Molecular Transducers of Physical Activity Consortium (MoTrPAC) – Adult

Consortium Steering Committee Chair & Principal Investigator: Wendy Kohrt, PhD

Funded by: NIH Common Fund

Version Number: 3.0

22 August 2023

Summary of Changes from Previous Versions:

Affected Section	Summary of Revisions Made	Rationale
4.1.2, 5.2.1	Remove COVID-related risk mitigation procedures	With the change in sIRB to Wake Forest IRB, COVID restrictions and procedures will be guided by local conditions and institutional policies rather than study-wide protocol requirements.
5.1.1, 5.1.2, 5.2.1, 5.2.2	Changes to study eligibility criteria (blood pressure, body mass index, thyroid stimulation hormone, antidepressant medications)	At the recommendation of the DSMB, eligibility criteria were evaluated and revised to enhance study recruitment while maintaining focus on participant safety and risk mitigation.
5.5.4	Increase participant remuneration to $2,000 for Sedentary participants and $1,000 for Highly Active participants	MoTrPAC is time-consuming for volunteers and an increase in participant stipends will improve both recruitment and retention.

Contents

STATEMENT OF COMPLIANCE

The trial will be carried out in accordance with the following:

- United States (**US**) Code of Federal Regulations (**CFR**) applicable to clinical studies (45 CFR Part 46, 21 CFR Part 50, 21 CFR Part 56, 21 CFR Part 312, and/or 21 CFR Part 812)

National Institutes of Health (**NIH**)-funded investigators and clinical trial site staff who are responsible for the conduct, management, or oversight of NIH-funded clinical trials have completed Human Subjects Protection and Good Clinical Practice (**GCP**) training.

1 PROTOCOL SUMMARY

1.1 SYNOPSIS

The goal of the Molecular Transducers of Physical Activity Consortium (**MoTrPAC**) is to assess molecular changes that occur in response to physical activity (**PA**). To achieve this aim, a mechanistic randomized controlled trial (**RCT**) is conducted, in which approximately 1,980 adult study participants are randomized to endurance exercise (**EE**) training (n=840), resistance exercise (**RE**) training (n=840), or no exercise Control (n=300) for a period of approximately 12 weeks. The overarching hypothesis is that there are discoverable molecular transducers that communicate and coordinate the effects of exercise on cells, tissues, and organs, which may initiate processes ultimately leading to the health benefits of exercise. Because this is a mechanistic trial, the main goal is not a health-related outcome. Rather, the goal is to generate a map of the molecular responses to exercise that will be used by the Consortium and by the scientific community at large to generate hypotheses for future investigations of the health benefits of PA. Study assessments are completed before and after the intervention period (exercise or control), and at specific interim time points during the course of the intervention. Assessments include measurements of cardiorespiratory fitness, muscular strength, and body composition (including total body bone mineral content) determined by dual-energy x-ray absorptiometry (**DXA**). There is also collection of blood and muscle and adipose tissue biospecimens, monitoring of free-living PA level using wearable devices, and completion of participant reported outcomes and health status by interview and/or questionnaire. An additional 300 highly active (**HA**) individuals currently active in either EE (**HAEE**) or RE (**HARE**) are recruited for a single acute exercise testing session of either endurance or resistance exercise and other study assessments. MoTrPAC participants are recruited, trained, and assessed via six adult Clinical Centers (**CC**), involving 9 clinical sites. As part of the MoTrPAC functions, participant data and biological samples are transferred from the clinical sites to the Consortium Coordinating Center (**CCC**) Data Management, Analysis and Quality Control Center (**DMAQC**) and to the Biological Sample Repository, and later analyzed by the Consortium Chemical Analysis Sites (**CAS**) and the Bioinformatics Center (**BIC**).

Biological samples collected in this project undergo molecular phenotyping, including metabolomic, lipidomic, proteomic, epigenomic, transcriptomic, and genomic analyses. These assays are done at the MoTrPAC CAS.

Overall coordination of the study and analyses occurs at 4 institutions which make up the CCC and the BIC.

1.2 STUDY DESIGN

MoTrPAC will enroll ~1,980 sedentary adults and ~300 highly active adults who participate primarily in endurance or resistance exercise. To enroll the ~1,980 sedentary adults and ~300 highly active adults, approximately 5,940 sedentary and 600 highly active adults may be consented for the study. After screening the sedentary volunteers to determine eligibility, participants will be randomized to one of three intervention groups: EE, RE, and Control. As depicted in **Figure 1a**, all sedentary participants will undergo phenotyping visits and familiarization visits to prepare them for the acute exercise/rest visit, which is when biospecimens will be collected for the characterization of molecular responses to exercise. After completing the acute exercise/rest visit, participants will undergo the exercise/no exercise control interventions. The phenotyping visits and acute exercise/rest visit with biospecimen collections will then be repeated.

Figure 1a. MoTrPAC study design for sedentary participants randomized to endurance exercise training (EE), resistance exercise training (RE), or no-exercise control groups

Screening and Phenotyping* (~2 months)	Randomized to:	Pre-intervention Testing (~3 weeks)			Intervention (~12 weeks)			Post-intervention Testing (~2 weeks)			
		Familiarization, washout**	Acute exercise test	Biospecimen collection+	Endurance exercise	Resistance exercise	Physical activity monitoring	Phenotyping*	Washout**	Acute exercise test	Biospecimen collection+
~990 women aged 18+ yr	EE, n=~840	X	X	X	X		X	X	X	X	X
	RE, n=~840	X	X	X		X	X	X	X	X	X
~990 men aged 18+ yr	Control, n=~300	X		X			X	X	X		X
Meet eligibility criteria											

* Phenotyping includes assessments of aerobic fitness, muscle strength, body composition, physical activity; some pre-intervention phenotyping visits occur before randomization
** Washout (no exercise or testing) before acute exercise test, biospecimen collection
+ Biospecimen collection is described in more detail in Figure 2a

Figure 1b. MoTrPAC study design for highly active endurance exercise (HAEE) and resistance exercise (HARE) participants

Screening (~2 months)	Enrollment	Testing (~2 weeks)			
		Phenotyping*	Familiarization, washout**	Acute exercise test	Biospecimen collection+
	HAEE, n=~150	X	X	X	X
	HARE, n=~150	X	X	X	X

* Phenotyping includes assessments of aerobic fitness, muscle strength, body composition, physical activity
** Washout (no exercise or testing) before acute exercise test and biospecimen collection
+ Biospecimen collection is described in more detail in Figure 2b

The HA groups will undergo largely the same baseline testing and phenotyping as the sedentary adults, including the acute exercise test with biospecimen collection, but their involvement will end at this point; they will not undergo the intervention or post-intervention testing (**Figure 1b**).

During the development of the clinical protocol, concern was raised by the MoTrPAC Data and Safety Monitoring Board (**DSMB**) that the original biospecimen collection plan, which included 4 muscle and 2 adipose biopsies at each acute exercise/rest visit, would be overly burdensome to participants. To address this, a Vanguard Phase (described in Section 4.1) to assess the feasibility of the biospecimen collection plan (and other aspects of feasibility) will be conducted. The study design was expanded, such that Sedentary participants in the three intervention groups will also be randomized to four biospecimen collection profiles: Early, Mid, Late, or All. Although the Vanguard Phase was not completed because of the COVID-19 pandemic, the experience of the research teams was that the All biospecimen collection profile was too burdensome for the Sedentary participants and had the potential to adversely affect retention. Therefore, this profile was eliminated and randomization of Sedentary participants in the AE, EE, and control groups will be restricted to the Early, Mid, or Late biospecimen collection profiles. The timing of the collection of

blood, muscle, and adipose samples for the sedentary participants is depicted in **Figure 2a**. The HAEE and HARE participants will all undergo the original tissue sampling plan, which is the same as the All sampling profile **Figure 2b**.

Figure 2a. Biospecimen collection during acute exercise/rest visit in sedentary participants before and after exercise/no exercise intervention. Participants are randomized to one of three intervention groups (EE, RE, Control) and one of three biospecimen collection profiles (Early, Mid, Late)

Intervention Group	Biospecimen Collection Profile	30-minute Rest	Biospecimen Collection Before Exercise/Rest	Biospecimen Collection During Exercise/Rest		Biospecimen Collection After Exercise/Rest			
				~20 min	~40 min	~10 min	~15-45 min	~3.5-4 h	~24 h
EE	Early		B M A	B	B	B	B M A		
	Mid		B M A	B	B	B	B	B M A	
	Late		B M A	B	B	B	B		B M A
RE	Early		B M A			B	B M A		
	Mid		B M A			B	B	B M A	
	Late		B M A			B	B		B M A
Control	Early		B M A	B	B	B	B M A		
	Mid		B M A	B	B	B	B	B M A	
	Late		B M A	B	B	B	B		B M A

EE= endurance exercise group, RE = resistance exercise group, B = blood, M = muscle, A = adipose

The acute exercise/rest testing visit is repeated after the intervention. The biospecimen collection plan is repeated with the exception that participants in the Control group have only single blood, muscle, and adipose collections after the 30-minute rest interval.

Figure 2b. Biospecimen collection during acute exercise visit in highly active endurance exercise (HAEE) and resistance exercise (HARE) participants

Highly Active Group	30-minute Rest	Biospecimen Collection Before Exercise	Biospecimen Collection During Exercise		Biospecimen Collection After Exercise			
			20 min	40 min	10 min	15-45 min	4 hr	24 hr
HAEE		B M A	B	B	B	B M	B M A	B M
HARE		B M A			B	B M	B M A	B M

B = blood, M = muscle, A = adipose

1.3 SCHEDULE OF ACTIVITIES

The schedule of activities for sedentary participants is summarized in **Table 1**. The schedule of activities for HAEE and HARE participants generally follows Table 1, with the following exceptions: **1)** there is no randomization; and **2)** their participation ends after the acute exercise test. **Table 2** summarizes the estimated participant burden and time commitment to the protocol.

Table 1. MoTrPAC clinic flow for sedentary participants

Initial screen
↓
Study orientation and informed consent
↓
Screening visits to establish eligibility*
↓
Randomization to Endurance Exercise, Resistance Exercise, or Control groups
↓
Phenotyping visits (aerobic fitness, muscle strength, body composition, physical activity) and familiarization visits for the acute exercise test*
↓
Washout (no exercise or testing)
↓
Biospecimen collection with acute exercise/rest
↓
Endurance exercise training, resistance exercise training, or no exercise intervention
↓
Repeat phenotyping visits*
↓
Washout (no exercise or testing)
↓
Repeat biospecimen collection with acute exercise/rest

* The number of visits to complete screening and phenotyping tests will be allowed to vary to accommodate the schedules of participants; initial phenotyping tests that determine eligibility will be performed before randomization

Table 2. Summary of participant involvement and group-specific burden for the endurance exercise (EE), resistance exercise (RE), and control groups

Phase	Activities	Group-specific Burden		
		EE	RE	Control
Screening 1 visit 1-2 total hours	Orientation Consent Medical history Blood draw Resting BP, ECG			
Screening and pre-intervention testing 2-4 visits* 2-4 total hours	Questionnaires Body composition (DXA) CPET/VO_2peak** Muscle strength Wear physical activity monitor	 Hand grip, isometric knee extension 7 days	 Hand grip, isometric knee extension, 1RM 7 days	 Hand grip, isometric knee extension 7 days
Familiarization 0-3 visits, up to 4.5 hours	Learn how to perform acute exercise test	2 visits, ~45-90 min each	3 visits, ~45-90 min each	0 visits
Washout	No testing or exercise			
Pre-intervention acute exercise test or rest 1 or 2 visits 4-8 total hours	No NSAID/acetaminophen use Overnight fast, 1 or 2 nights Acute exercise or rest Blood samples Muscle samples Adipose samples Diet record	7 days before test Up to 18 hr ~45 min cycling 5-6 2 2 1 or 2 days	7 days before test Up to 18 hr ~45 min RE 3-4 2 2 1 or 2 days	7 days before test Up to 18 hr ~45 min rest 5-6 2 2 1 or 2 days
Intervention ~12 weeks, 3 days/week Up to 75 min/day Physical activity monitor	Supervised exercise Three times during the intervention	Cycling and treadmill 7 days each, at ~weeks 4, 8, 12	Weight lifting 7 days each, at ~weeks 4, 8, 12	No exercise 7 days each, at ~weeks 4, 8, 12
Post-intervention testing 2-3 visits, 2-3 total hours	Repeat pre-intervention testing			
Washout	No testing or exercise			
Post-intervention acute exercise test or rest 1 or 2 visits 2-8 total hours	Overnight fast, 1 or 2 nights Acute exercise or rest Blood samples Muscle biopsies Adipose biopsies Match pre-intervention test diet	Up to 18 hr 45 min cycling 4-6 2 2	Up to 18 hr ~45 min RE 3-4 2 2	Up to 12 hr 1 1 1

BP = blood pressure, ECG = electrocardiogram, DXA = dual-energy x-ray absorptiometry, CPET = cardiopulmonary exercise test, VO_2peak = peak oxygen consumption, 1RM = 1-repetition maximum, NSAID = non-steroidal anti-inflammatory drugs

* The number of screening and pre-intervention study visits will be allowed to vary across Clinical Sites to accommodate the schedules of participants

** The CPET/VO_2peak test is a screening test to determine eligibility and a pre-intervention test to measure cardiorespiratory fitness

2 INTRODUCTION

2.1 BACKGROUND AND STUDY RATIONALE

There is a wealth of evidence from preclinical and clinical research to support the beneficial effects of PA on the structure and function of multiple physiological systems. Perhaps the strongest "A-level" evidence (i.e., large RCT) for the potent therapeutic benefits of lifestyle interventions was from the Diabetes Prevention Program, which was stopped early because of efficacy. In 3,234 adults at risk for type 2 diabetes (**T2D**), lifestyle modification was more effective than Metformin and Placebo in preventing the progression to T2D. While PA level was ~6-fold higher in the Lifestyle group than in Placebo and Metformin groups throughout the study, the Lifestyle intervention also included a nutritional focus (reduced calorie, low fat) to achieve and maintain a 7% weight loss. **MoTrPAC will be the first large trial to isolate the effects of PA** on potential molecular mechanisms underlying such health benefits.

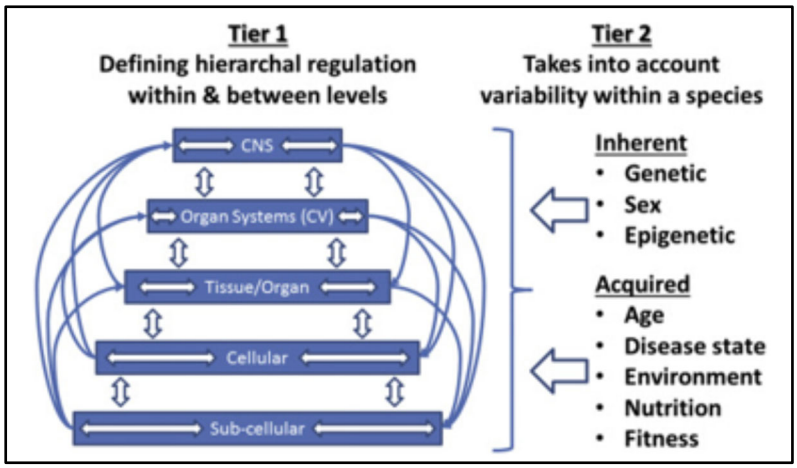
Three recent reviews summarized the extensive evidence for the health benefits of PA:

1. In 2007, the Department of Health and Human Services (**DHHS**) Secretary appointed 13 scientists to the Physical Activity Guidelines Advisor Committee to conduct a systematic review of the evidence on PA and health (US DHHS, 2008)[1]. The Federal Advisory Committee concluded there was strong evidence that physically active adults have lower rates of all-cause mortality, coronary heart disease (**CHD**), high blood pressure, stroke, T2D, metabolic syndrome, colon cancer, breast cancer, and depression. The Committee also concluded "*there remains a lack of data defining both the shape of the dose-response curve at the higher amounts and intensities of activity for most health outcomes and whether an upper limit of benefit exists.*" This report provided the evidence base for the *2008 Physical Activity Guidelines for Americans.*
2. Booth and colleagues [2] provided compelling evidence that PA contributes to the prevention of 35 chronic health conditions across the domains of metabolic syndrome, obesity, insulin resistance, T2D, non-alcoholic liver disease, cardiovascular diseases (**CVD**), cognitive functions and diseases, bone and connective tissue disorders, cancer, reproductive diseases, and diseases of the digestive tract, lungs, and kidneys. These authors addressed mechanisms by which PA contributes to the prevention of these conditions. In a few cases, subcellular mechanisms have been studied extensively (e.g., effect of PA on insulin signaling pathways in skeletal muscle that enhance glucoregulation and reduce risk for T2D), but in most cases either system-level mechanisms only are appreciated (e.g., prevention of CHD by PA is mediated by decreased systemic inflammatory cytokines) or mechanisms are unknown (e.g., mechanisms by which PA reduces postmenopausal breast cancer risk remain unclear).
3. Pedersen and Saltin presented evidence-based support for PA treatment of 26 different cardiovascular, metabolic, neurological, psychiatric, and pulmonary diseases.[3] Their overarching conclusion was that "*The evidence suggests that in selected cases exercise therapy is just as effective as medical treatment and in special situations more effective or adds to its effect. The accumulated knowledge is now so extensive that it has to be implemented."* Discussion of molecular mechanisms was considerable in some areas (e.g., insulin signaling and T2D) but superficial and speculative in others (e.g., cancer).

Several MoTrPAC investigators participated in activities that led to the MoTrPAC initiative, including working group meetings and the NIH workshop in 2014 on *Understanding the Cellular and Molecular*

Mechanisms of PA-induced Health Benefits. They contributed to the white paper that resulted from these activities.[4] As summarized in the white paper, PA is capable of disrupting homeostasis in essentially every organ system. Re-establishing homeostasis requires adaptive responses that are integrated across multiple systems and at multiple levels (i.e., Tier 1 in **Figure 3**), and also potentially modified by inherent or acquired factors (i.e., Tier 2 in Figure 3). It is the integrated adaptive responses that determine the impact on structure, function, and health status of a system. Despite the benefits of PA on numerous aspects of health, the molecular mechanisms by which such benefits occur and how they are influenced by such factors as age, sex, and genetic variation remain unclear. **By generating a map of molecular responses to PA, MoTrPAC will lay the foundation for a new era of biomedical research on Precision Exercise Medicine.**

Figure 3 [4]



The primary impetus for the MoTrPAC emanated from the broad consensus that many health benefits of PA are well documented, yet the mechanisms underlying their response are poorly understood. Although several small-scale studies have investigated various molecular exercise responses, a large unbiased and molecular mapping of exercise responses has never been undertaken. The molecular responses to exercise are likely numerous and diverse, including changes in the proteome, transcriptome, metabolome, lipidome, and epigenome within various tissues and cells.

Although MoTrPAC is clearly intended to produce novel data on various molecular transducers responses through 'omics' technologies, the Consortium will also perform key phenotyping of participants to characterize them and to measure their changes in key physiological outcomes linked to the molecular transducers. Cardiorespiratory fitness will be assessed by peak aerobic power (**VO_2peak**) during a graded cycling cardiopulmonary exercise test (**CPET**) and will be critical to characterize fitness levels of all participants and changes in cardiorespiratory fitness in the EE group. Muscle strength will be critical to assess in all participants and the responses to the RE intervention. Body composition, including fat-free mass, fat mass, and bone mass, will be assessed by DXA to yield key phenotypic data linked to a number of chronic disease risks, and will also provide key data on exercise intervention responses. Several other phenotyping variables, e.g., demographics, usual PA, health history, and medication history will be collected to link to and better interpret the molecular transducers responses to exercise.

2.1.1 RATIONALE FOR TIMING OF BIOSPECIMEN COLLECTION

The time course in which exercise induces changes in various types or classes of molecular transducers can vary from within the first few minutes of one acute bout of exercise to weeks or years of chronic exercise training. For example, post-translational modifications of proteins in muscle tissue (e.g., phosphorylation) or changes in circulating metabolites can occur very rapidly in response to acute exercise, whereas changes in gene expression may occur in the hours after a single exercise bout. Changes in protein concentrations may only be evident after weeks of exercise. Recent advances in several powerful high-throughput discovery approaches, including metabolomics, proteomics, genomics, transcriptomics, and epigenomics have now made it possible to perform profiling of many of these known (and possibly unknown) molecular signatures in response to acute and chronic exercise. Since it will be critical to decipher true exercise effects on these molecular transducers from other likely sources of biological and methodological variability, biospecimen collection in non-exercise Control participants over the course of minutes, hours and weeks will allow estimation of biological, seasonal, circadian and methodological variability. MoTrPAC will be the first large-scale trial to leverage these latest technologies to better understand the various and diverse molecular signatures in response to exercise anchored to physiological responses and adaptations underlying long-term health benefits.

One of the challenges for MoTrPAC is the timing of the biospecimen collection to capture as much information about the molecular responses to exercise as possible, while being sensitive to participant burden, as well as the practical considerations of cost and logistics. Although limited (***hence the impetus for MoTrPAC***), the current published literature suggests it will be important to collect biospecimens at the planned time points for the following purposes:

- **Fasted morning rested samples (blood, muscle, adipose)**
    - Will enable characterization of genetic and phenotypic molecular signatures unique to individuals and common among subgroups to explore the impact of the “starting material” on the acute molecular responses to exercise among RCT participants both before and after the intervention period, and the HAEE and HARE single bout responses:
        - Inter-individual response heterogeneity.
        - Commonalities and differences across age, sex, race, ethnicity, body composition, PA status, etc.
- **During and very early after exercise or no exercise**
    - ***Blood collection*** – ~20 min and ~40 min after initiation of exercise (EE group only) and ~10 min and ~30 min post-exercise, with comparative time points in no-exercise Controls – to capture secretion or appearance of exerkines, changes in metabolites, etc.
    - ***Muscle collection*** – ~15 to ~30 min post-exercise or post-rest – to capture early changes in the metabolome, phospho-proteome, and other posttranslational modifications of proteins etc. in response to exercise
    - ***Adipose collection*** – ~30 to ~45 min post-exercise or post-rest – to capture early changes in the transcriptome, metabolome, lipidome, phospho-proteome, and other posttranslational modifications of proteins etc. in response to exercise
- **~3.5 to 4 hours after exercise or no exercise**
    - ***Blood collection*** – to capture the longer-term secretion or appearance of exerkines, changes in metabolites, the content of exosomes, epigenetic modifications in blood cells, changes in

gene expression or the proteome in blood cells, changes in cell-free DNA, etc. in response to exercise

- ***Muscle collection*** – to capture epigenetic modifications, changes in gene expression, longer term changes in the metabolome, phospho-proteome, proteome, indices of protein and lipid metabolism, etc. in response to exercise
- ***Adipose collection*** – to capture epigenetic modifications, changes in gene expression, changes in the lipidome, metabolome, phospho-proteome, proteome, etc. in response to exercise

- **~24 hours after exercise or no exercise**
  - ***Blood collection*** – to capture the sustained or delayed appearance of potential molecular transducers in response to exercise
  - ***Muscle collection*** – to capture the sustained or delayed appearance of potential molecular transducers in response to exercise
  - ***Adipose collection*** – to capture the sustained or delayed appearance of potential molecular transducers in response to exercise

Mapping the time course of molecular responses will advance our understanding of processes that may play key roles in influencing: **1)** tissue adaptations, e.g., cell signaling, protein metabolism (synthesis, catabolism), indices of subcellular processes (e.g., mitochondrial biogenesis, ribosome biogenesis), angiogenesis, lipid trafficking, etc. and **2)** potential systemic factors driving inter-organ cross-talk.

We have included several references to support our plan to perform blood sampling and tissue biopsies over the time course from pre-exercise to ~24 hours post-acute exercise, and following training.

- Molecular expression can change from resting conditions immediately [5-10], hours [5-45], to even days after exercise [13, 18, 21, 22, 38, 41, 42, 46-50]
- The time course for each molecule is likely unique [6, 19, 20]
- The type and intensity of exercise will have an impact
  - Resistance versus endurance exercise [6, 19-22, 24]
  - High intensity versus low intensity exercise [7, 9, 13]
- Age and sex will influence the molecular responses to exercise [11, 25-34, 45]
- Training status will affect the acute exercise responses [10, 15, 17, 23, 24, 32, 33, 40]

Because the MoTrPAC program seeks to study women and men across a wide age spectrum in both the sedentary and trained states (for both RE and EE), it is impossible to predict the exact time points that will provide the most information regarding the molecular response to exercise.

Similarly to the adult data and biological samples, MoTrPAC stores and analyzes data and biological samples collected from a **separate protocol involving pediatric participants** enrolled at the University of California at Irvine. The separate Irvine pediatric protocol includes two interrelated studies: 1) a cross-sectional study, in which molecular transducers (obtained from blood sampling) are measured in response to an acute exercise challenge; and 2) an EE training study, in which the impact of chronic EE training on molecular transducers is measured. Pediatric data are transferred from the University of California at Irvine to the MoTrPAC DMAQC Center and biological samples are transferred to the

Biological Sample Repository, and will undergo analysis by the MoTrPAC CAS and the BIC. The Biological Sample Repository is blinded to identifying information about participants. Biological samples collected from pediatric participants at Irvine undergo molecular phenotyping, including metabolomic, lipidomic, proteomic, epigenetic, transcriptomic, and genomic analyses. These assays are done at the MoTrPAC CAS.

2.2 RISK/BENEFIT ASSESSMENT

2.2.1 KNOWN POTENTIAL RISKS

Please refer to Section 8.2 Safety and Other Assessments for known potential risks.

2.2.2 KNOWN POTENTIAL BENEFITS

The potential benefits to an individual participant in the study are not known. The majority of the participants will undergo EE or RE, which can trigger many beneficial adaptations, but individual responses and adaptations vary. All participants will acquire knowledge about their physiological status (e.g., muscle strength, cardiorespiratory fitness, body composition, etc.) that may be of personal benefit.

The potential benefits of the study to scientific and lay communities are expected to be vast. The near-term goal of MoTrPAC is to characterize the molecular responses to exercise as a strategy for meeting the longer-term goals, which are to better understand the molecular mechanisms by which PA improves health. MoTrPAC is expected to lead to cutting-edge translational research that will expand knowledge on how exercise can be used effectively to prevent or treat health conditions and diseases processes in women and men across the age spectrum.

2.2.3 ASSESSMENT OF POTENTIAL RISKS AND BENEFITS

MoTrPAC is not a therapeutic clinical trial and its aim is not to demonstrate efficacy; accordingly, there are no study-stopping criteria based on efficacy and there will be no formal assessment of benefits of the intervention. According to the 2018 Physical Activity Guidelines for Americans, the benefits of exercise far outweigh the risks [51]. However, the risks associated with participation in MoTrPAC are not only those related to exercise, but also those associated with tissue biopsies, blood sampling, maximal exercise testing, DXA measurements, and other risks described in section 8.2. These risks do not outweigh the advances in knowledge that MoTrPAC is expected to generate regarding the mechanisms by which exercise improves health. Safety and participation (recruitment progress, biopsy collection, and intervention adherence) outcomes will be monitored to determine whether the study approach should be modified.

3 OBJECTIVES AND ENDPOINTS

The overarching goal of MoTrPAC is to assemble maps of the molecular changes that occur in response to exercise in participants across the lifespan. We accomplish this goal by characterizing the molecular responses to an acute bout of EE or RE in phenotypically well-described sedentary and HA individuals at baseline, and after EE and RE interventions in the sedentary participants. The resulting *molecular maps* will be an invaluable resource, enabling the scientific community to accelerate both mechanistic research and clinical trials on the health benefits of exercise.

