Molecular Transducers of Physical Activity Consortium (MoTrPAC) – Pediatric Study

Consortium Coordinating Center Principal Investigator: Marco Pahor, MD Funded by: NIH Common Fund Draft Number: 8.12 14 March 2023

Summary of Changes from Previous Version:

Affected	Summary of Revisions Made	Rationale					
Section(s)							
5.2.2	Remove COVID vaccination requirement	Align with State of California move to no					
		longer mandate COVID vaccination for					
		school aged children					

Table of Contents

Statement of Compliance	4
1 Protocol Summary	5
1.1 Synopsis	5
1.2 Study Design	6
1.3 Schedule of Activities (SoA)	7
2 Introduction	8
2.1 Study Rationale	8
2.1.1 Rationale for Timing of Biospecimen Collection	9
2.2 Risk/Benefit Assessment	10
2.2.1 Known Potential Risks	10
2.2.2 Known Potential Benefits	10
2.3.3 Assessment of Potential Risks and Benefits	11
3 Objectives and Endpoints	12
4 Study Design	14
4.1 Overall Design	14
4.1.1 Cross Sectional Phase: Health and Fitness Phenotyping Visits	16
4.1.2 EE Intervention Phase: End of Study Health and Fitness Phenotyping Visits	17
4.1.3 Vanguard Phase	17
4.1.4 COVID-19 Pandemic	18
4.3 Justification for Dose	18

Molecular Transducers of Physical Activity Consortium–Pediatric Clinical Center Protocol

4.4 End of Study Definition	
5 Study Population	
5.1 Inclusion Criteria	19
5.2 Exclusion Criteria	19
5.3 Lifestyle Considerations	24
5.4 Screen Failures	24
5.5 Strategies for Recruitment and Retention	
5.5.1 Recruitment	24
5.5.2 Underrepresented Minorities and Age Groups	
5.5.3 Retention Strategies	
5.5.4 Participant Remuneration	
5.5.5 Monitoring Recruitment and Retention	
5.5.6 Efforts to maintain contact with non-adherent participants.	
6 Study Intervention	27
6.1 Study Intervention(s) Administration	
6.1.1 Endurance Training	
6.1.2 Control Group	
6.2 Measures to Minimize Bias: Randomization and Blinding	28
6.2.1 Measures to Minimize Bias	28
6.2.2 Encollment/Randomization/Masking Procedures	29
6.3 Study Intervention Compliance	29
6.3.1 Understanding and Promoting Adherence	29
6.3.2 Capturing and Monitoring Adherence in MoTrPAC	30
7 Study Discontinuation. Withdrawal	21
7 1 Discontinuation of Study Intervention	
7.2 Participant Discontinuation/Withdrawal from the Study	31
7 3 Lost to Follow-Up	31
8 Study Assessments and Procedures	32
8.1 Study Procedures and Outcomes	
8.1.1 Study Schedule	
8.1.2 Measures and Procedures Overview	32
8 2 Safety and Other Assessments	35
8.2.1 Event Assessment and Follow-up	35
8.2.2 Pre-Exercise Intervention and Safety Screening	36
8.2.3 Safety Considerations for Study Assessments	36
8.2.4 Safety Considerations for the Interventions and Acute Exercise Challenge	37
8.2.5 Safety Considerations Regarding BIOSPECIMEN COLLECTIONS	38
8.2.6 Safety Considerations Regarding Loss of Confidentiality	
8.2.0 Salety Considerations Regarding Loss of Confidentiality	
8.2 Adverse Events (AE) and Serious Adverse Events (SAE)	
8.2.1 Definition of Adverse Events (AE)	
0.3.1 Definition of Society Advorse Events (SAE)	
0.3.2 Deminicion of sentous Auverse Events (SAE)	
6.5.5 Classification of an Adverse Event	
8.3.4 Time Period and Frequency for Event Assessment and Follow-Up	
8.3.5 Adverse Event Reporting	
8.3.6 Serious Adverse Event Reporting	

Molecular Transducers of Physical Activity Consortium–Pediatric Clinical Center Protocol

8.3.7 Reporting Events to Participants	44
8.3.8 other reportable events	
8.4. Unanticipated Problems	46
8.4.1. Definition of Unanticipated Problems (UPS)	46
8.4.2. Unanticipated Problem Reporting	
8.4.3. Reporting Unanticipated Problems to Participants	
9 Statistical Considerations	
9.1 Statistical Hypotheses	
9.2 Sample Size Determination	
9.3 Populations for Analyses	
9.4 Statistical Analyses	
9.4.1. General Approach	
9.4.2. General Statistical Approaches	
9.4.3 Analysis of the Primary Efficacy Endpoint(s)	55
9.4.4 Analysis of the Secondary Efficacy Endpoint (s)	55
9.4.5. Safety Analyses	55
9.4.6. Baseline Descriptive Statistics	56
9.4.7. Planned Interim Analyses	56
9.4.8. Sub-Group Analyses	57
9.4.9. Tabulation of Individual Participant Data	57
9.4.10. Exploratory Analyses	57
10 Supporting Documentation and Operational Considerations	58
10.1 Regulatory, Ethical, and Study Oversight Considerations	
10.1.1 Informed consent (ICF) process and documentation	
10.1.2 Study Discontinuation and Closure	
10.1.3 Confidentiality and Privacy	59
10.1.4 Future Use of Stored Specimens and Data	61
10.1.5 Key Roles and Study Governance	62
10.1.6 Safety Oversight	65
10.1.7 Clinical Monitoring	66
10.1.8 Data Handling and Record Keeping	69
10.1.9 Protocol Deviations	76
10.1.10 Publication and Data Sharing Policy	76
10.1.11 Conflict of Interest Policy	78
10.2 Abbreviations	78
10.3 Protocol Amendment History	81
11 References	83

STATEMENT OF COMPLIANCE

The trial will be carried out in accordance with International Conference on Harmonization Good Clinical Practice (ICH GCP) and 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 ICH 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, studies will be conducted in adults and separately in children and adolescents. The UC Irvine MoTrPAC Pediatric Clinical Center oversees two interrelated study phases in children and adolescents:

1) A cross-sectional phase in which molecular transducers (obtained from blood sampling) are measured in response to an acute exercise challenge (n = 320);

2) An intervention phase is conducted as a mechanistic randomized controlled trial (**RCT**). Participants are recruited from the cross-sectional study phase and randomized to endurance exercise (EE) training (n = 120) or no exercise Control (n = 50) 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. An additional focus of the pediatric studies is to examine the impact of sex and developmental phase (self-reported pubertal stage) during childhood and adolescence on acute and chronic exercise responses.

Assessments include measurements of cardiorespiratory fitness, muscular strength, and body composition (including whole body bone mineral content) determined by dual-energy x-ray absorptiometry (**DXA**). There is also collection of blood, monitoring of free-living PA level using wearable devices, and completion of participant reported outcomes and health status by interview and/or questionnaire. As part of the MoTrPAC functions, participant data and biological samples are transferred from the Pediatric Clinical Site 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

The MoTrPAC Pediatric Clinical Center recruits for the cross sectional phase 320 participants consisting of ~270 low active endurance exercise (LAEE) and ~50 highly active endurance exercise (HAEE) children and adolescents. As shown in Table 1, after screening to determine eligibility, the participants are enrolled in the cross-sectional study phase. The enrollees undergo health and fitness phenotyping and familiarization visits to prepare them for the acute exercise visit of the cross sectional study phase that includes blood samples collected for the characterization of molecular responses to exercise.

Table 1. Pediatric MoTrPAC Overview											
Screening for eligibility	Cross-Sectional Phase Enrollment	Health and Fitness Phenotyping and Acute Exercise Challenge (~3 weeks)	Intervention Phase Enrollment (EE or Control, ~12 weeks)	Post Intervention Health and Fitness Assessments (~2 weeks)							
	135 LAEE females	 Aerobic fitness 	60 LAEE females	 Aerobic fitness 							
		 Muscle strength 		 Muscle strength 							
		 Habitual physical 		 Habitual physical 							
		activity		activity							
	135 LAEE males	 Dietary and 	60 LAEE males	 Dietary and 							
		behavioral		behavioral							
		questionnaires		questionnaires							
		 Preparation and 	25 control formalian	 Preparation for EE 							
	25 HALL TEMBLES	Familiarization with	25 control females	challenge							
		EE challenge									
		\checkmark		\checkmark							
	25 HAFF males	Washout EE	25 control males	 Washout 							
	25 HALL IIIdles	Challenge with	25 control males	\checkmark							
		blood sampling		• EE Challenge with							
				blood sampling							

The HAEE group will undergo the same baseline testing as the LAEE group, including the acute exercise test with blood sampling, but their involvement will end at this point; they will not undergo the intervention or post-intervention testing.

1.3 SCHEDULE OF ACTIVITIES (SOA)

The activities for participants in the cross sectional and intervention phases of the pediatric protocol are summarized in Table 2.

Table 2. MoTrPAC clinical flow for pediatric participants in the cross-sectional and intervention phases							
Initial screening							
\downarrow							
Study orientation and informed consent and assent							
\downarrow							
Screening to establish eligibility (health history, puberty and physical activity questionnaires, anthropometric, resting BP and ECG measurements, cardiorespiratory fitness (CPET))*							
\downarrow							
Eligible participants enrolled							
\downarrow							
Health and fitness phenotyping visits [muscle strength, body composition, habitual physical activity, health related behaviors questionnaires and familiarization with exercise tests]*							
\downarrow							
washout (no intense exercise or testing)							
\downarrow							
Acute endurance exercise challenge with blood sampling							
\downarrow							
Establish eligibility for intervention phase							
\downarrow							
Randomization to endurance exercise training or no exercise intervention comparison group							
\downarrow							
Intervention phase (approximately 12 weeks)							
\downarrow							
Repeat health and fitness phenotyping visits							
\downarrow							
washout (no intense exercise or testing)							
\downarrow							
Repeat acute endurance exercise challenge with blood sampling							
*The number of visits to complete screening and phenotyping assessments will be allowed to vary to accommodate the schedules of participants.							

2 INTRODUCTION

2.1 STUDY RATIONALE

As outlined in the MoTrPAC adult study protocol, 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. Like in adults, major gaps exist in our understanding the molecular pathways in children that link exercise to health. Exercise in children and adolescents is not merely play but is an essential component of growth and development [1–3]. Children are among the most spontaneously physically active human beings [4]. It is not surprising that participation in PA is a major determinant of health across the lifespan and health-related quality of life in both healthy children and in children with chronic diseases [5–8]. Despite this essential biologic role for PA, children have not been spared the relentless reduction in levels of PA that is creating a crisis in health care in our nation and throughout the world [9].

Recognition of the enormous morbidity and cost of physical inactivity-related diseases, such as atherosclerosis, type 2 diabetes, and osteoporosis, has spurred new policy initiatives targeting preventive medicine early in life [10–12]. The concept of pediatric origins of adult health and disease is gaining scientific merit [13–16], highlighting the need to transform existing notions of how to evaluate

health in a growing child. A physically inactive (even normal weight) child may have no symptoms of disease, but evidence of deterioration in vascular health may already be present [17–19]. Equally worrisome is that the deleterious health effects of physical inactivity and poor fitness are exacerbated in children with chronic disease and/or disabilities [20–25] or with environmental-lifestyle conditions like obesity [26].

Once a pattern of physical inactivity and a sedentary lifestyle is established, a vicious cycle ensues





(Figure 1), in which constraints on PA harm immediate health and contribute to lifelong health impairment ranging from cardiovascular and metabolic disease to osteoporosis [1,27–31]. Exactly what constitutes ideal physical fitness in a child with a chronic condition remains unknown.

