Effectiveness of the On the Move Group Exercise Program to Improve Mobility in Community-Dwelling Older Adults*

Protocol Number*: STUDY21070206

National Clinical Trial (NCT) Identified Number: NCT05126355

Principal Investigator*: Dr. Jennifer S. Brach

Sponsor: University of Pittsburgh

"Sponsor" indicates an institution, foundation, or individual who takes responsibility for and initiates a clinical investigation; often times this is the university with which the Principal Investigator is affiliated.

Grant Title: Effectiveness of the On the Move Group Exercise Program to Improve Mobility in Community-Dwelling Older Adults

Grant Number*: R01 AG071520

Funded by: NIA

Version Number: 1.0.0

22 October 2021

CONFIDENTIALITY STATEMENT

This document is confidential communication. Acceptance of this document constitutes agreement by the recipient that no unpublished information contained herein will be published or disclosed without prior approval of the Principal Investigator or other participating study leadership and as consistent with the NIH terms of award.

Table of Contents

STATEMENT OF COMPLIANCE 1
INVESTIGATOR'S SIGNATURE 2
1 PROTOCOL SUMMARY 3
1.1 Synopsis 3
1.2 Schema 4
1.3 Schedule of Activities 5
2 INTRODUCTION 7
2.1 Study Rationale 7
2.2 Background 9
2.3 Risk/Benefit Assessment 11
2.3.1 Known Potential Risks 11
2.3.2 Known Potential Benefits 12
2.3.3 Assessment of Potential Risks and Benefits 12
3 OBJECTIVES AND ENDPOINTS 13
4 STUDY DESIGN 14
4.1 Overall Design 14
4.2 Scientific Rationale for Study Design 15
4.3 Justification for Intervention 15
4.4 End-of-Study Definition 16
5 STUDY POPULATION 16
5.1 Inclusion Criteria 16
5.2 Exclusion Criteria 17
5.3 Lifestyle Considerations 17
5.4 Screen Failures 17
5.5 Strategies for Recruitment and Retention 17
6 STUDY INTERVENTION(S) OR EXPERIMENTAL MANIPULATION(S) 18
6.1 Study Intervention(s) or Experimental Manipulation(s) Administration 18
6.1.1 Study Intervention or Experimental Manipulation Description 18
6.1.2 Administration and/or Dosing 19
6.2 Fidelity 19
6.2.1 Interventionist Training and Tracking 19
6.3 Measures to Minimize Bias: Randomization and Blinding 19
6.4 Study Intervention/Experimental Manipulation Adherence 21
6.5 Concomitant Therapy 21
6.5.1 Rescue Therapy 21
7 STUDY INTERVENTION/EXPERIMENTAL MANIPULATION DISCONTINUATION AND PARTICIPANT DISCONTINUATION/WITHDRAWAL 21
7.1 Discontinuation of Study Intervention/Experimental Manipulation 21
7.2 Participant Discontinuation/Withdrawal from the Study 21
7.3 Lost to Follow-Up 22
8 STUDY ASSESSMENTS AND PROCEDURES 22
8.1 Endpoint and Other Non-Safety Assessments 22
8.2 Safety Assessments 25
8.3 Adverse Events and Serious Adverse Events 25
8.3.1 Definition of Adverse Events 25

8.3.2 Definition of Serious Adverse Events ... 25
8.3.3 Classification of an Adverse Event ... 26
8.3.4 Time Period and Frequency for Event Assessment and Follow-Up ... 27
8.3.5 Adverse Event Reporting ... 28
8.3.6 Serious Adverse Event Reporting ... 29
8.3.7 Reporting Events to Participants ... 29
8.3.8 Events of Special Interest ... 29
8.3.9 Reporting of Pregnancy ... 29
8.4 Unanticipated Problems ... 29
8.4.1 Definition of Unanticipated Problems ... 29
8.4.2 Unanticipated Problems Reporting ... 30
8.4.3 Reporting Unanticipated Problems to Participants ... 30
9 STATISTICAL CONSIDERATIONS ... 30
9.1 Statistical Hypotheses ... 30
9.2 Sample Size Determination ... 30
9.3 Populations for Analyses ... 31
9.4 Statistical Analyses ... 31
9.4.1 General Approach ... 31
9.4.2 Analysis of the Primary Endpoint(s) ... 32
9.4.3 Analysis of the Secondary Endpoint(s) ... 32
9.4.4 Safety Analyses ... 34
9.4.5 Baseline Descriptive Statistics ... 34
9.4.6 Planned Interim Analyses ... 34
9.4.7 Sub-Group Analyses ... 34
9.4.8 Tabulation of Individual Participant Data ... 34
9.4.9 Exploratory Analyses ... 35
10 SUPPORTING DOCUMENTATION AND OPERATIONAL CONSIDERATIONS ... 35
10.1 Regulatory, Ethical, and Study Oversight Considerations ... 35
10.1.1 Informed Consent Process ... 35
10.1.2 Study Discontinuation and Closure ... 36
10.1.3 Confidentiality and Privacy ... 36
10.1.4 Future Use of Stored Specimens and Data ... 36
10.1.5 Key Roles and Study Governance ... 37
10.1.6 Safety Oversight ... 37
10.1.7 Clinical Monitoring ... 38
10.1.8 Quality Assurance and Quality Control ... 38
10.1.9 Data Handling and Record Keeping ... 39
10.1.10 Protocol Deviations ... 39
10.1.11 Publication and Data Sharing Policy ... 40
10.1.12 Conflict of Interest Policy ... 40
10.2 Additional Considerations ... 41
10.3 Abbreviations and Special Terms ... 41
10.4 Protocol Amendment History ... 43
11 REFERENCES ... 44

