety of conducting CR in non obese humans. We also used this study to test the effects of 6-month CR on surrogate markers of longevity and some of the mechanisms implicated in the beneficial effect of CR.

#### 22.1.2 Methods

Healthy male (25-50y) and female (25-45y), overweight participants ( $25 \leq \text{BMI} \leq 30$ ) were recruited. Participants were excluded if they smoked, exercised more than twice per week, were pregnant, lactating or post-menopausal, had a personal history of obesity (BMI never greater than  $32\text{kg/m}^2$ ), cardiovascular disease, diabetes, or regularly used medications (except birth control). Five hundred ninety nine individuals were screened; 551 were excluded, 460 of these were ineligible and 91 withdrew during screening.

After stratification by sex and BMI, 48 subjects were randomized into one of four groups 6-months: Control (healthy diet for weight maintenance), CR = 25% calorie restriction of baseline energy requirements, CREX = 12.5% CR + 12.5% increase in total energy expenditure by structured exercise, LCD = low calorie diet until a 15% reduction in body weight, followed by maintenance of the new lower body weight. Weight was measured weekly and all tests were performed during 4-5-day inpatient stays at baseline, month 3 (M3) and month 6 (M6). Fasting blood samples were taken. Body composition measured by DEXA (Hologics, QDA 4500A Bedford, MA) and abdominal fat multislice CT. Muscle and liver lipid stores were determined by proton MRS. Sedentary EE was measured in a whole room indirect calorimeter. Fat cell size (FCS) and number were determined using the Multisizer-3 counter. Insulin sensitivity was determined by the insulin-modified frequently sampled intravenous glucose tolerance test. 24-h blood samples were collected to assess the diurnal rhythm of hormone and substrates.



#### 22.1.3 Results

**Hypothesis A:** Chronic CR (resulting in loss of weight and maintenance of energy balance at a new lower body mass) is associated with several metabolic adaptations, including lower absolute and relative rates of energy expenditure, lower body temperature, and evidence of lower tissue oxidative stress. **Aim A:** Measure and compare for difference: Energy expenditure (free living, sedentary 24-h, resting, exercise efficiency), body temperature and DNA, protein and lipid oxidative damage.

Weight change at M6 was  $-1.0 \pm 1.1\%$  (Control),  $-10.4 \pm 0.9\%$  (CR),  $-10.0 \pm 0.8$  (CREX),  $-13.9 \pm 0.7\%$  (LCD). At M6, fasting serum insulin was significantly reduced from baseline in CR, CREX and LCD groups ( $p < 0.01$ ), whereas DHEAS and glucose were unchanged. Core temperature was reduced from baseline in CR and CREX groups only ( $p < 0.05$ ). After adjustment for changes in body composition, sedentary 24h-EE was unchanged in controls ( $-18 \pm 52$  kcal/d;  $p = \text{NS}$ ), but significantly lower than predicted in CR ( $-135 \pm 42$  kcal/d,  $p = 0.002$ ), CREX ( $-117 \pm 52$  kcal/d,  $p = 0.008$ ) and LCD ( $-125 \pm 35$  kcal/d,



$p=0.006$ ) groups. These “metabolic adaptations” (~6% more than expected based on loss of metabolic mass) were statistically different from controls ( $p<0.05$ ). A significant decrease in DNA damage was also observed from baseline in CR, CREX and LCD groups at M6 ( $p\leq 0.002$ ).