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 both pediatric and adult 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**₂**peak**) during a graded cycling cardiopulmonary exercise test (**CPET**) and will be critical to characterize fitness levels of all participants and changes in cardiorespiratory fitness. 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, habitual PA, and health history) will be collected to link to and better interpret the molecular transducers responses to exercise. In addition, pubertal stage in children and adolescents is assessed using self-report assessment tools.

The challenges to research in exercise science outlined above for adults are even more true in efforts to understand the role of exercise in children and adolescents. Sample sizes in pediatric studies have tended to be small. Since children are a vulnerable population, certain procedures which may be feasible in adults (e.g., muscle biopsies in healthy individuals) are viewed with greater scrutiny in youth. In general, in pediatric exercise studies, relatively few efforts have focused on potential molecular pathways, and accurate assessments of key developmental stages during childhood (e.g., assessment of pubertal status) are not often part of the clinical trial. In summary, given the long-term impact of levels of physical fitness and activity during childhood, we still lack basic mechanistic information that can be used to benefit the health of children. By including pediatric studies in the MoTrPAC research portfolio, the emerging molecular maps will useful in understanding the impact of physical activity across the human lifespan.

2.1.1 RATIONALE FOR TIMING OF BIOSPECIMEN COLLECTION

In the pediatric protocol, in contrast to adults, we will only be collecting blood specimens. 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 and changes in circulating metabolites can occur very rapidly in response to acute exercise. Changes in circulating leukocyte gene expression have been documented to occur within 20 min of the onset of exercise [32–36] and have yet to be studied at earlier timepoints. The duration of these effects is not known in children or adolescents. 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, as well as 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 biological as well as methodological variability, biospecimen collection in non-exercise control subjects 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 in order to better understand the various and diverse molecular signatures in response to exercise anchored to physiological responses and adaptations underlying its 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 subject burden, as well as the practical considerations of cost and logistics. Although limited, the current

published literature suggests it will be important to collect blood at the planned time points for the following purposes:

A. Morning resting blood: Will enable characterization of genomics and phenotypic molecular signatures unique to individuals and common among subgroups to explore the impact of the "starting material" on the molecular responses to exercise with particular attention to: Inter-individual response heterogeneity; and commonalities and differences across age, maturational status, sex, race/ethnicity, body composition, exercise training status, etc.

B. During and early post-exercise blood collection:

- 18-20-min and 38-40-min after initiation of exercise
- 10-min and 30-min post-exercise

This will capture secretion/appearance of exerkines, changes in metabolites, and leukocyte transcriptomics, etc.

C. 4 hours post-exercise blood collection: Will capture the longer-term secretion/appearance of exerkines, changes in metabolites, the content of exosomes, epigenetic modifications in blood cells, changes in transcriptome or the proteome in blood cells, changes in cell-free DNA, etc.

Mapping the time course of molecular responses will advance our understanding of processes that may play key roles in influencing: (i) tissue adaptations [e.g., cell signaling, protein metabolism (synthesis/catabolism)], (ii) indices of subcellular processes (e.g., mitochondrial biogenesis), (iii) lipid trafficking, and (iv) potential systemic factors driving inter-organ cross-talk.

References to support our plan to perform blood sampling over the time course from pre-exercise to 4 hours post-acute exercise, and following training include:

- Molecular expression changes from resting conditions rapidly, following exercise [32–38]
- The time course for each molecule is likely unique [39]
- The type and intensity of exercise will have an impact on molecular responses [40]

 Maturational status and sex will influence the molecular responses to exercise [34,35,39,41]
 - Training status will affect the acute exercise responses [39,42,43]

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

All participants will acquire knowledge about their physiological status (e.g., aerobic capacity, muscle strength, body composition, etc.) that may be of personal benefit. About 40% of the participants will undergo endurance exercise training, which can trigger many beneficial adaptations, but not all individuals respond similarly to such adaptations.

The potential benefits of the study to scientific and lay communities are expected to be vast. The nearterm 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 physical activity benefits 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.

2.3.3 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 [44], the benefits of exercise far outweigh the risks. However, the risks associated with participation in MoTrPAC are not only those related to exercise, but also those associated with, blood sampling, maximal exercise testing, and other risks described in section 8.2. These risks do not outweigh the advances in knowledge MoTrPAC is expected to achieve regarding the mechanisms by which exercise improves health. Safety and participation (recruitment progress 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 a map 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 acute EE in phenotypically well-described healthy children across all pubertal stages at baseline, and after intervention (EE training vs. no exercise). The resulting *molecular map* is a resource that will enable the scientific community to accelerate both mechanistic research and clinical trials on the health benefits of exercise. Our major target tissue is the circulating blood. First, work in an increasing number of laboratories has clearly demonstrated the impact of exercise on gene expression in circulating leukocytes in children and adolescents [34,35]. Moreover, circulating blood, unlike biopsies of muscle or fat, is commonly used in child health research even in healthy children. Measures to be made in blood and blood cells include genetics, genomics, epigenetics, transcriptomics, metabolomics, lipidomics, and proteomics.

The MoTrPAC Consortium uses the data generated in the pediatric clinical center to address questions about the potential molecular mechanisms that mediate the health benefits of exercise training. Among them are the following:

- 1. Do the molecular responses to acute exercise differ in the untrained and trained states?
- 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, pubertal status)?
- 3. Which molecular response patterns are most closely associated with exercise training-induced changes in phenotype (e.g., cardiorespiratory fitness, neuromuscular strength, body composition, etc.)?
- 4. 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 atlarge via controlled and public access to the MoTrPAC data repository.

The major aims of the MoTrPAC pediatric protocol are the following:

- 1. To assess the response of molecular transducers primarily in the circulating blood to a single acute endurance exercise testing session in both low and highly active endurance exercise children and adolescents.
- 2. To assess the response of molecular transducers primarily in the circulating blood to a chronic endurance exercise training program.
- 3. To discover the dynamic interaction between the molecular transducers of growth and the molecular transducers of physical activity.
- 4. To determine how the molecular pathways change across the lifespan.

To achieve the above mentioned goals, participant's data and biological samples are transferred from the Clinical Sites to the Coordinating Center DMAQC and to the Biological Samples Repository, and later analyzed by the Consortium CAS and BIC.

4 STUDY DESIGN

4.1 OVERALL DESIGN

MoTrPAC is a discovery investigation to identify molecular transducers of the health effects of physical activity. The overall strategy of the pediatric component of MoTrPAC consists of two interrelated study phases activities:

- 1. A cross sectional health and fitness assessment phase to examine the effect of an acute exercise challenge on molecular transducers across the developmental spectrum (Pubertal stages 1-5, roughly ages 10-17 years).
- 2. A 12-week intervention phase (endurance exercise training versus no exercise) to determine the effect of chronic exercise on molecular transducers in children and adolescents.

The pediatric center at the University of California Irvine (UCI) is the only venue where children and adolescents are recruited in the overall MoTrPAC consortium. Two major objectives have guided the design of the pediatric study:

- 1. To ensure that the data obtained in children and adolescents can be compared as effectively as possible with the adult studies.
- 2. To ensure that there is adequate representation for both sexes across all pubertal stages. This is important because substantial changes in metabolism, immune response, and endocrine mediators occur during growth and maturation, a relatively short interval in children and adolescents, and these factors could influence molecular transducer responses to both acute and chronic exercise.

Because this is a discovery study, there is no primary hypothesis. The Pediatric Exercise and Genomics Research Center (PERC) will recruit a sufficient number of participants to enroll 320 healthy participants (10–17 y.o.) for the cross-sectional health and fitness assessment phase with the intent to complete data collection on 300 participants. Equal number of males and females will be enrolled and self-report of pubertal stage, a common, minimally obtrusive approach in many pediatric studies, will be used as a pre-screen "guide" to ensure an equitable distribution of pubertal stages across the 10-17 y.o. age range.

It is anticipated that about 16% (n=50) will be HAEE participants, estimated by their participation in structured/regular endurance sports (such as a club sport, ≥4 days/week in the last 9 months). This percentage of highly active pediatric participants is designed to parallel the proportion of highly active participants as determined in the adult clinical sites of MoTrPAC. The HAEE participants are excluded from the intervention phase. The recruitment targets for the cross-sectional phase are shown in Table 3.

Table 3. Cross-Sectional Phase: Targeted Enrollment of Males/Females and Pubertal Stage											
Pubertal Stage (tanner)	Males	(n=160)	Females (n=160)								
	LAEE	HAEE	LAEE	HAEE							
Early (1 & 2)	54	10	54	10							
Mid (3)	27	5	27	5							
Late (4 & 5)	54	10	54	10							

We will also ask 170 LAEE participants from the cross-sectional phase (50 HAEE participants are excluded) to participate in the intervention phase designed as a randomized, intent-to-treat, clinical trial. An otherwise eligible child might be excluded from the intervention phase for the following reasons:

- 1. The participant decides not to volunteer for the intervention phase.
- 2. The participant exhibits profound anxiety regarding either the blood draw or the performance of exercise itself.
- 3. In the course of the interactions with the participant during the cross-sectional health and fitness assessment phase, new information becomes apparent that would necessitate exclusion in the training study (e.g., we discover that the child is asthmatic (requiring medications), has an undisclosed cigarette or drug habit, or reveals other disqualifying discrepancies from the original screening questionnaires).
- 4. The target number of participants for the intervention phase has already been met.

Participants are randomly assigned to either an endurance exercise training group or a no-exercise control group. The randomization is stratified by sex with 120 participants assigned to the exercise intervention and 50 assigned to the no-exercise control group. The MoTrPAC adult protocol includes a greater proportion of sedentary participants randomized to the intervention than the control arm; thus, the pediatric protocol is designed to parallel the adult protocol. These 170 participants are equally distributed between males and females; in addition, we will strive to guide enrollment of participants to achieve an equitable distribution of pubertal stages. The recruitment targets for the cross-sectional phase are shown in Table 4.

Table 4. Intervention Phase: Targeted Enrollment of Males/Females										
Males (n=	85)	Females (n=85)								
Endurance Exercise	No Exercise	Endurance Exercise	No Exercise							
60	25	60	25							

After approximately 12 weeks of intervention (training vs. no-training control) which includes free-living physical activity monitoring for all participants, the volunteers then complete follow-up phenotyping assessments and repeat the acute exercise testing (Tables 5 and 6). Blood will be analyzed using genomics, epigenomics, transcriptomics, metabolomics, lipidomics, and proteomics assays.

4.1.1 CROSS SECTIONAL PHASE: HEALTH AND FITNESS PHENOTYPING VISITS

The key components of the cross sectional phase, health and fitness phenotyping are detailed in Table 5.

Table 5. Health and Fitness Phenotyping Visits											
Study Visits^	Activities	Approximate	Location								
		participant time									
Visit 1	Orientation and informed consent/assent	• 60 min	PERC								
Consent/ assent	Weight, height, waist circumference,	• 60 min									
Screening and	medical history, resting BP, resting ECG,										
familiarization	puberty and physical activity										
	questionnaires										
	Familiarization for laboratory exercise	• 25 min									
	testing										
	CPET clearance										
Visit 2	Body composition (DXA)	• 20 min	PERC								
Screening and	 CPET (Peak VO₂) 	• 45 min									
Health and fitness	Food Frequency Questionnaire (FFQ)	• 30 min									
phenotyping	Orientation for physical activity	• 10 min									
	monitoring										
Visit 3	Isometric knee extension strength and	• 45 min	PERC								
Health and fitness	grip strength										
phenotyping	Questionnaires	• 30 min									
and familiarization	• Familiarization and calibrating endurance	• 30 min									
	exercise challenge intensity										
Visit 4#	Familiarization and calibrating endurance	• 30 min	PERC								
Acute exercise	exercise challenge intensity										
challenge	1-day 24-hour dietary recall	• 15 min	Home								
familiarization											
Visit 5	Clinical blood labs										
Acute exercise	Acute endurance exercise challenge (with	• 6-7 h	PERC								
challenge with	blood sampling*)										
blood sampling											
AThe number of scre	ening and phenotyping study visits may yary to	accommodate the sc	hadulas of								

^AThe number of screening and phenotyping study visits may vary to accommodate the schedules of participants and the number of familiarization visits that are needed. #Occurs only if there is a need for further familiarization and calibration. *Blood will be collected before, during, and in multiple time points during the 4-h recovery following the endurance exercise challenge. Each blood draw ~ 20 ml. Total volume ~ 130 ml and not exceed the lesser of 150 ml or 5 ml per kg in an 8-week period per accepted standards for healthy children and adolescents.