STATEMENT OF COMPLIANCE

The trial will be carried out in accordance with International Council on Harmonisation 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.

The protocol, informed consent form(s), recruitment materials, and all participant materials will be submitted to the IRB for review and approval. Approval of both the protocol and the consent form(s) must be obtained before any participant is consented. Any amendment to the protocol will require review and approval by the IRB before the changes are implemented to the study. All changes to the consent form(s) will be IRB approved; a determination will be made regarding whether a new consent needs to be obtained from participants who provided consent, using a previously approved consent form.

INVESTIGATOR'S SIGNATURE

The signature below constitutes the approval of this protocol and provides the necessary assurances that this study will be conducted according to all stipulations of the protocol, including all statements regarding confidentiality, and according to local legal and regulatory requirements and applicable US federal regulations and ICH guidelines, as described in the *Statement of Compliance* above.

Principal Investigator:

Signed:  Date: 10/18/2021

Name*: Dr. Jennifer S Brach

Title*: Professor Department of Physical Therapy and Associate Dean for Faculty Affairs and Development, School of Health and Rehabilitation Sciences

Investigator Contact Information

Affiliation*: University of Pittsburgh

Address: Bridgeside Point I, Suite 500
100 Technology Drive
Pittsburgh, PA 15219

Telephone: 412-383-6533

Email: jbrach@pitt.edu

1 PROTOCOL SUMMARY

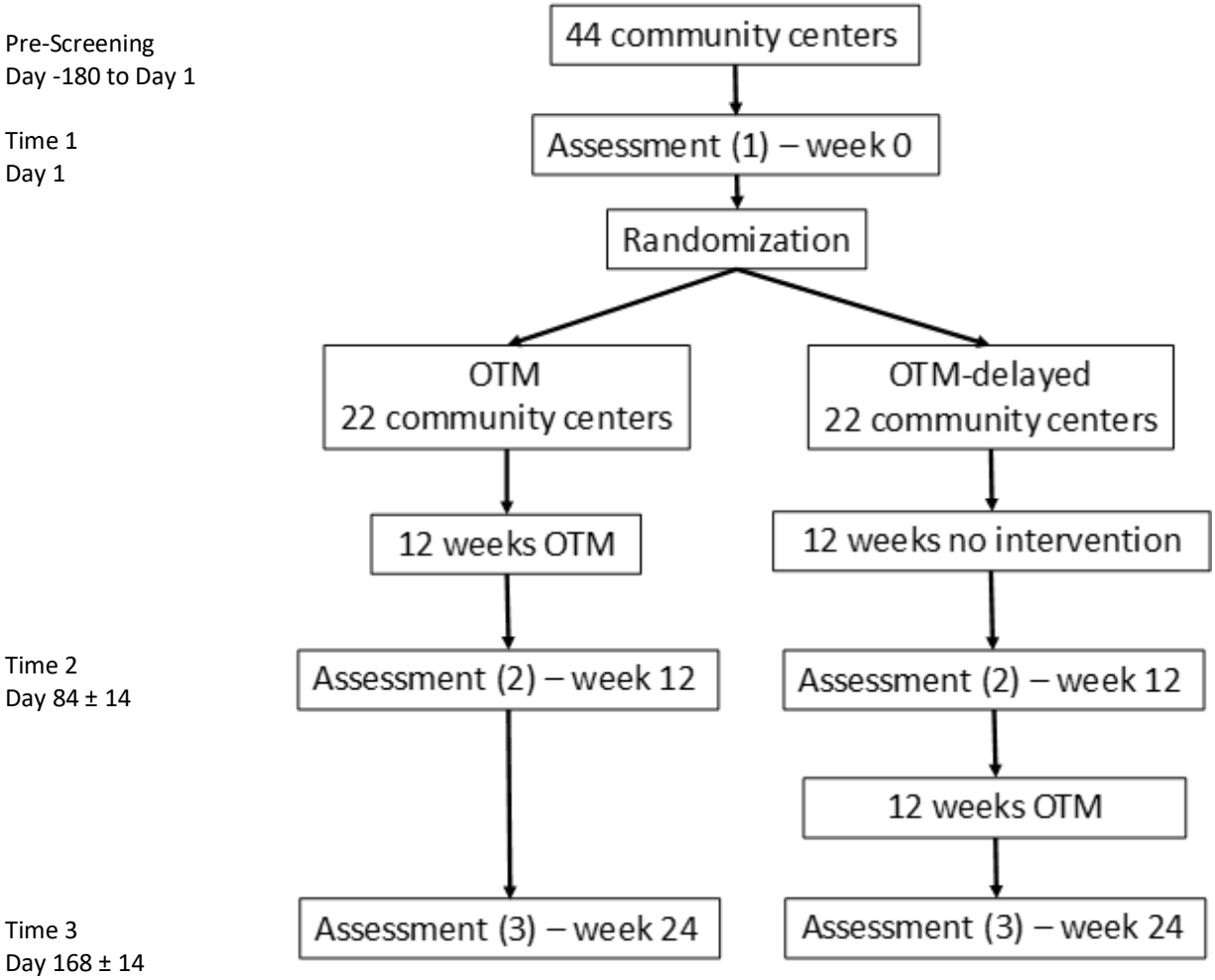
1.1 SYNOPSIS

Title:	Effectiveness of the On the Move group exercise program to improve mobility in community-dwelling older adults
Grant Number:	R01 AG071520
Study Description:	This is a Hybrid I trial, blending effectiveness and implementation research aims to: test the effectiveness of OTM, evaluate intervention fidelity, and determine the extent intervention fidelity moderates intervention effectiveness. The study design is a cluster randomized, single-blind, two-arm intervention trial to compare the effectiveness of OTM to a delayed intervention control in 502 community-dwelling older adults, across 44 senior community centers. We hypothesize Compared to a randomized control group followed for 12 weeks, the OTM group will have greater gains in gait speed (primary outcome) and self-reported mobility at 12 and 24 weeks of follow-up.
Objectives*:	Primary Objective: Establish the effectiveness of OTM on improvements in walking ability and post-intervention persistence of benefits. Secondary Objectives: Assess intervention fidelity (measures of adherence and competence) and identify the impact of organizational, instructor, and participant level factors on intervention fidelity. Tertiary Objectives: Determine the extent to which intervention fidelity, measured by adherence and competence, moderates the effectiveness of OTM.