**Hypothesis B.** Chronic CR improves surrogate markers (risk factors) for chronic diseases, including type-2 diabetes and cardiovascular disease. These adaptations are the same whether the energy deficit is produced by combining physical activity (PA) and CR or by CR alone. **Aim B:** Measure and compare for difference: B1) CVD risk factors (BP, lipid profile, hemostasis factors, homocysteine, endothelial function, and markers of inflammation). B2) Type 2 diabetes risk factors. (Insulin action and secretion)

At baseline, FCS was related to visceral adipose tissue (VAT) and intra-hepatic lipids (IHL) ( $p<0.05$ ), but not to intra-myocellular lipids (IMCL). FCS was the strongest determinant of insulin sensitivity ( $S_i$ ) ( $p<0.01$ ). At M6, VAT, FCS, % body fat and IHL were reduced in the three intervention groups ( $p<0.01$ ), but IMCL was unchanged.  $S_i$  was increased ( $p=0.05$ ) in the CREX ( $37\pm 18\%$ ) and LCD ( $70\pm 34\%$ ) groups ( $p<0.05$ ) but only tended to increase in the CR group ( $40\pm 20\%$ ,  $p=0.08$ ). The improvement in insulin sensitivity was related to loss in weight, fat mass and VAT, but not IHL, IMCL or FCS.

Triglyceride, cholesterol, and LDL concentrations were all reduced in the intervention groups whereas HDL increased; however the reverse effect was often noted in the control group over the 6 month intervention period indicating improvements in risk factors associated with the development of CVD.

### **22.1.4 Discussion**

The results of this study show that prolonged CR by diet or by a combination of diet and exercise was successfully implemented as evidenced by calculated adherence as well as reduced weight, fat mass, fasting serum insulin and body core temperature. Furthermore, we observed that “metabolic adaptation” develops in response to energy deficit in non-obese humans leading to reduced oxygen consumption per unit of fat-free mass. Finally, this study supports previous findings that calorie restriction results in a decline in DNA damage. However, longer studies are required to determine if these effects are sustained and impact the aging process. We also observed that calorie restriction by diet alone, or in conjunction with exercise leads to similar improvements in insulin sensitivity with concomitant reductions in  $\beta$ -cell secretion. The results of this study also provide support for the hypothesis that the underlying pathology of insulin resistance is related to an abnormal partitioning of fat between the adipose, hepatic, muscle and pancreatic tissues. In particular, we observed relationships between visceral fat, subcutaneous abdominal adipocyte size and lipid accumulation in liver supporting our hypothesis that ectopic fat deposition may be driven by the inability to create new fat cells leading to lipid “spill-over” in response to a state of positive energy balance. However, the finding that IMCL was not responsive to weight loss (despite improvements in insulin sensitivity) suggests that intracellular triglyceride accumulation is not a causal factor of insulin resistance in muscle. Other metabolites of fat may be more important.

## **22.2 Tufts University Study**

### **22.2.1 Introduction**

Caloric Restriction (CR) is associated with delayed biological aging and increased lifespan in a wide range of animal models, but it is not known whether similar benefits may occur in humans. The overall goal of the CALERIE study is to conduct a RCT to examine whether CR delays biological aging without unacceptable side effects in humans. The Tufts CALERIE pilot was therefore designed to prepare us to conduct a RCT of human CR, and the primary specific aim was to develop, refine and compare two different 1-year CR interventions (high glycemic load (HG) diet containing 60% carbohydrate, 20% protein and 20% fat, and low glycemic load diet (LG) containing 40% carbohydrate, 30% protein and 30% fat) providing 70% of baseline energy requirements for their ability to support sustained CR. The two 30% groups were using change in body weight and body fat as an initial indicator of adherence. Outcome measurements anticipated for the main CALERIE study were also compared between the diet groups, including metabolic rate, hunger and satiety and changes in insulin sensitivity, blood lipids and other parameters that typically are adversely affected by aging and improved by CR.

### 22.2.2 Methods

The Tufts CALERIE pilot was a parallel-group RCT testing two levels of CR (30% as the CR intervention and 10% CR as the control) and two types of dietary composition (HG and LG) in four groups of men and women with group randomizations balanced for BMI and gender. 47 individuals were enrolled who were eligible based on detailed inclusion (including age 20-42 yrs, BMI in the range 25-29.9 kg/m<sup>2</sup>, healthy based on self-report and standard medical screenings, availability for the entire study period and willingness to participate in all components of the study) and exclusion criteria (including use of medications that might interfere with ability to participate in the intervention or outcome testing, history of eating disorders, psychiatric disorders or chronic disease, inability to complete an accurate food record).