4.1.2 EE INTERVENTION PHASE: END OF STUDY HEALTH AND FITNESS PHENOTYPING VISITS

The key components of the EE interventional phase, post-intervention health and fitness phenotyping are detailed in Table 6.

Table 6: Post Intervention Phenotyping Visits										
Study Visits^	Activities	Approximate	Location							
		participant time								
Visit 1	Weight, height	• 5 min	PERC							
Post-intervention	 Body composition (DXA) 	• 20 min								
health and fitness	 CPET (Peak VO₂) 	• 45 min								
phenotyping	Questionnaires	• 30 min								
Testing and	•	•								
familiarization										
Visit 2	Isometric knee extension strength	• 45 min	PERC							
Post-intervention	and grip strength									
health and fitness	Questionnaires	• 30 min								
phenotyping testing	Calibrating endurance exercise	• 30 min								
and familiarization	challenge intensity									
Visit 3#	Calibrating endurance exercise	• 30 min	PERC							
Post-intervention	challenge intensity 1-day 24-hour									
acute exercise	dietary recall									
challenge		• 15 min	Home							
familiarization										
Visit 4	Acute endurance exercise challenge	• 6-7 h	PERC							
Post-intervention	(with blood sampling*)									
acute exercise										
challenge with blood										
draws										
^The number of pheno	typing study visits may vary to accommodate	e the schedules of partic	ipants and the							

^AThe number of phenotyping study visits may vary to accommodate the schedules of participants and the number of familiarization visits that are needed. #Occurs only if there is a need for further familiarization and calibration. *Blood will be collected before, during, and in multiple time points during the 4-h recovery following the endurance exercise challenge. Each blood draw ~ 20 ml. Total volume blood draw ~ 130 ml and no more not exceed the lesser of 150 ml or 5 ml per kg in an 8-week period per accepted standards for healthy children and adolescents.

4.1.3 VANGUARD PHASE

As noted above, the DSMB, in conjunction with NIH leadership, requested that the protocol for MoTrPAC include an initial Vanguard phase for the adult study, sedentary randomized participants. The Vanguard phase will last over the first 6-7 months of the study. The purpose of the Vanguard is to assess feasibility related to recruitment, participant and staff burden, and adherence to the protocol within randomized participants. The need for the Vanguard phase arose primarily from uncertainties in the adult protocol regarding whether the number and timing of muscle and fat biopsies would prove to be an excessive burden for many participants. Muscle and fat biopsies are not part of the pediatric studies, thus, these elements of the Vanguard protocol are not relevant to the pediatrics studies. While no parallel formal sample size analytics for the pediatric study was formulated for the pediatric study, we plan on carefully monitoring accrual and retention as well as protocol compliance, and work with the MoTrPAC DSMB and Steering committee should any modifications be necessary.

4.1.4 COVID-19 PANDEMIC

Recruitment of MoTrPAC pediatric participants started in November 22, 2019. Recruitment was suspended on March 16, 2020, because of the COVID-19 pandemic. Clinical visits were suspended on March 20. When it became apparent the suspension would last more than a few weeks, 14 pediatric participants enrolled in MoTrPAC at the time of the suspension who had not yet completed follow-up visits were administratively withdrawn from the study. For the adult study, it was decided the planned Vanguard Phase, which would have included the first 288 randomized adult participants, will not be completed and the analytical plan will not be carried out. Rather, the valuable information generated during the abbreviated Vanguard Phase was used to modify the adult protocol (v2.0). The data generated during the abbreviated Vanguard Phase will be integrated into the MoTrPAC database. For the pediatric study, it was decided the data generated prior to COVID-19 study suspention will be integrated into the MoTrPAC database.4.2 Scientific Rationale for Study Design

Please see Section 2.1 Study Rationale

4.3 JUSTIFICATION FOR DOSE

The doses of endurance exercise for the training interventions are those that have been demonstrated in numerous published studies to result in physiological adaptations to exercise. This approach is intended to reveal the molecular mechanisms underlying the adaptive responses.

4.4 END OF STUDY DEFINITION

It is expected that all activities related to the clinical protocol (e.g., study visits, processing and analysis of biospecimens, quality assurance 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

- Parent or legal guardian and participant are willing to provide informed consent and assent to participate in the MoTrPAC Study
- Must be able to read and speak English well enough to provide informed consent, assent and understand instructions
- Children and adolescents ages 10-17 (Pubertal stages 1-5)
- Determined to be in good health by pre-participation medical history review performed at PERC
- BMI %ile (≥5th, <95th)
- Weight ≥30 Kg (minimum required for blood collection)
- LAEE children defined as self-reported of no more than 2 days per week, lasting no more than 120 minutes per week, of regular (structured) intense endurance exercise (e.g., running, cycling, elliptical, soccer, swimming and rowing activity that results in feelings of substantially increased heart rate and rapid breathing and sweating with limited ability to talk) in the 3 months prior to study enrollment.
- HAEE children in this study is defined as:
 - self-reported participation in a structured/regular endurance sports (e.g., running, cycling, soccer, swimming and rowing activity that results in feelings of substantially increased heart rate and rapid breathing and sweating with limited ability to talk).
 - Participation in these activities is ≥4 times per week (≥240 min per week) for at least 9 months prior to study enrollment.

5.2 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, and/or clinician judgement as specified for each criterion.

• Diabetes (self-report and laboratory test)

Participants with self-reported use of treatment with any hypoglycemic agents or taking hypoglycemic drugs (e.g., metformin) for diabetes-unrelated reasons are excluded.

Laboratory tests at the last visit of the cross-sectional phase that confirm participant's HbA1c >6.4% and fasting glucose >125 mg/dL will exclude participants from being randomized to the intervention phase. Also, participants with these abnormal values will be flagged in the MoTrPAC database of possibly not meeting the MoTrPAC definition as healthy participants.

- HbA1c >6.4% (laboratory test; may reassess once if 6.5-6.7%)
- Fasting glucose >125 mg/dL (laboratory test; may reassess once)

Molecular Transducers of Physical Activity Consortium–Pediatric Clinical Center Protocol

• Abnormal Bleeding or Coagulopathy (self-report)

- History of a bleeding disorder or clotting abnormality
- Thyroid Disease (self-report)
 - Participants with thyroid abnormalities
- Pulmonary (self-report)
 - Participants with self-reported asthma diagnosis will be reviewed and those who are currently treated with asthma medications (on a frequent or daily basis) are excluded.
- Muscular/Skeletal (self-report)
 - Any connective tissue/autoimmune disease (such as lupus, scleroderma, and/or juvenile idiopathic arthritis)
 - Compression fractures or history of spinal surgery (e.g., scoliosis)

• 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 Resting Blood Pressure (BP) Readings (screening test)

Participants with hypertension are excluded with the following cut-offs based on age.

- Age ≥13 years: Systolic Blood Pressure (SBP) ≥ 140 mmHg; Diastolic Blood Pressure (DBP) ≥ 90 mmHg
- Age <13 years: SBP and DBP ≥ 95th percentile + 12 mmHg (see Tables below), or ≥140 mmHg; DBP ≥ 90 mmHg (whichever is lower)

Dioou	I Tessure re		JI OIII	b o j u	5° una	TIVISI	it pere	Unitite							
		SBP (mm Hg)						DBP (mm Hg)							
Age (y)	BP Percentile		Height Percentile or Measured Height							Height Percentile or Measured Height					
		5%	10%	25%	50%	75%	90%	95%	5%	10%	25%	50%	75%	90%	95%
10	Height (cm)	129.7	132.2	136.3	141	145.8	150.2	152.8	129.7	132.2	136.3	141	145.8	150.2	152.8
	95th + 12 mm Hg	125	126	126	128	129	131	132	87	87	88	88	88	88	88
11	Height (cm)	135.6	138.3	142.8	147.8	152.8	157.3	160	135.6	138.3	142.8	147.8	152.8	157.3	160
	95th + 12 mm Hg	127	128	129	130	132	135	136	88	89	89	89	89	89	89
12	Height (cm)	142.8	145.5	149.9	154.8	159.6	163.8	166.4	142.8	145.5	149.9	154.8	159.6	163.8	166.4
	95th + 12 mm Hg	130	131	132	134	136	137	138	90	90	90	90	91	91	91

Blood Pressure levels for Girls by age and Height percentile

Blood Pressure levels for Boys by age and Height percentile

Molecular Transducers of Physical Activity Consortium–Pediatric Clinical Center Protocol

		SBP (mm Hg)								DBP (mm Hg)					
Age (y)	BP Percentile		Height Percentile or Measured Height							Height Percentile or Measured Height					
		5%	10%	25%	50%	75%	90%	95%	5%	10%	25%	50%	75%	90%	95%
10	Height (cm)	130.2	132.7	136.7	141.3	145.9	150.1	152.7	130.2	132.7	136.7	141.3	145.9	150.1	152.7
	95th + 12 mm Hg	124	125	126	128	130	132	133	88	88	89	89	90	90	90
11	Height (cm)	134.7	137.3	141.5	146.4	151.3	155.8	158.6	134.7	137.3	141.5	146.4	151.3	155.8	158.6
	95th + 12 mm Hg	126	126	128	130	132	135	136	89	90	90	90	90	90	90
12	Height (cm)	140.3	143	147.5	152.7	157.9	162.6	165.5	140.3	143	147.5	152.7	157.9	162.6	165.5
	95th + 12 mm Hg	128	129	130	133	136	138	140	90	90	90	90	90	91	91

Reassessment of a participant's BP during screening is allowed to ensure resting values are repeatable

o Cardiovascular (self-report, screening test, and clinician judgement)

- Significant heart defects
- Congenital heart disease

Cardiopulmonary Exercise Test

- 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 (laboratory test at the end of the cross-sectional phase)

LAEE participants are not invited to participate in the intervention phase if:

- Fasting triglyceride > 500mg/dL and/or
- Low-Density Lipoprotein Cholesterol (LDL-C) >190mg/dL

For both HAEE and LAEE who are participating in the cross-sectional phase, participants with familial hyperlipidemia/dyslipidemia (triglyceride > 500mg/dL and/or Low-Density Lipoprotein Cholesterol (LDL-C) >190mg/dL) are flagged in the MoTrPAC database as possibly not meeting the MoTrPAC definition of healthy participants. Elevated lipid values are not an issue of exercise safety, rather an overall child health.

• Cancer (self-report)

• Participants are excluded if there is self-reported history of cancer treatment

• Chronic Infection (self-report)

- o Active or latent infections requiring chronic antibiotic or anti-viral treatment
- Chronic active infection whether on chronic antimicrobials or not
- Human immunodeficiency virus (HIV)
- o Active hepatitis B or hepatitis C undergoing antiviral therapy
- Individuals successfully treated for hepatitis C and virologically negative for at least 6 months are not excluded

• Hematocrit (laboratory test)

LAEE participants are excluded from the intervention phase if the blood test result for anemia is hematocrit >3 points outside the normal range of the local lab in children and adolescents.