Endpoints*:	Primary Endpoint: The primary endpoint is mobility as measured by gait speed and the PROMIS-mobility measure. Secondary Endpoints: Fidelity will be assessed by measuring adherence to content (inclusion of essential components of OTM), adherence to dose (frequency and duration of exercise classes), and competence of delivery (the quality of program delivery by instructor). *Acceptability of Intervention Measure (AIM)*.[1] The AIM will be used to measure acceptability of OTM among participants, instructors, and senior center managers. *Intervention Appropriateness Measure (IAM)*.[1] We will use the IAM to measure the perceived fit, relevance, or compatibility of the intervention for a given participant, instructor, and senior center manager. *Feasibility of Intervention Measure (FIM)*.[1] We will use the FIM to measure the extent to which OTM can be successfully used or carried out within a given agency or setting (i.e., feasibility). Tertiary Endpoints: We will explore whether there are any thresholds that elicit a large difference in mobility gains, which may indicate a level of fidelity above which effectiveness encounters diminishing returns (e.g., <80% vs ≥80% attendance).
Study Population:	The target population includes 502 community-dwelling older adults in Western Pennsylvania who are 60 years of age or older, who attend senior community centers, and can participate in group exercise programs.

Phase* or Stage:	Stage 4 effectiveness study; Hybrid I trial; based on NIH Stage Model for Behavioral Intervention Development.[2]
Description of Sites/Facilities Enrolling Participants:	We will include 44 senior community centers from six counties located in Western Pennsylvania (Allegheny, Erie, Fayette, Greene, Mercer, Washington, and Westmoreland). In some counties where the Area Agency on Aging (AAA) plays a significant role in overseeing senior community centers (i.e., Washington, Fayette, Greene, and Westmoreland Counties), we will work with the Agency administrator to select centers which best represent the county and facilitate diversity. We plan to enroll 10-15 participants from each center and conduct multiple sections of the class if the enrollment is over the maximum class size of 10.
Description of Study Intervention/Experimental Manipulation:	The study design is a cluster randomized, single-blind, two-arm intervention trial to compare the effectiveness of OTM to a delayed intervention control in 502 community-dwelling older adults, across 44 senior community centers. OTM group exercise classes will be taught by certified OTM instructors in the community twice weekly for 12 weeks. Centers randomized to the control arm will receive no intervention for the first 12 weeks and then will receive 12-weeks of OTM classes. Participants are assessed at baseline, 12 weeks, and 24 weeks by individuals masked to the treatment group.
Study Duration*:	60 months
Participant Duration:	Six months

1.2 SCHEMA

Study Design: *Cluster randomized control trial*



***Refer to Section 1.3, Schedule of Activities**

1.3 SCHEDULE OF ACTIVITIES

Schedule of Activities – Senior Centers

	Pre-screening (Pre-consent)	Time 1 Day 1	Time 2 Day 84 ±14	Time 3 Day 168 ±14	Unscheduled Visit
Study Set-up	X				
Senior Center Recruitment	X				
Contact (e.g., telephone and/or email)	X				