MoTrPAC uses the data generated to address questions about the potential molecular mechanisms that mediate the health benefits of exercise. Among them are the following:

1. Do the molecular responses differ between RE and EE?
2. Do the molecular responses to acute exercise differ in the untrained and trained states?
3. Are molecular signatures (genomic, transcriptomic, epigenomic, metabolomic, lipidomic, proteomic) at baseline and/or in response to acute exercise predictive of the inter-individual heterogeneity in exercise training responses and/or adaptations? And are these responses and adaptations influenced by non-modifiable factors (e.g., sex, age)?
4. Which molecular response patterns are most closely associated with exercise training-induced changes in phenotype (e.g., cardiorespiratory fitness, neuromuscular strength, body composition, etc.)?
5. Which previously unknown molecular response patterns reveal candidate pathways for future studies on the potential molecular mechanisms by which exercise training exerts benefits on health? Candidate pathways would be identified and pursued by the scientific community at-large via controlled and public access to the MoTrPAC data repository.

The major aims of MoTrPAC are:

1. To assess the response of molecular transducers to a single bout of EE or RE in both sedentary and trained (HAEE, HARE) individuals.
2. To assess the response of molecular transducers to a single exercise bout after a 12-week period of supervised EE or RE training.
3. To assess the relationships among the above-mentioned molecular transducers and phenotypic changes measured in the sedentary participants before and after the intervention period.

To achieve the aims participant data and biological samples are transferred from the clinical sites to the DMAQC and to the Biological Sample Repository, and later analyzed by the CAS and BIC.

4 STUDY DESIGN

4.1 OVERALL DESIGN

4.1.1 OVERVIEW

MoTrPAC is a mechanistic RCT to identify molecular transducers of PA. The overarching hypothesis is that there are discoverable molecular transducers that communicate and coordinate the effects of exercise, which may initiate processes ultimately leading to the health benefits of exercise to all organs of the body. Because this is a mechanistic trial, the main goal is not a health-related outcome. Rather, the goal is to generate a map of the molecular responses to exercise that will be used by the Consortium and by the scientific community at large to generate hypotheses for future investigations of the health benefits of PA. A total of approximately 1,980 sedentary, healthy adult women and men are randomized to EE (n=840), RE (n=840), or control (n=300) groups. The RCT is conducted in accordance with an intent-to-treat (**ITT**) design (Figure 1a). In addition, ~150 HAEE and ~150 HARE participants complete a single EE or RE bout as comparison groups (Figure 1b). All participants are recruited from the community, administered the informed consent form (**ICF**), and screened for eligibility. Randomization is stratified according to clinical site and sex, with one exception: during the Vanguard Phase (section 4.1.2), stratification will be on clinical site only. The sample size for the Vanguard Phase is not sufficiently large to stratify on sex and guarantee sites will gain experience with all experimental conditions.

In summary, for the RCT, qualifying sedentary participants are randomized to one of the following arms:

- Acute EE testing and ~12 weeks of EE training (n=840)
- Acute RE testing and ~12 weeks of RE training (n=840)
- Control involving no acute exercise testing or training (n=300)

After approximately 12 weeks of intervention, the RCT participants complete follow-up phenotyping assessments, and those randomized to EE or RE repeat the acute exercise bout with biospecimen collection. The Control group does not take part in the acute exercise testing, however this group does take part in PA monitoring, study phenotyping procedures, and biospecimen collection. Biospecimens will be analyzed using several technologies, including but not limited to genomics, epigenomics, transcriptomics, metabolomics, lipidomics, and proteomics.

Comparison groups of HAEE (n=~150) and HARE (n=~150) participants are recruited. After screening and familiarization, they will engage only in a single acute exercise testing session with biospecimen collection (Figure 1b), matching the EE and RE acute exercise protocols used with the sedentary participants.

4.1.1.1 INITIAL SCREENING, RANDOMIZATION, AND FAMILIARIZATION

Sedentary individuals who meet eligibility criteria and consent to participate complete baseline assessments during a series of clinic visits and/or phone or video calls (**Table 3a**). At the conclusion of the Screening Phase, they are randomized to EE, RE, or Control, and randomized to one of three biospecimen collection profiles (Figure 2a). Randomization is followed by a familiarization phase specific to each randomization arm. After a washout period (no testing, no exercise), participants in EE and RE complete the pre-intervention acute exercise test with biospecimen collections, and those randomized

to Control complete the biospecimen collections. The post-intervention testing visits are summarized in Table 3b.

Table 3a. Pre-intervention study visits

Study Visits	Activities	Approximate Participant Time	Notes
Visit 1 Screening	Orientation and informed consent	90 min	
Visit 2 Screening and pre-intervention testing*	Weight, height Medication inventory Medical history, exam Fasting blood draw Resting BP Resting ECG	5 min 15 min 20 min 10 min 15 min 15 min	 Must precede any testing Must precede any testing Must precede CPET Must precede CPET
Visit 3 Screening and pre-intervention testing*	Total body DXA Familiarization to cycle ergometer and VO_2 measurement CPET with monitoring of ECG and BP and measurement of VO_2peak** Orient to PA monitor	20 min 15 min 45 min 10 min	 Must precede CPET After completion of CPET
Visit 4 Screening and pre-intervention testing*	Grip strength Isometric knee extension strength Questionnaires	10 min 20 min 30 min	
RANDOMIZATION			
Visit 5 Familiarization	EE – practice cycling test RE – practice exercises	30 min 45 to 60 min	
Visit 6 Familiarization	EE – practice cycling test RE – practice exercises, perform 1RM	30 min 45 to 90 min	
Visit 7 Familiarization	RE – practice 10-RM work loads	45 to 90 min	
Visit 8 Acute exercise test/rest	2 muscle biopsies 2 adipose biopsies 3 to 6 blood samples ~45 minutes EE or RE or rest (control)	4 to 8 hours	Participants randomized to Late biospecimen profile have 2 visits on consecutive days (24-hr post-exercise samples)

BP = blood pressure, ECG = electrocardiogram, CPET = cardiopulmonary exercise test, DXA = dual-energy x-ray absorptiometry, VO_2peak = peak oxygen consumption, PA = physical activity, RM = repetition maximum, EE = endurance exercise, RE = resistance exercise

* **Number of screening and pre-intervention testing visits and activities performed at each visit will vary to accommodate the schedules of participants.**

** **CPET is a screening test to evaluate eligibility and a pre-intervention test to measure VO_2peak**

Table 3b. Post-intervention study visits

Study Visits	Activities	Approximate Participant Time	Notes
Visit 9 Post-intervention testing*	Total body DXA CPET/VO_2peak	20 min 45 min	
Visit 10 Post-intervention testing*	Grip strength Isometric knee extension strength Questionnaires	10 min 20 min 30 min	
Visit 11 Acute exercise test/rest and biospecimen collection	2 muscle biopsies 2 adipose biopsies 3 to 6 blood samples ~45 minutes EE or RE or rest (control)	4 to 8 hours	Participants randomized to Late biospecimen profile will have 2 visits on consecutive days (to collect 24-hr post-exercise samples)

DXA = dual-energy x-ray absorptiometry, VO_2peak = peak oxygen consumption, EE = endurance exercise, RE = resistance exercise
*** Number of post-intervention testing visits and activities performed at each visit will vary to accommodate the schedules of participants**

4.1.2 COVID-19 PANDEMIC

Recruitment of MoTrPAC participants started in September 2019 and all clinical sites were operational by November 2019. Recruitment was suspended on March 16, 2020, because of the COVID-19 pandemic. Over the subsequent 2 to 3 weeks, clinical visits were suspended at all clinical sites per local guidelines. When it became apparent the suspension would last more than a few weeks, all participants enrolled in MoTrPAC at the time of the suspension who had not yet completed follow-up visits were administratively withdrawn from the study. It was decided **the planned Vanguard Phase of the study (below), which would have included the first 288 randomized participants, will not be completed and the analytical plan (section 4.1.3.3) will not be carried out.** Rather, the valuable information generated during the abbreviated Vanguard Phase was used to modify the protocol (v2.0) and the data generated will be integrated into the MoTrPAC database.

4.1.3 VANGUARD PHASE

The DSMB, in conjunction with NIH leadership, requested that the protocol for MoTrPAC include an initial Vanguard phase for the sedentary randomized participants. The Vanguard does not involve the HA groups. The purpose of the Vanguard is to assess feasibility related to recruitment, participant and staff burden, and adherence to the protocol within randomized participants. There are two considerations for the design of the Vanguard. **First**, within the sedentary EE, RE, and Control arms, participants are randomized with equal probability to one of four biospecimen sampling profiles (Early, Mid, Late, All; Figure 2a) for the muscle, adipose, and blood collections.

Second, the Vanguard has two distinct phases. *Phase I* involves *evaluation of the protocol from recruitment through the first acute test*. It consists of 20 EE, 20 RE, and 8 Controls at each of the adult CCs (N=48) for a total of 120 EE, 120 RE, and 48 Controls (N=288). These sample sizes allow at least 1 Control and 2 EE and RE participants to be measured at each temporal mapping time point for CCs that have 2 sites. *There is no pause in recruitment after reaching this number*. Rather, this approach delineates the N we will use within CCs and sites to assess feasibility through the acute test in the following areas, among others:

- Calculate recruitment yields [48 (number in the Vanguard for each of the 6 adult CCs)/number contacted during phone screen at each of the 6 adult CCs].
- Perform qualitative reviews of recruitment information encompassing topics such as costs, reasons for refusal, etc.
- Calculate overall adherence to the study visits and the protocol for the acute test in which X denotes successful completion of the visits and/or completion of the specimen collection: X/288 for baseline; X/72 at 15-45 min, 4-hour, 24-hour and all time points, X/120 for EE and RE adherence across all temporal profiles, etc.
- Examine retention at various points in testing conducted prior to training.

If burden is not excessive, and adherence to the acute sampling assessments is acceptable, then there may be no further changes to the protocol for the initial acute test. However, *if burden is high, further modification of the protocol* based on analyses of Vanguard data may be implemented, such as relaxing the exclusion criteria. *It is important to emphasize that, although there is an analytical plan to evaluate the Vanguard and to provide objective data related to possible modifications of the protocol, we are not setting any prespecified criteria that would trigger such a response. Rather, MoTrPAC investigators, in collaboration with the DSMB and NIH, will use these data and their clinical experience to arrive at a collective decision on any protocol changes.*

Phase II will involve *following those participants in the Vanguard sample throughout the intervention period and for the remainder of assessments* to provide insight on feasibility, adherence, and retention. Conclusion of data collection for this phase will occur approximately 12 weeks after the last Vanguard participant is randomized.

4.1.3.1 JUSTIFICATION FOR THE VANGUARD SAMPLE SIZE

We chose the sample size for the Vanguard based on several considerations. First, as described in the timeline section, it is important to complete the Vanguard within the first 6- to 7-months of recruitment. Second, we wanted to ensure that clinical sites (three of the CCs have multiple performance sites) obtain experience with participants in each condition of the randomized study and at each temporal mapping time point. With the chosen sample size per CC, this will be possible with the exception of the Control group at the CC with three sites. If we assume that we will randomize 55-60 sedentary participants per CC during the initial 7 months of MoTrPAC, this would mean that we would randomize 24 or 25 participants to EE and RE conditions and 10-12 to Control because randomization is done at an approximate ratio of 8 EE: 8 RE: 3 Controls. With 10 possible Controls and 4 temporal sampling profiles, it is possible that a CC with 3 sites may not get Vanguard experience with controls in all temporal sampling profiles (see randomization section, below). Third, we want to balance recruitment across CCs to allow for an equally weighted comparison of experiences across CCs. We achieve this balance by considering the Vanguard to consist of the first 48 people randomized within each CC (See **Table 4**) that have a single site and divide this number across CCs that have 2 or 3 sites, respectively. Note that we will randomize more to some cells in the process of reaching these cell-specific minimum numbers.

Table 4. Allocation of participants by Clinical Site during the Vanguard Phase to the endurance exercise (EE), resistance exercise (RE), and control groups and Early, Mid, Late, and All biospecimen sampling profile

	Single-site Centers				Centers with 2 Sites				Center with 3 Sites			
	Early	Mid	Late	All	Early	Mid	Late	All	Early	Mid	Late	All
EE	N=5	N=5	N=5	N=5	N=2 or 3	N=2 or 3	N=2 or 3	N=2 or 3	N=1 or 2	N=1 or 2	N=1 or 2	N=1 or 2
RE	N=5	N=5	N=5	N=5	N=2 or 3	N=2 or 3	N=2 or 3	N=2 or 3	N=1 or 2	N=1 or 2	N=1 or 2	N=1 or 2
Control	N=2	N=2	N=2	N=2	N=1	N=1	N=1	N=1	N=0 or 1	N=0 or 1	N=0 or 1	N=0 or 1

4.1.3.2 VANGUARD RANDOMIZATION

The goal is to have Vanguard recruitment completed within approximately 6-7 months. We will not stratify the Vanguard recruitment on sex, but only on site. Stratifying on sex for the Vanguard period could cause the Vanguard recruitment to stretch out longer than 6-months due to using two parallel randomization lists (i.e., one for each sex) within site. To conceal allocation, we will perform this randomization in blocks that are slightly larger than the total Vanguard participants within a site.

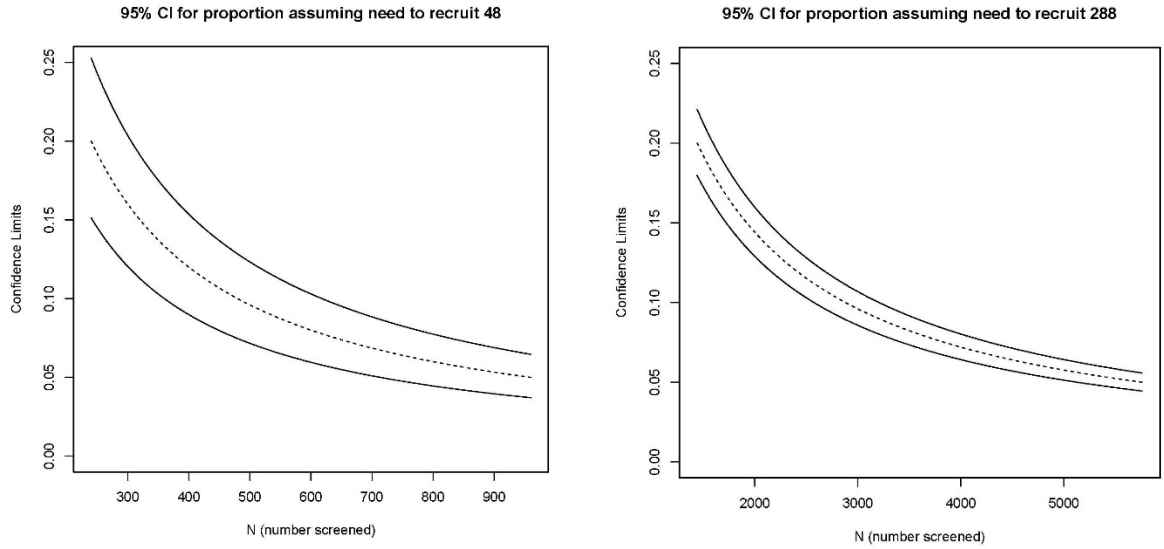
4.1.3.3 VANGUARD ANALYTIC PLAN

Analyses for the Vanguard will be conducted at both the Center level (N=48 randomized) and in the combined dataset (N=288 randomized). However, the appropriate denominator for specific metrics may vary.

The proposed analyses described below assume that we have available the numerator and denominator needed to estimate a specific proportion. We will use exact confidence intervals (**CIs**) or inverted score tests for proportions to estimate the confidence limits for the proportions listed among the metrics for the Vanguard.[52] The binomial distribution can be skewed when the sample size is small; additional correction for the skewness may be needed, especially when the observed proportions are further away from 0.5 for metrics computed at the site level. These approaches will also be used for categorical variables with more than 2 levels and for between Center comparisons.[53] The Vanguard data combined across Centers will be used to evaluate sex-, race-, and age-specific metrics; as shown in the Figures below, estimates of these metrics within Centers are likely to be imprecise (e.g., they will have wide confidence intervals), but will be calculated. Additional analyses will include time from study eligibility determination to randomization, which can include study exclusion prior to randomization as a competing event, in which case, we will estimate cause-specific cumulative incidence rates.

Important in these analyses are the width of confidence intervals that will be generated on estimates such as recruitment yields at each Center (e.g., 48/# screened) and overall (288/#screened), and estimates of adherence at the different temporal sampling time points (e.g., X/72 for each temporal sampling profile). **Figure 4** provides plots of confidence intervals of the estimated recruitment yield at an individual Center and overall as a function of the number of participant screened. These sample sizes will provide confidence intervals with widths less than approximately 5% for individual Centers and 2% for overall when the proportion randomized to screening is 10% or less.

Figure 4. 95% CI on Recruitment Yield



For adherence to procedures done on all participants (e.g., baseline blood), the width of 90% confidence intervals around the proportion is less than 8% overall when the proportion adherent is >0.80, but would be approximately 20% within Centers (**Table 5**); thus, the Vanguard will not provide very precise estimates of Center-specific adherence and retention. If we observe percent adherence of 90% or above, the lower bounds of the confidence intervals will be approximately 80% or higher, even at the Center level. There also may be interest in testing for overall differences between proportions observed across all Centers.

Table 5. Width of confidence intervals for estimated proportions based on sample sizes

Sample Size	Observed Proportion	90% Confidence Interval
48 estimated overall adherence at a Clinical Center	0.80	0.68, 0.89
	0.85	0.74, 0.93
	0.90	0.80, 0.96
	0.95	0.86, 0.99
72 estimated adherence for each biospecimen sampling profile	0.80	0.71, 0.87
	0.85	0.76, 0.91
	0.90	0.82, 0.95
	0.95	0.89, 0.98
120 EE or RE participants	0.80	0.73, 0.86
	0.85	0.79, 0.90
	0.90	0.84, 0.94
	0.95	0.90, 0.98
288 all participants	0.80	0.76, 0.84
	0.85	0.81, 0.88
	0.90	0.87, 0.93
	0.95	0.92, 0.97

For testing adherence between EE, RE and Control groups, across all temporal profiles, we assume that one group (EE, RE) has 90% or 95% adherence. The selected sample sizes provide at least 80% power to detect a significant difference between groups when the other group (EE, RE) has an adherence rate of 78% or lower (75% if looking at EE or RE vs Control), assuming alpha=0.10 and 90% adherence. If one group (EE, RE) has 95% adherence, we have 80% power to detect a significant difference between groups when the other group (EE, RE) has 86% adherence (84% when looking at EE or RE vs Control), or lower, assuming alpha=0.10.

For testing adherence between temporal profiles, we assume one group has 90% or 95% adherence. We have at least 80% power to detect a difference between temporal sampling groups when another group has an adherence rate of 74% or lower, assuming alpha=0.10 and 90% adherence. If one temporal sampling group has 95% adherence, then we have 80% power to detect a significant difference between groups when the other group has 82% adherence, or lower, assuming alpha=0.10.

4.1.3.4 VANGUARD TIMELINE

We believe that it is important to complete the Vanguard data collection within the first 6-7 months of recruitment. With randomization of ~1,980 participants across approximately 40 months, it is necessary to randomize at least 50/month assuming uniform recruitment. Assuming a slower start-up in the initial month, we assume that we would randomize 30 participants (5/Center) in the initial month and approximately 50-55/month (8 to 9 per Center) after that time point. Thus, after 7 months of operations, we will have randomized approximately 330-360 (55-60 per Center) participants. It is expected that analyses and vetting of Phase I Vanguard results will take approximately 2-3 months, indicating that any recommendations based upon the Vanguard period should be available within 10 months of the initiation of recruitment.

4.2 SCIENTIFIC RATIONALE FOR STUDY DESIGN

See section 2.1 Study Rationale.

4.3 JUSTIFICATION FOR DOSE

The prescribed doses of EE and RE have been demonstrated in numerous published studies to result in myriad physiological adaptations to exercise training. The current study aims to reveal molecular responses that may drive these adaptations.

4.4 END OF STUDY DEFINITION

It is expected that all activities related to the clinical protocol (e.g., study visits, final entry and initial analysis of phenotyping tests and all clinical data, processing and analysis of biospecimens, quality assurance (**QA**) of data) will be completed in the final year of the award period. The analysis and publication of MoTrPAC data are expected to continue for many years by investigators internal and external to the Consortium.

5 STUDY POPULATION

5.1 INCLUSION CRITERIA

5.1.1 ADULT PARTICIPANT INCLUSION CRITERIA – SEDENTARY PARTICIPANTS

- Willingness to provide informed consent to participate in the MoTrPAC Study
- Must be able to read and speak English well enough to provide informed consent and understand instructions
- **Aged** ≥18 y
- Body Mass Index (**BMI**) ≥19 to ≤37.5 kg/m^2
- **Sedentary** defined as self-reporting no more than 1 day per week, lasting no more than 60 minutes, of regular (structured) EE [e.g., brisk walking, jogging, running, cycling, elliptical, or swimming activity that results in feelings of increased heart rate, rapid breathing, and/or sweating] or RE (resulting in muscular fatigue) in the past year
    - Persons bicycling as a mode of transportation to and from work >1 day/week etc. are not considered sedentary
- Leisure walkers are included unless they meet the heart rate, breathing, and sweating criteria noted above
- Willingness to include de-identified individual-level data at low risk of re-identification (e.g., non-genomic data) in the MoTrPAC open-access database
- Only one member of a household can participate

5.1.2 ADULT PARTICIPANT INCLUSION CRITERIA – HIGHLY ACTIVE PARTICIPANTS

- Willingness to provide informed consent to participate in the MoTrPAC Study
- Must be able to read and speak English well enough to provide informed consent and understand instructions
- **Aged** ≥18 y
- **BMI** ≥19 to ≤37.5 kg/m^2
- **Comparator Participants**
    - **HAEE:** defined as ≥240 minutes/week of ET for ≥1 year; this can include running, walking (brisk, power), cycling, elliptical, etc. which (at a minimum) results in increased heart rate, rapid breathing and sweating
        - Must include cycling at least 2 days/week
        - RT in the past year must be limited to ≤2 days/week of upper body RE and ≤2 muscle groups of upper body RE and ≤1 day/week of lower body RE
    - **HARE:** defined as RT of ≥3 upper and ≥3 lower body muscle groups ≥2 times/week for ≥1 year; using a prescription sufficient to increase strength and muscle mass

- ET in the past year must be limited to ≤90 minutes/week of vigorous EE, with no limit on cycling days per week

- **Elite or Competitive Athletes:** can be included, if they meet HAEE or HARE inclusion criteria
  - Potential participants are informed that use of performance enhancing drugs in the last 6 months is exclusionary
- Willingness to include de-identified individual-level data at low risk of re-identification (e.g., non-genomic data) in the MoTrPAC open-access database
- In addition to meeting HAEE or HARE inclusion criteria, all HA participants must meet all other exclusion criteria defined in this protocol

5.2 EXCLUSION CRITERIA

5.2.1 ADULT PARTICIPANT EXCLUSION CRITERIA

Exclusion criteria are confirmed by self-report (i.e., medical and medication histories reviewed by a clinician), screening tests performed by the MoTrPAC study team at each clinical site, and/or clinician judgement as specified for each criterion.

- **Diabetes (self-report and screening tests)**
  - Treatment with any hypoglycemic agents (self-report) or A1c >6.4% (screening test; may reassess once if 6.5-6.7%)
  - Fasting glucose >125 mg/dL (screening test; may reassess once)
  - Use of hypoglycemic drugs for non-diabetic reasons (self-report)
- **Abnormal bleeding or coagulopathy (self-report)**
  - History of a bleeding disorder or clotting abnormality
- **Thyroid disease (screening test)**
  - Thyroid Stimulating Hormone (**TSH**) value > 5.9 IU/mL
    - Individuals with hypothyroidism may be referred to their primary care provider (**PCP**) for evaluation and retested; any medication change must be stable for ≥3 months prior to retesting
    - Individuals with hyperthyroidism are excluded, including those with normal TSH on pharmacologic treatment
- **Pulmonary (self-report)**
  - Clinical diagnosis of Chronic Obstructive Pulmonary Disease (**COPD**)
- **Metabolic bone disease (self-report)**
  - History of non-traumatic fracture from a standing height or less
  - Current pharmacologic treatment for low bone mass or osteoporosis, other than calcium, vitamin D, or estrogen
- **Pregnancy (screening test) and pregnancy-related conditions (self-report)**
  - Pregnant – pregnancy test performed on day of DXA scan in women of child-bearing potential
  - Post-partum during the last 12 months

- Lactating during the last 12 months
- Planning to become pregnant during the participation period

- **Elevated blood pressure readings (screening test)**
  - Resting Systolic Blood Pressure (**SBP**) ≥150 mmHg or resting Diastolic Blood Pressure (**DBP**) ≥95 mmHg
  - Reassessment of BP during screening will be allowed to ensure resting values are repeatable
- **Cardiovascular (self-report, screening test, and clinician judgement)**
  - Congestive heart failure, coronary artery disease, significant valvular disease, congenital heart disease, serious arrhythmia, stroke, or symptomatic peripheral artery disease (self-report, screening test)
  - Specific criteria used to determine whether a volunteer can undergo the screening CPET follow the American Heart Association (**AHA**) Criteria [54]
  - Inability to complete the CPET
  - Reassessment of the CPET may be allowed under some circumstances (e.g., test was not a maximal effort)
- **Abnormal blood lipid profile (screening test)**
  - Fasting triglycerides >500 mg/dL
  - Low-density lipoprotein cholesterol (**LDL-C**) >190mg/dL
- **Cancer (self-report)**
  - History of cancer treatment (other than non-melanoma skin cancer) and not “cancer-free” for at least 2 years
  - Anti-hormonal therapy (e.g., for breast or prostate cancer) within the last 6 months
- **Chronic active or latent infection (self-report)**
  - Active or latent infections requiring chronic antibiotic or anti-viral treatment
  - Chronic active infection whether on chronic antimicrobials or not
  - Human Immunodeficiency Virus
  - Active hepatitis B or C undergoing antiviral therapy
  - Individuals successfully treated for hepatitis C and virologically negative for at least 6 months are not excluded
- **Liver enzyme tests (alanine transaminase, aspartate transaminase) (screening test)**
  - >2 times the laboratory upper limit of normal
  - Reassessment during screening may be allowed under some conditions (e.g., recent use of acetaminophen)
  - Individuals may be referred to their PCP for evaluation; any medication change must be stable for ≥3 months prior to retesting
- **Chronic renal insufficiency (screening test)**
  - Estimated glomerular filtration rate <60 mL/min/1.73 m^2 from serum creatinine (mg/dL) by the Chronic Kidney Disease Epidemiology Collaboration equation

- Reassessment may be allowed under some conditions (e.g., questionable hydration status or other acute renal insult)

- **Hematocrit (screening test)**
  - Hematocrit >3 points outside of the local normal laboratory ranges for women and men
  - Reassessment may be allowed under certain conditions
  - Individuals may be referred to their PCP for evaluation; any medication change must be stable for ≥3 months prior to retesting
  - Individuals with known thalassemia trait may be included (despite having >3 points outside of the local normal laboratory ranges), upon approval from their PCP or a hematologist
- **Blood donation (self-report)**
  - Whole blood donation in the last 3 months or plans for blood donation during the entire protocol period
  - Platelet or plasma donation in the last week or plans for platelet or plasma donation during the entire protocol period
- **Autoimmune disorders (self-report)**
  - Individuals receiving any active treatment (including monoclonal antibodies) within the last 6 months
- **Alcohol consumption (self-report)**
  - More than 7 drinks per week for women
  - More than 14 drinks per week for men
  - History of binge drinking (≥5 drinks for males or ≥4 drinks for females in a 2-hour period more than once per month)
- **Tobacco (self-report)**
  - Current smokers: any tobacco or e-cigarette/e-nicotine products
  - Former smokers:
    - Stopped smoking <10 years at time of screening for those with a ≥20 pack-year smoking history
    - Stopped smoking <5 years at time screening for those with a <20 pack-year smoking history
- **Marijuana (self-report)**
  - Self-reported use ≥3 days/week in any form
- **Shift workers (self-report)**
  - Night shift work in the last 6 months
  - Planning night shift work during the study period
- **Cognitive status (screening)**
  - Unable to give consent to participate in and safely complete the protocol, as based on the judgement of the local investigators
- **Psychiatric illness (self-report and screening test)**
  - Hospitalization for any psychiatric condition within one year (self-report)