- One reassessment may be allowed under certain conditions
- Individuals with known thalassemia trait may be included (despite having >3 points outside the normal range of the local lab in children and adolescents), upon approval from their primary care provider or a hematologist

For both HAEE and LAEE who are participating in the cross-sectional phase, participants with laboratory evidence of anemia as defined by hematocrit >3 points outside the normal range of the local lab in children and adolescents are flagged in the MoTrPAC database of possibly not meeting the MoTrPAC definition as healthy participants.

- Blood Donation (self-report)
 - Whole Blood donation in the past 3 months (self-report) 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
 - Participant can be rescreened once window of exclusion for donation is passed
- Autoimmune Disorders (self-report)
 - Individuals receiving any active treatment (including monoclonal antibodies) within the last 6 months
- Alcohol Consumption (self-report)
 - Any regular use of alcohol (self-report)
- Tobacco (self-report)
 - \circ Self-reported regular use of tobacco or e-nicotine products ≥3 days/week
- Marijuana / Illegal drugs (self-report)
 - o Self-reported regular use of marijuana or illegal drugs in any form
- Cognitive Status (screening)
 - Unable to give consent/assent to participate in and safely complete the protocol, as based on the judgement of the local investigators
- 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
- COVID-19
 - 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 (without a negative antigen test) or symptom-free for at least 7 days with a negative antigen test on the day of the study visit. A PCR test can be used if antigen tests are not available.
- Other Conditions (clinician judgement)
 - Genetic metabolic disorders that could affect 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

• Medication Exclusions

Daily or weekly chronic use of any prescription medication in the past 3 months (excluding birth control).

- Continuous use for 7 days or more of a new drug (prescription or over-the-counter) 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 is an exclusion; changes in eye and ear drops are allowed.
- In these cases, a period of 3 months or more would need to pass before re-screening.
 Participants that take a non-exclusionary medication within 3 months, but not in daily/chronic fashion (less than 3 months), may enter MoTrPAC. However, a washout period of at least 5 half-lives must pass before any baseline sampling of blood or tissue.

• Special Medication 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

• Birth Control Medications

• Birth control pills, patch or shots are permitted.

The intervention phase has additional 2 exclusion criteria 1) children who meet the definition of HAEE defined above, 2) only one member of a household can participate; "household" is defined in the Manual of Procedures (MOP).

5.3 LIFESTYLE CONSIDERATIONS

Certain lifestyle factors such as physical activity level, smoking, alcohol intake, and body weight regulation are considerations for eligibility, and for each there are well-defined criteria for inclusion/exclusion. Free living physical activity and body weight will be monitored several times throughout 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

Effective recruitment and retention of study participants is critical for the success of MoTrPAC. The recruitment goal of MoTrPAC pediatric study is to enroll 320 participants and complete evaluation on 300 children (aged 10-17 years) for the Pediatric Cross Sectional Phase. A subgroup of 170 children are randomized to the intervention phase -- 12-week aerobic training (n = 120) or no training controls (n = 50).

The recruitment plans utilize a variety of approaches directed at population-based recruitment within the catchment area, including flyers/posters at community events, use of social media, and newspaper advertisements. A key component of our recruitment strategy is built on our long-term relationship with K-12 schools (elementary, middle, and high school) in our region, with whom we have successfully recruited several hundred 10-17 y.o. participants in a variety of studies. In addition, we will recruit from our database of past participants who gave consent for future contact, and our access to regional health centers, and regional health fairs.

All participants are enrolled with their parents and/or legal guardians present. PERC staff are well trained in objectively presenting the study's purpose and potential adverse effects. Enrollment can occur at the PERC site or at affiliated schools, or community venues. Informed consent and assent are obtained by the 2 PIs, Co-investigators, and trained designated PERC personnel from each subject's parent(s) or legal guardian(s) as well as assent obtained from the child. The consent and assent forms are reviewed with each subject's parent(s) or legal guardian(s) in accordance with the guidelines set forth by the IRB.

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, the pediatric site develops and implements an outreach strategy for recruitment of underrepresented minorities reflecting the population heterogeneity in our region.

5.5.3 RETENTION STRATEGIES

MoTrPAC employs several retention promotion strategies: 1) Provision of health-related study results to participants and, if approved by the participants' guardians, to their primary care providers; 2) Promoting easy access to and accommodating study assessment visit times; 3) Accommodating study training session days and times including weekends.

All participants receive useful information including a metabolic panel and lipid screening which is now recommended by the American Academy of Pediatrics for children between the ages of 9-11 years-old. In addition, each child receives information about their body composition and fitness. For those participants randomized to the control group in the intervention study, we maintain regular contact throughout the intervention period and plan to provide Control participants with an individualized exercise prescription, guidance and education materials following their participation. In addition to the above, high schools grant community service credit to students from their schools who participate in our studies.

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

- **Participant-staff relationship**. A key element contributing to participants 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 provide friendly and individualized interactions. The participant is reminded of the importance and benefits of study participation.
- **Convenience and accessibility**. An easily accessible study location and flexible hours all serve to facilitate study adherence. PERC make study visits as easy as possible for participants, a factor critical to the success of the study. Attendance is not compromised by unsuitable hours of clinic operation, or any similar circumstance.
- **Time in clinic**. Total research 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. Research staff is trained to take time to visit with and listen attentively to participants.

Molecular Transducers of Physical Activity Consortium–Pediatric Clinical Center Protocol

- Endurance exercise training study. A personalized training program is designed for each participant in the exercise training and will be monitored using activity logs and HR monitors. Trainers review performance and personal goals with participants on a weekly basis.
- **Control group.** 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 meet the study team at school or in PERC approximately every 4 weeks to have the data from the physical activity monitor downloaded and to have their height and weight checked.

5.5.4 PARTICIPANT REMUNERATION

The participants' remuneration is given for attendance at study activities. The participants receive up to \$280 for completing the cross-sectional health and fitness phenotyping assessments. For the intervention training study phase, the exercise group will be paid up to an additional \$570 and no-exercise control group up to an additional \$320.

5.5.5 MONITORING RECRUITMENT AND RETENTION

Screening, recruitment and retention yields will be continuously monitored, and reported by the pediatric 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 the intervention phase are encouraged to attend the assessment visits for intent-to-treat analysis.

6 STUDY INTERVENTION

6.1 STUDY INTERVENTION(S) ADMINISTRATION

All participants first engage in a single acute EE exercise challenge on a cycle ergometer.

Prior to and following the acute EE exercise challenge, blood samples are collected (before; 18-20 and 38-40 min into the exercise; 10 min, 30 min, and 4 hr following the exercise).

Following the pre-intervention acute exercise challenge and blood collections, a subgroup of participants completes approximately 12 weeks of either EE training as described below or continue their LAEE status as a Control participant. All participants are instructed to not change their levels of free-living PA outside of the active interventions, or change their habitual diet during the intervention period.

To monitor the expected growth and development of many of our participants, participants in EE intervention have standard anthropometric data (weight, height) and are regularly queried as to whether there has been any change in their level of activity outside of the training sessions. Non-compliance to remaining weight stable (in line with normal growth) and not engaging in structured endurance exercise outside of the supervised exercise sessions triggers an action plan to assist participants in achieving these goals. All participants in the intervention phase 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, participants (EE training and controls) repeat the acute exercise challenge with blood collection.

6.1.1 ENDURANCE TRAINING

Participants randomized to EE training engage in three school-based or center-based (PERC) EE training sessions each week for approximately 12 weeks; each session lasting roughly 70 min with a 45-50 minutes of a stimulus phase and the remaining time being used to warm-up and cool down. The pediatric EE training protocol is not precisely the same as for adults. In order to achieve the endurance training duration and intensity in the pediatric age group, we provide exercises that are appropriate to the pediatric participants' age and willingness to exercise for long periods of time. As noted in the recent statement of the 2018 Physical Activity Guidelines Advisory Committee Scientific Report [44] "it is important to encourage young people to participate in physical activities that are appropriate for their age, that are enjoyable, and that offer variety." Middle and high school participants are trained on a variety of ergometers (cycle, treadmill, elliptical and rowing machine) and physical activities appropriate for the participant's age, capability, and enthusiasm based on a predetermined target heart rate. The individual instructors are trained at PERC to ensure uniformity of the endurance exercise. Elementary school students are trained in a form of circuit training (e.g., endurance activity stations: cycle ergometer, steppers, individual jump rope, pacer, sliders) to keep the younger students engaged. Each week, the goal is for at least two of the sessions to involve a cycle ergometer. 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 in level and duration during the intervention: $65\% \pm 5$ bpm for weeks 1-4; $75\% \pm 5$ bpm for weeks 5-8; $80\% \pm 5$ bpm for weeks 9-10; and $85\% \pm 5$ bpm for weeks 11-12. Periodically during training sessions perceptual data from participants are recorded, which are used to track the subjective experience of participants and in interpreting adherence data. Personal goals and performance are reviewed with the participants on a weekly basis. 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 CONTROL GROUP

Participants randomly assigned to the control group do not participate in our exercise intervention during the 12-week period of the EE training. However, we do not restrict the members of the control group from participating in their planned exercise activities, outside of those activities specified in the exclusion criteria. We instruct all control participants not to change their physical activity levels. We implement two steps with the control group in order to maintain contact and to monitor changes in physical activity and nutrition. First, control participants are called regularly to check on their health and to query about any changes in their physical activity or diet. Second, the control participants will meet with study staff approximately every 4 weeks to have the data from the accelerometer (wearable physical activity monitors) downloaded and to have their height and weight checked. Participants who are identified as changing either their level of physical activity or dietary habits are reminded about the importance of the study goals and provided an action plan to assist in correcting the problem area(s). These calls and visits will also help with participant retention during the intervention phase.

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 (**AEs**). 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 various clinical/chemical sites, 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 is essential in capturing any systematic bias. Tools that estimate and correct for latent bias such as population principal component analyses (**PCAs**) 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 randomly assigned in the intervention phase. 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 intervention 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 their pre-enrollment levels of physical activity 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 randomly assigned 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 child and adolescent participants who join MoTrPAC do so because they are interested in helping to improve health of other people, and some children and adolescents do want to learn more about being physically fit. Although participants are informed prior to baseline testing that they may be randomly assigned to either EE 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 the training intervention.

Because randomization to the control group poses a risk for dropping out, pediatric participants assigned to this group will be contacted by study staff at regular intervals. In addition, it is essential that all participants in the control group 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). Study staff will review these data, identify substantial changes in physical activity levels, and discuss such changes with the participants.

Adherence to the exercise protocols is 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 participants 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 parents and caregivers 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 and their parent/guardian 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 (e.g., intensity, duration)) are stored for future statistical analyses 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 complete 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

Participants begin study participation at screening and orientation. Table 5 describes in detail the procedures that each pediatric participant enrolled in the cross-sectional study phase completes. Following this, a sub group will be randomly assigned to the 12-week EE training program or no exercise control group. Randomization to the intervention phase occurs no later than 30 days from cross-sectional acute bout and blood collection visit.

Total time for participants to complete post-intervention phenotyping assessments and acute exercise challenge/blood collection is approximately 10-12 hours. Phenotyping assessments can begin during the final week of training. The post-training acute exercise test with blood collection must occur after the intervention period and, for EE training participants, this session occurs approximately 2-3 days after the final exercise training session. In addition, we ensure that the post training assessment occur only if the last two weeks of training were successfully completed. In the event that a participant leaves the intervention before its conclusion, the participant is invited for the final assessments, consistent with the intent-to-treat study design.

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

8.1.2 MEASURES AND PROCEDURES OVERVIEW

Phenotyping measures are 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 AND ASSENT

Following an initial phone/email/web-based screening (e.g., for self-reported age, BMI, exercise history), potential participants who meet eligibility criteria and remain interested are scheduled for a study orientation session. After all questions have been addressed, written informed consent/assent 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 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

Review of medical history, medication inventory, height and weight (to assess weight and BMI percentile eligibility), pubertal staging by self-report, resting blood pressure, and resting electrocardiogram.