	Pre-screening (Pre-consent)	Time 1 Day 1	Time 2 Day 84 ±14	Time 3 Day 168 ±14	Unscheduled Visit
Consent	X				
Demographics	X				
Randomization	X				
Outcome Evaluation: **Senior Center**					
Acceptability of Intervention Measure (AIM)*	X	X	X	X	
Intervention Appropriateness Measure (IAM)*	X	X	X	X	
Feasibility of Intervention Measure (FIM)*	X				
Semi-structured interviews and focus groups*			X	X	
Consolidated Framework for Implementation Research (CFIR) – Inner Setting	X				
Adverse Events Reporting		X	X	X	X

Schedule of Activities – OTM Instructors

	Pre-screening (Pre-consent)	Time 1 Day 1	Time 2 Day 84 ±14	Time 3 Day 168 ±14	Unscheduled Visit
Study Set-up	X				
Instructor Recruitment	X				
Contact (e.g., telephone and/or email) and schedule trainings	X				
Consent	X				
Demographics	X				
Randomization	X				
Outcome Evaluation: **Instructor**					
Acceptability of Intervention Measure (AIM)*	X	X	X	X	
Intervention Appropriateness Measure (IAM)*	X	X	X	X	
Feasibility of Intervention Measure (FIM)*		X			
Semi-structured interviews and focus groups*			X	X	
Consolidated Framework for Implementation Research (CFIR) – Inner Setting	X				
Frequency count of instructor use – technical support		X	X	X	X
Adverse Events Reporting		X	X	X	X

Schedule of Activities - Participant

	Pre-screening (Pre-consent)	Time 1 Day 1	Time 2 Day 84 ±14	Time 3 Day 168 ±14	Unscheduled Visit
Study Set-up	X				
Presentation at Senior Center	X				
Participant Recruitment	X				
Contact (e.g., telephone and/or email) and schedule participants	X				
Informed Consent	X				
Screening Exam	X				
Demographics	X				
Randomization	X				
Outcome Evaluation: **Participant**					
Gait Speed		X	X	X	
Patient-Reported Outcomes Measurement Information System (PROMIS)		X	X	X	
Participant Satisfaction Survey*			X	X	
Acceptability of Intervention Measure (AIM)*	X	X	X	X	
Intervention Appropriateness Measure (IAM)*	X	X	X	X	
Semi-structured interviews and focus groups*			X	X	
Adverse Events Reporting		X	X	X	X

*These outcomes will be completed at either the 12-week assessment (intervention group) or the 24-week assessment (delayed intervention group).

2 INTRODUCTION

2.1 STUDY RATIONALE

OTM delivered by research staff is efficacious.[3] Establishing the effectiveness of OTM delivered by community instructors is necessary before wide-scale implementation, as interventions tested in highly controlled trials often do not yield the same outcomes when applied in real-life contexts (i.e., effectiveness). When implementing evidence-based interventions in real world practice, it is important to understand potential factors that may lead to smaller effects than what was reported in efficacy trials. Understanding the impact of intervention fidelity on effectiveness and the impact of potential organizational-, instructor- and participant-level factors (e.g., facilitation strategies, responsiveness of participants, instructors and management, intervention complexity, recruitment, and context) on intervention fidelity will help to inform future implementation efforts. In the preliminary data we demonstrate 1) that our team has experience with the intervention (OTM) and setting (senior community centers), 2) the feasibility of the proposed research design and methods, 3) the ability to train community OTM instructors, and 4) the engagement of stakeholders throughout the research process.

On the Move group exercise is superior to usual care. In our prior PCORI-funded study (CE-1304-6301), we conducted a cluster randomized, single-blind intervention trial. We compared the effectiveness of the OTM group exercise program to a seated strength, endurance, and flexibility program (i.e., 'Usual Care') taught by research staff instructors on walking ability of older adults residing in independent living facilities, senior apartment buildings, or attending senior community centers.[3,4] Thirty-two sites were randomized to OTM (16 sites; 153 participants) or Usual Care (16 sites; 146 participants). Participants (mean age 80.0±8.1 years) had multiple chronic conditions. Exercise classes were held 2 times/week for 12 weeks. At the end of 12 weeks, the OTM group had greater improvements than the Usual Care group in walking ability: gait speed (0.05±0.13 vs -0.01±0.11 m/s; adjusted difference=0.05±0.02; p=0.0008). The between-group difference was adjusted for baseline outcome values and represents a meaningful difference. A meaningful improvement in gait speed (0.05m/s)[5] was achieved by 64 of 152 (42.1%) in the OTM group and 36 of 146 (24.7%) in the Usual Care group, resulting in a number needed to treat of 5.7. Effects were greatest for participants in senior community centers.[6] This prior work demonstrates the efficacy of OTM and our ability to carry-out a large cluster randomized trial in senior community settings (i.e., feasibility of proposed design).