- Center for Epidemiological Studies-Depression Scale (**CESD**) score ≥16 [55] (screening test)

- **Weight change (self-report)**
  - Weight change (intentional or not) over the last 2 months of >5% of body weight
  - Plan to lose or gain weight during the study
- **Lidocaine or other local anesthetic (self-report)**
  - Known allergy to lidocaine or other local anesthetic
- **COVID-19 infection**
  - Hospitalization for COVID-19 infection in the past 12 months
  - Individuals who tested positive for COVID-19 but were not hospitalized must be symptom-free at least 14 days
- **Other (clinician judgement)**
  - Genetic metabolic disorders that could effect metabolomic results (e.g., phenylketonuria)
  - Any other cardiovascular, pulmonary, orthopedic, neurologic, psychiatric, metabolic, or other conditions that, in the opinion of the local clinician, would preclude participation and successful completion of the protocol
  - Any other illnesses that, in the opinion of the local clinician, would negatively impact or mitigate participation in and completion of the protocol

5.2.2 EXCLUSIONS FOR MEDICATION USE

- Continuous use for 7 days or more of a new drug (prescription or over-the-counter; additional guidance in the MOP) in the last 3 months; eye and ear drops are allowed regardless of when they were started
- Dose change for any chronic-use drug in the last 3 months; changes in eye and ear drops are allowed
- Cardiovascular
  - Beta blockers and centrally acting anti-hypertensive drugs
  - Anticoagulants
  - Antiarrhythmics
  - Antiplatelet drugs (other than aspirin $\leq$100 mg/day)
- Psychiatric
  - Chronic use of medium- or long-acting sedatives and hypnotics, including all benzodiazepines; short-acting non-benzodiazepine sedative-hypnotics are allowed
  - Mood stabilizers
  - Antiepileptic drugs
  - Stimulants, Attention-Deficit/Hyperactivity Disorder (**ADHD**) drugs
  - Anti-psychotic drugs
- Pulmonary, inflammation
  - Chronic oral steroids; inhaled steroids are allowed

  - Burst/taper oral steroids more than once in the last 12 months; inhaled steroids are allowed
  - B2-agonists
    - allowed if on stable dose at least 3 months
- Genitourinary
  - 5-alpha reductase inhibitors
  - Daily phosphodiesterase type 5 inhibitor use
- Hormonal
  - Androgenic anabolic steroids
  - Anti-estrogens, anti-androgens
    - Estrogens and/or progestins used for reasons other than birth control or menopause
  - Growth hormone, insulin like growth factor-I, growth hormone releasing hormone
  - Any drugs used to treat diabetes mellitus or to lower blood glucose
  - Metformin for any indication
  - Any drugs used specifically to induce weight loss
  - Any drugs used specifically to induce muscle growth/hypertrophy or augment exercise-induced muscle hypertrophy
- Pain/inflammation
  - Narcotics and narcotic receptor agonists
  - Regular use of non-steroidal anti-inflammatory drugs (**NSAIDs**) or acetaminophen
  - Muscle relaxants ≥3 days per week
  - Oral/sublingual cannabidiol or similar in any formulation
- Other
  - Chronic systemic antimicrobials (antibiotic, antiviral, antifungal, antiparasite, etc) for any reason
  - High-potency topical steroids if ≥10% of surface area using rule of 9s
  - Continuous/chronic use of antibiotics or other anti-infectives for treatment or prevention
  - Monoclonal antibodies
  - Anti-rejection medications/immune suppressants
- Any other medications that, in the opinion of local clinicians, would negatively impact or mitigate full participation and completion

5.2.3 SPECIAL CONSIDERATIONS

NSAIDs or acetaminophen

- Consumption of NSAIDs or acetaminophen will be allowed occasionally (i.e., ≤2 days in a given week and not every week) for minor aches and pains (e.g., headache, menstrual cramps); participants should not take NSAIDS for 7 days prior to, during, and 3 days following the

biospecimen collection period; participants should not take acetaminophen for 7 days prior to and during the entire biospecimen collection period

Aspirin

- Consumption of low dose aspirin chronically (e.g., <100 mg/d) will be allowed; participants should not take aspirin for 7 days prior to, during, and 3 days following the biospecimen collection period

Vitamins, minerals, and non-Food and Drug Administration (FDA)-regulated supplements

- Vitamins, minerals, and non-FDA-regulated supplements will not be exclusionary but will be recorded and monitored
    - Individuals taking very high dose vitamin D (>50,000 IU per week or equivalent) at the time of screening will not be excluded, but must reduce the dose to ≤50,000 IU per week (or equivalent) throughout the study.
    - Individuals taking 1 mg or more of biotin (B7) supplements per day should not take the biotin supplement for 72 hours prior to any blood draw.

5.3 LIFESTYLE CONSIDERATIONS

Certain lifestyle factors such as PA level, alcohol intake, and body weight regulation are considerations for eligibility, and for each there are well-defined criteria for inclusion/exclusion. Free living PA and body weight will be monitored during the study using objective methods (accelerometry to monitor free living activity and scales to assess body weight).

5.4 SCREEN FAILURES

Participants are considered screening failures if they do not meet all defined study eligibility criteria.

5.5 STRATEGIES FOR RECRUITMENT AND RETENTION

5.5.1 RECRUITMENT STRATEGIES

Effective recruitment and retention of study participants is critical for the success of MoTrPAC. The recruitment goal of MoTrPAC is to enroll and randomize ~1,980 sedentary adults (aged 18+ years) and to enroll an additional ~300 HA adults.

All clinical sites have been successful in recruiting volunteers to participate in rigorous exercise intervention trials. Each site has a site-specific recruitment plan to accommodate the variability across locations in catchment area characteristics, media market outlets, and pool of potential participants. The recruitment plans utilize a variety of approaches directed at population-based recruitment within the catchment area, including direct mass mailing of study ads/brochures, radio and television ads, flyers/posters at community events, use of social media, and newspaper advertisements. A commercial recruiting mechanism, such as TrialFacts or TrialSpark, may be utilized across all clinical sites. As another centralized recruitment mechanism, the MoTrPAC website will have a link to a short questionnaire for potential volunteers. Contact information for individuals who live within 40 miles of a clinical site will be provided to the clinical site team.

5.5.2 UNDERREPRESENTED MINORITIES AND AGE GROUPS

MoTrPAC targets recruiting at least 30% racial and ethnic minorities and 50% females. In concert with the Clinical Center Recruitment and Retention (**CCRR**) subcommittee, each site develops and implements an outreach strategy for recruitment of underrepresented minorities. Additionally, the overall study targets are to achieve approximately balanced recruitment across the age groups of 18-39, 40-59, and 60+.

5.5.3 RETENTION STRATEGIES

MoTrPAC employs several retention promotion strategies for both the sedentary Control group and the EE and RE training groups, including: **1**) Provision of health-related study results to participants; **2**) Promoting easy access to and accommodating study assessment visit times; **3**) Maintaining regular contact with Control participants throughout the intervention period and providing Control participants with an individualized exercise prescription, guidance and education materials following their participation; and **4**) Contact of the participant or the participant's listed emergency contact, if necessary.

We employ the following strategies to promote retention to the protocol in terms of completion of assessment visit attendance.

- **Participant-staff relationship**. A key element contributing to continued commitment to the trial involves fostering positive, respectful relationships between study participants and individual members of the staff.
- **Participant-staff communications**. Good and consistent communication is essential. We provide clear instructions and friendly and individualized interactions.
- **Convenience and accessibility**. An easily accessible clinic location and convenient clinic hours all serve to facilitate study adherence. Clinical sites make study visits as easy as possible for participants, a factor critical to the success of the study. All sites will take steps to ensure that clinic attendance is not compromised by unsuitable hours of clinic operation, or any similar circumstance. Some sites may also have mechanisms to provide transportation to the clinic assessment visits.
- **Time in clinic**. Total clinic visit time is kept to a minimum. If waiting is necessary, the situation is explained to the participant and, if possible, an offer is made for the participant to see another staff member, or to reschedule the visit. On the other hand, participants are not rushed or made to feel unwelcome. Clinic staff is trained to take time to visit with and listen attentively to participants.

5.5.4 PARTICIPANT REMUNERATION

Remuneration for time and travel will be provided to participants. Sedentary participants (RE, EE, Control) will receive up to $2,000 and Highly Active (HAEE, HARE) participants will receive up to $1,000. The remuneration plan is fair and commensurate with the study demands and time commitment, is consistent across all adult clinical sites while permitting a modest degree of flexibility to account for site-specific factors (i.e., travel time, parking), and is prorated to provide partial payment for completion of discreet study events, milestones or phases**.**

5.5.5 MONITORING RECRUITMENT AND RETENTION

Screening, recruitment and retention yields will be continuously monitored, and reported to each clinical site on a regular basis.

5.5.6 EFFORTS TO MAINTAIN CONTACT WITH NON-ADHERENT PARTICIPANTS

MoTrPAC has the goal of maintaining some form of contact (e.g., phone, e-mail) with participants who are unable to continue full engagement in the study and to foster some form of continued contact (e.g., even an agreement to allow future contact) with participants who are inactive in the study interventions. The greatest importance is given to attending assessment visits; even participants who are unwilling to continue attending intervention sessions are encouraged to attend the assessment visits.

6 STUDY INTERVENTION

6.1 STUDY INTERVENTIONS ADMINISTRATION

Participants randomized to EE or RE first engage in a single acute exercise test of either EE (on a cycle ergometer) or RE, consistent with their random assignment. The Control group does not engage in any acute exercise testing protocol.

Prior to, during, and following these acute exercise tests in EE and RE participants, blood, muscle and adipose samples are collected as described in Figure 2a and in Section 8. For participants in the Control group, blood, muscle and adipose samples are collected following the same time sampling scheme before the intervention; only single blood, muscle, and adipose samples are collected after the intervention period. Each RCT participant is randomized to 1 of 3 biospecimen collection profiles (see Figure 2a in Section 1.2).

Following the pre-intervention acute exercise test and biospecimen collections, participants complete approximately 12 weeks of either EE or RE training as described below or continue their sedentary status as a Control participant. All participants are instructed to not change their levels of free-living PA outside of the two active interventions, or change their habitual diet during the intervention period. This will result in little to no change in body weight during the study period (although changes in body composition are likely to occur).

To help achieve these goals, EE and RE participants are weighed and regularly queried as to whether there has been any change in their level of free-living exercise participation (i.e., outside of the exercise training sessions). Non-compliance to remaining weight stable (+/-5% of baseline weight) and not engaging in structured exercise outside of the supervised exercise sessions triggers an action plan to assist participants in achieving these goals. All participants (EE, RE, Control groups) will complete a health status update at ~4-week intervals during the intervention, which will query changes in health and medications.

At the conclusion of the intervention period, EE and RE participants repeat the acute exercise bout with biospecimen collection, at each individual's initial randomized assignment to collection profile (Early, Mid, Late). RCT participants randomized to no-exercise Control undergo a single, fasted morning blood draw, muscle biopsy, and adipose biopsy at the conclusion of the intervention period (see Figure 2a in Section 1.2).

6.1.1 ENDURANCE TRAINING

Participants randomized to EE engage in three center-based EE sessions each week for approximately 12 weeks; each session lasts roughly 75 minutes with a 50-60 minute stimulus phase and the remaining time being used to warm up and cool down. Each session will be split between cycle ergometer and treadmill exercise; other modes (e.g., elliptical machine) will be allowed in some circumstances. During all sessions, heart rate is monitored to ensure that participants maintain the target exercise intensity during training. Intensity is set as a percentage of heart rate reserve and increases during the intervention: 60% ± 5 bpm for weeks 1-4; 70% ± 5 bpm for weeks 4-8; 75% ± 5 bpm for weeks 9-10; and 80% ± 5 bpm for weeks 11-12. Self-reported perceptual data are recorded periodically during exercise sessions, which are used to track the subjective experience of participants and in interpreting adherence data. At approximately 4-week intervals during the interventions (i.e., weeks 4, 8 and 12), participants will wear an activity monitor for ~7 days. The distribution and return of monitors will occur at exercise training visits.

6.1.2 RESISTANCE TRAINING

Participants randomized to RE engage in three center-based RE sessions each week for approximately 12 weeks; each session lasts roughly 75 minutes with a 50-60 minute stimulus phase and the remaining time being used to warm up and cool down. Exercise sessions are recommended to be held on non-consecutive days to allow for adequate rest. The prescription includes 3 or 4 sets of 8 exercises performed at a 10-repetition maximum (**10-RM**) intensity (i.e., able to perform only 8 to 12 repetitions per set). Load increases when a participant is able to perform 12 repetitions for 2 of 3 sets of an exercise. Heart rate is monitored continuously during exercise sessions. Self-report perceptual data are recorded periodically during exercise sessions, which are used to track the subjective experience of participants and in interpreting adherence data. At approximately 4-week intervals during the interventions (i.e., weeks 4, 8 and 12), participants will wear an activity monitor for ~7 days. The distribution and return of monitors will occur at exercise training visits.

6.1.3 CONTROL GROUP

Participants randomly assigned to the Control group do not engage in any exercise intervention during the 12-week intervention period. Control participants are instructed to not change their PA levels. The importance of remaining weight stable throughout the study period is emphasized. Two steps are implemented with the Control group to maintain contact and to monitor changes in structured exercise or dietary patterns. First, Control participants are called about every two weeks to query about any changes in their structured exercise or diet patterns. Second, the Control participants will come to the clinic for a brief visit approximately every 4 weeks to have the data from the activity monitor downloaded and to have their weight checked. At these visits, participants will again be queried about any changes in their structured exercise or dietary patterns and changes in health status or medication use. Participants who are identified as changing either their level of structure exercise or dietary patterns that could result in an improvement in fitness or weight loss are reminded about the importance of the study goals and provided an action plan to assist in correcting the problem area(s).

6.2 MEASURES TO MINIMIZE BIAS: RANDOMIZATION AND BLINDING

6.2.1 MEASURES TO MINIMIZE BIAS

The MoTrPAC CCC and BIC take a vigorous lead in assuring the quality of study databases. The quality and eventual acceptance of all studies depend on issues such as: maintaining randomization integrity, accurately assessing participant eligibility, recording dropouts and adherence, measuring variables without bias, monitoring and assessing protocol adherence, and avoiding biases in the analysis of the results.

MoTrPAC maintains a measurement database with access rights that are completely separate and distinct from the safety monitoring system for the intervention groups. All MoTrPAC participants are systematically queried at clinical site visits to capture study data, medical events, or adverse events (**AE**s). Because the same staff are performing assessments and intervention procedures, MoTrPAC is not able to mask assessment staff.

A major challenge of the MoTrPAC project is preventing, identifying, and correcting bias and batch effects. Several examples include clinical/chemical site, dominant leg, technician, season/time of day, assay batch, and equipment differences. All efforts are made to randomize assays across conditions to avoid confounding due to batch effects. This is a challenge with some assays that have smaller batch sizes such as metabolomics/proteomics. The use of appropriate reference samples, device calibration,

and replicates across sites are essential in capturing any systematic bias. Tools that estimate and correct for latent bias such as population Principal Component Analyses (**PCA**) or expression Probabilistic Estimation of Expression Residuals (**PEER**) factors will be investigated. From the bioinformatics perspective, the important aspect is that the metrics are captured in the metadata and are reviewed early and regularly for bias. This type of process sensitivity coupled with robust analyses are main priorities of the BIC throughout the project.

6.2.2 ENROLLMENT/RANDOMIZATION/MASKING PROCEDURES

The randomization protocol and the randomization process are prepared and executed by investigators from the DMAQC through the secure web-based data management system, so that eligibility is automatically confirmed and records are current. The study statisticians have developed the randomization protocol to ensure that sufficient baseline and eligibility data are entered and validated before participants are randomized into the study. In addition, the study statisticians closely monitor the intervention allocations by clinical site to ensure protocols are being adhered to and balance is being maintained.

6.3 STUDY INTERVENTION COMPLIANCE

Adherence to EE and RE within the MoTrPAC RCT is critical to interpretation of the discovery objectives set forth in this initiative. Adherence involves both attendance to scheduled sessions and the quantity and quality of the exercise that occurs during each session. Additionally, since there is a Control group, it is important that participants in this group maintain a sedentary lifestyle and that they return for follow-up assessments. Below we describe steps to exclude participants prior to randomization who are unlikely to be adherent, and problem solve barriers to adherence and to develop strategies to promote adherence once participants are randomized to treatment. In addition, we review methods for capturing and monitoring adherence during the course of study.

6.3.1 UNDERSTANDING AND PROMOTING ADHERENCE

It is reasonable to assume that many participants who join MoTrPAC do so because they are interested in learning how to exercise properly and that the supervision and structure provided by study staff helps them to become physically active. Although participants are informed prior to baseline testing that they may be randomized to EE, RE, or Control groups, it is certainly expected that participants want to receive exercise training and believe that they have a good chance of ending up in either EE or RE.

Because randomization to Control poses a risk for dropping out, upon completion of the post-intervention period assessments, those individuals assigned to Control are given an exercise prescription, guidance, and education materials. In addition, it is essential that all participants in the RCT (EE, RE, Controls) maintain their pre-randomized levels of free-living PA throughout the approximately 12-week intervention period, and this will be monitored as described in Section 6.1. Moreover, to provide objective assessment of free-living PA that will be used for data analysis, participants are provided with an accelerometer that will be worn for four approximately 7-day intervals (pre-intervention and at approximately 4, 8, and 12 weeks of the intervention).

Adherence to the exercise protocols are affected by personal, social, and environmental factors. At a personal level, it is important to understand what goals participants hope to achieve by joining MoTrPAC, making certain that these goals are realistic, and confronting unreasonable expectations.

We recognize that social interactions between staff and participant as well as the broader social environment surrounding the MoTrPAC sites influence adherence to scheduled training sessions. All staff are instructed to engage with and be supportive toward all participants.

Finally, the physical environment and access to training facilities influence adherence. This includes such factors as ease of parking, transportation, and accommodations made for friends or significant others who may come to training facilities to support participants via transportation, etc. At the time of baseline assessments, staff discuss these topics with each participant to be certain there are no significant barriers in this domain and/or to problem solve potential challenges that exist.

6.3.2 CAPTURING AND MONITORING ADHERENCE IN MOTRPAC

Sites enter attendance data and information related to exercise mode, intensity and duration during each training session via a web-based tracking system. At the individual level, the tracking system enables staff to monitor the progress of participants across time and serves as a means of providing individualized feedback and motivation at regular intervals. Individual adherence data (both attendance and session quality (intensity, duration, etc.) are stored for future statistical analysis approaches that may incorporate adherence in the model(s).

7 STUDY DISCONTINUATION, WITHDRAWAL

7.1 DISCONTINUATION OF STUDY INTERVENTION

There are no study-stopping criteria for either efficacy or safety.

7.2 PARTICIPANT DISCONTINUATION, WITHDRAWAL FROM THE STUDY

Participation in the main MoTrPAC study is considered completed when one of following criteria is met:

- Participant completes the final study procedure within the study windows,
- Participant elects to withdraw from the study,
- Participant is considered lost to follow-up (**LTFU**), or
- An Investigator decides to withdraw the participant due to safety or other considerations.
    - Participant may be withdrawn due to a COVID-19 diagnosis after consenting to participate.

7.3 LOST TO FOLLOW-UP

Participant is considered LTFU when the study window for completion of procedures is closed and no additional study-related information can be gathered.

8 STUDY ASSESSMENTS AND PROCEDURES

8.1 STUDY PROCEDURES AND OUTCOMES

8.1.1 STUDY SCHEDULE

Sedentary participants begin study participation at orientation and/or screening. We estimate that the total time for participant commitment for orientation, pre-randomization screening, pre-randomization phenotyping assessments, and pre-training familiarization to be approximately 3-10 hours, dependent in part on randomized assignment. Following the familiarization sessions, participants will have a wash-out period (no testing, no exercise) prior to baseline acute exercise testing. The individual clinical site completes all screening and pre-randomization assessments and performs randomization within approximately 8 weeks of obtaining informed consent. The individual clinical site is permitted to determine the flow of the orientation, pre-randomization screening, and pre-randomization phenotyping assessment visits to optimize individual participant needs and minimize participant burden.

Following all screening and pre-intervention phenotyping, sedentary participants will be randomized and go through familiarization sessions based on group assignment. Following a standardized rest/wash-out period, participants complete the pre-intervention acute exercise test and biospecimen collection. Participants then undergo approximately 12 weeks of exercise training or control.

HA participants complete the same screening and pre-intervention study visits as sedentary participants, but do not proceed to the intervention phase.

Total time for sedentary participants to complete post-intervention phenotyping assessments and acute exercise testing/biospecimen collection is approximately 4-11 hours. Phenotyping assessments can begin during the final week of the intervention period (approximately 12 weeks; determined on an individual basis).

8.1.2 MEASURES AND PROCEDURES OVERVIEW

Phenotyping measures were selected to meet the overarching goal of linking molecular maps to key physiologic adaptations and health benefits of PA. These select measures provide high quality, reproducible phenotyping and can be standardized across all MoTrPAC clinical sites.

8.1.2.1 INFORMED CONSENT

Following an initial telephone verbal or web-based consent, and screen (e.g., self-reported age, BMI, exercise history), potential participants who are not excluded, and remain interested, are scheduled for a study orientation session. After all questions have been addressed, written informed consent is obtained before each participant undergoes further screening, familiarization, and testing.

Web-based screening will be conducted on a national platform hosted at Wake Forest School of Medicine (**WFSM**) on the central MoTrPAC website. This system implements industry standard security technology to ensure the security and integrity of the data collected. The web platform is structured to have a public facing web server that is located behind an external firewall/intrusion detection appliance. Firewall conduits allow the public facing webserver to talk to our relational database management system, which resides on the WFSM private network. Data are encrypted during transmission using the latest Transport Layer Security (**TLS**) encryption protocols, using 2048 bit Secure Socket Layer (**SSL**) certificates. Any data captured by the web screener would be stored in the Relational Database

Management System (**RDBMS**) with access control lists limiting data access to only those individuals who should have access to it. Prior to public launch, and annually thereafter, websites are tested for known vulnerabilities by testing from the WFSM Information Technology (**IT**) Security Department. Because the web screener will be accessed by unknown individuals, we will implement trusted algorithms to detect whether the activity is likely from a human or not and take the appropriate safeguards to protect the integrity of the system.

8.1.2.2 SCREENING ASSESSMENTS

All Participants (Sedentary, HAEE, HARE): During screening, the review of medical history and medication inventory and clinical assessments, including a blood panel, to provide important screening and phenotyping data will be completed. These include height and weight (to assess BMI eligibility), resting blood pressure, morning fasted blood profiles (Complete Blood Count, chemistry and metabolic panels, lipid profile, TSH, and hemoglobin A1c), CESD, and resting electrocardiogram.

8.1.2.3 PHYSIOLOGIC MEASURES

8.1.2.3.1 CARDIORESPIRATORY FITNESS (PEAK OXYGEN CONSUMPTION RATE [VO_2PEAK])

Sedentary Participants: Before and after the intervention period, cardiorespiratory fitness is determined by a CPET on an electronically braked cycle ergometer per the American College of Sports Medicine (**ACSM**) Testing Guidelines. Oxygen consumption and carbon dioxide production are measured by indirect calorimetry. 12-lead ECG recordings and blood pressure are monitored, and a clinician interprets for contraindications to exercise prior to enrollment. AHA criteria are used to halt exercise testing.[54]

HAEE and HARE: This assessment will occur only once before the acute exercise test.

8.1.2.3.2 BODY COMPOSITION AND ANTHROPOMETRICS

Sedentary Participants: Before and after the intervention period, total and regional body composition and bone mineral content (**BMC**) and density (**BMD**) are determined by whole body DXA, and waist circumference is measured. Body weight is measured weekly for EE and RE participants and every 4 weeks for Control participants.

HAEE and HARE: These assessments will occur only once before the acute exercise test.

8.1.2.3.3 ISOMETRIC STRENGTH

Sedentary Participants: Before and after the intervention period, participants are evaluated for knee extensor strength via isometric maximum voluntary contraction at approximately 60° knee flexion.

HAEE and HARE: This assessment will occur only once before the acute exercise test.

8.1.2.3.4 ONE-REPETITION MAXIMUM (**1RM**) DYNAMIC STRENGTH

Sedentary RE Participants: Before and after the intervention period, participants are tested for dynamic voluntary strength (i.e. 1-RM) on one upper body and two lower body exercises.

HARE: These assessments will occur only once before the acute exercise test.

Sedentary EE, Sedentary Control, and HAEE: These assessments will not be performed.

8.1.2.3.5 GRIP STRENGTH

Sedentary Participants: Before and after the intervention period, isometric hand grip strength of the dominant hand is tested using a handgrip dynamometer.

HAEE and HARE: This assessment will occur only once before the acute exercise test.

8.1.2.3.6 POTENTIAL ANCILLARY MEASURES

Ancillary measures may be added as part of approved ancillary studies to MoTrPAC (e.g., stool sample collection). An ancillary study may be conducted at some or all clinical sites, as defined in the approved ancillary study protocol. Site participation in each ancillary is at the discretion of the site. For each ancillary study, the target recruitment number of participants is defined for each site and may involve only a subset recruited under a separate informed consent process.

8.1.2.4 BEHAVIORAL MEASURES

8.1.2.4.1 PARTICIPANT-CENTERED OUTCOME MEASURES

Sedentary Participants: Before and after the intervention period, a limited battery of Patient-Reported Outcomes Measurement Information System (**PROMIS**) self-reported health outcome measures is administered [56]. Four measures were selected from the mental health domain (depression, anxiety, social isolation, and positive affect), two from the physical health domain (fatigue and sleep disturbance), and one from the social health domain (satisfaction with social roles and activities). All measures are administered via computer. Time to complete each measure ranges from 1-3 minutes. Thus, we anticipate that the total time for administration of the PROMIS battery ranges of approximately 8-15 minutes at both the pre-intervention and post-intervention testing.

HAEE and HARE: These assessments will occur only once before the acute exercise test.

8.1.2.4.2 DIETARY INTAKE ASSESSMENT

All Participants (Sedentary, HAEE, HARE): To characterize and describe the typical dietary intake of all participants, a standard semi-quantitative food frequency questionnaire (web-based Diet History Questionnaire III (**DHQ-III**)) is administered during the pre-randomization phase to estimate energy, macronutrient, and micronutrient intakes.

All Participants (Sedentary, HAEE, HARE): In addition, a diet record is maintained by all participants on the day before the acute exercise test/rest visit. For the HA participants and Sedentary participants assigned to the Late profile (i.e., those who must return for 24-h post-exercise sample collections), a diet record is also maintained on the day of the acute exercise test/rest and entered into the Automated Self-Administered 24-Hour (**ASA24®**) Dietary Assessment Tool website (https://epi.grants.cancer.gov/asa24/).

Sedentary EE and RE Participants: To minimize the influence of changes in dietary intake from before to after the intervention, participants are asked to replicate their pre-intervention dietary intake schedule on the day before the post-intervention biospecimen collection period and, for participants in the Late collection profile, the day of biospecimen collection.

Sedentary Control Participants: Accounting for dietary intake at the post-intervention time point only involves the 24-hour period prior to the single biospecimen collections (1 muscle, 1 adipose, 1 blood).