Following acceptance into the study, the subjects underwent a period of baseline testing when usual energy needs were assessed and baseline measurements were made of outcome variables. Subjects were then randomized to one of the four groups, and all food was provided for 6 months and then subjects were instructed to maintain their dietary randomization at home for a further 6 months. Behavioral support groups and individuals meetings with a research dietician throughout the study period supported the intervention. Subjects were blinded to their dietary randomization for the first 3 months of the intervention, and research staff involved in outcomes testing were blinded for the entire study period. Primary outcomes included body weight, body fat and metabolic rate. Secondary outcomes were hunger, satiety, risk factors for chronic disease and aging, bone mineral density and incidence of adverse events (adverse events and adherence are detailed in the section above). Changes over time in variables were compared between the two 30% groups using mixed model analysis of variance and *t*-tests.

### 22.2.3 Results

A total of 10,138 individuals expressed an interest in the study following advertisements in the print and visual media in the Greater Boston area and were sent a package of information about the study and a health history form to complete. Out of these individuals, 365 subjects screened for the study, and 47 were enrolled (85% Caucasian, 4% African American, 11% Asian and 2% other) and 39 completed the 12 month assessment at the end of the study. There was no difference in drop out between any of the groups.

Baseline characteristics of the population were: mean± SD, BMI 27.85± 1.5 kg/m<sup>2</sup>, age 35± 5 years, height 169.45±10.1cm, weight 80.23± 10.4 kg and % body fat 35.04± 7.3. Changes over time in key variables are shown in the table below for the 30% groups. Although weight and other variables were significantly changed over time within groups (*P*<0.001) there was no significant difference between the HG and LG groups in any of the variables shown below.

**Table 22.1: Summary of Results from the Tufts Phase 1 Study**

<b>Outcome</b>	<b>30%CR - HG</b>	<b>30% CR - LG</b>	<b>p-value</b>
<b>Weight change 6 months (kg)</b>	-6.95± 4.5	-7.5±4.0	0.95
<b>Weight change 12 months (kg)</b>	-6.40±3.7	-6.08±4.0	0.91
<b>% body fat change 6 months</b>	-16.33±11.7	-21.16±16.8	0.26
<b>% body fat change 12 months</b>	-14.80±8.8	-17.90±12.5	0.37
<b>BMR (kcal/d) 6 months</b>	-96.17±94.5	-88.04±145.5	0.51
<b>BMR (kcal/d) 12 months</b>	-64.10±111.1	-42±125.3	0.98
<b>Fasting insulin 6 months</b>	8.68±3.0	8.76±3.1	0.07
<b>Fasting glucose 6 months</b>	81.2±4.9	82.64±5.3	0.63
<b>Total-cholesterol 6</b>	148.80±19.5	152.71±28.4	0.14

Outcome	30%CR - HG	30% CR - LG	p-value
months			
LDL-cholesterol 6 months	83.30±17.9	92.35.±24.8	.079
Satisfaction with amount of study food (weeks 1-12)	62.38±13.4	63.35±14.5	0.85
Desire to eat other foods (weeks 1-12)	37.11±15.5	40.78±15.7	0.53

In addition to the results presented above comparing the two 30% CR groups, data were also collected from the 10% CR control group. The main finding among the outcomes for this group was that weight loss over time did not differ between the 10% and 30% groups and mean % weight change values were very similar (P=0.68-0.82).

#### **22.2.4 Discussion**

Results from this study indicated no difference in weight loss and body fat loss, and therefore no difference in anticipated adherence to CR between the two different dietary compositions, suggesting that a flexible approach to CR is feasible from the perspective of adherence to CR. In addition, the results indicated that 10% CR is not a suitable control for a 30% CR intervention because subjects appeared to have equivalent levels of adherence based on body weight loss. For Phase 2 we are planning a flexible dietary approach to CR and an assessment-only control group to learn from the experience of this pilot study.