8.1.2.3 PHYSIOLOGIC MEASURES

8.1.2.3.1 CARDIORESPIRATORY FITNESS (PEAK AND SUBMAXIMAL GAS EXCHANGE AND HEART RATE DATA)

At baseline (for all pediatric participants) and post intervention (EE training or control) cardiorespiratory fitness is determined by cardiopulmonary exercise testing (CPET) on an electronically braked cycle ergometer per the American College of Sports Medicine (**ACSM**) Testing Guidelines. Oxygen uptake and carbon dioxide production are measured using standard techniques. Resting 12-lead ECG recordings and blood pressure are monitored, and a clinician interprets for any contraindications to exercise prior to enrollment.

8.1.2.3.2 BODY COMPOSITION AND ANTHROPOMETRICS

At baseline (for all pediatric participants) and post intervention (EE training and control), total and regional body composition and bone mineral content (**BMC**), and density (**BMD**) are determined by whole body DXA scan. Standard measures of height, weight, and waist circumference will be made using calibrated stadiometers scales and measuring tapes, respectively. Body weight and height are measured every 4 weeks for participants in the intervention phase.

8.1.2.3.3 ISOMETRIC STRENGTH

At baseline (for all pediatric participants) and post-intervention (EE training and control), participants are evaluated for knee extensor strength via isometric maximum voluntary contraction at approximately 60° knee flexion.

8.1.2.3.4 ONE-REPETITION MAXIMUM (1RM) DYNAMIC STRENGTH

This assessment will not be performed in the pediatric participants as there will not be a resistance training arm of the intervention.

8.1.2.3.5 GRIP STRENGTH

At baseline (for all pediatric participants) and post-intervention (EE training and control), isometric hand grip strength is tested using a handgrip dynamometer and standardized methods.

8.1.2.3.6 POTENTIAL ANCILLARY MEASURES

Ancillary measures may be added as part of approved ancillary studies to MoTrPAC (e.g., sleep studies). 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

At baseline (for all pediatric participants) and post-intervention (EE training and control) a limited battery of self-reported health outcome is measured using the Patient-Reported Outcomes Measurement Information System (PROMIS) [45,46]. PROMIS measures are selected from the pediatric inventory that correspond to those used in the MoTrPAC adult studies. For example, depression, anger, and anxiety from the mental health domain, fatigue and pain from the physical health domain, and peer relationships from the social health domain are selected. All measures are administered via computer. Time to complete each measure is range from 1-3 minutes. Thus, we anticipate that the total time for administration of the PROMIS battery to range of approximately 8-15 minutes at both the baseline and follow-up testing.

8.1.2.3.2 DIETARY INTAKE ASSESSMENT

The DHQ3 (used in the adult MoTrPAC studies) is not appropriate for pediatric participants (less than 19 y.o.) because the food lists and portion-sizes are based on adult data. To characterize and describe the typical dietary intake of pediatric participants, a standard semi-quantitative food frequency questionnaire designed and validated specifically for children and adolescents (Block FFQ for children and adolescents) [47] is administered during the Health and Fitness Phenotyping Visits to estimate energy, macronutrient, and micronutrient intakes. In addition, a 24-hour dietary recall is recorded for the 24-hour period prior to the first blood collection and entered into the Automated Self-Administered 24-Hour (ASA24[®]) Dietary Assessment Tool website [48]. To minimize the influence of dietary intake, participants are then asked to replicate this same dietary intake schedule for the post-intervention blood collection period.

8.1.2.4.3 PHYSICAL ACTIVITY

Free-Living Activity is assessed using wearable activity monitors (accelerometers) that enable us to quantify the intensity and patterns of usual activity and sedentary behavior, including sleep, and will be worn for approximately 7 days in the cross-sectional phase and at weeks 4, 8, and 12 during the intervention phase.

8.1.2.5 BIOSPECIMEN COLLECTION

All study participants are asked to not perform intense physical exercise a few days prior to study visits involving blood collection. Participants arrive at the pediatric site in the morning after an overnight fast (~8-10 hours after consuming a meal the prior evening). Participants are encouraged to consume water the evening before and morning of the tests. An IV is then placed in the antecubital vein for metabolic screening (e.g., metabolic panel and lipids). The child then receives a light, standardized snack.

After 30 minutes of rest, a baseline blood sample is drawn. During and following the exercise challenge, additional blood samples are collected at multiple time points up to 4 h into recovery. Each blood draw ~ 20 ml. Total volume ~ 130ml. Additional blood samples may be collected for safety or other study needs when necessary,but <u>no more</u> than 5ml/kg per accepted standards for healthy children and adolescents. All blood 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 ACUTE EE CHALLENGE (CROSS-SECTIONAL PHASE AND POST EE TRAINING PHASE)

After the pre-exercise baseline blood sample is obtained, the participant performs a brief warm-up consisting of light cycling on a cycle ergometer. After the warm-up, the pediatric participants cycle at pre-determined target workload that corresponds to ~65% $\dot{V}O_2$ peak based on the true $\dot{V}O_2$ vs. power regression. This target workload is held constant throughout the protocol: two bouts of 20 minutes at a constant target work rate exercise with about a 2-minute rest between the 2 bouts, unless it has to be reduced due to participants' inability to complete the acute bout. If this occurs, the exact same load reduction will occur on the post-intervention test.

Following completion of the acute exercise test and 4-hour post exercise blood collection, the participant is provided with a light post procedure meal. Participants are given detailed instructions about care of the phlebotomy sites.

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 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 AND SAFETY SCREENING

To maximize the participants' safety, we follow a standardized screening protocol which is appropriate for healthy children. Accordingly, a brief medical history is performed by qualified study personnel. Exclusion and inclusion criteria are adhered.

8.2.3 SAFETY CONSIDERATIONS FOR STUDY ASSESSMENTS

All study assessments are done 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, or feeling faint, lightheaded or dizzy the assessment or procedure is stopped. Onsite clinical site staff are trained to provide basic life support and to provide immediate care when faced with medical emergencies. Institutional and community emergency services are activated if needed.

8.2.3.1 DXA

DXA scans to determine body composition and bone density 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 CPET

Peak aerobic capacity (peak $\dot{V}O_2$) is measured during CPET to exhaustion and used for prescription of exercise intensity (at the exercise challenge) and as an outcome measure of cardiorespiratory fitness. We have extensive experience measuring peak $\dot{V}O_2$ in children and adolescents.

Maximal exercise testing is a common procedure with risks, including fainting, dizziness, chest pain, or irregular heartbeat. The test is monitored by an exercise physiologist or a trained health professional and is stopped if problems occur. Blood pressure, heart rate and rhythm, and breathing are closely monitored before, throughout, and after the test.
8.2.4 SAFETY CONSIDERATIONS FOR THE INTERVENTIONS AND ACUTE EXERCISE CHALLENGE

The entire exercise intervention is conducted on-site (or in the schools) and all sessions are supervised by trained exercise interventionists, who monitor potential adverse experiences and symptoms. During the exercise sessions, on-site trained staff are available to deal with medical emergencies. Institutional and community emergency services are activated if needed. Participants are taught the importance and proper method of warming-up prior to and cooling-down following structured activity sessions, instructed on correct exercise techniques and supervised during their activity sessions. Heart rate is monitored throughout the intervention sessions. If at any point during an exercise session, participants develop chest pain, significant shortness of breath beyond what is expected during exercise, nausea or dizziness, they are instructed to rest and to contact their physicians if these symptoms persist or recur with further physical activity.

If for any reason the participant reports an injury, chest pain, shortness of breath, dizziness, or other sign or symptom of a cardiorespiratory event, they are referred to their doctor (or the study physician is to call the doctor or other health care provider). In addition, in the MOP, specific criteria for suspending or stopping physical activity are developed to adjust the program for inter-current illness. The exclusion criteria eliminate individuals who are high risk.

8.2.4.1 SAFETY CONSIDERATIONS FOR SUPERVISED EE 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
- Nausea or abdominal cramping
- Muscle, joint strains and soreness
- Soft tissue injury, falls and fractures
- Irregular heart beats
- Triggering of an asthma episode
- 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:
- Mental alerts include a change in mental status and/or mood.
- Acute shortness of breath.

If such health changes are encountered, clinical site staff are trained to evaluate the participant and take specific actions such as stopping exercise, calling study physician, cancelling training, immediate response to falls, and calling 911 if necessary. Clinical staff trained in Cardio-Pulmonary Resuscitation

(CPR) and the use of an Automated External Defibrillator (AED), are available to deal with medical emergencies. Also, institutional or community emergency services is activated if needed.

8.2.5 SAFETY CONSIDERATIONS REGARDING BIOSPECIMEN COLLECTIONS

8.2.5.1 BLOOD SAMPLE COLLECTION

Blood is drawn only by trained and experienced phlebotomists, to minimize the discomfort and risk as much as possible.

Venipunture 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

8.2.6 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.

8.2.7 COVID-19 SAFETY CONSIDERATIONS

A COVID-19 appendix to the MOP, which includes recommended strategies for minimizing exposure risk for participants and research staff, will be approved by the MoTrPAC DSMB. It is expected that each MoTrPAC clinical site will also be required to follow local policies and guidelines.

8.3 ADVERSE EVENTS (AE) AND SERIOUS ADVERSE EVENTS (SAE)

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 UCI 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 blood samples (i.e., bruising and tenderness) and the performance of exercise (i.e., muscle and joint discomfort).

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., tenderness and bruising during phlebotomy) and from the performance of exercise (i.e., muscle or joint discomfort) 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 UCI IRB prompt reporting guidelines. 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),
- Unresolved participant complaints that involve unexpected risks that can't be resolved by the research team.

Та	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			ed		Expected			Ur	Unexpected		
		Relatedness	Definitely	Not	Р/Р	Def	Not	P/P	Def		Not	P/P	Def	Not	P/P	Def
		Collect on study visit CRFs	Yes	No	No	No	No	No	No		No	No	No	No	No	No
	Actions Required	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 such as exercise-related muscle and joint soreness. These will be collected on various site 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 Problemsj, depending on whether they affect participant safety.

8.3.1 DEFINITION OF ADVERSE EVENTS (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 visit to a physician because it alters the participant's ability to exercise
- An event that occurs as a result of a study procedure which is not listed in the Risks section of the consent

8.3.2 DEFINITION OF SERIOUS ADVERSE EVENTS (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
- 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/SAEs reported by participants are likely to qualify as expected events, including injuries and accidents. Expected events are occurrences that are listed in the consent form or those which are unrelated to exercise, but are to be expected in the study population. The determination of expectedness is made by a clinician at the clinical site and confirmed by the Medical and Clinical Safety Committee.

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 or the patient population being studied, or normal consequences of an intervention. These conditions that may be unpleasant and bothersome to the participant, such as sore muscles, muscle or joint pain, fatigue or breathing problems, do not require discontinuing the study intervention or components of the intervention.

8.3.3.1.2 DEFINING UNEXPECTED ADVERSE EVENTS

Unexpected adverse events (UAEs) are defined as events that occur during participation in the study but do not commonly occur in the study population and are not listed in the informed consent or protocol and are not an expected natural progression of a participant's pre-disposing risk factor profile.

8.3.3.2 RELATEDNESS

The classification of potential relationship to the intervention is as follows.

- Not Related AE is clearly not related to the intervention or 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).
- Probably/Possibly Related AE that follows reasonable sequence from administration of intervention or procedure, but that could readily have been produced by several 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 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 side effects from the intervention and assessments to be used in MOTRPAC have been well defined in previous studies. However, following the guidance from the NIAMS, MoTrPAC sites collect, enter, and report to the NIAMS and the DSMB all AEs and SAEs disclosed by the participant.