Fidelity can be measured and quantified. We monitored intervention fidelity in both our NIA and PCORI funded trials. In our NIA funded trial (R01 AG45252), we monitored the fidelity of an individual timing and coordination intervention.[7] A study team member reviewed the intervention data sheets biweekly, observed interventions being delivered quarterly and conducted annual meetings with the physical therapist interventionists, with the goal of ensuring adherence to the specific protocols. Treatment fidelity data provided information on the adherence of the physical therapists to the interventions, and the adherence of the participants to the prescribed activities. Monitoring treatment fidelity provided an opportunity to detect and address potential study problems early, optimizing potential for benefit and ensuring accurate interpretation of findings. In the PCORI trial, intervention fidelity was assessed via in-person site visits or review of video-tape recordings of the exercise class using a standardized intervention fidelity checklist.[6] Our experiences in these two trials have informed the development of our tools to measure fidelity.

We are ready to train community instructors to deliver OTM. Prior to receiving the designation as an Evidence-Based Program, we created the resources necessary for delivering OTM on a wider scale. We created the OTM website (https://www.onthemove.pitt.edu/), which provides information about the program to interested participants, instructors, and senior community sites and serves as a communication hub for OTM instructors and support staff. We developed the instructor certification process, which includes a one-day course, an instructor manual that describes the program in detail, a practical examination, and a written examination. To date, we have held four instructor training courses, three in Pennsylvania and one in Colorado, certifying over 40 OTM instructors.

We value stakeholder engagement. We involved stakeholders throughout the development and testing of OTM.[4] We continued to involve stakeholders in the development of the current proposal. We have developed critical collaborations with the Area Agencies on Aging and senior community centers (see letters of support), demonstrating that we have the commitment of the senior community centers needed to conduct this study. The centers express a strong interest in offering evidence-based programs to their members but note great difficulty in identifying programs and finding the certified instructors needed to lead the programs. Two stakeholders are key members of the research team: Ms. Hallie Zeleznik, PT, NCS and Mr. John Zanardelli, MPH, FACHE. Ms. Zeleznik is the Assistant Director of the Outpatient Neuro/Vestibular Program at the Centers for Rehabilitation Services (CRS). In the preparation phase, we worked with Ms. Zeleznik to identify a source of potential OTM instructors. With Ms. Zeleznik's assistance we conducted an environmental scan of ~300 CRS therapists and assistants to gauge interest in becoming an OTM instructor. Over 100 individuals indicated interest, and the majority indicated willingness to drive

up to 20 miles. Based on these findings and additional inquiries we have received about becoming an OTM instructor, we believe we will be successful in identifying and training 44 instructors in a timely manner. Mr. Zanardelli is an Executive Consultant and the former CEO of Asbury Heights a Senior Care Community in Pittsburgh, PA; he is well-connected to individuals who provide care and services for older adults in Western PA. In the preparation phase, Mr. Zanardelli met with many of the managers of the Area Agencies on Aging and was instrumental in soliciting their interest. Throughout the project, he will coordinate stakeholder involvement, including establishing and providing direction for the Stakeholder Advisory Board. The Stakeholder Advisory Board will meet semi-annually to provide input and guidance on recruitment, retention, interpretation, and dissemination of results.

2.2 BACKGROUND

Our team developed OTM, a group exercise program to improve walking in older adults by targeting the timing and coordination of walking.[8] The program challenges the brain to match the timing and sequences of movements with posture to improve the smoothness and efficiency of walking. To achieve this 'better walking,' OTM includes a variety of progressive stepping and walking patterns that are based on the biomechanics of gait. OTM classes meet for 50 minutes, twice a week for 12 weeks. The program includes a warm-up, progressive walking and stepping patterns, lower extremity strengthening focused on the muscles important for walking, and a flexibility/cool down component. A PCORI-funded cluster randomized trial demonstrated that the 12-week OTM program was superior to a usual care exercise program for improving walking ability (see preliminary studies section for greater detail)[3,4,9] This trial was conducted in the community, but the intervention was delivered by OTM instructors who were research staff. The next step in the development of OTM as a fully implementable intervention is to conduct an effectiveness (Stage IV) study in which the intervention is delivered in community settings by community providers.