### **22.3 Washington University Study**

#### **22.3.1 Methods**

The Phase 1 Calorie Study at Washington University involved healthy sedentary volunteers in the 50 to 60 yr age range, with BMI's in the 23.5 to 29.9 range (Avg. 27.3). There were three groups: 20% CR for 12 mo; 20% increase in caloric expenditure by means of exercise while keeping caloric intake constant for 12 mo; and an unchanged energy balance, no intervention group practicing a Healthy Lifestyle.

#### **22.3.2 Results**

A total of 379 individuals were assessed for eligibility and 321 were excluded because of overweight, health problems or disinterest. Following screening, 58 were randomized; of these 11 were not enrolled because they decided they could not make the time commitment or because of family issues. Nineteen started the CR intervention, 19 started the exercise intervention and 10 entered the healthy lifestyle group (uneven randomization, 2:2:1 ratio). One participant discontinued the CR, and one discontinued the exercise intervention. No participants were lost to follow-up.

Over the 12 mo, the CR group lost  $8.2 \pm 4.8$  kg (10.3% of initial weight); the Exercise group lost  $6.6 \pm 5.5$  kg (8.6% of initial wt); and the Healthy Lifestyle (HL) group lost  $1.2 \pm 2.1$  kg (1.6% of initial wt). Fat loss averaged  $6.2 \pm 0.7$  kg in the CR, 5.5 kg in the Exercise and 0.4 kg in the HL group. Lean tissue loss averaged  $1.6 \pm 0.5$  kg in the CR,  $0.9 \pm 0.5$  in the Exercise and  $0.8 \pm 0.8$  in the HL group. The decrease in energy intake, estimated from DLW measurements, averaged 302 kcal/d (11.9%) over the 12 mo in the CR group.

Both the CR and the Exercise groups showed significant improvements in a range of risk factors for atherosclerosis, insulin action, glucose tolerance, markers of inflammation and leptin (Table 1). The only change that was unique to the CR group was a significant decrease in plasma cortisol level. IGF-1, other growth factors and lymphocyte count did not decrease significantly, suggesting that more prolonged CR and greater fat loss are needed to induce the reduction in growth factors and cell proliferation that have been observed in studies of CR in animals.

#### **22.3.3 Discussion**

We conclude that CR is feasible in free-living human volunteers but that a longer period of CR than 12 months is needed to induce the decrease in growth factors and cell proliferation seen in response to CR in studies on animals. It also seems prudent to study normal, rather than overweight, humans, in

order to attain the low body fat content/energy stores that are, likely, responsible for some of the major effects of CR on aging.