8.1.2.4.3 PHYSICAL ACTIVITY

Free-Living PA is assessed using wearable activity monitors (accelerometers) that quantify the intensity and patterns of usual activity and sedentary behavior, including sleep. Following all accelerometry assessments, participants will be asked to complete the International Physical Activity Questionnaire [57] (IPAQ), a self-report measure that captures activity over the past week in 4 different areas of life: **1)** job-related; **2)** transportation-related; **3)** housework, house maintenance, and caring for family; and **4)**

recreation, sport and leisure time. It also assesses time spent sitting. Following completion of the initial IPAQ only, sedentary and high active participants will be administered the Paffenbarger Physical Activity Questionnaire [58] to assess activity across the past year.

Sedentary Participants: PA will be assessed in participants with an accelerometer worn on the wrist. Some participants will be asked to also wear an accelerometer worn on the thigh that captures both PA and sedentary time. The accelerometers (thigh and wrist) are worn for approximately 1 week: **1)** before the acute exercise test as part of the phenotyping measures at baseline, **2)** during the intervention at approximately the 4- and 8-week marks, and **3)** at the end of the 12-week intervention.

HAEE and HARE: PA will be assessed in participants with an accelerometer worn on the wrist. The accelerometers (thigh and wrist) are worn for approximately 1 week after enrollment.

8.1.2.5 BIOSPECIMEN COLLECTION

The randomized biospecimen collection profiles (Early, Mid, Late) are described in Figures 2a and 2b.

Sedentary Participants: For sedentary participants, the pre-intervention biospecimen collection occurs after a washout interval of no exercise to allow any acute effects to subside. Sedentary participants randomized to EE or RE abstain from exercise for a few days after the final exercise training session, prior to the post-intervention study acute exercise test and biospecimen collection. Sedentary Control participants are asked to abstain from regular exercise throughout the intervention period.

HAEE and HARE: HAEE and HARE participants abstain from exercise for a few days prior to the acute exercise test and biospecimen collectionto match the same wash-out period applied to EE and RE at the post-intervention time point.

All Participants (Sedentary, HAEE, HARE): All participants arrive to the clinical site in the morning after a ~10- to 12-hour fast; water is encouraged. They will be instructed to consume a standardized shake, which will be provided, the prior evening. After about 30 minutes of supine rest, a blood sample (approximately 25 mL), a muscle sample (ideally >150 mg), and an adipose sample (ideally >600 mg for RCT participants and >200 mg for HARE and HAEE participants) are collected. A muscle biopsy of the vastus lateralis is performed under local anesthesia (i.e., lidocaine or an alternative anesthetic if lidocaine is in short supply) using previously published methods.[59] A subcutaneous adipose tissue biopsy is obtained under local anesthesia from the periumbilicus region using previously published methods.[60-66] Trained investigators experienced with biopsy procedures perform all tissue biopsies. Additional blood samples are obtained during the acute exercise test for EE and HAEE groups and during the rest interval for the Control group. Blood, muscle, and adipose samples are also collected for up to 24 hours after exercise. The timing of sample collection after exercise is different for the Early, Mid, and Late sampling groups (Figures 2a, b). Additional blood samples may be collected for safety reasons when necessary.

Conducting serial blood, muscle, and adipose samples before, during, and after exercise is an essential approach for MoTrPAC to study the molecular transducer responses to acute exercise. This is an approach that has been used by many MoTrPAC investigators and other investigators in a safe and well-tolerated manner (see citations in Section 2.1). All blood and tissue samples are processed and temporarily stored locally at the clinical sites until they are shipped to the MoTrPAC Biological Sample Repository using bar-coded cryovials. The samples will be analyzed by CAS using genetics, genomics, epigenetics, transcriptomics, metabolomics, lipidomics, and proteomics assays.

8.1.2.5.1 EE TEST

Sedentary Participants Randomized to EE and HAEE: After resting blood and tissue samples are collected, the participant performs a ~5-minute warm-up consisting of light cycling on a cycle ergometer. After the warm-up, participants cycle for ~40 minutes at a pre-determined workload that corresponds to ~65% VO_2peak based on the VO_2 vs power regression obtained during the CPET to determine VO_2peak on the cycle ergometer. This target workload is held constant throughout the ~40 minutes of cycling, unless it has to be reduced due to participant inability to complete the acute bout. If this occurs, the exact same load reduction will occur on the post-intervention test. Blood samples are collected during exercise and blood, muscle, and adipose samples are collected after exercise (Figure 2a, b).

8.1.2.5.2 RE TEST

Sedentary Participants Randomized to RE and HARE: After the resting blood and tissue samples are collected, the participant performs a 5-minute warm-up on a treadmill or cycle ergometer. After the warm-up, participants perform specified resistance exercises at target loads for each exercise and in a specified order. No blood samples are collected during exercise because of the concern of having an antecubital IV catheter in place during the performance of upper extremity resistance exercises. Blood, muscle, and adipose samples are collected after exercise (Figures 2a, b).

8.1.2.5.3 CONTROL TEST

Sedentary Participants Randomized to Control: After the resting blood and tissue samples are collected, the participant proceeds to a second rest period lasting ~40 minutes. Blood samples are collected during this rest period and blood, muscle, and adipose samples are collected after the rest period to coincide with the schedule for sedentary EE participants (Figure 2a, b).

8.1.2.5.4 SPECIAL CONSIDERATIONS

Following the last tissue and blood sample collection on any EE, RE, or control test day, participants are provided with a light meal or snack before they leave the facility. All participants are given detailed instructions about care of biopsy sites. Participants randomized to the Late biospecimen collection profile and HA participants are instructed to return to the clinical site the following day after a ~10-12-hour fast for the 24-hour post-exercise biospecimen collections.

8.2 SAFETY AND OTHER ASSESSMENTS

8.2.1 EVENT ASSESSMENT AND FOLLOW-UP

The documentation and collection of AEs begins after informed consent is obtained and ends after procedures for the final visit have been completed. Participants are contacted and interviewed regarding possible AEs in a standardized fashion about every four weeks after randomization through their final visit. If a participant has died or cannot be reached, their emergency contact is contacted. AEs are investigated through the collection of data during the interview process. If needed, copies of medical records, Emergency Medical Services (**EMS**) reports, death certificates, and coroner and/or autopsy reports are obtained. In addition to fixed time points, events may be disclosed spontaneously (between the fixed time points) by the participant or a member of the clinical site staff. Source and timing of serious adverse events (**SAE**s) and AE reports are recorded to allow for sensitivity analyses for potential ascertainment bias. All clinical sites take appropriate action regarding the event and are responsible for documenting, collecting, processing, classifying, and coding AEs within the online safety collection system on the secure MoTrPAC website.

The Clinical Site Principal Investigator (**PI**), or a designee, submits all safety updates and periodic reports to the Regulatory authority(ies), as required by applicable local regulatory requirement(s). Individual safety reports, if required, are created by the clinical site PI. Central, study-wide SAE listings and an annual safety report are available to the individual clinical sites for their reporting needs.

It is anticipated that some AEs will occur during the course of the study during assessments and intervention. The following is a summary of a plan of action based on level of acuity of the problem.

8.2.2 PRE-EXERCISE INTERVENTION SAFETY SCREENING

To maximize participant safety, a standardized screening protocol is followed. Accordingly, a thorough medical history and limited clinician exam are performed by a qualified clinician during the screening process. Exclusion and inclusion criteria are adhered to strictly. All potential participants undergo screening for CVD and other major diseases by means of a medical history form, medication inventory, laboratory measures, and other ancillary assessments including ECG and blood pressure. These are initially collected by clinical site staff and reviewed by a qualified clinician prior to randomization into the study.

8.2.3 SAFETY CONSIDERATIONS FOR STUDY ASSESSMENTS

All study assessments are completed by trained and certified clinical site staff. Safety precautions are taken during all assessments by applying standardized stopping criteria. If the participant reports chest pain, tightness or pressure, significant shortness of breath or difficulty breathing, feeling faint, lightheaded or dizzy, or any other clinically concerning symptom, a test is stopped. Onsite clinical site staff are trained to provide basic life support and to provide immediate care when faced with medical emergencies. Institutional or community EMS services are activated if needed.

8.2.3.1 DXA

DXA scans to determine body composition, including BMC and BMD, involve small amounts of radiation exposure. The amount of radiation exposure from each whole body DXA scan is $\leq$2.5 mRem per scan or $\leq$5.0 mRem for two scans. Although the potential long-term risk from these radiation doses is uncertain, such doses are exceedingly low and have never been associated with any definite adverse effects.

8.2.3.2 CARDIORESPIRATORY FITNESS (PEAK AEROBIC POWER [VO2PEAK])

VO_2peak is measured during a CPET to exhaustion and used for prescription of exercise intensity (at baseline) and as an outcome measure of cardiorespiratory fitness. The first CPET is also used as a screening test for eligibility. The clinical sites have extensive experience measuring VO_2peak in participants of all ages, including older adults.

Maximal exercise testing is a common procedure with risks, including fainting, dizziness, chest pain, irregular heartbeat, or heart attack, although the latter is extremely rare (<1 death in 10,000 tests) in people with no history of heart disease. The test is monitored by an exercise physiologist or a clinician and is stopped if problems occur. Blood pressure and ECG are monitored before, throughout, and after the test.

8.2.4 SAFETY CONSIDERATIONS FOR THE INTERVENTIONS AND ACUTE EXERCISE TESTING

The exercise intervention is conducted on-site and all sessions are supervised by trained exercise interventionists, who monitor adherence to the protocol and potential adverse experiences and symptoms. During the exercise sessions, an Automated External Defibrillator (**AED**) and on-site trained staff are available to deal with medical emergencies. Institutional or community EMS are activated, if

needed. Participants are taught the importance and proper method of warming-up before and cooling-down after exercise and instructed on correct exercise techniques. Heart rate is monitored during the exercise sessions. If at any point during an exercise session a participant develops chest pain, shortness of breath, dizziness, or other concerning symptoms they will inform the exercise interventionist immediately and appropriate action will be taken.

If a participant reports chest pain, shortness of breath, dizziness, or other sign or symptom of a cardiovascular event, a follow-up evaluation by the participant's doctor or a cardiologist will be required before they can continue the exercise intervention. In addition, guidelines suspending or stopping PA for inter-current illness are in the MOP. The exclusion criteria eliminate individuals who are high risk.

8.2.4.1 SAFETY CONSIDERATIONS FOR SUPERVISED ENDURANCE TRAINING

Potential risks associated with EE are explained to each participant by appropriately trained clinical site staff during the initial informed consent, screening, and study orientation.

EE risks include:

- Difficulty breathing and/or shortness of breath
- Muscle, joint strains and soreness
- Soft tissue injury, falls and fractures
- A transient increase in the risk of sudden death, stroke, and acute myocardial Infarction (**MI**) occurring during a bout of vigorous exercise has been reported, especially in previously sedentary individuals
- Orthostatic hypotension
- Irregular heart beats
- Exercise-induced asthma
- Exacerbation of arthritis or other joint conditions
- Falls during exercise or exercise visits

Procedures to minimize injury during exercise include warm-up and cool-down activities:

- Staff are trained to respond to changes in health status such as
    - A change in mental status and/or mood
    - Elevated resting or exercise blood pressure
    - Chest pain during exercise or newly reported since the last visit
    - Acute shortness of breath, visual disturbances, stroke symptoms, dizziness during exercise, altered resting or irregular heart rate, heart palpitations, leg swelling or edema

If such health changes are encountered, clinical site staff are trained to evaluate the participant and take appropriate actions. Clinical site staff are trained in Cardiopulmonary Resuscitation (**CPR**) and the use of an AED. Institutional or community EMS is activated if needed.

8.2.4.2 SAFETY CONSIDERATIONS FOR SUPERVISED RESISTANCE TRAINING

Potential risks associated with RE are explained to each participant by appropriately trained clinical site staff during the initial informed consent, screening, and study orientation.

RE risks include:

- Difficulty breathing and/or shortness of breath
- Muscle, joint strains and soreness

- Soft tissue injury, falls and fractures
- Increasing intraocular and systemic pressures associated with use of the Valsalva maneuver to levels that may cause injury (detached retina, hernia, conjunctival hemorrhage), which is greatly reduced by instruction in proper exercise technique
- Hernia
- Exacerbation of arthritis or other joint conditions
- Lumbar disk hernia
- A small increase in the risk of sudden death, stroke, and acute MI occurring during a bout of vigorous exercise has been reported, especially in previously sedentary individuals
- Orthostatic hypotension
- Exercise-induced asthma

The same safety considerations as outlined for the EE intervention are adhered to for the RE intervention.

8.2.4.3 SAFETY CONSIDERATIONS FOR HIGHLY ACTIVE EE AND RE GROUPS

The same safety considerations outlined for the EE and RE training are utilized for the HAEE and HARE groups.

8.2.5 SAFETY CONSIDERATIONS REGARDING BIOSPECIMEN COLLECTIONS

8.2.5.1 BLOOD SAMPLE COLLECTION

Blood is drawn only by trained and experienced personnel via venipuncture and/or intravenous cannulation to minimize the discomfort and risks.

Blood sampling risks include:

- Bleeding
- Pain
- Becoming faint or vaso-vagal response
- Developing a bruise or bump
- Slight risk of infection at the site where blood was drawn

8.2.5.2 MUSCLE TISSUE COLLECTION

There may be temporary discomfort and bleeding from the muscle biopsy. There is often brief discomfort when the local anesthetic is administered. A small amount of bruising can occur. There is a slight risk of infection, but this procedure is done under sterile conditions to protect against this.

Muscle tissue sampling risks include:

- Pain at the site of the procedure
- Becoming faint or vaso-vagal response
- Bleeding/hematoma
- Scar formation of ~5 mm for each incision or keloids
- Skin irritation at the site of the bandage
- Numbness in the area
- Reaction to the medication used to numb the area (e.g., Lidocaine)
- Infection at the site where the sample was collected

8.2.5.3 ADIPOSE TISSUE COLLECTION

There may be temporary discomfort and bleeding from the adipose biopsy. There is often brief discomfort when the local anesthetic is administered. Bruising can occur. There is a slight risk of infection, but this procedure is done under sterile conditions to protect against this. Adipose tissue sampling risks include:

- Pain at the site of the procedure
- Becoming faint or vaso-vagal response
- Bleeding/hematoma
- Scar formation of ~5 mm for each incision or keloids
- Skin irritation at the site of the bandage
- Numbness in the area can rarely occur
- Reaction to the medication used to numb the area (e.g., Lidocaine)
- Infection at the site where the sample was collected

There may be some temporary discomfort and bleeding from the adipose biopsy. There is often brief discomfort when the local anesthetic is administered. Bruising can occur. There is a slight risk of infection, but this procedure is done under sterile conditions to protect against this.

8.2.6 SAFETY CONSIDERATIONS REGARDING LOSS OF CONFIDENTIALITY

A breach of confidentiality may result in psychological harm to individuals (in the form of embarrassment, guilt, stress, etc.) or in social harm.

8.3 ADVERSE EVENTS AND SERIOUS ADVERSE EVENTS

The safety of study participants is a high priority of the study. To ensure the safety of MoTrPAC participants, we standardize the study operating procedures and the safety and reporting procedures to be in full compliance with NIH policies and Institutional Review Board (**IRB**) prompt reporting requirements. Monitoring the safety of participants includes: evaluation of safety for inclusion at screening, safety during testing- and intervention-related activities, and continuation of the assessment or intervention after a safety event has occurred.

Key safety-related eligibility criteria are exclusions from participation in MoTrPAC. Adherence to eligibility criteria addresses safety and ensures the inclusion of appropriate participants in the study. Once participants have agreed to participate and have provided consent, all participants are monitored for safety issues potentially related to participation in the study.

AEs are monitored using a variety of standardized methods including: structured solicitation of hospitalizations and questionnaires for events. In addition, the intervention and assessment procedures are monitored for safety events, including common, expected signs and symptoms associated with the collection of tissue samples (i.e., discomfort during the procedure, bruising, tenderness, or drainage at the biopsy site) and the performance of vigorous exercise (i.e., muscle and joint discomfort, lightheadedness/dizziness, nausea/vomiting, vasovagal response).

Table 7 provides an overview of the categorization and reporting of safety events. It should be noted that some common expected mild signs or symptoms from the tissue collection procedures (i.e., discomfort during the procedure, tenderness, bruising, or drainage at the biopsy site) and from the performance of exercise (i.e., muscle or joint discomfort, lightheadedness/dizziness, nausea/vomiting, vasovagal response) will be collected on study visit case report forms (**CRFs**) rather than on AE log CRFs to lessen the burden on study teams. However, these events will be included in the non-expedited,

aggregated regulatory reports. Reporting timeframes will adhere with the sIRB prompt reporting guidelines. All deaths are reported within 72 hours of discovery. The following events are reported within 10 working days of discovery:

- Unanticipated problems (**UPs**) involving risks to participants or others (described in Section 8.4),
- Serious or continuing non-compliance on the part of a clinical site or investigators (described in Section 8.3.8),
- Unexpected adverse device effects, which are handled like SAEs,
- Potential breaches of confidentiality (described in section 8.3.8),
- Incarceration of a participant where the study is not approved to involve prisoners,
- Unresolved participant complaints that involve unexpected risks that can't be resolved by the research team.

Table 7. Categorization and reporting of adverse events

		Type of Event													
		Common, mild signs/symptoms*	Non-serious Adverse Event (AE)						Serious Adverse Event (SAE)						
	Expectedness	Expected	Expected			Unexpected			Expected			Unexpected			
	Relatedness	Definitely	Not	P/P	Def	Not	P/P	Def	Not	P/P	Def	Not	P/P	Def	
Actions Required	Collect on study visit CRFs	Yes	No	No	No	No	No	No	No	No	No	No	No	No	
	Collect on AE log CRF**	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
	Expedited regulatory reporting of individual event	No	No	No	No	No	No	No	No	No	No	Yes	Yes	Yes	
	Non-expedited, aggregated regulatory reporting	Yes	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
	Considered as Unanticipated Problem***	No	No	No	No	No	Y/N	Y/N	No	No	No	No	Yes	Yes	

* Common, mild signs and symptoms include 1) discomfort during biopsies and tenderness, bruising, or drainage at the site, and 2) exercise-related muscle or joint discomfort, lightheadedness/dizziness, nausea/vomiting, or vasovagal response. These will be collected on various study visit CRFs and reported only in aggregate at continuing review (IRB) and twice annual meetings (DSMB).

** Prompts are used on CRFs to elicit completion of the AE log for expected and unexpected events discovered during queries performed at regular intervals. Spontaneously reported events are also entered on the AE log.

*** Non-serious AEs that are possibly/probably or definitely related to the study may or may not be Unanticipated Problems, depending on whether they affect participant safety.

8.3.1 DEFINITION OF AE

An AE is defined as any health-related unfavorable or unintended medical occurrence temporally associated with the procedures used within the study, whether or not related to the study. Examples of AEs include, but are not limited to, the following:

- A clinically significant laboratory or clinical test result at follow up assessments
- An event that requires a clinical consult because it alters the participant's ability to exercise
- An event that occurs as a result of a study procedure

8.3.2 DEFINITION OF SAE

SAEs are defined as events that may be harmful to the participant and/or may be serious enough to warrant either temporary or permanent discontinuation of the study intervention or study procedures,

either because they are intolerable or because they are judged to be potentially harmful. Consistent with NIH guidelines, SAEs are AEs that meet any of the following criteria:

- Results in death
- Is life-threatening (places the participant at immediate risk of death from the event as it occurred)
- Requires inpatient hospitalization
- Prolongs an existing hospitalization
- Results in a persistent or significant disability/incapacity
- Important medical events that investigators judge to represent significant hazards or harm to participants, and may require medical or surgical intervention to prevent one of the other SAEs listed in this definition (e.g., hospitalization, death, persistent disability)
- Result in a congenital anomaly/birth defect

8.3.3 CLASSIFICATION OF AN ADVERSE EVENT

Classifications of AEs will be completed by a clinician at the clinical site and monitored by the Medical/Clinical Studies and Safety (**MCSS**) subcommittee or a designated working group of the subcommittee.

8.3.3.1 EXPECTEDNESS

In view of the nature of the study and the study population, many AEs are likely to qualify as expected events, including injuries and accidents. The determination of expectedness is made by a clinician at the clinical site and confirmed by the MCSS or a designated working group of the subcommittee.

8.3.3.1.1 DEFINING EXPECTED ADVERSE EVENTS

Expected AEs are defined as expected events based on the prior experience with the assessments and/or intervention that are listed in the participant consent and protocol; these can be attributed to an underlying health condition, the participant population being studied, or expected consequences of the procedures or intervention. These conditions may be unpleasant and bothersome to the participant (e.g., sore muscles, muscle or joint pain, fatigue), but do not require discontinuing the study intervention or components of the intervention.

8.3.3.2 RELATEDNESS

The classification of relationship to the study is as follows.

- Not Related – AE is clearly not related to the intervention or a procedure (e.g., temporal sequence of events is not consistent with administration of intervention or procedures; other causes are more plausible or causal relationship is implausible)
- Possibly/Probably Related – AE that follows reasonable sequence from administration of intervention or procedure, but that could readily have been produced by other factors
- Definitely Related – AE is clearly related to the intervention or procedure (e.g., follows reasonable temporal sequence from administration of the study intervention or procedure, confirmed by improvement after stopping intervention or procedure, and reappearance on repeated exposure to intervention or procedure, and cannot be explained by participants clinical or health status).

8.3.3.3 SEVERITY OF EVENT

The classification of the level of severity is as follows:

- Mild – Events require minimal or no treatment and do not interfere with the participant's daily activities.
- Moderate – Events result in a low level of inconvenience or concern with the therapeutic measures. Moderate events may cause some interference with functioning.
- Severe – Events interrupt a participant's usual daily activity and may require systemic drug therapy or other treatment. Severe events are not usually potentially life-threatening or incapacitating.

8.3.4 TIME PERIOD AND FREQUENCY FOR EVENT ASSESSMENT AND FOLLOW-UP

The burden of collecting and reporting data on every possible AE in MoTrPAC is considerable and common mild signs and symptoms from the intervention and and some procedures have been well defined in previous studies. However, following the guidance from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (**NIAMS**), MoTrPAC sites collect all AEs and SAEs disclosed by participants, and report all SAEs and most AEs following the reporting requirements of the sIRB, NIAMS, and DSMB. The events that are not reported are those that are categorized as non-serious and not related to the study (Table 7).

Participants are instructed to report SAEs immediately to the clinical site team. After randomization (sedentary participants) or enrollment (HA participants), participants can also disclose AEs at any time, and sedentary participants are queried for AEs in a standardized fashion about every four weeks through their final visit. There are two timeframes of AE collection and reporting. For the description below, we have considered "enrollment" of the HAEE and HARE groups to be the point at which they are eligible to do the acute exercise test. The two timeframes are described below:

First timeframe – before randomization (sedentary) or enrollment (HA only):

After providing informed consent and before randomization of sedentary participants or enrollment of HA participants:

- **AEs/SAEs Not Collected or Reported**: Pre-existing conditions and pre-planned procedures (surgeries or therapies) scheduled prior to signing the ICF are not considered AEs or SAEs. **These will not be collected or reported.**
- **Expedited Reporting of SAEs**: The study team will report within 48 hours or 10 days (see Section 8.3.6 for reporting requirements) of becoming aware of the event any SAEs that **1**) result in death, **2**) are life threatening, or **3**) are unexpected AND definitely or possibly/probably related to the study procedures.
- **DSMB Reporting**: All AEs/SAEs **definitely or possibly/probably related** to study procedures are collected and reported in the aggregate DSMB Report during the DSMB meetings, typically held biannually.

Second timeframe – after randomization or enrollment (HA only):

After randomization of sedentary participants or enrollment of HA participants and through the last contact (after intervention and second acute test for sedentary participants and after acute test for HA participants):

- **AEs/SAEs Collected**: All AEs/SAEs are collected regardless of expectedness or relatedness to the study.

- **Expedited Reporting of SAEs**: The study team will report within 48 hours or 10 days of becoming aware of the event any SAEs that **1**) result in death, **2**) are life threatening, or **3**) are unexpected AND definitely or possibly/probably related to the study interventions.
- **DSMB Reporting**: All AEs/SAEs are collected and reported in the aggregate DSMB Report during the DSMB meetings, typically held biannually.

8.3.5 ADVERSE EVENT REPORTING

The Clinical Site coordinator collects required information on all AEs as disclosed by the participant or identified otherwise. Any non-serious AEs, regardless of relationship to participation in the study, may require Medical Safety Officer (**MSO**) assessment for implications for participant safety. Any non-serious AEs that are **unexpected** will require collecting the same information as SAEs. Clinical sites are not required to collect further information for **non-serious, expected** AEs or AEs that are **not related** to the study. The MSO is responsible for near-real time assessment and action on these AEs and communicating this information back to the clinical site coordinator. SAEs that are unexpected and possibly/probably related to study participation must be reported to the sIRB within 10 days of the investigator becoming aware of the event. All AEs that are non-serious and possibly/probably or definitely related to the study are reported to the NIAMS and the DSMB at the time of the DSMB meetings, typically bi-annually. In the closed session, these will be separated by intervention group.

8.3.6 SERIOUS ADVERSE EVENT REPORTING

The clinical site coordinator collects required information on all SAEs. Any SAEs that result in death, are life threatening, or are unexpected AND definitely related to the study procedures or study interventions are reported within 48 hours of the site becoming aware of the event. In addition, all SAEs that are unexpected AND probably/possibly related to the study procedures or study interventions must be reported to the sIRB within 10 days of discovery. Deaths must be reported to the sIRB within 72 hours of discovery. These detailed reports including an event narrative, start and end dates, severity, Medical Dictionary for Regulatory Activities (**MedDRA**) coding, and other relevant medical information are completed on all SAEs that meet the expedited report criteria via the MoTrPAC website. The clinical site coordinator collects the required information on these SAEs and reports the SAEs to the clinical site MSO. The MSO is responsible for real-time assessment and action on these SAEs and communicating this information back to the clinical site coordinator.

Through the MoTrPAC website, the clinical site coordinator enters as much SAE report information as collected. The expedited detailed SAE report is generated and an email notifying Navitas Clinical Research, Inc (NCR)/DSMB Safety Officer and NIAMS is sent within 48-hours of the site becoming aware of the event. The email indicates that the detailed SAE report is available online for review. As additional information is received the clinical site coordinator updates the report information online. In addition, the MCSS subcommittee reviews SAE reports regularly to help ensure that all sites report/classify events in the same manner. A summary of all SAEs, regardless of relatedness, or expectedness are reported to the NIAMS and the DSMB at the time of the DSMB meetings, typically bi-annually (in the closed session report, these will be separated by intervention group).

8.3.7 REPORTING EVENTS TO PARTICIPANTS

8.3.7.1 INFORMING PARTICIPANTS OF UNEXPECTED SERIOUSNESS OR FREQUENCY OF EVENTS

If either the DSMB or IRB has concerns that AEs are more frequent or more serious than expected, this information will be conveyed to study participants through a mechanism recommended and approved by the DSMB and IRB (e.g., reconsenting).

8.3.7.2 REPORTING INDIVIDUAL INCIDENTAL FINDINGS TO PARTICIPANTS

Clinically significant incidental findings from a study visit will be provided to study participants. When appropriate, participants will be encouraged to share the results with their health care provider.

8.3.8 OTHER REPORTABLE EVENTS

8.3.8.1 PROTOCOL DEVIATIONS

Protocol deviations (**PDs**) are reported and classified as major or minor. A PD is a departure from the approved protocol's procedures made with or without prior IRB approval. It is the responsibility of the clinical site PI to determine whether a deviation from the protocol is emergent, major,or minor/administrative and to ensure proper reporting. Emergency deviations, such as a PD to immediately protect the physical well-being of a participant, require reporting to the sIRB as soon as possible, but no later than 5 days after the situation occurred. Major, non-emergent deviations are planned deviations that are non-emergent and represent a major change in the approved protocol. These deviations are changes that the sIRB must approve before the proposed change is implemented. All emergency and major, non-emergent deviations require both immediate reporting into the MoTrPAC web-based data system and an assessment of the implications for the continuation of the study and/or modification of the consent form or protocol.