Why is it important to improve walking in community-dwelling older adults? Walking is important for maintaining one's independence and is highly valued by older adults. Walking difficulty, which can be quantified using gait speed, is a common, costly problem in older adults.[10] Gait speed is a strong predictor of poor health outcomes, including death.[11,12] Improvements in gait speed predict better survival.[13] Also, those who develop walking difficulty have greater healthcare costs ($1,128 more/year) than those who do not have difficulty.[10] Improving gait speed and preventing walking difficulty could have a substantial impact on quality of life, morbidity, mortality, and healthcare costs.[14] In secondary, exploratory analyses of data from our PCORI trial, point estimates suggest On the Move (OTM) participation resulted in a 18% decrease in incidence rate of hospitalization compared to usual care (adjusted IRR=0.82; 95% CI=0.56-1.21).[15]

The evidence-practice gap for exercise in older adults is significant. Despite the growing evidence in well-controlled research trials (Stage III efficacy studies) that physical activity and exercise are beneficial for the health of older adults, few of these interventions have been adopted in clinical or community practice.[16-19] Most of the Centers for Disease Control and Prevention (CDC) compendium exercise interventions have only been studied in efficacy trials conducted under controlled conditions.[16] True effectiveness trials for exercise are rare but necessary to determine whether it makes sense to implement a program in real-world settings. In order to eliminate the evidence-practice gap there is a need to transition efficacy research findings into effectiveness studies in community settings.[16,17] Clinical and research programs are available to a select group of individuals. Transitioning to a community-based program delivered by community instructors will greatly increase the reach of OTM. By meeting older adults where they live (urban, rural, high/low resource areas), OTM will have a much wider impact on walking and health.

OTM is an evidence-based program. In February 2018, OTM was designated as an evidence-based program by the National Council on Aging (NCOA). The NCOA determines if the program qualifies for evidence-based designation based on the scientific evidence. This designation is important because senior community centers receive government funding to support evidence-based programming. Having obtained the evidence-based designation for OTM promotes the sustainability of the program, as centers can utilize government funding to pay for the delivery of the program. Part of the process for obtaining evidence-based status included developing a "packaged program" with a variety of supportive materials so the program can be consistently delivered with high fidelity (see preliminary studies section for greater detail).

Diminished fidelity can reduce intervention effects. Interventions tested in highly controlled trials often do not yield the same outcomes when applied in real life contexts (i.e., effectiveness). One reason for this dilution of effect is diminished intervention fidelity. Poor fidelity can reduce intervention potency, leading to faulty conclusions about effectiveness (type III error). To minimize type III error when transitioning from more controlled to less controlled environments, it is imperative to monitor and measure intervention fidelity. We use a modified version of the **Conceptual Framework for Implementation Fidelity**[20], a well-accepted model with over 460 citations in the scientific literature, to guide our investigation (**Figure 1**). The Framework depicts the vital elements of intervention fidelity and their relationship to one another. Carroll et al defines intervention fidelity as "the degree to which programs are implemented as indicated by the program developers."[20] Intervention fidelity can be measured by adherence to content and dose and competence in the quality of delivery of the intervention. Carroll proposes that many of these elements interact and may influence the association between intervention, fidelity, and outcomes. For our investigation, we will quantify intervention fidelity by measuring instructor adherence (to content and dose of OTM) and competence (quality of delivery). Adaptations or modifications to the evidence-based program occur during the implementation process to address differences between the context in which the intervention was originally designed and tested and the one into which it is ultimately implemented.[21] Some adaptations are planned/purposeful and occur with the intention to retain fidelity where others are unplanned and in reaction to a specific circumstance. While some modifications may facilitate implementation and sustainability, they can also erode treatment integrity. Therefore, is it critical to monitor both fidelity and adaptations to gain a more complete understanding of the implementation process and the levels of fidelity necessary to promote desired outcomes. We will record, define, and classify adaptations to OTM to inform