**Table 22.2: Summary of Results from the Washington University Phase 1 Study**

<b>Outcome</b>	<b>CR (n=18)</b>	<b>Exercise (n=18)</b>	<b>Healthy Lifestyle (n=10)</b>	<b>p-value</b>
<b>Weight, kg</b>	-8.2 ± 4.08*	-6.6 ± 5.5*	-1.2 ± 2.1	0.0003
<b>BMI, kg/m<sup>2</sup></b>	-2.9 ± 1.7*	-2.3 ± 1.7*	-0.5 ± 0.8	0.0002
<b>Total fat mass, kg</b>	-6.3 ± 3.8*	-5.6 ± 4.9*	-0.4 ± 1.7	0.001
<b>Abdominal visceral fat %</b>	-36 ± 14*	-39 ± 13*	-9 ± 8	0.0004
<b>Fat Free Mass, kg</b>	-1.6 ± 0.5*	-0.9 ± 0.5	-0.8 ± 0.8	NS
<b>Fasting insulin µU/ml</b>	-2.5 ± 3.9*	-2.7 ± 5.0*	1.3 ± 3.2	0.01
<b>Insulin AUC x 10<sup>2</sup> µU·ml<sup>-1</sup>·min<sup>-1</sup></b>	-1.3 ± 2.1*	-3.4 ± 3.0	1.4 ± 2.1	0.001
<b>Glucose AUC x 10<sup>3</sup> mg·dl<sup>-1</sup>·min<sup>-1</sup></b>	-2.5 ± 2.5*	-2.4 ± 2.5*	-0.7 ± 1.5	NS
<b>ISI</b>	2.0 ± 3.9*	3.0 ± 2.7*	0.3 ± 1.4	0.001
<b>Adiponectin, µg/ml</b>	1.9 ± 4.0	2.2 ± 4.7	-1.9 ± 2.6	0.005
<b>Leptin, ng/L</b>	-5.4 ± 1.6*	-3.6 ± 1.6*	-2.1 ± 2.1	NS
<b>TNF<math>\alpha</math>, pg/ml</b>	-0.20 ± 0.26*	-0.25 ± 0.36*	-0.06	NS
<b>Cortisol, µ/ml</b>	-4.0 ± 4.1*	-0.1 ± 4.4	-2.2 ± 4.2	NS
<b>Total Cholesterol, mg/dl</b>	-22 ± 18*	-10 ± 25	2.3	NS
<b>LDL Cholesterol, mg/dl</b>	-19 ± 16	-17 ± 23	7.1 ± 19	0.01
<b>HDL Cholesterol, mg/dl</b>	3.2 ± 8.4*	5.7 ± 11*	0.4 ± 3.6	NS
<b>T Chol./HDL Chol ratio</b>	-0.6 ± 0.5*	-0.5 ± 0.6*	0.1 ± 0.3	0.002
<b>Triglycerides, mg/dl</b>	-30 ± 30*	-3.7 ± 29	-26 ± 31	0.03
<b>hs CRP mg/L</b>	-0.07 ± 1.2*	-0.5 ± 3.1	0.4 ± 15	
<b>Systolic BP, mmHg</b>	-5.2 ± 14	0.3 ± 17	1.3 ± 10	NS
<b>Diastolic BP, mmHG</b>	-3.2 ± 9.2	1.1 ± 9.8	1.3 ± 7.1	NS
<b>10 yr CHD Risk</b>	-2.4 ± 4.1*	-0.9 ± 2.6	0.6 ± 1.7	NS

Values are means ± SD. \*Change, P<0.05

### 23. APPENDIX 2: SCHEDULE OF EVALUATIONS

CALERIE Evaluation	Screening	Baseline	Follow-up Month						
			1	3	6	9	12	18	24
<b>Screening Procedures (Section 9):</b>									
Telephone Screen: Age; Height, Weight ⇒ BMI Inclusion Criteria	X								
Screening Informed Consent and HIPAA Authorization	X								
Demographic Information	X								
<b>Exclusion Criteria (Section 6.3):</b>									
<i>Medical and Medication History (Section 16.13)</i>	X								
<i>Abbreviated Medical and Medications History</i>		X							
<i>Screening Questionnaires (Eating Inventory, MAEDS, SCID II, Personality Q're, BDI, GHQ, Body Morph test, IDED-IV)</i>	X								
<i>Serum Pregnancy Test for Women (Section 11.1.3)</i>	X								
<i>Barriers to Participation, Diet Preferences, Allergies, Special Conditions</i>	X								
<i>14-day Food Record</i>	X								
<b>Study Informed Consent and HIPAA Authorization</b>		X							
<b>Vital Signs (Section 16.15)</b>		X	X	X	X	X	X	X	X
<b>Clinic Height (Section 12.8.2)</b>		X							
<b>Clinic Weight (Section 11.4, 12.8)</b>		X	X	X	X**	X	X**	X**	Mo.17, 18,23,24
<b>Home Diaries, i.e., Weight, Food, Symptoms (Section 7 &amp; 16.1)</b>		X	<b>Daily</b>						
<b>Energy Metabolism (Section 12.1):</b>									
TEE by DLW		XX			X*		X	X*	X
RMR		XX			X*		X	X*	X