Participants are instructed to spontaneously report SAEs immediately to the site PI. Post-enrollment, participants can also disclose AEs/SAEs between visits and enrolled participants are queried for SAEs and AEs in a standardized fashion about every four weeks post-randomization (intervention phase) through their final visit. There are two timeframes of AE/SAE collection and reporting. The two timeframes are described below:

First timeframe:

After signing informed consent and prior to initial exercise challenge:

- **AEs/SAEs Not Collected or Reported**: Pre-existing conditions and pre-planned procedures (surgeries or therapies) scheduled prior to signing the Informed Consent are not considered AEs or SAEs. **These will not be collected or reported.**
- **Expedited Reporting of SAEs**: The study team will report in an expedited fashion (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 probably/possibly related to the study procedures.
- **DSMB Reporting**: All AEs/SAEs **definitely or probably/possibly related** to study procedures are collected and reported in the aggregate DSMB Report during the DSMB meetings, typically held biannually.

Second timeframe – Post-randomization:

After randomization until their last contact (after intervention and second exercise challenge):

- **AEs/SAEs Collected or Reported**: All AEs/SAEs are collected regardless of expectedness or relatedness to the study.
- **Expedited Reporting of SAEs**: The study team will report in an expedited fashion (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 probably/possibly 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 UCI IRB 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 UCI IRB within 10 days of discovery. Deaths must be reported to the UCI IRB 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 KAI Research, Inc. (KAI)/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

Incidental findings from study assessments will be provided to the participants. If the DSMB or IRB finds that adverse events associated with participation are more serious or more frequent than expected, this information will be conveyed to study participants through a mechanism recommended and approved by the DSMB and IRB (e.g., reconsenting).

8.3.7.1 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. For UCI purposes, 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 UCI IRB 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 UCI IRB 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 UCI IRB 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 UCI IRB at the time the continuing review application is submitted.

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 UCI 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 UCI IRB and the UCI 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:

- 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.
- 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)

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 a 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 unanticipated problem (UP) are reported to the UCI IRB in accordance to their policy, within 10 days of participant disclosure to the clinical site. All other events are reported at the time of the clinical site's annual report.

8.4.3. REPORTING UNANTICIPATED PROBLEMS TO PARTICIPANTS

Unanticipated Problems can be either **1)** Unanticipated AE/SAEs, which are unexpected events that relate directly to participant safety, or **2)** PDs that place patient privacy at risk or place patients 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).

Power analyses were performed using the power REG procedure in STATA [49] and then cross validated with G*power 3.1 [50]. The approach was to determine the level of power available to detect within group differences, between group differences and interactions for a range of effect sizes. We controlled for the False Discover Rate (FDR) at the 0.05 level, (assuming 10% of hypotheses fall under the

alternative hypothesis), with a threshold alpha=0.005263. The cross-sectional study will recruit a sample size of 320. The goal of the crosssectional study is to establish a wellcharacterized cohort of adolescent males

and females

measured before and after acute exercise.

Table 8. Distribution of peak \dot{VO}_2 normalized to weight (kg)							
	Females 11-17 y.o. n=55 (median age=14)	Males 12-17 y.o. n=50 (median age=15)					
Percentile	VO₂/ml/kg	VO₂/ml/kg					
100	58.14	65.78					
95	51.52	60.75					
75	41.19	51.85					
50	34.86	45.92					
25	30.15	38.47					
5	23.56	31.12					
0	21.90	25.54					
Mean (SD)	36.65 (8.5)	45.67 (9.3)					

Analysis of physical fitness will consist of sex stratified estimation of many variables including but not limited to peak VO₂/ml/kg obtained during the progressive exercise test.

Sample size/power analyses were centered on peak $\dot{V}O_2$ due to available data and its history as key outcome in the field. Distributions of normalized peak $\dot{V}O_2$ from past research demonstrate a wide range of fitness levels to be captured in this target age range (Table 8).

These observed distributions are comparable to the norms presented for an adolescent population in the Netherlands with a larger range and standard deviation, and slightly lower mean/median values

(especially for females).	Table 9. Power for biomarkers with effect sizes (ES) less than peak $\dot{V}O_2$							
Regression	Total Sample	Estimated Power						
models will be	Size	Moderate Effect	Small-to-moderate	Small Effect				
constructed to		5% Additional	Effect	3% Additional				
evaluate mean		Variability	4% Additional	Variability				
differences			Variability					
between the	260	0.95	0.87	0.72				
sexes, main	280	0.96	0.90	0.76				
puberty status	300	0.98	0.92	0.80				
(pubertal	320	0.98	0.93	0.83				

stage), as well as an interaction of puberty and sex. Previous analyses by PERC scientists of 55 females and 50 males in the target age range indicated a change in R² of 0.15 attributed to age. It is expected that pubertal stage will have approximately equivalent effect as age, with both covariates together having an R² = 0.25-0.30. Assuming this effect of puberty and sex for a model that includes: sex, puberty, and an additional factor; the target sample size of 320 participants will have approximately 98% power to detect an additional R² = 5% for the additional factor. The results in Table 9 provide power to detect a range of explained variability, which could be observed in other exercise biomarkers.

The training study is planned as a randomized parallel repeated-measures design (measures evaluated pre- and post-intervention). A control group (no exercise training intervention) was deemed necessary due to the changes in growth that can occur even in a 12-week period during normal adolescence. To the extent that pubertal processes are related to fitness, it will be important to consider that changes in fitness can be related to the intervention and not simply maturation. Thus, randomization will be stratified by sex and we will strive to distribute enrollment of participants across pubertal stages. We expect to have an analytic sample of 170 participants, 120 in the training arm and 50 in the control arm. Each arm will contain approximately 20% male pubertal stages 1-2, 10% male pubertal stage 3, 20% pubertal stages 4-5, 20% female pubertal stages 1-2, 10% female pubertal stage 3, and 20% female pubertal stages 4-5.

Initial power calculations were based on a mixed linear model with adjustment for baseline, time (if there are >1 post-baseline measures), sex, and treatment. Other baseline variables may be included for some analyses. Primary interest will be in testing contrasts at 12 weeks. We will also test for differential subgroups by adding in an indicator variable for the subgroup and the subgroup by treatment interaction. Sex and baseline puberty status will be included as covariates. Previous data from training in healthy adolescents done in our laboratory has shown an average 14.5% change in peak $\dot{V}O_2max$ over 8-week of EE intervention. The average pre-post change score was 5.7 ml $O_2/kg/min$ with a pooled standard deviation of 9.1. The correlation between the two repeated measures was 0.8. This translates to a within-subjects effect size (i.e., mean change / SD of change) of 0.99. A larger effect size of 1.02 was seen among 11 participants whom exhibited lower baseline fitness, suggesting the

improvement can be expected across fitness levels. The proposed sample sizes will yield greater than 99% power to detect within-group effect sizes as large as these even when a sub group is as small as 35 participants.

Investigation of sex and Tanner effects on levels of fitness and fitness gains (i.e., degree of change) will require additional exploratory modeling. Power analyses depend on the sample sizes in each group and the correlation between paired measurements. In Table 10, we have included detectable effects comparing the 120 in the training arm compared with the 50 in the control group, and for investigating subgroup effects within the exercise condition (e.g., sex n=60/group).

	Table 10. Detectable Differences in Mean Change						
	80%	Power	90% Power				
Intra-Individual		N=120/50		N=120/50			
Correlation (r)/SD of Change	N=60/60		N=60/60				
.5 / 1 SD	0.67	0.62	0.76	0.69			
.6 / 0.89 SD	0.60	0.55	0.68	0.62			
.7 / 0.78 SD	0.52	0.48	0.59	0.54			
.8 / 0.63 SD	0.43	0.39	0.48	0.44			

9.3 POPULATIONS FOR ANALYSES

Both intent to treat (ITT) and per-protocol (PP) analyses will be performed with the data collected in MoTrPAC. Analyses will be performed across all participants and within subgroups, which will include, but will not be limited to, age, sex, pubertal stage, 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, percentiles, proportions, etc.) for participant characteristics and outcome measures by intervention, actual versus projected accrual, and 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. These models are used to account for over-dispersion typically observed in count data. 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 [51,52]. 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 [53]. 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 [54].

To alleviate these issues mediation hypotheses from two families of causal mediation models can be used: structural mean models REF and principal stratification REF. 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 [53] discusses practical graphical methods that can be easily implemented to address the point above.

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 that were randomized, regardless of adherence. Participants are only considered to be "off-study" if they have withdrawn consent in writing (where possible) for all future contact or have died. Participants who discontinue the intervention are encouraged to continue study assessment procedures. Per protocol 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 TWAS-based approaches where genetic information is used to predict missing gene expression measurements [55].

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 generalized estimating equations 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 of the exercise 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 site 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, multiomic 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 analyses for long-term experiments 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 etc. using fewer parameters.

The molecular analyte (transcript, protein, metabolite) at a given time point in a specific subject 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 subject 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 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 [56]. 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 [57] 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, and molecular assay. 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) [58,59].

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 generalized linear models 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 VO2max, change in strength, change in body composition, change in heart rate, change in cognition, change in insulin sensitivity, 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 [60] 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, from many subjects, 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 [61,62] compartmental models [63] 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. 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 patient. 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 is 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 [64]. Naturally, we shall 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 broken down into two types: (1) quantifying association in the presence of time and repeated measures, and (2) normalization of the correlation matrices, both inter- and intra-omics.

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. DREM [65,66], for example, uses input-output 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 TF-target networks as these data, but in essence, any gene-based exogenous information can be used, including other omics [67]. 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, which can be used to augment the original models with more information.

Another example is a tool called TWIGS REF that was developed for detecting differential abundance modules in longitudinal datasets that have many subjects. It is more flexible in that it allows a module to be only partially represented in a subject (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 can be addressed by more advanced network analyses. For example, network propagation approaches [68] 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 multiomics 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 [69].

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 EFFICACY ENDPOINT (S)

There are no pre-specified secondary endpoints.

9.4.5. SAFETY ANALYSES

AEs/SAEs are coded using the MedDRA. We summarize frequency of all AEs/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, standard deviation, 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 is made for pre-specified safety endpoints as this is considered a conservative approach; interpretation of significance levels on safety endpoints that may arise are done cautiously.

In general, between-group differences in the proportion of patients with predefined adverse events are assessed using either chi-square or Exact Tests (when expected cell counts are small). The proportion of patients with other adverse experiences are estimated within intervention groups, and expressed as relative effects via odds ratios (95% Confidence Intervals [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 including: 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 pubertal status and ethnic/racial minorities. However, given the sample size of the study some of these 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 Medical/Clinical Studies Safety Committee (MCSSC) 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 (ICF) PROCESS AND DOCUMENTATION

Informed consent/assent are obtained in compliance with the principles of the Declaration of Helsinki, UCI IRB regulations, and ethical standards. The IRB reviews and approves all informed consent documents before they are presented to the volunteers.

- 1. All participants and their legal guardian must be informed both written and orally.
- 2. All participants and legal guardians must give their informed consent/assent before the screening and enrollment (although some initial pre-screening via telephone/email is allowed prior to giving written informed consent/assent (e.g., age, sex, BMI, exercise history)).
- 3. Informed consent/assent is obtained from the volunteer by the clinical site investigator and/or the designated research staff. The MoTrPAC research staff fully informs 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 informed consent form (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 IRB or Ethics Committee.

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):
 - o 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
 - o 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.
 - o 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.
 - o 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
- BIC will have access to dates of measurements for some types of data that are directly transferred to the BIC (e.g., raw accelerometry files, metabolic cart files); however, all files transferred from the DMAQC to the BIC will convert dates to days from initial registration
- o CAS will have access to days from initial registration
- **Age:** There is no planned aggregation of ages among pediatric site participants.
 - o 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
 - o 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
 - o 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 Johns Hopkins (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 Institute 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 monitors subcontracts with external providers, administers opportunity funds, matches timelines and deliverables, 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 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.