According to the sIRB procedures "All minor or administrative deviations are those which do not affect the scientific soundness of the research plan or the rights, safety, or welfare of human subjects." All deviations of this type must be recorded in a Minor Deviation Log and submitted as required to the sIRB at the time the continuing review application is submitted.

Clinical site investigators may have to also report PDs to the local IRB. When required, local guidelines will be followed.

8.3.8.2 BREACHES OF CONFIDENTIALITY

The privacy rights of all participants, including maintaining the confidentiality of participant information and research records, are protected under the Health Information Portability and Accountability Act (**HIPAA**) Privacy Regulations and other applicable laws. A breach of confidentiality is defined by JMH as "any unauthorized disclosure of a subject's personally identifiable information." Disclosures of research Protected Health Information (**PHI**) that are made by mistake or to the wrong person must be immediately reported to the sIRB and the local site privacy officer.

8.3.8.3 NON-COMPLIANCE

Non-Compliance is defined as the failure to follow the research protocol, federal, state, or local laws or regulations governing human subjects research, institutional policies, or the requirements or determinations of the IRB. Only incidents that may qualify as serious or continuing non-compliance must be promptly reported:

1. Serious non-compliance is defined as non-compliance that either **a)** significantly harms or poses an increased risk of significant harm to participants or others, or **b)** significantly compromises the rights and welfare of the participants or the integrity of the human research protection program.

2. Continuing non-compliance is defined as a pattern of non-compliance that significantly compromises the scientific integrity of the study or the rights and welfare of the participants or the integrity of the human research protection program.

8.3.8.4 PREGNANCY

If a pregnancy does occur, participation in the MoTrPAC study is discontinued (including exercise intervention, procedures, and visits), the pregnancy documented and collected via the MoTrPAC website, and every effort is made to follow-up until the pregnancy outcome is available.

8.4 UNANTICIPATED PROBLEMS

8.4.1 DEFINITION OF UNANTICIPATED PROBLEMS

UPs can be either **1)** unexpected AE/SAEs that relate directly to participant safety or **2)** PDs that place participant privacy at risk or place participants at risk in some way that does not have an impact on their health and safety.

UPs are defined as any incident, experience, or outcome that meets all of the following criteria:

1. Unexpected, in terms of nature, severity, or frequency, given
    a) the research procedures that are described in the protocol and protocol-related documents and
    b) the characteristics of the study population;
2. Related (definitely or probably/possibly) to participation in the research, meaning there is a reasonable chance that the incident, experience, or outcome may have been caused by the procedures involved in the research;
3. Suggests that the research places participants or others at a greater risk of harm (including physical, psychological, economic, or social harm) than was previously known or recognized.

8.4.2 UNANTICIPATED PROBLEM REPORTING

In MoTrPAC, clinical sites are required to report UPs (safety events and deviations). All UPs are recorded in the MoTrPAC web-based data system by clinical site staff as an AE or PD. These events are entered into the appropriate system as soon as possible and if the event meets the definition of an UP it will trigger a 48 hour expedited report to NIAMS and the DSMB. Unanticipated safety events are reported through the AE Reporting procedures. Unanticipated PDs are reported through the online deviation log. In addition, all events that meet the requirements of an UP are reported to the sIRB in accordance to their policy, within 10 days of discovery of the event. All other events are reported at the time of a clinical site's annual report.

8.4.3 REPORTING UNANTICIPATED PROBLEMS TO PARTICIPANTS

UPs can be either **1**) Unanticipated AE/SAEs, which are unexpected events that relate directly to participant safety, or **2**) PDs that place participant privacy at risk or place participants at risk in some way that does not have an impact on their health and safety. Such events will be discussed with the participant(s) affected, by a member of the local study team.

9 STATISTICAL CONSIDERATIONS

9.1 STATISTICAL HYPOTHESES

This is a randomized, mechanistic intervention trial, not an efficacy trial, that has a goal of discovering molecular responses to exercise. There are no primary or secondary efficacy endpoints and no *a priori* hypotheses.

9.2 SAMPLE SIZE DETERMINATION

Effect Size Considerations

The overarching goal of MoTrPAC is to generate a data resource for future analyses. There are no *a priori* hypotheses or primary outcomes to guide power calculations. As such, sample size calculations have focused on reducing what may ultimately be complex analyses to relatively simple hypothesis testing approaches (i.e. correlation coefficients, differences between means), whereby detectable effect sizes have been calculated for 90% power for several different within and between group analyses, controlling the False Discovery Rate (**FDR**) at the 0.05 level.

We assumed totals of 840 participants per randomized intervention group, 300 randomized Controls, and 150 in each HA group. Detectable effect sizes (with 90% power controlling for a 0.05 FDR) are presented for:

- Within-group change (N=840, N=300, N=280, N=150, N=100; **Table 8**),
- Within-group correlations (N=840, N=300, N=280, N=150, N=100; **Table 9**),
- Between-group differences in change (840 vs. 840; 840 vs. 300; 840 vs. 150; 280 vs 280; 280 vs 150; 280 vs 100; **Table 10**),
- Between-group differences in correlations (840 vs. 840; 840 vs. 300; 840 vs. 150; 280 vs 280; 280 vs 150; 280 vs 100; **Table 11**).

We also calculated the detectable effect sizes assuming 10% of those enrolled do not have follow-up measurements. All calculations assume statistical independence of measurements made on different participants. Note that reference to samples sizes for within- and between-group comparisons in the following tables may refer to change within an acute test (e.g., the multiple blood samples in HARE, HAEE; random temporal samples within EE/RE), or change between pre-intervention samples to post-intervention (e.g., muscle samples taken before the acute exercise test before vs after the intervention among EE, RE), among other comparisons. Allocation of 300 to the Control group permits detection of between-group effects versus the intervention groups that are roughly 35% larger, with 90% power, in comparison to detectable effect sizes for comparisons between the two intervention groups (Table 10).

Note that the alpha for controlling the FDR depends on the percent of outcomes assumed to fall under the alternative hypothesis[67]. Calculations were performed in PASS 14 Power Analysis and Sample Size Software (2015). NCSS, LLC. Kaysville, Utah, USA, ncss.com/software/pass.

Table 8. Detectable effect size (mean/SD) with 90% power for within-group changes

Approximate Group Size (Description)	Controlling FDR at 5%, assuming 10% of outcomes fall under the alternative hypothesis: alpha = 0.005263158	Controlling FDR at 5%, assuming 5% of outcomes fall under the alternative hypothesis: alpha = 0.002493075
N=100 (Controls temporal sample*)	0.415	0.441
N=150 (HAEE or HARE)	0.337	0.357
N=280 (EE or RE temporal sample*)	0.245	0.259
N=300 (All Controls)	0.237	0.251
N=840 (All EE or RE)	0.141	0.149
Assuming Approximate 10% LTFU		
N=90 (Controls temporal sample*)	0.439	0.466
N=135 (HAEE or HARE)	0.356	0.377
N=250 (EE or RE temporal sample*)	0.260	0.275
N=270 (All Controls)	0.250	0.264
N=756 (All EE or RE)	0.148	0.157

SD = standard deviation, FDR = false discovery rate, HAEE = highly active endurance exercise, HARE = highly active resistance exercise, EE = endurance exercise, RE = resistance exercise

* Temporal sample subgroups are those randomized to EARLY, MID, and LATE biospecimen collection profiles

Table 9. Detectable correlation and percent variation explained* with 90% power for within-group correlations

Approximate Group Size (Description)	Controlling FDR at 5%, assuming 10% of outcomes fall under the alternative hypothesis: alpha = 0.005263158		Controlling FDR at 5%, assuming 5% of outcomes fall under the alternative hypothesis: alpha = 0.002493075	
	Correlation	% Variation explained	Correlation	% Variation explained
N=100 (Controls temporal sample**)	0.390	15.2%	0.411	16.9%
N=150 (HAEE or HARE)	0.323	10.4%	0.341	11.6%
N=280 (EE or RE temporal sample**)	0.240	5.8%	0.253	6.4%
N=300 (All Controls)	0.232	5.4%	0.245	6.0%
N=840 (All EE or RE)	0.140	2.0%	0.148	2.2%
Assuming Approximate 10% LTFU				
N=90 (Controls temporal sample**)	0.410	16.8%	0.431	18.6%
N=135 (HAEE or HARE)	0.340	11.6%	0.358	12.8%
N=250 (EE or RE temporal sample**)	0.253	6.4%	0.267	7.1%
N=270 (All Controls)	0.244	6.0%	0.257	6.6%
N=756 (All EE or RE)	0.147	2.2%	0.156	2.4%

FDR = false discovery rate, HAEE = highly active endurance exercise, HARE = highly active resistance exercise, EE = endurance exercise, RE = resistance exercise

* Percent variation explained from linear regression is the square of the detectable correlation coefficient. For large samples sizes, the detectable correlation is approximately equal to the detectable effect size based on means in Table 6.

** Temporal sample subgroups are those randomized to EARLY, MID, and LATE biospecimen collection profiles

Table 10. Detectable effect size ((mean 1 – mean 2)/SD) with 90% power for between-group differences in means

Approximate Group 1 Size (Description)	Approximate Group 2 Size (Description)	Controlling FDR at 5%, assuming 10% of outcomes fall under the alternative hypothesis: alpha = 0.005263158	Controlling FDR at 5%, assuming 5% of outcomes fall under the alternative hypothesis: alpha = 0.002493075
N=840 (EE or RE overall)	N=150 (HAEE or HARE)	0.362	0.383
	N=300 (All Controls)	0.274	0.290
	N=840 (All EE or RE)	0.199	0.210
N=280 (EE or RE temporal sample)	N=100 (Controls temporal sample*)	0.477	0.505
	N=150 (HAEE or HARE)	0.414	0.438
	N=280 (EE or RE temporal sample*)	0.345	0.365
Assuming Approximate 10% LTFU			
N=756 (EE or RE overall)	N=135 (HAEE or HARE)	0.381	0.403
	N=270 (All Controls)	0.289	0.306
	N=756 (All EE or RE)	0.210	0.222
N=250 (EE or RE temporal sample)	N=90 (Controls temporal sample*)	0.503	0.533
	N=135 (HAEE or HARE)	0.437	0.463
	N=250 (EE or RE temporal sample*)	0.366	0.387

FDR = false discovery rate, HAEE = highly active endurance exercise, HARE = highly active resistance exercise, EE = endurance exercise, RE = resistance exercise
* Temporal sample subgroups are those randomized to EARLY, MID, and LATE biospecimen collection profiles

Table 11. Detectable effect size (Corr1 – Corr2) with 90% power for between-group differences in correlations

Approximate Group 2 Size (Description)	Approximate Group 1 Size (Description)	Controlling FDR at 5%, assuming 10% of outcomes fall under the alternative hypothesis: alpha = 0.005263158	Controlling FDR at 5%, assuming 5% of outcomes fall under the alternative hypothesis: alpha = 0.002493075
N=840 (EE or RE overall)	N=150 (HAEE or HARE)	(Corr2 – Corr1) = 0.342 ($Corr2^2 – Corr1^2$)† = 15.1%	(Corr2 – Corr1) = 0.360 ($Corr2^2 – Corr1^2$)† = 16.5%
	N=300 (All Controls)	(Corr2 – Corr1) = 0.264 ($Corr2^2 – Corr1^2$)† = 9.6%	(Corr2 – Corr1) = 0.278 ($Corr2^2 – Corr1^2$)† = 10.5%
	N=840 (All EE or RE)	(Corr2 – Corr1) = 0.194 ($Corr2^2 – Corr1^2$)† = 5.7%	(Corr2 – Corr1) = 0.205 ($Corr2^2 – Corr1^2$)† = 6.2%
N=280 (EE or RE temporal sample)	N=100 (Controls temporal sample*)	(Corr2 – Corr1) = 0.436 ($Corr2^2 – Corr1^2$)† = 23.3%	(Corr2 – Corr1) = 0.457 ($Corr2^2 – Corr1^2$)† = 25.4%
	N=150 (HAEE or HARE)	(Corr2 – Corr1) = 0.385 ($Corr2^2 – Corr1^2$)† = 18.6%	(Corr2 – Corr1) = 0.404 ($Corr2^2 – Corr1^2$)† = 20.3%
	N=280 (EE or RE temporal sample*)	(Corr2 – Corr1) = 0.327 ($Corr2^2 – Corr1^2$)† = 13.9%	(Corr2 – Corr1) = 0.343 ($Corr2^2 – Corr1^2$)† = 15.2%
Assuming Approximate 10% LTFU			
N=756 (EE or RE overall)	N=135 (HAEE or HARE)	(Corr2 – Corr1) = 0.359 ($Corr2^2 – Corr1^2$)† = 16.5%	(Corr2 – Corr1) = 0.377 ($Corr2^2 – Corr1^2$)† = 18.0%
	N=270 (All Controls)	(Corr2 – Corr1) = 0.278 ($Corr2^2 – Corr1^2$)† = 10.5%	(Corr2 – Corr1) = 0.292 ($Corr2^2 – Corr1^2$)† = 11.5%
	N=756 (All EE or RE)	(Corr2 – Corr1) = 0.204 ($Corr2^2 – Corr1^2$)† = 6.2%	(Corr2 – Corr1) = 0.215 ($Corr2^2 – Corr1^2$)† = 6.8%
N=250 (EE or RE temporal sample)	N=90 (Controls temporal sample*)	(Corr2 – Corr1) = 0.456 ($Corr2^2 – Corr1^2$)† = 25.4%	(Corr2 – Corr1) = 0.478 ($Corr2^2 – Corr1^2$)† = 27.6%
	N=135 (HAEE or HARE)	(Corr2 – Corr1) = 0.403 ($Corr2^2 – Corr1^2$)† = 20.3%	(Corr2 – Corr1) = 0.423 ($Corr2^2 – Corr1^2$)† = 22.1%
	N=250 (EE or RE temporal sample*)	(Corr2 – Corr1) = 0.344 ($Corr2^2 – Corr1^2$)† = 15.3%	(Corr2 – Corr1) = 0.362 ($Corr2^2 – Corr1^2$)† = 16.7%

FDR = false discovery rate, HAEE = highly active endurance exercise, HARE = highly active resistance exercise, EE = endurance exercise, RE = resistance exercise

* Temporal sample subgroups are those randomized to EARLY, MID, and LATE biospecimen collection profiles

† Additional explained variability defined as the R^2 from Group 2 simple linear regression minus R^2 from Group 1 simple linear regression. Assumed small, positive correlation (Corr1=0.05) in controls, "highly active" and one intervention group when directly comparing intervention groups. Because the correlation in Group 1 is defined as 0.05, the R^2 for this model in Group 1 is 0.25%.

9.3 POPULATIONS FOR ANALYSES

Both ITT and per-protocol (**PP**) analyses will be performed by members of the Consortium with the data collected in MoTrPAC. Analyses will be performed across all adults and within subgroups, which will include, but will not be limited to, age, sex, and ethnic/racial minorities.

9.4 STATISTICAL ANALYSES

9.4.1 GENERAL APPROACH

The MoTrPAC project collects a wide range of clinical phenotype data, including morphometric, physiologic, demographic, genetic, metabolic, behavioral, and psychosocial measurements. Because MoTrPAC is primarily aimed at the development of a molecular map, there are no primary outcomes or hypotheses; however, a number of within-group and between-randomized-group comparisons are of interest for these measures. In this section, we provide general approaches used to perform within- and between-group comparisons; more specific details are included in the analysis plans associated with each publication/report. Both ITT and PP analyses are possible with the data collected in MoTrPAC.

9.4.2 GENERAL STATISTICAL APPROACHES

Descriptive reports consist of summary statistics (means, standard deviations (**SDs**), percentiles, proportions, etc.) for participant characteristics and outcome measures by intervention, actual versus projected accrual, and quality control (**QC**) information (retention, missing data, etc.). Correlation analyses are used to characterize associations between- and within-group and their changes in measures. Estimation and hypothesis testing for most clinical measures are analyzed using standard generalized linear models (**GLM**) approaches (i.e., linear regression, logistic regression, Poisson regression). Regression diagnostics and residual plots are reviewed to determine if transformations are needed to ensure the linearity, homogeneity of variances, and normality assumptions of the residuals are met.

When repeated outcomes are available, continuous outcomes are analyzed using linear mixed effects models (**LMMs**) to estimate the intervention effects. Count data (e.g., number of events) are analyzed using Poisson or negative-binomial regression with a log link, where the natural log of the follow-up time are included in models as an offset term. GLMs, including generalized estimating equations (**GEE**) for repeated measures, are used to analyze binary outcomes with a logit link function. For between group comparisons of randomized groups, constrained mixed-models can provide more efficient estimates of post-randomization treatment differences when either baseline or post-randomization measures are missing[68, 69]. Where appropriate, contrasts are used within the framework of these LMMs to test the effect of the intervention at specific time points and to estimate within group changes.

Mediation analyses. Mediation analyses may be desired to reduce bias in cases where adherence is likely to be a significant factor of observed effect sizes. Direct methods that use a continuous mediating variable can reduce bias while preserving high power[70]. However, use of mediation analytical techniques in a randomized trial setting have traditionally assumed that the mediation factor is random among individuals (i.e., the sequential ignorability assumption), making such analyses vulnerable to the effect of unmeasured confounders, which can lead to biased inference[71].

To alleviate these issues mediation approaches from two families of causal mediation models can be used: structural mean models[72-74] and principal stratification[75, 76]. For example, counterfactual frameworks that use regression analysis to reduce the indirect effects even when unobserved confounders are present can be used to assess how post-randomization levels of adherence and other

intermediate measures may account for the indirect effects of the intervention on clinical outcomes within the framework of the models used to measure the direct effect of the intervention on these outcomes. Shpitser discusses practical graphical methods that can be easily implemented to address the point above[70].

Missing Data. Several procedures have been incorporated into MoTrPAC for addressing the issues of missing measurement data.

1. Prevention: First and foremost, every effort is made to minimize missing measurements
2. Definitions of being off-study: One approach to analyses is focused on randomized comparisons and incorporate ITT considerations that include all participants who were randomized, regardless of adherence. Participants are only considered to be "off-study" if they have withdrawn from the study in writing (where possible) for all future contact or have died. Participants who discontinue the intervention are encouraged to continue study assessment procedures. PP analyses also form an important component of analyses and are based on predefined definitions of intervention adherence.
3. Analytical plans routinely follow the recommendations of the 2010 National Academy of Sciences report. (National Research Council, 2010) Where possible, we use all observed outcome data through the use LMMs, thus taking advantage of statistical properties that account for "missingness" that may be dependent on the levels of previously collected outcomes (termed Missing at Random (**MAR**) in the statistical literature). Sensitivity analyses, described below, are performed to investigate how estimates (and conclusions) may be altered under various assumptions about the missing data process.
4. Imputation of molecular phenotypes can be applied to account for some types of missing information. A common example is in genotyping, where missing variants can be imputed using genetic correlation. For molecular phenotypes, measurements can also be imputed when significant correlation structure is present. A common example is through Transcriptome-wide association studies (**TWAS**)-based approaches where genetic information is used to predict missing gene expression measurements.[77]

The mixed effects model based on maximum likelihood estimation is unbiased if missing data are unrelated to outcomes, i.e., if the data are considered MAR or missing completely at random (**MCAR**). Nevertheless, because it is not always known whether missing data are ignorable and because missing observations have the potential to alter the results of analyses, the pattern of missing data and dropouts are examined among the intervention groups.

We also examine whether missing outcomes are related to prior values of the outcome. Different association tests can be used here. For example, by using logistic regression models to determine if the outcome measure at the follow-up times preceding the missed visit predicts that the next value is missing and if baseline value predicts monotone missing outcomes (i.e., non-intermittent). Covariates in such models may be defined *a priori*, or selected based on statistical significance or magnitude of effects. This type of exploration is critical to the extension of GEEs to account for MAR missingness through inverse probability weighting or the inclusion of propensity scores based on baseline covariates in a wide array of analysis techniques.

Adherence and Retention Analyses. Adherence is measured in terms of attendance at structured intervention sessions, the quality and quantity of exercise performed during those sessions, and through wearable activity monitors. Summary measures of adherence (e.g., percent of expected sessions attended, average heart rate intensity, percent of time in target heart rate, etc.) are calculated and

summarized by clinical site, intervention group and follow-up time, among other divisions. Retention summary measures include statistics focused on percent of each measure obtained, percent loss to follow-up, and percent withdrawn. Summaries are tabulated by clinical site, intervention group, and follow-up visit.

Multiple Comparisons/Multiplicity. In keeping with MoTrPAC's overarching aim of discovery, we advocate a general philosophy toward adjustments for multiple hypothesis testing, whether this involves multiple outcomes, multiple comparisons for the same outcome, or multiple exploratory analyses. Individual future published reports using the MoTrPAC data should pre-specify how multiple comparisons are accounted for in separate analysis plans. For investigations focused on a limited number of clinical endpoints, this may involve control for the family-wise error rate; whereas, for questions focused on discovery involving a large number of comparisons, control for the FDR may be more appropriate.

9.4.2.1 BIOINFORMATIC STATISTICAL AND ANALYTICAL PLANS

The analysis plan at the BIC consists of several components. Initial analysis includes preprocessing, feature extraction, and QC of the raw and processed data, with emphasis on second-tier QC and batch effect correction steps. Preprocessing and feature extraction are undertaken at each CAS for local QC purposes and again centrally using agreed upon pipelines. Subsequent analyses include differential abundance and temporal mapping, with analyses becoming progressively and adaptively more complex. Such advanced analyses can include the use of quantitative trait loci (**QTLs**), predictive models, multi-omic integration, and time-series analysis. The different steps are taken with the end goal of integrating the results into a molecular map of significant transducers.

Interim data are continually analyzed to evaluate signal to noise for each molecular assay. Re-sampling of existing data are performed to see if effects are robust and variances have stabilized. Extrapolation through simulation helps to determine the point at which variances stabilize and assess if results from a specific assay are near saturation.

Several classes of analyses are used to explore associations. Below we provide brief descriptions of these.

Differential Abundance. Data from quantitative molecular assays are tested for contrasts across time within individuals and compared across intervention groups. Preliminary analyses will compare all acute testing time points against their respective baseline for both pre- and post-intervention sessions. Other analyses will perform similar explorations of long-term and may include exercise type as a covariate. As discussed above, GLMs or generalized linear mixed models (**GLMMs**) are useful tools for detecting differential abundance while accounting for covariates, interactions, and time. GLMs estimate fixed effects, whereas GLMMs can be more powerful as they incorporate random effects such as those introduced by batches, sites, etc. using less parameters.

The molecular analyte (transcript, protein, metabolite) at a given time point in a specific participant is the response variable. An appropriate model family is selected for its distribution. For example, count-based data, such as RNA-seq and ATAC-seq, are considered to have a negative binomial distribution, as implemented in the deseq2 R package. Time may be modeled as a continuous or discrete ordinal variable. Contrasts can be extracted for specific time comparisons as required. Interaction will be tested to examine differences in molecular effects across participant groups such as intervention arms. GLMMs will be used to directly model batches or to perform a meta-analysis of analyte effects over batches.

While the models above can use a phenotype of interest as a covariate in the model and estimate its effect, they are less suitable for prediction or when multiple analytes are used. For this goal, a more suitable approach is to use analytes as predictor variables and the trait of interest as the response, with suitable link functions within GLMs or GLMMs.

Time Series. A primary objective of MoTrPAC is generation of a "molecular map". The univariate analyses above utilize time and the results can be used for developing an initial map. For example, using time as an ordered and/or categorical covariate, "maps" for both transient and non-transient analyte responses can be developed. Graphical displays of results will be frequently used to further understand emerging patterns. More advanced inference methods are discussed in the next section.

Briefly, development of the map will involve leveraging repeated molecular measurements in time at both short (hours) and long-term scales (weeks). We will explore methods for leveraging time series molecular data as they become available. Linking different time trajectories using clustering or alignment methods may provide ways to add dynamic features to the learned map.

Multi-omics and multi-tissue integration. We approach integration across 'omics from a number of perspectives based on evolving best practices. Initial steps include mapping all molecular analytes to a common reference ontology, and learning simple correlation networks between and within 'omics[78]. Another approach is to use multi-omics factor analysis (**MOFA**) to explain main axes of variation in the data in a way that links the different 'omics[79]. Such analyses allow linking of metabolites to proteins and corresponding genes and transcripts.

QTL. Whole genome sequencing (**WGS**) data are generated in the second half of the project as sequencing costs decrease. At that time, QTL analysis are performed. Broadly speaking, this analysis involves modeling quantitative molecular features such as transcript and metabolite levels (or change due to perturbation) as a function of genomic variation, usually in the form of single-nucleotide polymorphisms (**SNPs**).

Multiple testing considerations of the molecular data analysis. Most of the described analyses are applied iteratively across polymorphisms, genes, molecular assay, and tissue. Each strata and even gene have different power to detect differences based on the signal to noise of the underlying assay/biology. Moreover, certain assays have dramatically different multiple testing spaces resulting in highly variable FDR corrections. In general, power is expected to be lowest for single-nucleotide polymorphism analyses and increase with transcriptomics, proteomics, and be maximal in metabolomics. We will initially use standard FDR-based procedures such as the Benjamini and Hochberg or Benjamini and Yekutieli methods. Note that unlike Bonferroni correction, FDR can be powerful even when the number of tests is in the order of millions. Nevertheless, we will take advantage of advances in the field to control error rates; i.e., stratified FDR and hypothesis weighting using 'omics-specific known covariates that can help improve detection power (e.g., using minor allele frequencies as weights for SNP *p*-values)[79].

9.4.2.2 USING BIOINFORMATICS TO LINK CLINICAL AND MOLECULAR MEASUREMENTS

There are simple and complex modeling and analysis questions that should be addressed at different time points during the trial. The most fundamental at early stages is looking at differential abundance for contrasts across specific conditions. Contrasts include time points surrounding acute testing, before and after chronic training, and between different intervention groups. These basic analyses can be performed using simple t-tests (or Wilcoxon tests) for certain molecular elements or GLMs for count data, such as RNA-Seq under a Poisson or a negative binomial with adjustment for relevant covariates.

Response QTLs and response to PA modeling are two additional examples of analyses that leverage both clinical and molecular data. Response QTLs look to associate polymorphisms to change in phenotype, i.e., gene fold-change, change in VO_2peak, change in strength, change in body composition, change in heart rate, change in metabolic metrics, etc. This can be extended to test for association with a combination of clinical and intervention naive molecular signals that are predictive of intervention response.

9.4.2.3 SYSTEMS BIOLOGY APPROACHES

Systems biology approaches[80] are used to integrate clinical and molecular measurements into a map of molecular transducers of physical activities. Omics and clinical data are collected on many scales (molecular, tissue, and whole body levels), from various organs, many participants, and at many times. Such data are used for systematic reconstruction of the network of molecular transducers across temporal and spatial scales. Two categories of systems biology approaches are used: hierarchical models and dynamics models.

Hierarchical models are used to address the granularity of biosystems. Typical approaches include agent-based models[81, 82], compartmental models[83], and hybrid models. Dynamic models are employed to describe the information propagation across multiple temporal and spatial scales corresponding to the measurements at different time points and from different tissues. These methods are used to study how clinical interventions trigger biomedical events through the molecular transducer network and finally drive the different clinical outcomes at endpoints. Hierarchical models provide the conceptual vision of the molecular transducer networks, while the dynamic models describe the temporo-spatial changes of key biomedical traits for each participant. Together, these systems biology approaches bridge the gap between multi-omics data and clinical interventions and outcomes.