\* CR Intervention only

\*\* Multiple clinic weights during the DLW periods

CALERIE Evaluation	Screening	Baseline	Follow-up Month						
			1	3	6	9	12	18	24
Core Body Temperature		X			X		X		X
<b>Cardiovascular Risk Factors (Section 12.2):</b>									
Resting Blood Pressure		X	X	X	X	X	X	X	X
Lipids, Markers of Inflammation, CRP, TGF, etc.		X					X		X
<b>Glucose Tolerance and Insulin (Section 12.3):</b>									
OGTT, insulin and C-peptide		X					X		X
<b>Immune Function:</b>									
DTH (Section 12.4.1)		X					X		X
Antibody Response to Vaccines (Section 12.4.3)		X							Mo.17, 18,23,24
<b>Endocrine Response and Growth Factors (Section 12.5):</b>									
Norepinephrine, DHEA, Sex Hormones (for men), Thyroid Hormones, Adipokines, Angiotensin II, Growth Hormone & Growth Factors		X					X		X
<b>QoL, Psychological, Cognitive Function (Section 12.6)</b>		<b>X</b>			<b>X</b>		<b>X</b>		<b>X</b>
<b>Physical Activity Measurements (Section 12.7):</b>									
Stanford 7-day PAR	X	XX			X*		X	X*	X
Maximal Oxygen Uptake (VO <sub>2max</sub> )		X					X		X
Muscular Strength and Endurance		X					X		X
<b>Body Composition (Section 12.9):</b>									
Waist Circumference		X	X	X	X	X	X	X	X
DXA, including BMD and BMC		XX			XX*		X	X*	X
<b>Markers of Bone Turnover (Section 12.10)</b>		<b>X</b>			<b>X</b>		<b>X</b>		<b>X</b>
<b>Nutrient Intake (Section 12.11)</b>	<b>X</b>	<b>XX</b>			<b>X*</b>		<b>X</b>	<b>X*</b>	<b>X</b>

\* CR Intervention only

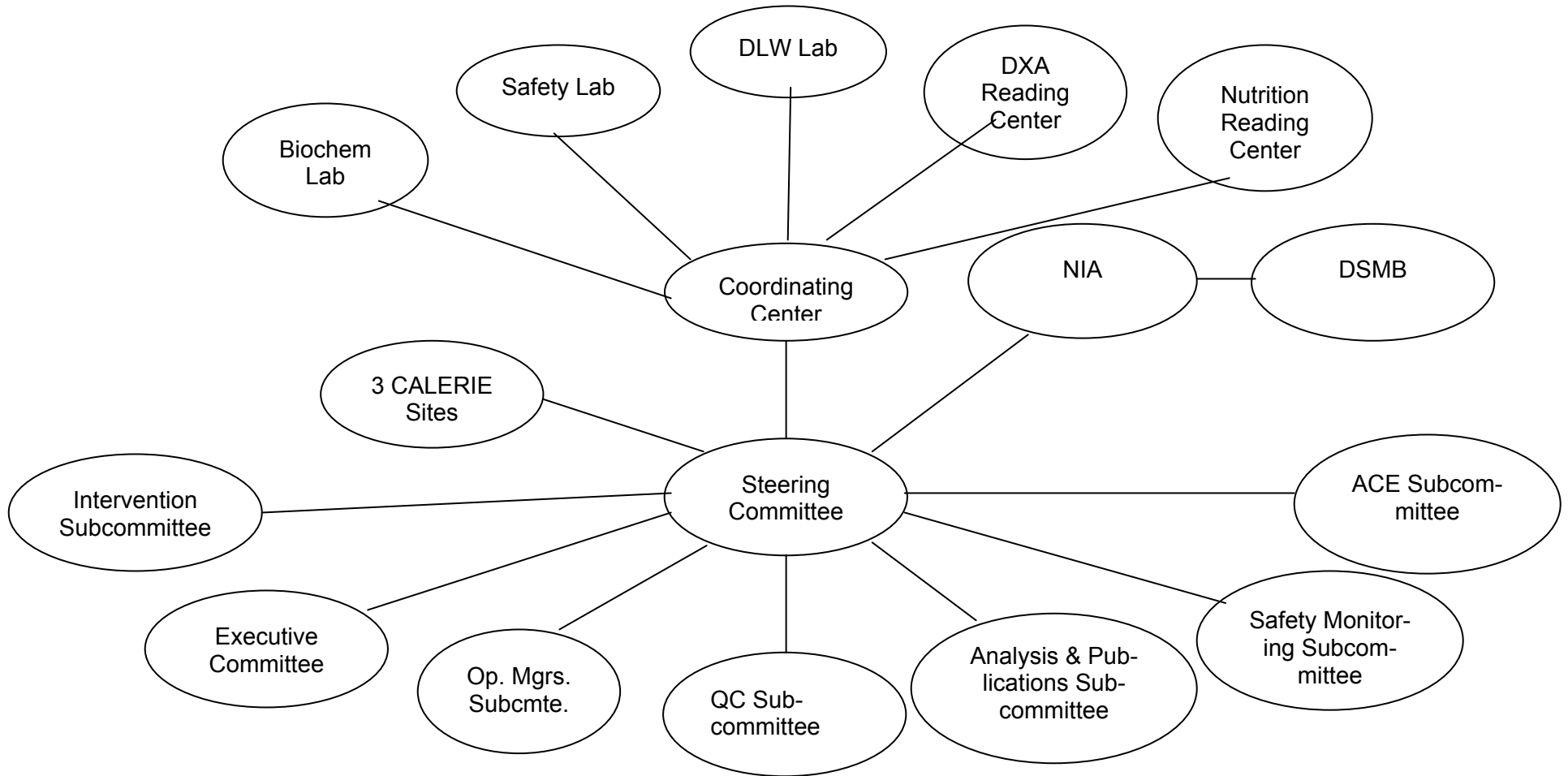
\*\* Multiple clinic weights during the DLW periods