10.1.5.2 CLINICAL STUDY TIMELINE (TABLE 12)



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 IRB.

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; clinical 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 IRB 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 us 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 clinical 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

Molecular Transducers of Physical Activity Consortium–Pediatric Clinical Center Protocol

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. In the pediatric study, block randomization will be used for sex. 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 has access. 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 are stored at the DMAQC via the secure and encrypted website. Access to the website and to the data on the website, as well as various areas of the website, are 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.
- 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 10 years after the formal discontinuation of MoTrPAC per the Standard Operating Policies and Procedures of the UCI IRB. 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).
Molecular Transducers of Physical Activity Consortium–Pediatric Clinical Center Protocol

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 assessment, 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 multionic 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 <u>BCF</u>. 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 [70,71]. In addition, we have recently described how our metadata objects are modeled and updated [72].

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 [73].

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 (<u>https://goo.gl/7JrqRy</u>). 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 [74]. 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 [74]. 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 [75,76].

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 who 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. Identifying data such as genotypes or raw sequencing are considered high risk and, as such, have limited accessibility even within the Consortium. Highly summarized data such as differential abundance or genetic associations are considered low risk and are 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 targeting local and national newspapers, television, and radio outlets;
- 4. Production of the research summary document and facts sheet targeting the general public which clearly and concisely summarize the key conclusions of the study;

- 5. Production of professionally designed flyers, posters, brochures, and research briefs targeting 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 targeting 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 submits 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 reidentification, 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 are coded with linked identifiers to controlled-access metadata labels/tables that facilitate coordinated identification across **1**) samples, **2**) individuals, **3**) interventions, and **4**) databases. All human sample-derived sequencing-based data are submitted through the database of Genotypes and Phenotypes (**dbGaP**) and the associated protected Sequence Read Archive (**SRA**) instance.

Data undergo different levels of processing, as defined at <u>https://osp.od.nih.gov/wp-</u>

<u>content/uploads/Supplemental Info GDS Policy.pdf</u>. 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) are 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) are provided through the BIC portal.

The timing of external data releases is 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 evaluates 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 are removed and indirect identifiers are 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

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
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			
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			
KAI	Kunitz and Associates, Inc.			
LAEE	Low Active Endurance Exercise			
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				
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				
РСР	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				
SAE	Serious Adverse Event				
SBP	Systolic Blood Pressure				
SC	Steering Committee				
SD	Standard Deviation				
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				
UCI	University of California Irvine				
UP	Unanticipated Problem				
US	United States				
VO ₂ peak	Peak Aerobic Power				
WGS	Whole Genome Sequencing				

10.3 PROTOCOL AMENDMENT HISTORY

Version	Date	Sections Affected	Description of Change	Brief Rationale
	09/18/2019	1.2, 1.3	Edits to wash-out time intervals	Removed testing window and put in MOP
		4.1.3	Edits to vanguard phase	Not completed per protocol due to COVID-19 suspension and not resuming
		New 4.1.4	Edits to COVID-19 details,	Add description of suspension
		5.1	Edits to Inclusion criteria BMI %ile >5 th to BMI %ile ≥5th	Per CDC definition should be ≥
		5.2	Edits to Exclusion criteria - Changed triglyceride to ≥ 500mg/dL; Added COVID-19 infection; Clarified that only one member of a household can be randomized to the LAEE training phase	Similar to adult protocol; minimize effects of shared exposures
		8.1.2.1	Eliminated statement that orientation visit must be in- person	Allow teleconference or videoconference orientation visits
		8.1.2.5	Additional blood samples	As needed, for safety reasons
		8.1.2.5.1	Eliminated information on monitoring exercise intensity	Detailed plans are in the MOP
		8.2.7	COVID-19 safety considerations	Specific guidelines added to the MOP
		10.1.5	Eliminated NIBIB as a managing institute	No longer serving in this role
8.9	12/7/21	5.2	Updated definition of fully vaccinated for COVID-19 align with CDC guidelines	Update to align with MoTrPAC adult protocol
		7.2	Add that participant may be withdrawn due to a COVID-	Updated to align with MoTrPAC adult protocol

			19 diagnosis after consenting to participate.	
8.10	1/11/22	5.2	Revise eligibility criteria to exclude individuals who tested positive for COVID- 19 but were not hospitalized must be <u>symptom-free at least 7</u> <u>days</u>	With the exponential increase in COVID-19 cases related to the omicron variant, revised to allow continued study recruitment
		10.1.5	Updates to listing of Adult Clinical Sites and Preclinical Animal Study Sites	To reflect changes in clinical and pre-clinical sites

11 REFERENCES

- [1] Klentrou P. Influence of Exercise and Training During Critical Stages of Bone Growth and Development. Pediatr Exerc Sci 2016. doi:10.1123/pes.2015-0265.
- [2] Shay CM, Ning H, Daniels SR, Rooks CR, Gidding SS, Lloyd-Jones DM. Status of cardiovascular health in US adolescents: prevalence estimates from the National Health and Nutrition Examination Surveys (NHANES) 2005-2010. Circulation 2013;127:1369–76. doi:10.1161/CIRCULATIONAHA.113.001559.
- [3] Gates PE, Strain WD, Shore AC. Human endothelial function and microvascular ageing. Exp Physiol 2009;94:311–6. doi:10.1113/expphysiol.2008.043349.
- [4] Mielgo-Ayuso J, Aparicio-Ugarriza R, Castillo A, Ruiz E, Ávila JM, Aranceta-Batrina J, et al. Physical Activity Patterns of the Spanish Population Are Mostly Determined by Sex and Age: Findings in the ANIBES Study. PLoS One 2016;11:e0149969. doi:10.1371/journal.pone.0149969.
- [5] Herman KM, Hopman WM, Sabiston CM. Physical activity, screen time and self-rated health and mental health in Canadian adolescents. Prev Med (Baltim) 2015;73:112–6. doi:10.1016/j.ypmed.2015.01.030.
- [6] Maher CA, Toohey M, Ferguson M. Physical activity predicts quality of life and happiness in children and adolescents with cerebral palsy. Disabil Rehabil 2016;38:865–9. doi:10.3109/09638288.2015.1066450.
- [7] Mutlu EK, Mutlu C, Taskiran H, Ozgen IT. Association of physical activity level with depression, anxiety, and quality of life in children with type 1 diabetes mellitus. J Pediatr Endocrinol Metab 2015;28:1273–8. doi:10.1515/jpem-2015-0082.
- [8] Pacheco DRR, Silva MJB, Alexandrino AMS, Torres RMT. Exercise-related quality of life in subjects with asthma: a systematic review. J Asthma 2012;49:487–95. doi:10.3109/02770903.2012.680636.
- [9] Kohl HW, Craig CL, Lambert EV, Inoue S, Alkandari JR, Leetongin G, et al. The pandemic of physical inactivity: global action for public health. Lancet 2012;380:294–305. doi:10.1016/S0140-6736(12)60898-8.
- [10] Nishi A, Milner DA, Giovannucci EL, Nishihara R, Tan AS, Kawachi I, et al. Integration of molecular pathology, epidemiology and social science for global precision medicine. Expert Rev Mol Diagn 2016;16:11–23. doi:10.1586/14737159.2016.1115346.
- [11] Briggs AM, Cross MJ, Hoy DG, Sànchez-Riera L, Blyth FM, Woolf AD, et al. Musculoskeletal Health Conditions Represent a Global Threat to Healthy Aging: A Report for the 2015 World Health Organization World Report on Ageing and Health. Gerontologist 2016;56 Suppl 2:S243-55. doi:10.1093/geront/gnw002.
- [12] Chu P, Pandya A, Salomon JA, Goldie SJ, Hunink MGM. Comparative Effectiveness of Personalized Lifestyle Management Strategies for Cardiovascular Disease Risk Reduction. J Am Heart Assoc 2016;4:e002737. doi:10.1161/JAHA.115.002737.
- [13] Sata F. Developmental Origins of Health and Disease (DOHaD) and Epidemiology. Nihon Eiseigaku Zasshi 2016;71:41–6. doi:10.1265/jjh.71.41.
- [14] Hardy R, Lawlor DA, Kuh D. A life course approach to cardiovascular aging. Future Cardiol 2015;11:101–13. doi:10.2217/fca.14.67.
- [15] Blaisdell CJ, Weinmann GG. NHLBI viewpoint: Lung health and disease prevention research starting in childhood. Pediatr Pulmonol 2015;50:604–6. doi:10.1002/ppul.23198.
- [16] Barker DJP, Osmond C, Forsén TJ, Thornburg KL, Kajantie E, Eriksson JG. Foetal and childhood growth and asthma in adult life. Acta Paediatr 2013;102:732–8. doi:10.1111/apa.12257.
- [17] Idler N, Teuner CM, Hunger M, Holle R, Ortlieb S, Schulz H, et al. The association between physical activity and healthcare costs in children--results from the GINIplus and LISAplus cohort studies. BMC Public Health 2015;15:437. doi:10.1186/s12889-015-1721-6.
- [18] Stary HC. Lipid and macrophage accumulations in arteries of children and the development of atherosclerosis. AmJClinNutr 2000;72:1297S-1306S.
- [19] Cayres SU, de Lira FS, Machado-Rodrigues AM, Freitas Junior IF, Barbosa MF, Fernandes RA. The mediating role of physical inactivity on the relationship between inflammation and artery thickness in prepubertal adolescents. JPediatr 2015;166:924–9.
- [20] Akber A, Portale AA, Johansen KL. Pedometer-assessed physical activity in children and young adults with CKD. Clin J Am Soc Nephrol 2012;7:720–6. doi:10.2215/CJN.06330611.
- [21] Lunt D, Briffa T, Briffa NK, Ramsay J. Physical activity levels of adolescents with congenital heart disease. Aust J Physiother 2003;49:43–50.