Network Biology Approaches. Correlation-based network and modularity analysis are useful for several applications including predicting function, detecting hierarchical structures in complex pathways, and simplifying the data such that new hypotheses can be generated. For example, the association structure of the analytes can be used to identify ions derived from the same metabolite, or even detect associations between environmental exposures and cellular molecules[84]. We will explore simple clustering analyses (e.g., hierarchical clustering) followed by enrichment analysis as a tool to get initial results. Expected issues that will require special attention in our dataset can be categorized into three types: **1)** quantifying association in the presence of time and repeated measures, and **2)** normalization of the correlation matrices, both inter- and intra-omics, and **3)** integrating the results across tissues.

Dynamic Models. Several graphical models have been suggested previously to handle longitudinal biomedical data. Most methods were developed to answer specific questions, mostly in modeling gene expression and regulation. The Dynamic Regulatory Events Miner (**DREM**) [85, 86] for example, uses input-output Hidden Markov Models (**HMMs**) that allow modeling events of differential co-expression over time. Given a partition of genes into different modules in a specific time point, a classifier is fitted as an attempt to explain what drove the observed expression change. DREM uses transcription factor-target networks as these data, but in essence, any gene-based exogenous information can be used, including other ‘omics[87] DREM and its more recent variants will be tested, especially as a tool for deciphering likely regulatory events. A unique property of our study is that we will be able to test if the detected models are consistent with results from other ‘omics and tissues, which can be used to augment the original models with more information.

Another example is a tool called **TWIGS** (three-way module inference via Gibbs sampling) [88] that was developed for detecting differential abundance modules in longitudinal datasets that have many

participants. It is more flexible in that it allows a module to be only partially represented in a participant (e.g., ones that have lower level responses but in the same biological pathways).

Interpretation and Enrichments. Enrichment analysis will be critical for evaluating detected modules. Standard analyses such as overlap enrichment of Gene Ontology (**GO**) terms or known pathways will be used regularly. Integration of the detected results (e.g., across tissues) can be addressed by more advanced network analyses. For example, network propagation approaches[89] can be used to test, for example, if there is a significant flow of information between two gene sets in an underlying network. Another approach to handle multi-omics datasets is to use network summary methods. For example, these methods can be used to integrate different correlation networks (e.g., under sedentary or control conditions vs. responses after exercise or from different 'omics) and summarize them in a map of linked modules. Such analysis can detect novel patterns, and fine grained modules of molecules that are likely to work together [90].

9.4.3 ANALYSIS OF THE PRIMARY EFFICACY ENDPOINT(S)

There are no pre-specified primary efficacy endpoints.

9.4.4 ANALYSIS OF THE SECONDARY ENDPOINT(S)

There are no pre-specified secondary endpoints.

9.4.5 SAFETY ANALYSES

AEs and SAEs are coded using the MedDRA. We summarize frequency of all SAEs by system organ class, preferred term, and intervention group, with the number of participants and percentage reporting the event. We summarize withdrawals due to SAEs for each body system and preferred term by intervention group.

Descriptive statistics (number of observations, mean, SD, minimum, proportions, median, and maximum values) are calculated for clinical laboratory tests at baseline. The number and percentage of subjects with abnormal laboratory results are provided.

The exact presentation of analysis results for safety are developed in collaboration with the DSMB. No adjustment for multiplicity are made for pre-specified safety endpoints, because this is considered a conservative approach; significance levels on safety endpoints that may arise are interpreted cautiously.

In general, between-group differences in the proportion of participants with predefined AEs are assessed using either chi-square or Exact Tests (when expected cell counts are small). The proportion of participants with other adverse experiences are estimated within intervention groups, and expressed as relative effects via odds ratios (95% CIs) or as absolute effects (95% CIs) for between-treatment differences in proportions. Continuous safety measures are compared between groups using t-tests and 95% CIs. Differences between means are compared using t-tests assuming homogeneity of variance, where appropriate. Should this assumption appear inappropriate, Satterthwaite's t-test accounting for unequal variances is used. Trends over time in continuous outcomes are investigated using mixed effects models.

9.4.6 BASELINE DESCRIPTIVE STATISTICS

The adequacy of randomization is assessed by comparing the distribution of baseline characteristics among the intervention groups: demographics, physical examination and medical history measures, symptoms, and medications. Continuous variables are examined for skewness, outliers, or other departures from a normal distribution graphically and by summary statistics. F tests for equality of

variances are used to determine the appropriate statistical test for continuous measures. Categorical variables are examined by calculating frequency distributions.

9.4.7 PLANNED INTERIM ANALYSES

Data are generated over five years with pre-planned internal and external timely releases. All analyses and resulting data and results are shared in compliance with the NIH Genomic Data Sharing (**GDS**) policy and DSMB requirements for the randomized study. More details on data sharing are in Section 10.1.10.2.

9.4.7.1 PHASE I

Phase I focuses on data collected early in MoTrPAC and primarily consists of internal analysis. During this phase, emphasis is analysis of molecular data (raw and processed) without making clinical associations. Phase I is considered to focus on high-resolution molecular profiling of all samples from a limited number of participants. This is the map building phase. Some activities planned for phase I include the following.

- QC analyses focused on estimating technical variation across the large number of batches (i.e., site, device, assay). These analyses will help identify any technical issues at the clinical or chemical sites, with issues that arise communicated back to the Consortium for review and adjustment.
- Analyses to review information content and signal saturation for different analytical endpoints. First and foremost, this investigates signal to noise estimates (i.e., coefficient of variation and intra-class correlation among replicates) at different sampling time points, among different tissues, interventions, and molecular assays. Variance associated with population differences, inter-person variability, environment, study design, and latent batch are thoroughly examined. This analysis will begin with a minimum of 30 participants per time point because some variances are expected to become stable at that sample size. These results will be presented for discussion among the Consortium to assess the merits of adjusting sampling depths for certain time points or molecular assays in building a molecular map.

9.4.7.2 PHASE II

Phase II focuses on data collected in the later years and consists of basic exploratory analyses. Phase II is considered a shallow dive that focuses on population variability and gaining an understanding of individual differences.

- As sample sizes meet minimum analysis milestones to be specified in an interim analysis plan, then specific analyses are completed where appropriate. The most basic analysis includes differential abundance over both short (acute testing) and long (chronic training) time scales. Differential abundance systematically compares relevant contrasts such as different time points in the acute testing or between training naive and trained acute tests.

9.4.8 SUB-GROUP ANALYSES

Exploratory analyses will be performed within subgroups, which will include, but will not be limited to, age, sex, and possibly ethnic/racial minorities. However, given the sample size of the study some of these analyses will be underpowered.

9.4.9 TABULATION OF INDIVIDUAL PARTICIPANT DATA

Individual response data are monitored using multiple approaches. Web-based reports (both tabulations and graphical summaries) provide feedback to sites on potential outliers in real time. Individual longitudinal measures are explored with spaghetti plots, allowing identification of trends and/or unexpected changes through time. Safety reports contain descriptions of events to be reviewed by the MCSS subcommittee and the DSMB.

9.4.10 EXPLORATORY ANALYSES

As a randomized, mechanistic trial aimed at discovery, MoTrPAC is expected to generate many exploratory analyses, both by investigators in the Consortium and by the scientific community at large.

10 SUPPORTING DOCUMENTATION AND OPERATIONAL CONSIDERATIONS

10.1 REGULATORY, ETHICAL, AND STUDY OVERSIGHT CONSIDERATIONS

10.1.1 INFORMED CONSENT PROCESS AND DOCUMENTATION

Informed consent is obtained in compliance with the principles of the Declaration of Helsinki , sIRB regulations, local regulations, and ethical standards of the applicable IRBs of the MoTrPAC clinical sites. The sIRB reviews and approves all informed consent documents before they are presented to the volunteers.

1. All volunteers must be informed both written and orally.
2. All volunteers must give their informed consent before the screening and randomization (although some initial pre-screening via telephone/email is allowed prior to giving written informed consent (e.g., age, sex, BMI, exercise history)).
3. Informed consent is obtained from the volunteer by the clinical site investigator and/or by the designated research staff. The MoTrPAC research staff fully inform the volunteer of all pertinent aspects of the MoTrPAC study including the objectives, inclusion, intervention, study procedures, significance, risks, and implications of the study. All volunteers are informed to the fullest extent possible, in language and terms they are able to understand. Volunteers are given ample time and opportunity to inquire about details of the study and to decide whether to participate in MoTrPAC. Volunteers are informed that their participation is voluntary and that they may withdraw consent to participate at any time.
4. Prior to a volunteer's involvement in the trial, the written ICF is signed, name filled in, and personally dated by the participant and by the person who conducted the informed consent discussion. A copy of the signed and dated written ICF is provided to the volunteer. The original signed ICF is stored in the volunteer's individual file, held by the investigator.
5. The form used for obtaining the volunteer's informed consent must be the current version that has been reviewed and approved by the appropriate sIRB or Ethics Committee.

Teleconsenting is used in place of in-person consenting when possible to reduce unnecessary in-person encounters for the consent procedure. For the teleconsent process, participants are given a copy of the ICF before the teleconsent meeting via email, fax, or mail or at an in-person visit.

Participants have adequate time to consider the research study and ask questions before signing the ICF. The consent designee verifies the participant physically signed the ICF by viewing via video conference, obtaining a photo of the signed consent document, or obtaining verbal confirmation from the participant that he/she signed the ICF or agreed to participate electronically. The participant signs and dates the ICF and then emails, faxes, or mails the ICF to the consent designee. The participant is instructed to return the original signed ICF at the first in-person visit. If the original ICF is mailed, the consent designee signs and dates the copy in their possession after the participant has acknowledged signature on their copy. When the original is received by the consent designee, any copies are attached to make a single document. In all other instances, the consent designee signs and dates the ICF when it is received. After the consenting process is completed, the ICF is filed in the research record.

10.1.2 STUDY DISCONTINUATION AND CLOSURE

The NIH may terminate or curtail the study (or an individual award) in the event of **a)** failure to implement the study protocol, **b)** a substantial shortfall in participant recruitment, follow-up, data reporting and dissemination, QC, or other major breach of the protocol, **c)** substantive changes in the agreed-upon protocol with which the NIH does not concur, **d)** reaching a major study objective substantially before schedule with persuasive statistical evidence, **e)** emergence of new information external to the study, such as scientific developments that impact the planned conduct of the study, and/or **f)** human subject safety and ethical issues that may dictate a premature termination.

10.1.3 CONFIDENTIALITY AND PRIVACY

10.1.3.1 CONFIDENTIALITY AND SECURITY OF DATA

Confidentiality of data is maintained by using research identification numbers that uniquely identify each individual. Safeguards are established to ensure the security and privacy of study records. Appropriate measures are taken to prevent unauthorized use of study information. The research Participant Identification (**PID**) number is used. The research records are kept in a locked room in the clinical site. Only study personnel have access to these files. After the study is completed, local data are stored with other completed research studies in a secured storage vault.

In compliance with HIPAA and the Standards for Privacy of Individually Identifiable Health Information of the DHHS, MoTrPAC accesses PHI after receiving signed informed consent. If the collection of outside medical records is required, a medical release authorization is obtained prior to creating a records request. If a participant's medical records are obtained, they are reviewed and abstracted. Such records are stored in a locked cabinet.

The following participant identifiers will be accessible by each of the different MoTrPAC operational components:

- **Names and contact information (including telephone number and email address):**
    - Clinical sites will have access only to information on their own participants
    - DMAQC will maintain a repository of contact information on all MoTrPAC participants on a secure server, with this information only accessible to authorized members of the DMAQC
    - BIC, Biological Sample Repository, and CAS will not have access to this information
- **Geographic identifiers:** Site identifiers are essential for performing analyses and masked site identifiers (i.e., different codes than are used internally) that allow analysts to group observations by site will be released broadly. We note that data from participants from the pediatric site (see section 2.1 of adult protocol for information about pediatric site data sharing) will be identifiable based on the age of the participants.
    - DMAQC will have access to the unmasked site identifiers for all collected data
    - BIC will also be able to identify the actual geographic site for some types of data that are directly transferred to the BIC (e.g., raw accelerometry files, metabolic cart files), which will enable the BIC to identify which PIDs are associated with particular geographic sites and which data are from participants at the pediatric site, but will not have any other identifying information for a participant

- Biological Sample Repository will know what samples are shipped to them from geographic sites, but will not have any other identifying information for a participant
- CAS will receive only masked site identifiers

- **Dates:** For data distributed from the DMAQC to the BIC, dates will be converted to days from initial registration for internal and external releases.
  - Clinical sites will have access to all dates for their own participants
  - DMAQC will have access to actual dates, including but not limited to dates of birth, visit dates, and dates of health events
  - As necessitated by the SC approved genomic data sharing policy, BIC will have access to all dates of consent and measures, including visit dates, dates of health events and some types of data that are directly transferred to the BIC (e.g., raw accelerometry files, metabolic cart files);
  - CAS will have access to days from initial registration
- **Age:** Age >89 years will be aggregated to 90 for data transferred from the DMAQC to the Biological Sample Repository. There is no planned aggregation of ages among pediatric site participants.
  - Clinical sites will have access to actual ages for all participants at their sites
  - DMAQC will have access to actual ages for all participants at all sites
  - BIC will have actual ages for all participants aged <90 years and those 90+ years will be listed as 90 years in data received from the DMAQC
  - Biological Sample Repository will not have access to age
  - The CAS will have the same access as the BIC
- **Medical Record Number:** Clinical sites will have access to participant medical record numbers from their own site. No other MoTrPAC entities will have access to medical record numbers.
- **Other:** The following information will not be collected in MoTrPAC:
  - Social security number, passport number, driver's license number, health plan beneficiary numbers, account numbers
  - Device identifiers
  - Identifiable images
  - URL identifiers
  - IP addresses
  - Biometric identifiers

The BIC will scrub potential PHI/personally identifiable information (**PII**) from all data transferred from other MoTrPAC entities upon ingress/attempted ingress by collaborating entities.

All metadata will be housed in a PHI/PII compliant database, and any corresponding data files will be stored in isolated PHI/PII compliant cloud storage buckets (encrypted on disk) in the unlikely event that the automated scrubbing tooling fails to appropriately detect and scrub identifiers.

The limited data set of identifiers (location-based signatures, e.g. those contained in raw data files and dates related to raw data files) will be stored in an encrypted and PHI compliant MoTrPAC database that will be vetted by the Stanford Privacy Office and audited by an external information security consultant.

Prior to public release of MoTrPAC data, all potential identifiers will be removed and indirect identifiers will be vetted by the Stanford Privacy Office per University policy.

10.1.3.2 CONFIDENTIALITY AND SECURITY OF BIOLOGICAL SAMPLES

The MoTrPAC Biological Sample Repository complies with the Office for Human Research Protections requirements and guidelines related to the research use of stored biological samples as stated in "Issues to consider in the research use of stored data or tissues" from the Office of Protection from Research Risks. (http://ohrp.osophs.dhhs.gov/humansubjects/guidance/reposit.htm). In addition, we are members of the International Society for Biological and Environmental Repositories (**ISBER**) and are compliant with ISBER and NIH Biospecimen Repository guidelines. All laboratory specimens, evaluation forms, reports, and other records that are processed in laboratories, used in analyses, or shared with laboratory staff will be identified only by coded number to maintain participant confidentiality. All biospecimen containers will be labeled only with a pre-specified code linked to the participant. As detailed in Section 10.1.3.1, there will be no direct link to participant identifying information without access to protected files containing the identifying information linking the specimens to a given participant. Access to linked identifiers will be limited.

Biospecimen safety and QC are ensured by several mechanisms: **a**) a preventive maintenance program, with daily inspections, and 24/7 certified freezer repair service; **b**) locations are monitored 24/7 with generator backup emergency power; **c**) backup freezer storage space (currently housing >160 freezers); **d**) sample handling in deep-frozen state, with well-developed Biological Sample Repository Standard Operating Procedures (**SOPs**), including dry ice-based local transport and handling and, when needed, standardized thawing protocols; **e**) all label stock is optimized to withstand ultracold temperatures, water baths, etc.; **f**) all technicians are certified annually in handling biohazardous materials; **g**) shipping is generally done on dry ice following International Air Transport Association shipping and packing regulations. Sufficient quantities of dry ice is key, as is a solid working arrangement with overnight shipping contractors; this latter condition is enhanced by the large volume of samples shipped in and out of the Biological Sample Repository on a daily basis. Freezer inspection, maintenance, and cleaning are performed routinely. Storage and locations of samples are tracked in computer databases. Barcode technology and inventory software are used to enhance tracking of sample details including receipt, location, and storage conditions. If hard copies of any documents are required, they are kept in secure file cabinets in locked rooms with limited, authorized-only access. All electronic data are kept in secure, password-only databases on institutional servers behind extensive firewall protection.

10.1.4 FUTURE USE OF STORED SPECIMENS AND DATA

Biospecimens will be collected in a manner that will promote appropriate data sharing with investigators across MoTrPAC and external to MoTrPAC, as specified in the data sharing plan. Stored biospecimens and data can be used for integration into larger databases (e.g., use of common data elements: http://cde.nih.gov/), consistent with achieving the goals of MoTrPAC. Data will be collected in a manner to allow other researchers to use the stored biospecimens for analysis, as well as for analysis of the 'omics and phenotypic data, not being conducted by MoTrPAC investigators, including conducting meta-analyses.

The BIC will submit data to external databases (see Section 10.1.10.2) on behalf of the Consortium, with the Stanford University Institutional Signing Official as signing authority, to certify compliance with the NIH GDS policy and to handle Institutional Certification with NIAMS as the Institute and the NIAMS Genomic Program Administrator (**GPA**) to facilitate this process. This will require coordination among the sIRB (adult), University of California Irvine (pediatric), and Stanford IRBs, the GPA, and the Stanford institutional signing official to ensure the coordination of IRB approval and Institutional Certification.

10.1.5 KEY ROLES AND STUDY GOVERNANCE

Funding
The NIH Common Fund through cooperative agreements managed by:

- National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
- National Institute of Diabetes and Digestive and Kidney Diseases (**NIDDK**)
- National Institute on Aging (**NIA**)

Participating Institutions:

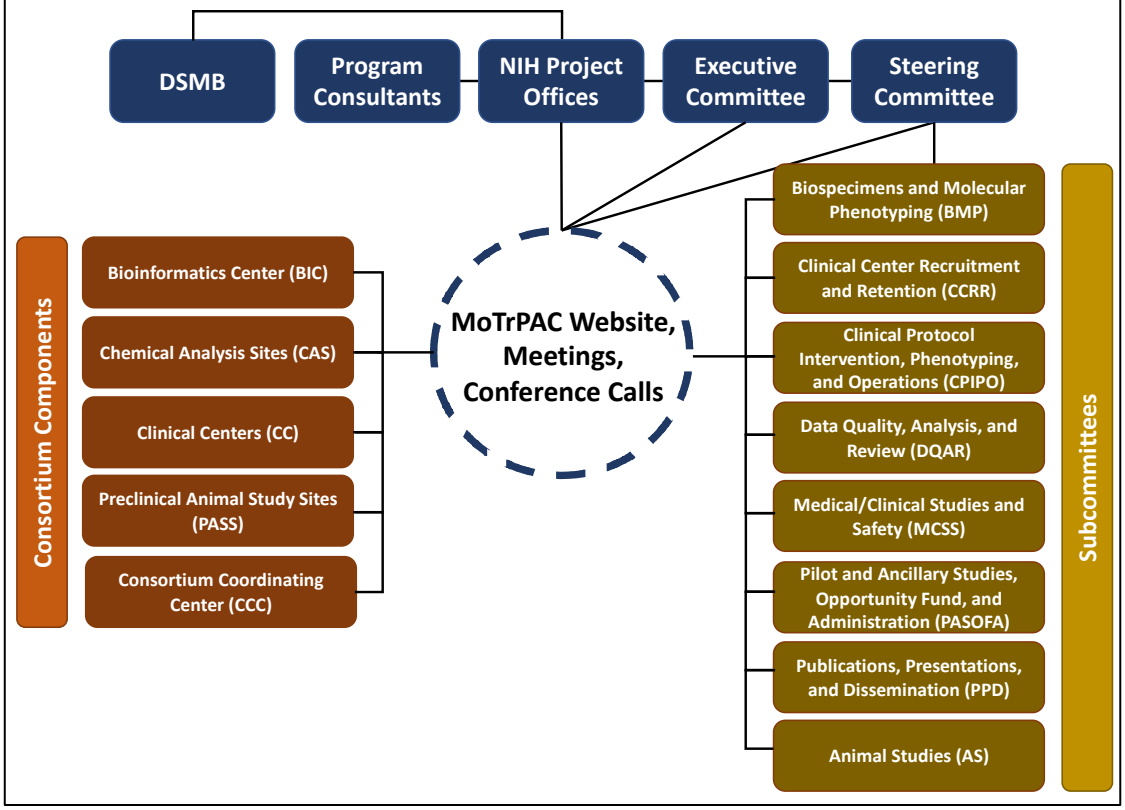
1. **Bioinformatics Center**
   - Stanford University
2. **Chemical Analysis Sites**
   - Broad Institute
   - Duke University
   - Emory University
   - Georgia Intitute of Technology
   - Icahn School of Medicine at Mount Sinai
   - Mayo Clinic Rochester
   - Pacific Northwest National Laboratory
   - Stanford University
   - University of Michigan
3. **Adult Clinical Centers**
   - Exercise & Physical Activity Collaborative Team Clinical Center
     - Ball State University
     - AdventHealth Orlando
     - University of Alabama at Birmingham
   - North Carolina Collaborative Clinical Center
     - Duke University
     - East Carolina University
   - Pennington Biomedical Research Center
   - University of Colorado Denver
   - University of Texas Health Science Center, San Antonio
   - University of Texas Medical Branch, Galveston
4. **Consortium Coordinating Center**
   - University of Florida
   - Wake Forest School of Medicine
   - Wake Forest University
   - University of Vermont
5. **Preclinical Animal Study Sites**
   - Joslin Diabetes Center

- Univeresity of California, Los Angeles
- University of Florida
- University of Iowa
- University of Kansas
- University of Missouri
- University of Virginia

10.1.5.1 MOTRPAC ORGANIZATION (**FIGURE 5**)

The Steering Committee (**SC**) is charged with the overall governance of study design and conduct to ensure that all aspects of the protocol and the analysis plans contribute to the MoTrPAC goals. The Executive Steering Committee (**ESC**) deals with various administrative issues and other governance aspects. The DSMB is described in detail below. The **CCC** has four components: **1)** The Administrative Coordinating Center (**ACC**) facilitates communication within the study, develops and monitor subcontracts with external providers, administers opportunity funds, matches timelines and deliverables, facilitates communications with and submissions to sIRB and the DSMB, review unmasked safety data to be reported to the DSMB, and oversees all administrative matters (e.g., scheduling meetings and conference calls, taking minutes, arranging training and site visits, producing study documents etc.); **2)** The DMAQC is responsible for data acquisition, management, initial analysis, and QC; **3)** The Biological Sample Repository oversees the collection, storage, tracking, shipping and QC of the biological samples; and **4)** The Exercise Intervention Core (**EIC**) coordinates the intervention protocols and implements intervention QC. The CCs, after **s**IRB and DSMB approval, recruit study participants, administer the interventions, ensure adherence and retention, conduct assessments, and collect tissue biospecimens. The BIC oversees data standardization, integration, and storage and implements data sharing and computational tools for the integrated analysis of clinical and molecular data. The Preclinical Animal Study Sites (**PASS**) will collect samples and data from rodents in parallel with the CCs to interrogate biospecimens beyond blood, muscle, and adipose. There are eight subcommittees that oversee MoTrPAC operations: **1)** CCRR designs and implements recruitment and retention strategies for study participants; **2)** Clinical Protocol Intervention, Phenotyping, and Operations (**CPIPO**) develops study eligibility criteria, designs and implements the exercise protocol, monitors and promotes intervention adherence, and designs and implements the participant's phenotyping and assessments; **3)** MCSS oversees safety, addresses IRB and clinical issues, promptly reviews severe AEs, and provides feedback to the CCs on follow-up measures; **4)** Publications, Presentations, and Dissemination (**PPD**) encourages production of high quality publications and presentations and assures maintenance of a database on study publications; **5)** Pilot and Ancillary Studies and Opportunity Fund Administration (**PASOFA**) monitors the cutting-edge science, and stimulate and review ancillary studies. If Opportunity Funds are available, PASOFA issues and publicizes requests for applications outlining the eligibility, scope, application instructions, and review criteria for ancillary studies; **6)** Biospecimens and Molecular Phenotyping (**BMP**) plans and implements molecular analyses; **7)** Data Quality, Analysis, and Review (**DQAR**) plans and implements data analyses and data QC procedures; and **8)** Animal Studies (**AS**) oversees the rodent studies that complement the clinical protocol by providing tissues for analysis from exercised animal models that cannot be obtained from humans. The targeted timeline for the clinical study is described in **Table 12**.

Figure 5. MoTrPAC Organization



10.1.5.2 CLINICAL STUDY TIMELINE

Table 12. MoTrPAC Clinical Study Target Timeline

Calendar Year	2017				2018				2019				2020				2021				2022				2023			
Grant Year	1				2				3				4				5				6				NCE			
Quarter	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Protocol development																												
MOP, CRF development																												
Clinical site training																												
Vanguard phase																												
Clinical visits, interventions																												
Biospecimen analysis																												
Manuscripts																												

MOP = manual of procedures, CRF = case report forms
Black entries indicate suspension of all clinical activities due to COVID-19
Hatched entries indicate partial suspension of clinical activities (some sites)

10.1.6 SAFETY OVERSIGHT

Each clinical site PI is responsible for ensuring participant safety. The DSMB acts in an advisory capacity to the NIAMS to monitor participant safety, evaluate the progress of the study, and review procedures for maintaining the confidentiality of data, quality of data collection, management, and analyses.

A Data and Safety Monitoring Plan (**DSMP**) is implemented to ensure the safety of all participants involved in the study and to ensure the validity and integrity of the data. The PIs with the advice and assistance of the MCSS and the SC monitor all aspects of safety. The MCSS reviews all SAEs and AEs and makes recommendations to the SC for any changes in reporting, consent, or study activities.

A DSMB is established with responsibility to monitor all aspects of the study, including those that require access to any masked data. The DSMB and its chair are named and approved by the NIAMS. The DSMB meets by conference call or in person as determined by the DSMB and the NIAMS. The DSMB has access to all study data, documents, and progress.

The DSMB has the following charges:

- Review the study protocols and the ICFs with regard to recruitment, randomization, assessment, procedures, interventions, participant safety, data management, plans for auditing of participant records, and QC and analysis plans, and to identify needed modifications prior to the start of the study.
- Identify the relevant data parameters and the format of the information to be regularly reported.
- Review data (including masked data) over the course of the trial relating to efficacy, recruitment, randomization, compliance, retention, protocol adherence, trial operating procedures, forms completion, intervention effects on safety parameters, sex and ethnic minority inclusion, and participant safety.
- Identify problems relating to safety over the course of the study and inform the CCC via written report, which in turn ensures that all CC PIs receive this report.
- Identify needs for additional data relevant to safety issues and request these data from the study investigators.
- Propose appropriate analyses and periodically review developing data on safety and endpoints.
- Make recommendations regarding recruitment, intervention effects, retention, compliance, safety issues, and continuation of the study.
- Send study investigators written reports following each DSMB meeting. The CCC submits these reports to the sIRB.

At any time, the DSMB may recommend discontinuation of any component/intervention group of the study for any of the following reasons:

1. Compelling evidence from this (or any other) study of an adverse effect of the study interventions or procedures that is sufficient to override any potential benefits of the study.
2. A very low probability of addressing the study goals within a feasible time frame.