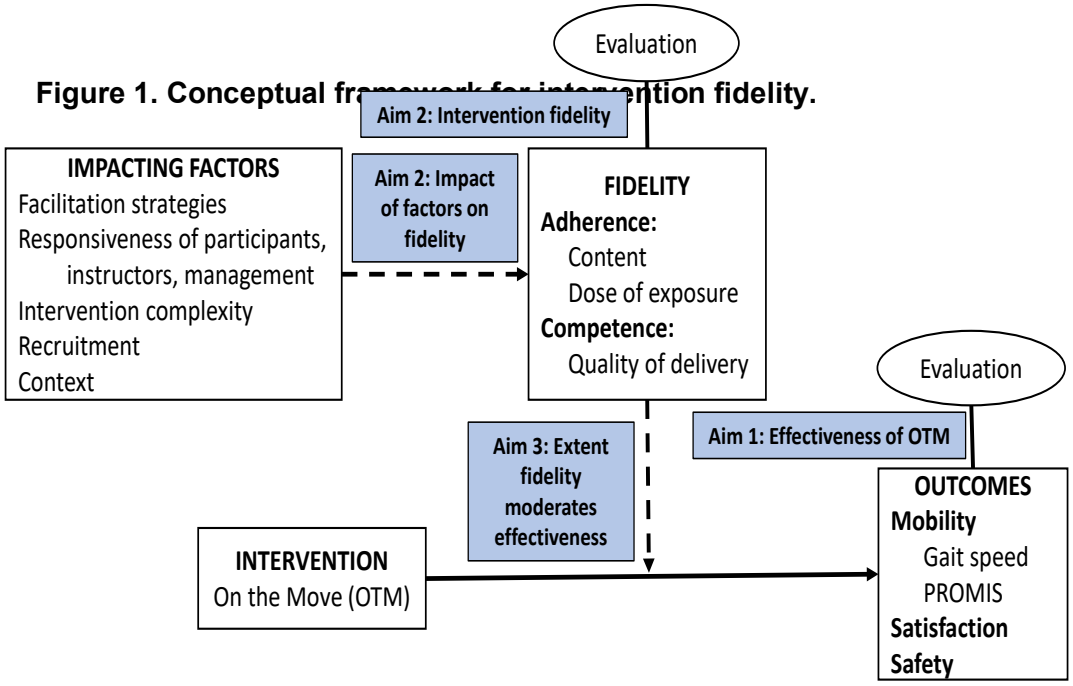


Figure 1. Conceptual framework for intervention fidelity.

future implementation efforts.

Intervention fidelity may be influenced by other factors.

Hasson proposes that the level of fidelity achieved may be influenced by factors such as facilitation strategies, responsiveness of participants, instructors and management, intervention complexity, recruitment, and context.[22] These factors can be measured at the organizational, instructor, and participant level. A description of these factors and a rationale as to why these factors may influence fidelity can be found in **Table 1**.

Table 1. Factors that may impact intervention fidelity

Factor	Rationale
Facilitation strategies. Training/support to optimize fidelity. Includes manuals, guidelines, training, monitoring and feedback.	Adequate facilitation strategies and access to these strategies (use) increase opportunities for higher and more standardized fidelity.
Responsiveness. How well participants respond to, or are engaged by, an intervention. Refers to individuals receiving or delivering (instructor) the intervention as well as upper management that supports the intervention. Includes measures of acceptability, appropriateness, feasibility, & adoption.	Participants who accept the intervention, and believe it is relevant, are more likely to be engaged and have greater intervention adherence. If the organization is not committed to the intervention, the response of the instructors is less positive (Roger's Diffusion of Innovation Theory).[23]
Intervention Complexity. The instructor's perspective of the intervention, including items as number of concepts covered and the degree which exercises change over time.	Complex interventions that are vaguely described are assumed to be more difficult to implement with high fidelity than simple interventions (Roger's "complexity").[23]
Recruitment. Reasons for nonparticipation, subgroups less likely to participate, and consistency of recruitment procedures among potential participants.	Can contribute to correct generalizability of findings.
Context. Surrounding social systems, such as structures, and cultures of organizations and groups, and historical and concurrent events.	Without adequate resources (staff or space) cannot deliver program with high fidelity. Positive experiences with previous programs more likely to "buy in" to program and deliver it with fidelity (Roger's "compatibility").[23]

We propose to investigate how these potential factors influence intervention fidelity to inform future implementation efforts (Aim 2). We will also explore the extent to which intervention fidelity moderates the effectiveness of OTM (Aim 3). Understanding the impact of fidelity on outcomes is critical when transitioning from controlled trials to real-life contexts. Unless fidelity is evaluated, it cannot be determined whether a lack of intervention effect is due to poor implementation or inadequacies inherent to the program itself (i.e., Type III error). Only by understanding and measuring whether OTM was implemented with fidelity can we gain a better understanding of how OTM works, and the extent to which outcomes can be improved.

2.3 RISK/BENEFIT ASSESSMENT

2.3.1 KNOWN POTENTIAL RISKS

As with any experimental procedure, there may be adverse events or side effects that are currently unknown, and certain of these unknown risks could be permanent, severe, or life-threatening. The possible risks of this research study are mostly due to the testing of your walking and/or your participation in the exercise program. The risks of participating in the testing and exercise program for walking are minimal. However, there are certain risks and discomforts of which you should be aware.

The potential risks are:

1. You may get tired during the testing sessions or exercise classes. If you become tired, you can stop the testing or exercise and resume when you are rested.
2. Side effects from testing or exercise, such as muscle soreness, are expected to be infrequent.
3. There is a rare risk that you may fall during testing or exercise. However, members of the research staff are trained in exercise testing and supervision of exercise programs. They will be present during all tasks to render assistance in the event of an emergency and will watch you as you walk and stop you from continuing if you appear to be having difficulty or will help steady you if you appear to be falling.
4. There is a rare risk that you may have heart problems during the testing or exercise. We will conduct a screening to determine if you have any conditions which would put you at increased risk. If you have certain medical conditions you will be unable to participate. The personnel completing the exercise sessions may not be medical personnel. It is important that you report to the research staff or exercise leader any potentially heart related symptoms that you experience such as shortness of breath, heart racing, lightheadedness, etc. From reviewing previous studies reported and our own experience with very similar exercise treatment programs, we do not know of an incidence of serious heart problems in a supervised testing and exercise situation such as this.
5. There is a rare risk that your confidentiality could be broken. All of the records involving your participation in this study will be kept in locked file cabinets and/or password-protected files. All of the investigators listed on the first page of this form and any additional staff that assist with management of your files are trained in the privacy and confidentiality regulations that govern research.
6. Because this is a research study, it is also possible that there are risks of participation in this study that are not currently known.