- [22] Houghton KM, Tucker LB, Potts JE, McKenzie DC. Fitness, fatigue, disease activity, and quality of life in pediatric lupus. Arthritis Rheum 2008;59:537–45. doi:10.1002/art.23534.
- [23] Hoffman MC, Mulrooney DA, Steinberger J, Lee J, Baker KS, Ness KK. Deficits in physical function among young childhood cancer survivors. J Clin Oncol 2013;31:2799–805. doi:10.1200/JCO.2012.47.8081.
- [24] Troutman WB, Barstow TJ, Galindo AJ, Cooper DM. Abnormal dynamic cardiorespiratory responses to exercise in pediatric patients after Fontan procedure. J Am Coll Cardiol 1998;31:668–73.
- [25] Moser C, Tirakitsoontorn P, Nussbaum E, Newcomb R, Cooper DM. Muscle Size and Cardiorespiratory Response to Exercise in Cystic Fibrosis. AmJRespirCrit Care Med 2000;In Press:1823–7. doi:10.1164/ajrccm.162.5.2003057.
- [26] Jiménez-Pavón D, Kelly J, Reilly JJ. Associations between objectively measured habitual physical activity and adiposity in children and adolescents: Systematic review. Int J Pediatr Obes 2010;5:3–18. doi:10.3109/17477160903067601.
- [27] Booth FW, Roberts CK, Laye MJ. Lack of exercise is a major cause of chronic diseases. ComprPhysiol 2012;2:1143–211.
- [28] Magnussen CG, Smith KJ, Juonala M. When to prevent cardiovascular disease? As early as possible: lessons from prospective cohorts beginning in childhood. CurrOpinCardiol 2013;28:561–8. doi:10.1097/HCO.0b013e32836428f4.
- [29] Cunningham SA, Kramer MR, Narayan KMV. Incidence of childhood obesity in the United States. NEnglJMed 2014;370:403–11. doi:10.1056/NEJMoa1309753.
- [30] Chung A, Backholer K, Wong E, Palermo C, Keating C, Peeters A. Trends in child and adolescent obesity prevalence in economically advanced countries according to socioeconomic position: a systematic review. Obes Rev 2015;17:276–95. doi:10.1111/obr.12360.
- [31] Lobstein T, Jackson-Leach R, Moodie ML, Hall KD, Gortmaker SL, Swinburn BA, et al. Child and adolescent obesity: part of a bigger picture. Lancet 2015;385:2510–20. doi:10.1016/S0140-6736(14)61746-3.
- [32] Radom-Aizik S, Zaldivar FP, Haddad F, Cooper DMM. Impact of brief exercise on peripheral blood NK cell gene and microRNA expression in young adults. JApplPhysiol 2013;114:628–36. doi:10.1152/japplphysiol.01341.2012.
- [33] Radom-Aizik S, Zaldivar Jr. F, Leu S-YY, Adams GRR, Oliver S, Cooper DMM. Effects of exercise on microRNA expression in young males peripheral blood mononuclear cells. Clin TranslSci 2012;5:32–8. doi:10.1111/j.1752-8062.2011.00384.x.
- [34] Radom-Aizik S, Zaldivar Jr. F, Leu S-YY, Cooper DMM. Brief bout of exercise alters gene expression in peripheral blood mononuclear cells of early- and late-pubertal males. Pediatr Res 2009;65:447–52. doi:10.1203/PDR.0b013e3181993473.
- [35] Radom-Aizik S, Zaldivar Jr. F, Leu S-YY, Cooper DMM. A brief bout of exercise alters gene expression and distinct gene pathways in peripheral blood mononuclear cells of early- and late-pubertal females. J Appl Physiol 2009;107:168–75. doi:10.1152/japplphysiol.00121.2009.
- [36] Radom-Aizik S, Zaldivar FP, Haddad F, Cooper DM. Impact of brief exercise on circulating monocyte gene and microRNA expression: Implications for atherosclerotic vascular disease. Brain Behav Immun 2014;39. doi:10.1016/j.bbi.2014.01.003.
- [37] Radom-Aizik S, Leu S-Y, Cooper DM, Zaldivar Jr. F. Serum from exercising humans suppresses t-cell cytokine production. Cytokine 2007;40. doi:10.1016/j.cyto.2007.08.008.
- [38] Radom-Aizik S, Zaldivar Jr. F, Oliver S, Galassetti P, Cooper DMM. Evidence for microRNA involvement in exercise-associated neutrophil gene expression changes. J ApplPhysiol 2010;109:252–61. doi:10.1152/japplphysiol.01291.2009.
- [39] Gjevestad GO, Holven KB, Ulven SM. Effects of Exercise on Gene Expression of Inflammatory Markers in Human Peripheral Blood Cells: A Systematic Review. Curr Cardiovasc Risk Rep 9:34. doi:10.1007/s12170-015-0463-4.
- [40] Meckel Y, Nemet D, Bar-Sela S, Radom-Aizik S, Cooper DM, Sagiv M, et al. Hormonal and inflammatory responses to different types of sprint interval training. J Strength Cond Res 2011;25. doi:10.1519/JSC.0b013e3181dc4571.
- [41] Abbasi A, de Paula Vieira R, Bischof F, Walter M, Movassaghi M, Berchtold NC, et al. Sex-specific variation in signaling pathways and gene expression patterns in human leukocytes in response to endotoxin and exercise. J Neuroinflammation 2016;13:289. doi:10.1186/s12974-016-0758-5.

Molecular Transducers of Physical Activity Consortium–Pediatric Clinical Center Protocol

- [42] Liu D, Wang R, Grant AR, Zhang J, Gordon PM, Wei Y, et al. Immune adaptation to chronic intense exercise training: new microarray evidence. BMC Genomics 2017;18:29. doi:10.1186/s12864-016-3388-5.
- [43] Lu KD, Radom-Aizik S, Haddad F, Zaldivar F, Kraft M, Cooper DM. Glucocorticoid Receptor Expression on Circulating Leukocytes Differs between Healthy Male and Female Adults. J Clin Transl Sci n.d.
- [44] Piercy KL, Troiano RP. Physical Activity Guidelines for Americans From the US Department of Health and Human Services. Circ Cardiovasc Qual Outcomes 2018;11:e005263. doi:10.1161/CIRCOUTCOMES.118.005263.
- [45] Irwin DE, Gross HE, Stucky BD, Thissen D, DeWitt E, Lai J, et al. Development of six PROMIS pediatrics proxy-report item banks. Health Qual Life Outcomes 2012;10:22. doi:10.1186/1477-7525-10-22.
- [46] PROMIS: Health Measures Transforming How Health is Measured n.d. http://www.healthmeasures.net/index.php?option=com_content&view=category&layout=blog&id=147&It emid=806 (accessed May 8, 2019).
- [47] Hunsberger M, O'Malley J, Block T, Norris JC. Relative validation of Block Kids Food Screener for dietary assessment in children and adolescents. Matern Child Nutr 2015;11:260–70. doi:10.1111/j.1740-8709.2012.00446.x.
- [48] ASA24[®] Dietary Assessment Tool | EGRP/DCCPS/NCI/NIH n.d. https://epi.grants.cancer.gov/asa24/ (accessed May 8, 2019).
- [49] Odunaiya NA, Grimmer K, Louw QA. High prevalence and clustering of modifiable CVD risk factors among rural adolescents in southwest Nigeria: implication for grass root prevention. BMC Public Health 2015;15:661. doi:10.1186/s12889-015-2028-3.
- [50] Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analysis using G*Power 3.1: Tests for correlation and regression analyses. Behav Res Methods 2009;41:1149–60.
- [51] Liang KY. Longitudinal data analysis of continuous and discrete responses for pre-post design. Indian J Stat 2000;62:134–48.
- [52] Lu K. On Efficiency of Constrained Longitudinal Data Analysis Versus Longitudinal Analysis of Covariance. Biometrics 2010;66:891–6. doi:10.1111/j.1541-0420.2009.01332.x.
- [53] Shpitser I. Counterfactual graphical models for longitudinal mediation analysis with unobserved confounding. Cogn Sci 2013;37:1011–35. doi:10.1111/cogs.12058.
- [54] Lynch KG. Casual mediation analyses for randomized trials. Heal Serv Outcomes Res Methodol 2008;8:57– 76.
- [55] Gamazon ER, Wheeler HE, Shah KP, Mozaffari S V, Aquino-Michaels K, Carroll RJ, et al. A gene-based association method for mapping traits using reference transcriptome data. Nat Genet 2015;47:1091–8. doi:10.1038/ng.3367.
- [56] Price ND, Magis AT, Earls JC, Glusman G, Levy R, Lausted C, et al. A wellness study of 108 individuals using personal, dense, dynamic data clouds. Nat Biotechnol 2017;35:747–56. doi:10.1038/nbt.3870.
- [57] Argelaguet R, Velten B, Arnol D, Dietrich S, Zenz T, Marioni JC, et al. Multi-Omics Factor Analysis—a framework for unsupervised integration of multi-omics data sets. Mol Syst Biol 2018;14:e8124. doi:10.15252/MSB.20178124.
- [58] Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc B 1995;57:289–300.
- [59] Reiner A, Yekutieli D, Benjamini Y. Identifying differentially expressed genes using false discovery rate controlling procedures. Bioinformatics 2003;19:368–75.
- [60] Clancy CE, An G, Cannon WR, Liu Y, May EE, Ortoleva P, et al. Multiscale Modeling in the Clinic: Drug Design and Development. Ann Biomed Eng 2016;44:2591–610. doi:10.1007/s10439-016-1563-0.
- [61] Su J, Zhang L, Zhang W, Choi DS, Wen J, Jiang B, 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:e85059. doi:10.1371/journal.pone.0085059.
- [62] Tao Peng T, Huiming Peng H, Dong Soon Choi DS, Jing Su J, Chung-Che Chang C-CJ, Xiaobo Zhou X. Modeling Cell–Cell Interactions in Regulating Multiple Myeloma Initiating Cell Fate. IEEE J Biomed Heal Informatics 2014;18:484–91. doi:10.1109/JBHI.2013.2281774.
- [63] Sun X, Su J, Bao J, Peng T, Zhang L, Zhang Y, et al. Cytokine combination therapy prediction for bone remodeling in tissue engineering based on the intracellular signaling pathway. Biomaterials 2012;33:8265– 76. doi:10.1016/j.biomaterials.2012.07.041.

- [64] Uppal K, Walker DI, Liu K, Li S, Go Y-M, Jones DP. Computational Metabolomics: A Framework for the Million Metabolome. Chem Res Toxicol 2016;29:1956–75. doi:10.1021/acs.chemrestox.6b00179.
- [65] Ernst J, Vainas O, Harbison CT, Simon I, Bar-Joseph Z. Reconstructing dynamic regulatory maps. Mol Syst Biol 2007;3:74. doi:10.1038/msb4100115.
- [66] Schulz MH, Devanny WE, Gitter A, Zhong S, Ernst J, Bar-Joseph Z. DREM 2.0: Improved reconstruction of dynamic regulatory networks from time-series expression data. BMC Syst Biol 2012;6:104. doi:10.1186/1752-0509-6-104.
- [67] Schulz MH, Pandit K V., Lino Cardenas CL, Ambalavanan N, Kaminski N, Bar-Joseph Z. Reconstructing dynamic microRNA-regulated interaction networks. Proc Natl Acad Sci 2013;110:15686–91. doi:10.1073/pnas.1303236110.
- [68] Cowen L, Ideker T, Raphael BJ, Sharan R. Network propagation: a universal amplifier of genetic associations. Nat Rev Genet 2017;18:551–62. doi:10.1038/nrg.2017.38.
- [69] Amar D, Shamir R. Constructing module maps for integrated analysis of heterogeneous biological networks. Nucleic Acids Res 2014;42:4208–19. doi:10.1093/nar/gku102.
- [70] Malladi VS, Erickson DT, Podduturi NR, Rowe LD, Chan ET, Davidson JM, et al. Ontology application and use at the ENCODE DCC. Database 2015;2015. doi:10.1093/database/bav010.
- [71] Hong EL, Sloan CA, Chan ET, Davidson JM, Malladi VS, Strattan JS, et al. Principles of metadata organization at the ENCODE data coordination center. Database 2016;2016:baw001. doi:10.1093/database/baw001.
- [72] Hitz BC, Rowe LD, Podduturi NR, Glick DI, Baymuradov UK, Malladi VS, et al. SnoVault and encodeD: A novel object-based storage system and applications to ENCODE metadata. PLoS One 2017;12:e0175310. doi:10.1371/journal.pone.0175310.
- [73] Chen L, Liu P, Evans TC, Ettwiller LM. DNA damage is a pervasive cause of sequencing errors, directly confounding variant identification. Science (80-) 2017;355:752–6. doi:10.1126/science.aai8690.
- [74] De Domenico M, Solé-Ribalta A, Omodei E, Gómez S, Arenas A. Ranking in interconnected multilayer networks reveals versatile nodes. Nat Commun 2015;6:6868. doi:10.1038/ncomms7868.
- [75] Broman KW, Keller MP, Broman AT, Kendziorski C, Yandell BS, Sen S, et al. Identification and Correction of Sample Mix-Ups in Expression Genetic Data: A Case Study. G3: Genes|Genomes|Genetics 2015;5:2177–86. doi:10.1534/g3.115.019778.
- [76] Westra H-J, Jansen RC, Fehrmann RSN, te Meerman GJ, van Heel D, Wijmenga C, et al. MixupMapper: correcting sample mix-ups in genome-wide datasets increases power to detect small genetic effects. Bioinformatics 2011;27:2104–11. doi:10.1093/bioinformatics/btr323.