The DSMB may convene an executive session at any time. The NIH makes the final decision on whether or not to accept the DSMB recommendations about discontinuation of the entire study or any of its

components. Any SAEs that might be due to the study intervention are reported to the DSMB, the IRB, and the NIH.

10.1.7 CLINICAL MONITORING

Data integrity and quality are among the highest priorities in clinical studies. There are two primary purposes for QC: to document the level of quality of data collection and to provide feedback to the clinical and laboratory centers in order to maintain and improve the quality of the study data over the course of the study. This chapter outlines the QA and QC activities that are conducted in the MoTrPAC Study. Two expressions are used: **1**) QA, which refers to documenting the quality of the data via the manuals and procedures that are in place to assure the integrity of the data; and **2**) QC, which describes the monitoring and analytic activities that assess performance during data collection and its processing.

The Data Quality, Analysis, and Review (DQAR) subcommittee establishes guidelines for and oversees QC and QA activities for the study overall, integrating input from other subcommittees. Reports are presented to DQAR on a regular basis and any areas of concern are presented to the SC for consideration.

The CCC has the primary responsibility for operational aspects of clinical QC monitoring, working directly with clinical sites on their performance. The MoTrPAC website provides necessary reports allowing multiple study groups, including the CCC, subcommittees and clinical sites, to ensure that all aspects of the study are being carried out in accordance with the protocol.

KAI is the clinical monitoring company for NIAMS and will carry out additional monitoring as needed per the direction of NIAMS.

10.1.7.1 MANUAL OF PROCEDURES

Standardization of study procedures is essential. The MOP includes the detailed descriptions of all study procedures. The MOP is used for training purposes and as a reference throughout the duration of the study for all clinical site investigators and staff. The MOP is updated as necessary throughout the duration of the study and updates are communicated to clinical sites via the MoTrPAC website.

Essential study procedures are standardized and a written description provided in the MOP. This includes procedures such as administration of standard forms and questionnaires, appropriate process for measuring vital signs, biospecimen collection and storage, and other data collection procedures. Furthermore, standard safety event definitions and event validation procedures are used.

10.1.7.1.1 CLINICAL SITE TRAINING

Training for clinical site investigators and staff is crucial to standardizing procedures and assuring high data quality. MoTrPAC uses several different training models that have been proven to be effective in previous clinical studies: central training for clinical site staff, web-based training, on-site training by CCC personnel, and a train-the-trainer approach. Phone calls and web-based platforms are used for periodic refresher trainings.

Some study procedures and forms require certification prior to clinical staff conducting clinic visits. Certifications are tracked centrally and staff are periodically recertified.

10.1.7.1.2 CLINICAL SITE INITIATION

Clinical site initiation to screen and randomize participants is dependent upon completion of a series of preliminary tasks. These tasks include completion of appropriate regulatory approvals and documents; letters of agreement; slinical site staff training and certifications where necessary; receipt of all study

supplies; and the development of a site-specific recruitment plan. The CCC provides the appropriate assistance toward these ends as needed.

10.1.7.1.3 REGULATORY APPROVAL

Clinical sites must have sIRB approval prior to site initiation and participant enrollment. The CCC works closely with the clinical sites to facilitate this approval and provide all documentation needed for submission. An ICF template is developed by the CCC, which is then be adapted for use at local clinical sites.

10.1.7.2 CLINICAL SITE MONITORING

Clinical site monitoring is an important aspect of QC and standardization of protocol procedures and are conducted by personnel from the CCC or other study leadership as necessary. Monitoring takes place remotely by the CCC and through in-person site visits. The DMAQC and BIC provide the DSMB with requested clinical reports as required.

10.1.7.2.1 CLINICAL DATA MONITORING

DQAR, with input from other MoTrPAC subcommittees and the CCC, develops key performance indicators, both to document data quality and to provide feedback to individual clinical sites on their performance, which are tracked in QC reports. All reports are generated by the DMAQC and available on the secure MoTrPAC website. The CCC is responsible for reviewing reports on study progress and site-level quality metrics as requested by DQAR and/or the SC, as well as providing feedback to clinical sites on individual study performance.

10.1.7.2.2 DSMP, DSMB REPORTS

A DSMP is designed to ensure the safety of all participants involved in the study and to ensure the validity and integrity of the data. An independent DSMB has primary responsibility for monitoring the accumulating study data for signs of adverse trends in morbidity/mortality and treatment-related SAEs.

The data and safety monitoring report includes, but is not limited to, the following:

- a protocol synopsis,
- summaries of past meetings and protocol changes,
- updates on study status including screening and randomization rates,
- summaries of PDs,
- demographics of randomized participants,
- retention rates,
- data quality information
- key study measures to evaluate the effectiveness and integrity of the intervention,
- and various safety reports including AEs and SAEs

Most reports are tabulated for the study overall and by site. Each report includes information to be presented during an open session where blinded investigators may attend, and a closed session where aggregate safety data are unblinded (per NIAMS 2017 technical report). Other reports that are specifically requested by the DSMB are also prepared.

10.1.7.2.3 CLINICAL SITE MONITORING VISITS

Clinical site monitoring visits are important to maintain QC and standardization of protocol procedures. These visits are conducted by personnel from the CCC and other study leadership, and may include appropriate staff from the NIAMS and KAI. Monitors visit slinical sites periodically for the purpose of

assuring that the study is being conducted in accordance with the protocol. It is expected that the clinical site PI be present or available for consultation during such scheduled monitoring visits. Site monitors must be given access to all data pertaining to participant participation in this clinical investigation, provided that participant confidentiality is maintained in accordance with local requirements. The scope of these visits is broad and can include, but is not limited to: review of all regulatory documents, study communications, site initiation, site staffing, ICFs, inclusion/exclusion criteria, data verification, and general site performance.

Site visits may also be conducted to evaluate performance deficits in one or more critical area, such as consistent departures from the protocol or MOP. Site visits are also an opportunity for refresher training and/or training of new staff.

The CCC staff prepare a written summary of the site visit, an overview of action items, and list any PDs for the clinical site PI.

A sample of site visit reports and follow-up letters may be reviewed by the SC or other MoTrPAC subcommittees with recommendations for follow-up actions or reporting changes as needed.

10.1.7.3 DATA QUERIES

QA concepts are employed during the development of CRFs. Web-based data entry screens are developed from CRFs, and enable the incorporation of range and logical checks at the time of data entry. These features contribute to QA. Clinical site staff review each set of completed CRFs for accuracy and completeness.

The DMAQC is responsible for data checking, which includes checks for missing data, unrealistic values, and cross checks for inconsistencies. Data are checked on CRF submission and any additional data queries are presented to the data entry clinical staff for immediate resolution, if possible. The DMAQC produces data query reports on the website that summarize the number and types of queries by clinic. Clinical site staff are responsible for reviewing and resolving the data queries in a timely manner.

CRF data may also be reviewed by CCC monitors at clinical site monitoring visits. Source document verification is performed per the monitoring plan. Once data are concluded to be complete and accurate, the CRFs are locked, meaning that the forms become read-only. It is expected that site investigators maintain adequate supervision and oversight such that they can attest to the quality of data collection at their clinical site.

10.1.8 DATA HANDLING AND RECORD KEEPING

10.1.8.1 DATA COLLECTION AND MANAGEMENT RESPONSIBILITIES

To provide a seamless and secure transmission of results from the clinical sites and CAS, and allow the Consortium members and the general public to have access to data, the DMAQC, CAS, and the BIC work together to integrate clinical and molecular data.

The DMAQC is responsible for clinical site participant and intervention data acquisition and standardization, data management, data transfer, and QC analyses. The BIC is responsible for data standardization, ingress, and storage and implements data sharing and computational tools for the integrated analysis of clinical and molecular data.

10.1.8.2 CLINICAL DATA MANAGEMENT

10.1.8.2.1 STUDY WEBSITE OVERVIEW

All clinical sites use the World Wide Web to enter MoTrPAC data collected from participants seen within the clinical sites. Each clinical site has a password protected area on the MoTrPAC home page through which data are entered. Documentation of the data entry system is maintained at the DMAQC.

10.1.8.2.2 DATA COLLECTION

Each participating clinical site maintains appropriate medical and research records for this study, in compliance with federal regulatory and institutional requirements for the protection of confidentiality of participants. Each clinical site also maintains documentation that all members of the research team have completed training requirements. As part of participating in an NIH-sponsored study, clinical records for the purposes of QA reviews, audits, and evaluation of study safety, progress, and data validity are available as required. Data collection is the responsibility of the clinical site staff under the supervision of the clinical site PI. The clinical site PI is responsible for ensuring the accuracy, completeness, legibility, and timeliness of the data reported.

Data are collected in multiple ways at all participant contacts, including electronic CRFs (**eCRF**) or hard copy CRFs and automatically generated machine data. Clinical site staff are expected to review hard copy CRFs for accuracy and completeness and resolve any data issues prior to data entry. Clinical data (including AEs, concomitant medications, clinical laboratory result data) are entered into the MoTrPAC website, a 21 CFR Part 11-compliant data capture system provided by the DMAQC.

10.1.8.2.2.1 DATA ENTRY, VERIFICATION, QC, AND METADATA

During data entry, a variety of programmed error checks are performed for key variables, such as automatic range checks and logical consistency checks, to identify data that are inconsistent, incomplete, or inaccurate. When these edit checks fail, data may be flagged for further review or prevented from becoming part of the study database. At regular intervals, data queries are carried out on the computerized databases to perform consistency checks on key variables and other data. Metadata of the date, person, programmed edit check results, as well as the creation, modification, deletion, transfer, aggregation, and derivation of data are collected and documented.

10.1.8.2.2.2 MACHINE GENERATED DATA

Clinical site-generated measurements utilizing instrumentation are primarily stored at clinical sites. Working with the clinical sites and the BIC, the DMAQC is responsible for standardizing metadata captured for machine-generated data and evaluating the proper storage of machine generated data at clinical sites. Datasets that are not readily abstracted by the clinical sites for input into the eCRFs on the MoTrPAC website are uploaded to the BIC or DMAQC. Clinical sites archive data until the BIC or DMAQC are able to upload and verify data from machines at clinical sites. Data are archived through the end of the study and/or as required by local and national regulations.

10.1.8.2.3 RANDOMIZATION

MoTrPAC uses an internet-based, web-based randomization procedure. Clinical sites access the randomization application through the study web site. Access to this application is password protected and its communications are encrypted. Once security requirements are satisfied, the eligibility of the participant is verified and randomization occurs. When the session is complete, an e-mail indicating that the participant has been properly randomized and appended to the database is sent to the clinical site coordinator, the ACC, and the DMAQC.

10.1.8.2.4 CLINICAL SITE TRACKING

The MoTrPAC website maintains a clinical site Tracking System where all tools used to track various aspects of the study reside. This includes tools for tracking and monitoring recruitment, reporting and monitoring safety, monitoring adherence, and monitoring regulatory activities. The system includes a fully integrated tracking and notification system that advises clinical site staff about participant follow-up windows, and projects clinic and laboratory workload. Tracking a participant begins at screening and continues automatically throughout the study by integrating participant follow-up data with a schedule of target dates for each of the participant contacts.

10.1.8.2.5 SECURITY AND DATA PROTECTION

Data security in the web-based data system uses 2048-bit encryption and SSL. Once data are received at the DMAQC, recovery from disasters such as natural phenomenon (water, fire, or electrical) is possible through the ability to reconstruct both the database management system and the data up to the last back-up through the use of nightly backups. This process ensures optimal recovery of data systems in the event of a disaster.

10.1.8.2.5.1 CLINICAL SITE DATA SECURITY

Paper and/or electronic records for participants are stored at the clinical sites. All records receive the same care, as would ordinary medical records. Access to the data in any local MoTrPAC database is controlled by a system of user identification names and passwords to ensure only authorized staff can enter. Each clinical site staff member is authorized before being given a username, password, and staff to use the data system on the MoTrPAC website.

10.1.8.2.5.2 DMAQC DATA SECURITY

All data collected by the DMAQC are identified only by PID number and is stored at the DMAQC via the secure and encrypted website. Access to the website, privileges to various areas of the website, and to the data on the website is managed by the DMAQC Project Manager in consultation with the clinical site coordinator.

Confidentiality of information within the DMAQC is protected through a variety of procedures and facilities:

1. The confidential nature of the data collected, processed, and stored at the DMAQC is explained to all new personnel.
2. All access to DMAQC office space containing data is controlled through a single door, which is locked and only accessible by key or security badge.
3. All participant data sent to the DMAQC are encrypted as described above.
4. All participant data stored on the WFSM computers are likewise encrypted. In addition, all such databases are protected by passwords that must be supplied before the data can be accessed. Passwords are released only to DMAQC staff with a need to use the particular file, and are changed on a regular schedule.
5. All printouts, plots, and reports containing individually identifiable data are produced on printers and plotters within the DMAQC's secure office space.
6. A Certificate of Confidentiality for MoTrPAC prevents researchers from being forced to disclose identifying information by certain legal proceedings.

PHI such as participant name, addresses, contact information and other identifiers of concern, if collected and data entered are securely stored separately from the main clinical data on eCRFs. Access to these data is limited to a few primary people at the DMAQC and only the PID number connects these data, if required.

10.1.8.2.6 RECORDS RETENTION

Documents pertaining to the study should be retained for a minimum of 2 years after the formal discontinuation of MoTrPAC. These documents should be retained for a longer period, however, if required by local regulations. No records are destroyed. It is the responsibility of study leadership to inform the PI when these documents no longer need to be retained.

10.1.8.3 BIOINFORMATIC (MOLECULAR) DATA MANAGEMENT

The BIC and CAS sites are primarily responsible for molecular assay data management.

10.1.8.3.1 DATA COLLECTION

All chemical data are deposited in the BIC database after primary QA/QC at the CAS. This includes genomic, transcriptomic, epigenomic, metabolomics and proteomics data. Both raw and processed data are deposited. All data uploads include a Consortium defined metadata file that describes the experiment, sample, subject and analysis. Each individual site processes their respective data according to institutional best practices, with key metrics identified by the Consortium.

Clinical data from the DMAQC are synchronized with the BIC for integrative analysis, archiving, and dissemination. A key distinction between the two sites is that efforts are made to minimize PHI and PII from the BIC (see Section 10.1.10.2). Exceptions that cannot be fully anonymized could include those generated in central labs or activity trackers that flow directly to the BIC. These data are stored in a secure partition in MoTrPAC's cloud infrastructure referred to as the "bridge" server where it is anonymized for inclusion within the general data store (see Section 10.1.3.1).

Upon receiving the data, the BIC processes the data in the series of steps outlined below:

1. Transaction Tracking and Validation - the file identity, data format, and metadata schema are validated
2. Primary Processing
3. Primary QC
4. Integration - the data are archived as storage object and the metadata are loaded into the snoVault database
5. Secondary Processing
6. Secondary QC

To optimize file storage, the BIC applies compression strategies of genomic sequencing quality scores to simultaneously reduce file size and improve signal characteristics.

10.1.8.3.2 DATA VERIFICATION AND QC

Our QC model is based on the Encyclopedia of DNA Elements Consortium (**ENCODE**) Data Coordination Center project. QC begins during the initial transaction when data are uploaded from a CAS site (primary QC) and during checkpoints as data generation milestones are met across multiple samples and assays (secondary QC). All external uploads are initially housed in a staging environment where a series of QC checks are applied including assessment of file identifier integrity, data format, and metadata schema validation.

Beyond primary validation of files, samples, assays and subject, there are ongoing analyses to look for consistency of data types over time. It is expected that there are significant batch effects some of which will be readily apparent. The BIC in partnership with the DMAQC exhaustively record metadata to identify and correct these biases.

Both primary QC within uploaded batch and secondary QC across archived batches result in reports that are communicated back to the different sites. This feedback is crucial in addressing any systematic issues at the clinical, chemical, and analytical level.

Transaction. We confirm files are not corrupted (i.e*.,* consistent file identifier keys, MD5 checksum, and file sizes) to ensure fidelity of the upload transaction. Additionally, all globally unique identifiers for the uploaded data are confirmed against the local cache of the DMAQC database. Any records missing are queried and retrieved before the data are ingested.

Identity. A key aspect of molecular QC is ensuring that samples come from the appropriate subject and tissue of origin, and that these samples are not contaminated. Substantial efforts are used to prevent such issues. Genetic data are tested for sex, blood type, and ethnicity to match the donor subject. Duplicates and cryptic relatedness are assessed via identity by descent. Contamination is assessed by heterozygosity levels and allele frequency estimates. Other repeated measure assays from the Genomics, Epigenomics, Transcriptomics (**GET**) sites (i.e., transcriptomics data) are compared for identity using extracted SNPs. Tissues are confirmed using a combination of deconvolution methods that provide tissue proportion estimates and hierarchical clustering. Metabolomic and proteomic datasets are tested primarily using hierarchical clustering within tissue and the individual participant to prevent sample mix-ups.

Checkpoints. We define QC checkpoints at multiple locations within our global workflow. Four examples are: **1)** the initial file upload, **2)** aggregate analysis, **3)** secondary time-point, and **4)** integrative multi-omic analytics. QC checkpoints trigger a set of appropriate validation steps to ensure integrity of the data before they pass to the next step of processing or are released to the community.

Validation. Each data type then undergoes file specification validation using a set of *a priori* rules (https://goo.gl/FCHP7S) and the University of California, Santa Cruz validate files utility (https://goo.gl/hoYpeK). Currently formats including FASTQ, BAM, BED, VCF, bigWig, and PEAKS are supported. We propose to extend functionality to better support the more robust VCF format BCF. gVCF format specifications are also supported (recording a call for every position in a genome sequence) in order to improve incremental calling and simplified extraction of common genotypes for QC operations such as ethnicity, relatedness and QTL analysis. All BAM files are checked against FASTQs for lossless representation. Specific file formats for some genomic features depend on the analytical pipelines implemented by the chemical cores. We work with each core to make this seamless.

Metadata. For each data transaction, a corresponding meta-data file is required. Date, time, feature types, tool versions, sample identifiers, and experimental conditions are some of the generic requirements. Each metadata .json file is validated against a schema including controlled language as described in our ENCODE papers[91, 92]. In addition, we have recently described how our metadata objects are modeled and updated[93].

Batch Effects. The BIC works to identify computational procedures to reduce batch effects across sites and uses these data to conduct multi-omic QC analyses beyond the scope of any single CAS. Data generating cores typically only process small sets of data with little ability to track subtle quality tolerance thresholds across long periods of time. For example, as reagent lots, hardware, personnel, software versions, or SOPs evolve, subtle biases can creep into data. This is particularly evident the more

heavily processed a feature set is (VCF, RPKMs). We propose two strategies. First, we estimate the effects over time to assess drift from tolerance with submitted data and communicate this back to the chemical core. Second, we pursue joint or batch processing of larger sets of samples. This helps to identify biases from sequencing that are apparent from larger sample sets such as contamination[94].

Metabolomics, Proteomics. We follow QA/QC best practices for experimental conditions including replicates, spike-ins, quantity / cell count normalization and appropriately matched controls (matched in tissue, demographics, and measurement technology). The primary QA/QC steps that we currently implement are described in the following document (https://goo.gl/7JrqRy). These steps are not a substantial part of ENCODE or University of California, Santa Cruz infrastructure so we work closely with the CAS to operationalize this step. A key secondary analytic operation for metabolomics and proteomic QC is longitudinal monitoring of patterns (i.e., sensitivity), joint normalization, pre-processing, and identity checking across multiple time points. Deviations from expected patterns help identify problematic samples including mislabeling.

Multi-omic Integrative QC. We propose the use of integrated multi-layer networks to unify our systems understanding of multi-omic data including the time domain[95]. This unified approach offers great opportunity to develop new methods to identify subtle technical errors including contamination, failed or biased assays, and sample mis-attribution. We propose to leverage the inverse of multi-layer centrality to identify samples that are inconsistent across dimensions both within subject and across features[96]. We compare this systems approach to more conventional brute force QTL analysis. As examples: **1)** learn the QTL relationships for the different data types (genotype vs expression, epigenetic, proteomic, metabolomic); **2)** for each data type, sum the residuals across multiple QTLs but within samples. Those assay-sample observations that are maximal outliers in QTL space are questionable and are reviewed. If there are multiple outliers, label-swapping permutation can be applied to see if an alternate genomic background better fits the observed QTLs for each omic data type[97, 98].

10.1.8.3.2.1 CHEMICAL ANALYSIS SITE DATA TRANSFER

CAS directly upload to the BIC Google Cloud. Transfers are encrypted in transit for security. All data are uploaded to the staging “bridge” storage bucket. All new uploads then trigger appropriate data QC, processing, and ingestion workflows. If either the data or metadata fails validation criteria, then the communication is initiated with the CAS representative and the process repeated. Automation evolves over the course of the study.

Representatives from each of the CAS have been selected for accounts within the Google Cloud. They are responsible for the uploading of data and have access to raw data downloading. To avoid unnecessary costs and maximize security, all efforts are made to centralize processing on the cloud.

10.1.8.3.2.2 MACHINE-GENERATED DATA

There are several data flow exceptions that are outside direct transfer from a CAS or DMAQC to BIC. These exceptions include machine-generated data from the metabolic carts, activity monitors, and potential ancillary study data including imaging.

10.1.8.3.3 SECURITY AND DATA PROTECTION

All data transferred to the BIC are encrypted in transit and then reside in the Google Cloud Platform. Egress of raw data to Consortium collaborators are similarly securely transferred and logged. Processed low risk data such as summary statistics or analysis outcomes that are released to the public are not specifically encrypted.

Importantly, Google and Stanford have signed a Business Associate Agreement extending HIPAA liability into the cloud. All efforts are taken to follow this policy even though PHI/PII is purged from the BIC storage.

User management is strictly controlled and data provenance logged. Stanford users have institutional accounts that are centrally controlled. Only persons with full Stanford accounts have access to the core project code and data. External sites within the Consortium designate several representatives that are given access to storage buckets for data transfer but not the main data storage bucket or codebase. Data prepared for public releases are hosted on a separate bucket for broader sharing.

Data exist in different pools according to processing level and risk. Data which could be considered identifying such as genotypes or raw sequencing reside in high risk group and as such have limited access even within the Consortium. Highly summarized data such as differential abundance or genetic associations are placed in the low risk group and be prepared for public data releases (see Section 10.1.10.2). Security audits are performed to ensure that processes and workflows are secure.

10.1.8.4 CLINICAL AND BIOINFORMATIC DATA INTEGRATION

The core linking of all chemical and clinical data are through strict tracking of unique identifiers and linked metadata. Special focus on verification of identifiers are implemented in QA/QC processes for all data transactions.

10.1.8.4.1 DATA TRANSFER

DMAQC data are transferred as a direct database dump that has undergone anonymization and purging of PHI/PII. This dump is incrementally loaded into the BIC database for synchronization. The DMAQC database includes clinical data, CRFs on phenotypes, sample/participant metadata, biorepository processing information, and all identifiers. All clinical data are locked from editing on the BIC side and only changed based on updates directly accessed from the DMAQC.

10.1.8.4.2 DATA ANONYMIZATION

Even though, in general, there is no PHI at the BIC there may be some PII for certain data elements (imaging, labs, wearables). Any such data are effectively quarantined for anonymization before they are incorporated into the larger data pool.

10.1.9 PROTOCOL DEVIATIONS

Adherence to the study protocol is crucial to collection of high quality data and to the internal validity of the study. See Section 8.3.8.1 for reporting requirements.

10.1.10 PUBLICATION AND DATA SHARING POLICY

10.1.10.1 STUDY RESULTS AND DISSEMINATION

Widespread dissemination may involve the following:

1. Presentations at lay and scientific meetings;
2. Publications in scientific journals;
3. Media coverage through press releases and interviews targeted to local and national newspapers, television, and radio outlets;
4. Production of the research summary document and facts sheet targeted to the general public which clearly and concisely summarizes the key conclusions of the study;

5. Production of professionally designed flyers, posters, brochures, and research briefs targeted to broad audiences;
6. Use of new media and social networking approaches to widely disseminate videos of the studies as appropriate;
7. Use of concise policy briefs to advocate for legislative and policy change at local, state, and national levels, which focus on how new evidence has implications for a particular policy as appropriate regarding the MoTrPAC studies;
8. Study newsletters targeted to research study participants;
9. Distribution of dissemination materials to community agencies, professional societies, and health-related websites and listserves;
10. Hosting and attending seminars, conferences, community forums, and health fairs;
11. Mailing personal thank you letters to research study participants; and
12. Posting information and documents on the MoTrPAC public website

All MoTrPAC publications and presentations are submitted to the PPD subcommittee in accordance with the PPD policy. Final manuscripts are required to be submitted to the NIH National Library of Medicine PubMed central for archiving upon acceptance for publication.

10.1.10.2 DATA SHARING PLAN

We are familiar with the NIH policies regarding data sharing and the study complies with local, state, and federal laws, such as the Privacy Rule, a Federal regulation under the HIPAA.

MoTrPAC will submit human sample-derived data to both unrestricted and controlled access databases. Two tiers of access are planned: **1)** levels of risk of data, and **2)** potential for participant re-identification, while recognizing both participant consent choices and data risk category, and in keeping with generally accepted practices across data types. All individual-level deposited data will be coded with linked identifiers to controlled-access metadata labels/tables that will facilitate coordinated identification across **1)** samples, **2)** individuals, **3)** interventions, and **4)** databases. All human sample-derived sequencing-based data will be submitted through the database of Genotypes and Phenotypes (**dbGaP**) and the associated protected Sequence Read Archive (**SRA**) instance.

Data will undergo different levels of processing, as defined at https://osp.od.nih.gov/wp-content/uploads/Supplemental_Info_GDS_Policy.pdf. To minimize risk of re-identification, the sequencing data (including initial sequence reads, data subsequent to an initial round of cleaning or basic quality analysis, and analyses to identify genetic variants, gene expression patterns; as described at the above link and contained in Table 1 of the MoTrPAC GDS) will be released under controlled access and de-identified by standards consistent with both HIPAA and the Common Rule. Aggregate data analyses, genomic summary results, and summary statistics (i.e., unrestricted data) will be provided through the BIC portal.

The timing of external data releases will be developed taking into consideration the rate of recruitment, time for processing and analysis of samples, QC of clinical data, and MoTrPAC external data sharing policies. For all releases, the BIC will evaluate minimally viable datasets for particular data types to ensure that rare characteristics are not personally-identifiable when released in isolation. Consistent with Section 10.1.3.1, prior to public release of MoTrPAC data, all potential identifiers will be removed and indirect identifiers will be vetted by the Stanford Privacy Office per University policy.

10.1.11 CONFLICT OF INTEREST POLICY

The independence of this study from any actual or perceived influence is critical. Therefore any actual conflict of interest of persons who have a role in the design, conduct, analysis, publication, or any aspect of this study is disclosed and managed. Furthermore, persons who have a perceived conflict of interest are required to have such conflicts managed in a way that is appropriate to their participation in the trial. The study leadership in conjunction with the NIH has established policies and procedures for all study group members to disclose all conflicts of interest and establish a mechanism for the management of all reported dualities of interest.