CALERIE Evaluation	Screening	Baseline	Follow-up Month						
			1	3	6	9	12	18	24
<b>Archive Materials (Section 12.12):</b>									
Blood		X		X	X		X	X	X
Urine		X					X		X
Muscle and Abdominal Fat Biopsy		X					X		X
<b>Concomitant Medications (Sections 11.2, 11.3, 11.4)</b>		X	X	X	X	X	X	X	Mo.17, 23,24
<b>Participant Safety:</b>									
Blood Chemistry, Hematology and Urinalysis (Section 16.4)	X	X	X	X	X	X	X	X	X
Urine Pregnancy Test for Women (Sections 11.2 and 16.4.1)		X			X*		X	X*	Mo.17, 23, 24
Adverse Events (Section 11.4, 16.1)		X	X	X	X	X	X	X	Mo.17, 23,24
Serious Adverse Events (Section 16.2)			As they occur with an expedited reporting protocol						
Physical Examination (Section 16.14)	X	X					X		X
Potassium Surveillance, ECG, Creatine Phosphokinase (Section 0)	X	X	X	X	X	X	X	X	X
Anemia Screening / Surveillance (Section 16.6)	X	X	X	X	X	X	X	X	X
Cholesterol Surveillance (Sections 6.3 and 16.7)	X	X					X		X
Eating Disorders Screening / Surveillance (Section 16.8)	X	X		X	X		X	X	X
BMI Screening / Surveillance (Section 16.9)	X		X	X*	X*	X*	X*	X*	X*
BDI Screening / Surveillance (Section 16.10)	X	X	X	X	X	X	X	X	X
DXA scans of hip and spine		X			X*		X		X

\* CR Intervention only

\*\* Multiple clinic weights during the DLW periods

**24. APPENDIX 3: ORGANIZATIONAL STRUCTURE FOR CALERIE**



\* CR Intervention only

\*\* Multiple clinic weights during the DLW periods



## 25. APPENDIX 4: INFORMED CONSENT TEMPLATES