10.2 ABBREVIATIONS

10-RM	10-Repetition Maximum
1RM	1-Repetition Maximum
ACC	Administrative Coordinating Center
ACSM	American College of Sports Medicine
ADHD	Attention-Deficit/Hyperactivity Disorder
AE	Adverse Event
AED	Automated External Defibrillator
AHA	American Heart Association
ASA24®	Automated Self-Administered 24-Hour
BIC	Bioinformatics Center
BMC	Bone Mineral Content
BMD	Bone Mineral Density
BMI	Body Mass Index
BMP	Biospecimens & Molecular Phenotyping
CAS	Chemical Analysis Sites
CC	Clinical Center
CCC	Consortium Coordinating Center
CCRR	Clinical Center Recruitment and Retention
CFR	Code of Federal Regulations
CHD	Coronary Heart Disease
CI	Confidence Interval
COPD	Chronic Obstructive Pulmonary Disease
CPET	Cardiopulmonary Exercise Test
CPIPO	Clinical Protocol Intervention, Phenotyping, and Operations
CPR	Cardiopulmonary Resuscitation
CRF	Case Report Form
CVD	Cardiovascular Disease
dbGAP	database of Genotypes and Phenotypes
DBP	Diastolic Blood Pressure
DHHS	Department of Health and Human Services
DHQ-III	Diet History Questionnaire III
DMAQC	Data Management, Analysis, and Quality Control
DQAR	Data Quality, Analysis & Review
DSMB	Data and Safety Monitoring Board
DSMP	Data and Safety Monitoring Plan
DXA	Dual-energy X-ray Absorptiometry

eCRF	electronic Case Report Form
EE	Endurance Exercise
EIC	Exercise Intervention Core
EMS	Emergency Medical Services
ENCODE	Encyclopedia of DNA Elements Consortium
ESC	Executive Steering Committee
FDA	Food and Drug Administration
FDR	False Discovery Rate
FFQ	Food Frequency Questionnaire
GCP	Good Clinical Practice
GDS	Genomic Data Sharing
GEE	Generalized estimating equations
GEO	Gene Expression Omnibus
GET	Genomics, Epigenomics, Transcriptomics
GLM	Generalized linear model
GLMM	Generalized linear mixed models
GO	Gene Ontology
GSR	Genomic Summary Results
HAEE	Highly Active Endurance Exercise
HARE	Highly Active Resistance Exercise
HIPAA	Health Information Portability and Accountability Act
ICF	Informed Consent Form
IRB	Institutional Review Board
ISBER	International Society for Biological and Environmental Repositories
IT	Information Technology
ITT	Intent-to-treat
LDL-C	Low-Density Lipoprotein Cholesterol
LMMs	Linear mixed effects models
LTFU	Lost to follow-up
MAR	Missing at Random
MCAR	Missing Completely at Random
MCSS	Medical/Clinical Studies Safety
MedDRA	Medical Dictionary for Regulatory Activities
MI	Myocardial Infarction
MOFA	Multi-omics factor analysis
MOP	Manual of Procedures
MoTrPAC	Molecular Transducers of Physical Activity Consortium
MSO	Medical Safety Officer
NCR	Navitas Clinical Research, Inc.
NIA	National Institute on Aging
NIAMS	National Institute of Arthritis and Musculoskeletal and Skin Diseases
NIDDK	National Institute of Diabetes and Digestive and Kidney Diseases
NIH	National Institutes of Health
NSAID	Non-steroidal anti-inflammatory drugs
PA	Physical Activity
PASOFA	Pilot and Ancillary Studies, and Opportunity Fund Administration
PCP	Primary Care Provider

PD	Protocol Deviation
PHI	Protected Health Information
PI	Principal Investigator
PID	Participant Identification
PII	Personally identifiable information
PP	Per-protocol
PPD	Publications, Presentations, and Dissemination
PROMIS	Patient-Reported Outcomes Measurement Information System
QA	Quality Assurance
QC	Quality Control
QTL	Quantitative trait locus
RCT	Randomized Controlled Trial
RE	Resistance Exercise
RM	Repetition Maximum
SAE	Serious Adverse Event
SBP	Systolic Blood Pressure
SC	Steering Committee
SD	Standard Deviation
sIRB	Single Institutional Review Board
SNP	Single-nucleotide Polymorphism
SOP	Standard Operating Procedures
SRA	sequence read archive
SSL	Secure Socket Layer
T2D	Type 2 Diabetes
TSH	Thyroid Stimulating Hormone
TWAS	Transcriptome-wide association studies
TWIGS	three-way module inference via Gibbs sampling
UP	Unanticipated Problem
US	United States
VO_2peak	Peak Aerobic Power
WGS	Whole Genome Sequencing

10.3 PROTOCOL AMENDMENT HISTORY

Version	Date	Sections Affected	Description of Change	Brief Rationale
1.1	4/18/20	5.2.1, 5.2.2	Eliminated incomplete lists of exclusionary drugs and made other minor corrections	Revised per IRB recommendation
1.2	9/11/20	4.1.1.1 Initial Screening, Randomization, and Familiarization	Allow flexibility for some screening assessments to be conducted by phone or video calls.	Reduction of in person contacts during COVID-19 pandemic.
1.3	11/9/20	10.1.3.1	Enable the DMAQC to transfer dates to the BIC	Dates are needed to provide complete information on participant consent and to be consistent with the genomic

				data sharing document. This is for both release of low-risk and de-identified data and withdrawal of external data sharing consent.
2.0	2/25/21	1.3, 4.1.1.1, 8.1.1, 8.1.2.5	Remove details on testing windows	Testing windows are detailed in the MOP and the informed consent forms
		4.1.3	Modified Vanguard Phase	Was not completed per protocol because of the COVID-19 suspension; will not resume after the suspension
		New 4.1.2	Added brief description of COVID-19 suspension	
		5.1.2	Clarified HA inclusion criteria	Minimize inclusion of cross-trained individuals
		5.1.1, 5.1.2	New inclusion criterion related to data sharing	Volunteers must agree to have de-identifed individual-level data at low risk of re-identification in the MoTrPAC open access database
		5.2.1	Clarified exclusion criteria	Address questions raised during screening
		5.2.2	Clarified exclusions for medication use	Address questions raised during screening
		7.2	Added individual stopping criterion for COVID-19 diagnosis	Participants with a new COVID-19 diagnosis will be withdrawn from the study because of the duration of quarantine required and the unknown consequences of the infection
		8.1.2.1, 10.1.1	Eliminated statement that orientation visit must be in-person	Allow teleconference or videoconference orientation and consenting visits
		8.1.2.4.3	Added new physical activity questionnaires and a second accelerometer	New questionnaires will obtain a more comprehensive assessment of physical activity habits. The second accelerometer (ActivPal) will better capture sedentary time at baseline and after the intervention.
		8.1.2.5	Additional blood samples	As needed, for safety reasons

		8.1.2.5.4	Clarified when a light meal is provided	Provided to all participants after final biospecimen collection on acute test days
		8.3	Added to the description of commonly occurring signs and symptoms of tissue biopsies and exercise	This modification clarified the common signs and symptoms associated with tissue biopsies and exercise that are captured on the visit CRF rather than an AE form.
		8.3.3.1.2	Removed definition of unexpected adverse event	Because the categorization of expectedness is binary, events that do not meet the definition of expected will be categorized as unexpected
		10.1.5	Eliminated NIBIB as a managing institute	No longer serving in this role
		10.1.5.1	Added roles of Administrative Coordinating Center	Describe role of Coordinating Center
2.1	4/20/21	1.2, 6.1, 8.1.2.4.2, 8.1.2.5, 8.1.2.5.4	Eliminated ALL biospecimen profile for sedentary participants (changed throughout the document)	Determined to be too burdensome
		5.1.1	Restricted sedentary enrollment to one member of a household	Minimizes effects of shared exposures
		5.2.1	Specified that the CPET can be repeated	Allows the test to be repeated under certain conditions (e.g., was not a maximal effort, instrument was not calibrated, etc)
		8.1.2.4.3	Clarified assessment of physical activity	Assessment of PA and sedentary time with a second accelerometer will be conducted on only a subset of participants because of limited funding
		8.1.2.5.1, 8.1.2.5.2	Eliminated details on monitoring exercise intensity	Detailed instructions are in the MOP
		9.2	Updated detectable effect sizes	Changed as a result of eliminating the All biospecimen sampling profile
2.4	1/11/22	1.1, 10.1.5	Updates throughout in list/number of clinical and pre-clinical sites	Reflects changes in active clinical and pre-clinical sites

		5.1.2	Modified criteria for Highly Active Resistance Exercisers	Per DSMB recommendation, expand catchment of the Highly Active group
		5.2.1	Revise eligibility criteria to exclude individuals who tested positive for COVID-19 but were not hospitalized must be symptom-free at least 7 days	With the exponential increase in COVID-19 cases related to the omicron variant, revised to allow continued study recruitment
		5.2.2	Eliminated exclusion for use of lipid-lowering medications	Per DSMB recommendation, expand catchment in an effort to help boost enrollment across all three age bins
		7.2	Revise withdraw criteria to reflect that participants may be withdrawn due to a COVID-19 diagnosis	To allow continued participation in participants to experience a mild or asymptomatic COVID-19 infection
2.5	2/3/22	5.2.1	Revise eligibility criterion for COVID-19 to include antigen testing	Recommended by IRB
2.6	3/16/22	1.2	Updated study design with the approximate number of individuals that may be consented in order to meet the target sample size	Request from IRB
2.7	7/14/2022	5.2.1	Revised eligibility criterion to exclude current and former smokers	Exclude potential participants at higher risk for heart disease
2.8	4/6/2023	Title page	Update Lead Study Investigator to Wendy Kohrt, PhD	Marco Pahor, MD will be leaving his position at the University of Florida and the MoTrPAC Consortium.

11 REFERENCES

1. Committee, P.A.G.A., *Physical activity guidelines advisory committee report*, U.S.D.o.H.a.H. Services, Editor. 2008: Washington, DC.
2. Booth, F.W., C.K. Roberts, and M.J. Laye, *Lack of exercise is a major cause of chronic diseases.* Compr Physiol, 2012. **2**(2): p. 1143-211.
3. Pedersen BK, S.B., *Exercise as medicine - evidence for prescribing exercise as therapy in 26 different chronic diseases.* Scand J Med Sci Sports, 2015. **25** (Suppl 3): p. 1-72.
4. Neufer, P.D., et al., *Understanding the Cellular and Molecular Mechanisms of Physical Activity-Induced Health Benefits.* Cell Metab, 2015. **22**(1): p. 4-11.
5. Kiens, B. and E.A. Richter, *Utilization of skeletal muscle triacylglycerol during postexercise recovery in humans.* Am J Physiol, 1998. **275**(2 Pt 1): p. E332-7.
6. Louis, E., et al., *Time course of proteolytic, cytokine, and myostatin gene expression after acute exercise in human skeletal muscle.* J Appl Physiol (1985), 2007. **103**(5): p. 1744-51.
7. Egan, B., et al., *Exercise intensity-dependent regulation of peroxisome proliferator-activated receptor coactivator-1 mRNA abundance is associated with differential activation of upstream signalling kinases in human skeletal muscle.* J Physiol, 2010. **588**(Pt 10): p. 1779-90.
8. Little, J.P., et al., *An acute bout of high-intensity interval training increases the nuclear abundance of PGC-1alpha and activates mitochondrial biogenesis in human skeletal muscle.* Am J Physiol Regul Integr Comp Physiol, 2011. **300**(6): p. R1303-10.
9. Barres, R., et al., *Acute exercise remodels promoter methylation in human skeletal muscle.* Cell Metab, 2012. **15**(3): p. 405-11.
10. Pilegaard, H., B. Saltin, and P.D. Neufer, *Exercise induces transient transcriptional activation of the PGC-1alpha gene in human skeletal muscle.* J Physiol, 2003. **546**(Pt 3): p. 851-8.
11. Stec, M.J., D.L. Mayhew, and M.M. Bamman, *The effects of age and resistance loading on skeletal muscle ribosome biogenesis.* J Appl Physiol (1985), 2015. **119**(8): p. 851-7.
12. Wallace, M.A., et al., *Striated muscle activator of Rho signalling (STARS) is a PGC-1alpha/oestrogen-related receptor-alpha target gene and is upregulated in human skeletal muscle after endurance exercise.* J Physiol, 2011. **589**(Pt 8): p. 2027-39.
13. Schwarz, N.A., et al., *Effect of resistance exercise intensity on the expression of PGC-1alpha isoforms and the anabolic and catabolic signaling mediators, IGF-1 and myostatin, in human skeletal muscle.* Appl Physiol Nutr Metab, 2016. **41**(8): p. 856-63.
14. Reidy, P.T., et al., *Postexercise essential amino acid supplementation amplifies skeletal muscle satellite cell proliferation in older men 24 hours postexercise.* Physiol Rep, 2017. **5**(11).
15. Perry, C.G., et al., *Repeated transient mRNA bursts precede increases in transcriptional and mitochondrial proteins during training in human skeletal muscle.* J Physiol, 2010. **588**(Pt 23): p. 4795-810.
16. Dickinson, J.M., et al., *Leucine-enriched amino acid ingestion after resistance exercise prolongs myofibrillar protein synthesis and amino acid transporter expression in older men.* J Nutr, 2014. **144**(11): p. 1694-702.
17. Egan, B., et al., *Time course analysis reveals gene-specific transcript and protein kinetics of adaptation to short-term aerobic exercise training in human skeletal muscle.* PLoS One, 2013. **8**(9): p. e74098.
18. Neubauer, O., et al., *Time course-dependent changes in the transcriptome of human skeletal muscle during recovery from endurance exercise: from inflammation to adaptive remodeling.* J Appl Physiol (1985), 2014. **116**(3): p. 274-87.

19. Raue, U., et al., *TWEAK-Fn14 pathway activation after exercise in human skeletal muscle: insights from two exercise modes and a time course investigation.* J Appl Physiol (1985), 2015. **118**(5): p. 569-78.
20. Yang, Y., et al., *Time course of myogenic and metabolic gene expression in response to acute exercise in human skeletal muscle.* J Appl Physiol (1985), 2005. **98**(5): p. 1745-52.
21. Phillips, S.M., et al., *Mixed muscle protein synthesis and breakdown after resistance exercise in humans.* Am J Physiol, 1997. **273**(1 Pt 1): p. E99-107.
22. Miller, B.F., et al., *Coordinated collagen and muscle protein synthesis in human patella tendon and quadriceps muscle after exercise.* J Physiol, 2005. **567**(Pt 3): p. 1021-33.
23. Mayhew, D.L., et al., *Eukaryotic initiation factor 2B epsilon induces cap-dependent translation and skeletal muscle hypertrophy.* J Physiol, 2011. **589**(Pt 12): p. 3023-37.
24. Coffey, V.G., et al., *Interaction of contractile activity and training history on mRNA abundance in skeletal muscle from trained athletes.* Am J Physiol Endocrinol Metab, 2006. **290**(5): p. E849-55.
25. Fry, C.S., et al., *Aging impairs contraction-induced human skeletal muscle mTORC1 signaling and protein synthesis.* Skelet Muscle, 2011. **1**(1): p. 11.
26. Fry, C.S., et al., *Skeletal muscle autophagy and protein breakdown following resistance exercise are similar in younger and older adults.* J Gerontol A Biol Sci Med Sci, 2013. **68**(5): p. 599-607.
27. Mayhew, D.L., et al., *Translational signaling responses preceding resistance training-mediated myofiber hypertrophy in young and old humans.* J Appl Physiol (1985), 2009. **107**(5): p. 1655-62.
28. Kumar, V., et al., *Age-related differences in the dose-response relationship of muscle protein synthesis to resistance exercise in young and old men.* J Physiol, 2009. **587**(1): p. 211-7.
29. Drummond, M.J., et al., *Skeletal muscle amino acid transporter expression is increased in young and older adults following resistance exercise.* J Appl Physiol (1985), 2011. **111**(1): p. 135-42.
30. Kim, J.S., J.M. Cross, and M.M. Bamman, *Impact of resistance loading on myostatin expression and cell cycle regulation in young and older men and women.* Am J Physiol Endocrinol Metab, 2005. **288**(6): p. E1110-9.
31. Kim, J.S., et al., *Resting and load-induced levels of myogenic gene transcripts differ between older adults with demonstrable sarcopenia and young men and women.* J Appl Physiol (1985), 2005. **99**(6): p. 2149-58.
32. Kosek, D.J. and M.M. Bamman, *Modulation of the dystrophin-associated protein complex in response to resistance training in young and older men.* J Appl Physiol (1985), 2008. **104**(5): p. 1476-84.
33. Kosek, D.J., et al., *Efficacy of 3 days/wk resistance training on myofiber hypertrophy and myogenic mechanisms in young vs. older adults.* J Appl Physiol (1985), 2006. **101**(2): p. 531-44.
34. Merritt, E.K., et al., *Heightened muscle inflammation susceptibility may impair regenerative capacity in aging humans.* J Appl Physiol (1985), 2013. **115**(6): p. 937-48.
35. Bajpeyi, S., et al., *Effect of exercise intensity and volume on persistence of insulin sensitivity during training cessation.* J Appl Physiol (1985), 2009. **106**(4): p. 1079-85.
36. Heath, G.W., et al., *Effects of exercise and lack of exercise on glucose tolerance and insulin sensitivity.* J Appl Physiol Respir Environ Exerc Physiol, 1983. **55**(2): p. 512-7.
37. Bamman, M.M., et al., *Cluster analysis tests the importance of myogenic gene expression during myofiber hypertrophy in humans.* J Appl Physiol (1985), 2007. **102**(6): p. 2232-9.
38. Bickel, C.S., et al., *Time course of molecular responses of human skeletal muscle to acute bouts of resistance exercise.* J Appl Physiol (1985), 2005. **98**(2): p. 482-8.
39. Cartoni, R., et al., *Mitofusins 1/2 and ERRalpha expression are increased in human skeletal muscle after physical exercise.* J Physiol, 2005. **567**(Pt 1): p. 349-58.

40. Kim, J.S., et al., *Load-mediated downregulation of myostatin mRNA is not sufficient to promote myofiber hypertrophy in humans: a cluster analysis.* J Appl Physiol (1985), 2007. **103**(5): p. 1488-95.
41. Mahoney, D.J., et al., *Analysis of global mRNA expression in human skeletal muscle during recovery from endurance exercise.* Faseb j, 2005. **19**(11): p. 1498-500.
42. Ortega, J.F., et al., *Higher Insulin-sensitizing Response after Sprint Interval Compared to Continuous Exercise.* Int J Sports Med, 2015.
43. Trappe, T.A., et al., *Effect of ibuprofen and acetaminophen on postexercise muscle protein synthesis.* Am J Physiol Endocrinol Metab, 2002. **282**(3): p. E551-6.
44. Peterson, J.M., et al., *Ibuprofen and acetaminophen: effect on muscle inflammation after eccentric exercise.* Med Sci Sports Exerc, 2003. **35**(6): p. 892-6.
45. Thalacker-Mercer, A.E., et al., *Differential genomic responses in old vs. young humans despite similar levels of modest muscle damage after resistance loading.* Physiol Genomics, 2010. **40**(3): p. 141-9.
46. Khassaf, M., et al., *Time course of responses of human skeletal muscle to oxidative stress induced by nondamaging exercise.* J Appl Physiol (1985), 2001. **90**(3): p. 1031-5.
47. Morton, J.P., et al., *Time course and differential responses of the major heat shock protein families in human skeletal muscle following acute nondamaging treadmill exercise.* J Appl Physiol (1985), 2006. **101**(1): p. 176-82.
48. Mikines, K.J., et al., *Effect of physical exercise on sensitivity and responsiveness to insulin in humans.* Am J Physiol, 1988. **254**(3 Pt 1): p. E248-59.
49. Perseghin, G., et al., *Increased glucose transport-phosphorylation and muscle glycogen synthesis after exercise training in insulin-resistant subjects.* N Engl J Med, 1996. **335**(18): p. 1357-62.
50. Peake, J.M., et al., *The effects of cold water immersion and active recovery on inflammation and cell stress responses in human skeletal muscle after resistance exercise.* J Physiol, 2017. **595**(3): p. 695-711.
51. Committee, P.A.G.A., *2018 Physical Activity Guidelines Advisory Committee Scientific Report.* 2018, U.S. Department of Health and Human Services: Washington, DC.
52. Agresti, A., *Approximate is better that "exact" for interval estimation of binominal proportions.* The American Statistician, 1998. **52**(2): p. 119-126.
53. Hepworth, G., *Confidence intervals for proportions estimated by group testing with groups of unequal sizes.* Journal of Agricultural, Biological, and Envrionmental Statistics, 2005. **10**(4): p. 478-497.
54. Fletcher, G.F., et al., *Exercise standards for testing and training: a scientific statement from the American Heart Association.* Circulation, 2013. **128**(8): p. 873-934.
55. Vilagut, G., et al., *Screening for Depression in the General Population with the Center for Epidemiologic Studies Depression (CES-D): A Systematic Review with Meta-Analysis.* PLoS One, 2016. **11**(5): p. e0155431.
56. *Health Measures: Transforming how health is measured*. http://www.healthmeasures.net/ 2018 2/9/18].
57. Meriwether, R.A., et al., *Physical activity assessment: validation of a clinical assessment tool.* Am J Prev Med, 2006. **31**(6): p. 484-91.
58. Paffenbarger, R.S., Jr., et al., *Measurement of physical activity to assess health effects in free-living populations.* Med Sci Sports Exerc, 1993. **25**(1): p. 60-70.
59. Bergstrom, J., *Muscle electrolytes in man.* Scan J Clin Lab Invest, 1962. **68**: p. 1–110.
60. Bastard, J.P., et al., *Percutaneous adipose tissue biopsy by mini-liposuction for metabolic studies.* JPEN J Parenter Enteral Nutr, 1994. **18**(5): p. 466-8.

61. Pasarica, M., et al., *Reduced adipose tissue oxygenation in human obesity: evidence for rarefaction, macrophage chemotaxis, and inflammation without an angiogenic response.* Diabetes, 2009. **58**(3): p. 718-25.
62. Sparks, L.M., et al., *Relation of adipose tissue to metabolic flexibility.* Diabetes Res Clin Pract, 2009. **83**(1): p. 32-43.
63. Karastergiou, K., et al., *Distinct developmental signatures of human abdominal and gluteal subcutaneous adipose tissue depots.* J Clin Endocrinol Metab, 2013. **98**(1): p. 362-71.
64. Alam, M. and M.S. Kaminer, *Liposuction*, in *Procedures in cosmetic dermatotlogy: Body conturing*, B.E. Katz and N.S. Sadick, Editors. 2010, Saunders Elsevier: Edinburgh.
65. Iyer, S. and B.I. Raskin, *Liposuction with local tumescent anesthesia and microcannula technique*, in *Liposuction principles and practice*, M.A. Shiffman and A.D. Giuseppe, Editors. 2006, Springer Berlin Heidelberg: New York.
66. Klein, J.A. and D.R. Jeske, *Estimated Maximal Safe Dosages of Tumescent Lidocaine.* Anesth Analg, 2016. **122**(5): p. 1350-9.
67. Tong, T. and H. Zhao, *Practical guidelines for assessing power and false discovery rate for a fixed sample size in microarray experiments.* Stat Med, 2008. **27**(11): p. 1960-72.
68. Liang KY, Z.S., *Longitudinal data analysis of continuous and discrete responses for pre-post design.* Sankhya: Indian J Stat (Series B), 2000. **62**: p. 134-148.
69. Lu, K., *On efficiency of constrained longitudinal data analysis versus longitudinal analysis of covariance.* Biometrics, 2010. **66**(3): p. 891-6.
70. Shpitser, I., *Counterfactual graphical models for longitudinal mediation analysis with unobserved confounding.* Cogn Sci, 2013. **37**(6): p. 1011-35.
71. Lynch, K.G., et al., *Causal Mediation Analyses for Randomized Trials.* Health Serv Outcomes Res Methodol, 2008. **8**(2): p. 57-76.
72. J, R., *Correcting for non-compliance in randomized trials using structural nested mean models.* Comm Statist Theory Meth, 1994. **23**: p. 2379–2412.
73. Ten Have T, E.M., Zanutto E, Datto C, *Causal models for randomized physician encouragement trials in treating primary care depression.* J Am Statist Assoc, 2004. **99**: p. 8–16.
74. Have, T.R., et al., *Causal mediation analyses with rank preserving models.* Biometrics, 2007. **63**(3): p. 926-34.
75. Frangakis, C.E., et al., *Methodology for Evaluating a Partially Controlled Longitudinal Treatment Using Principal Stratification, With Application to a Needle Exchange Program.* J Am Stat Assoc, 2004. **99**(465): p. 239-249.
76. Frangakis, C.E. and D.B. Rubin, *Principal stratification in causal inference.* Biometrics, 2002. **58**(1): p. 21-9.
77. Gamazon, E.R., et al., *A gene-based association method for mapping traits using reference transcriptome data.* Nat Genet, 2015. **47**(9): p. 1091-8.
78. Price, N.D., et al., *A wellness study of 108 individuals using personal, dense, dynamic data clouds.* Nature Biotechnology, 2017. **35**: p. 747.
79. *MOFA: Multi-Omics Factor Analysis*. https://github.com/PMBio/MOFA 2018 [cited 2018.
80. Clancy, C.E., et al., *Multiscale Modeling in the Clinic: Drug Design and Development.* Ann Biomed Eng, 2016. **44**(9): p. 2591-610.
81. Su, J., et al., *Targeting the biophysical properties of the myeloma initiating cell niches: a pharmaceutical synergism analysis using multi-scale agent-based modeling.* PLoS One, 2014. **9**(1): p. e85059.
82. Peng, T., et al., *Modeling cell-cell interactions in regulating multiple myeloma initiating cell fate.* IEEE J Biomed Health Inform, 2014. **18**(2): p. 484-91.

83. Sun, X., et al., *Cytokine combination therapy prediction for bone remodeling in tissue engineering based on the intracellular signaling pathway.* Biomaterials, 2012. **33**(33): p. 8265-76.
84. Uppal, K., et al., *Computational Metabolomics: A Framework for the Million Metabolome.* Chem Res Toxicol, 2016. **29**(12): p. 1956-1975.
85. Ernst, J., et al., *Reconstructing dynamic regulatory maps.* Mol Syst Biol, 2007. **3**: p. 74.
86. Schulz, M.H., et al., *DREM 2.0: Improved reconstruction of dynamic regulatory networks from time-series expression data.* BMC Syst Biol, 2012. **6**: p. 104.
87. Schulz, M.H., et al., *Reconstructing dynamic microRNA-regulated interaction networks.* Proc Natl Acad Sci U S A, 2013. **110**(39): p. 15686-91.
88. Amar, D., et al., *A hierarchical Bayesian model for flexible module discovery in three-way time-series data.* Bioinformatics, 2015. **31**(12): p. i17-26.
89. Cowen, L., et al., *Network propagation: a universal amplifier of genetic associations.* Nat Rev Genet, 2017. **18**(9): p. 551-562.
90. Amar, D. and R. Shamir, *Constructing module maps for integrated analysis of heterogeneous biological networks.* Nucleic Acids Res, 2014. **42**(7): p. 4208-19.
91. Malladi, V.S., et al., *Ontology application and use at the ENCODE DCC.* Database (Oxford), 2015. **2015**.
92. Hong, E.L., et al., *Principles of metadata organization at the ENCODE data coordination center.* Database (Oxford), 2016. **2016**.
93. Hitz, B.C., et al., *SnoVault and encodeD: A novel object-based storage system and applications to ENCODE metadata.* PLoS One, 2017. **12**(4): p. e0175310.
94. Chen, L., et al., *DNA damage is a pervasive cause of sequencing errors, directly confounding variant identification.* Science, 2017. **355**(6326): p. 752-756.
95. De Domenico M, G.C., Porter MA, Arenas A, *The physics of spreading processes in multilayer networks.* Nature Physics, 2016. **12**: p. 901–906.
96. De Domenico M, S.-R.A., Omodei E, Gómez S, Arenas A, *Ranking in interconnected multilayer networks reveals versatile nodes.* Nature Communications, 2015. **6**: p. 6868.
97. Broman, K.W., et al., *Identification and Correction of Sample Mix-Ups in Expression Genetic Data: A Case Study.* G3 (Bethesda), 2015. **5**(10): p. 2177-86.
98. Westra, H.J., et al., *MixupMapper: correcting sample mix-ups in genome-wide datasets increases power to detect small genetic effects.* Bioinformatics, 2011. **27**(15): p. 2104-11.