2.3.2 KNOWN POTENTIAL BENEFITS

There is the potential for you to benefit from participating in an exercise program. However, there is no guarantee that you will receive such a benefit.

The potential benefits are:

1. Improving your heart and lung function.
2. Circulation.
3. Strength and endurance.
4. The potential to improve your walking abilities.
5. The potential that information learned from this study may benefit individuals in the future to help improve their walking abilities.

2.3.3 ASSESSMENT OF POTENTIAL RISKS AND BENEFITS

The potential benefits of improving walking and preventing future disability outweigh the minimal risks of participating in a group exercise class, which include muscle soreness, fatigue, and a rare risk of falls.

3 OBJECTIVES AND ENDPOINTS

OBJECTIVES	ENDPOINTS	JUSTIFICATION FOR ENDPOINTS
Primary		
To establish the effectiveness of OTM on improvements in walking ability and post-intervention persistence of benefits.	The primary mobility outcome is gait speed. We will also examine self-reported mobility using the Patient-Reported Outcomes Measurement Information System (PROMIS) measure (PROMIS-Mobility)	We selected gait speed measure because of its strong psychometric properties, its significant association with morbidity and mortality, and the continuous nature of the variable, which has excellent distributional properties without ceiling or floor effects in our target population. We selected the PROMIS measure as it was developed and validated with state-of-the-science methods to be psychometrically sound. The measure is relevant in patients with a wide variety of chronic diseases and conditions for the assessment of symptoms and functions. The measure has been studied extensively, is easy to administer and has a low participant burden.
Secondary		
To assess intervention fidelity (measures of adherence and competence) and identify the impact of organizational, instructor, and participant level factors on intervention fidelity.	The secondary endpoint is fidelity, assessed by measuring adherence to content (inclusion of essential components of OTM), adherence to dose (frequency and duration of exercise classes), and competence of delivery (the quality of program delivery by instructor). In addition, Acceptability of Intervention Measure (AIM)[1] will be used to measure acceptability of OTM among participants, instructors, and senior center managers. Intervention Appropriateness Measure (IAM)[1] will measure	Interventions tested in highly controlled trials often do not yield the same outcomes when applied in real life contexts (i.e. effectiveness). One reason for this dilution of effect is diminished intervention fidelity. Poor fidelity can reduce intervention potency, leading to faulty conclusions about effectiveness (type III error). To minimize type III error when transitioning from more controlled to less controlled environments, it is imperative to monitor and measure intervention fidelity. Fidelity is defined as the degree to which an intervention is implemented as prescribed in the

OBJECTIVES	ENDPOINTS	JUSTIFICATION FOR ENDPOINTS
	the perceived fit, relevance, or compatibility of the intervention for a given participant, instructor, and senior center manager. Feasibility of Intervention Measure (FIM)[1] will measure the extent to which OTM can be successfully used or carried out within a given agency or setting (i.e., feasibility). effectiveness.	original protocol, or as intended by the program developers.[20,24] Intervention fidelity can be measured by adherence to content and dose and competence in the quality of delivery of the intervention. For our investigation, we will quantify intervention fidelity by measuring instructor adherence (to content and dose of OTM) and competence (quality of delivery).
Tertiary/Exploratory		
To determine the extent to which intervention fidelity, measured by adherence and competence, moderates the effectiveness of OTM.	We will explore whether there are any thresholds in intervention fidelity that elicit a large difference in mobility gains, which may indicate a level of fidelity above which effectiveness encounters diminishing returns (e.g., <80% vs ≥80% attendance).	Intervention fidelity may influence the effectiveness of OTM. Understanding the consequences of differing levels of intervention fidelity on effectiveness will inform future implementation efforts.

4 STUDY DESIGN

4.1 OVERALL DESIGN

The study design and outcomes were selected to support our conceptual framework. We propose a Hybrid I trial, blending effectiveness and implementation research aims to: test the effectiveness of OTM, evaluate intervention fidelity, and determine the extent intervention fidelity moderates intervention effectiveness. The study design is a cluster randomized, single-blind, two-arm intervention trial to compare the effectiveness of OTM to a delayed intervention control in 502 community-dwelling older adults, across 44 senior community centers. OTM group exercise classes will be taught by certified OTM instructors in the community twice weekly for 12 weeks. Centers randomized to the control arm will receive no intervention for the first 12 weeks and then will receive 12-weeks of OTM classes. Participants are assessed at baseline, 12 weeks, and 24 weeks by individuals masked to the treatment group. Effectiveness outcomes are performance-based (i.e., gait speed) and self-reported (i.e., PROMIS) measures of mobility. We will also measure participant satisfaction and monitor adverse events (safety). Intervention fidelity will be assessed by instructor self-report (intervention diaries) and observation (fidelity checklist). Organizational, instructor, and participant level factors which may potentially impact intervention fidelity will be assessed using mixed methods (questionnaires, focus groups, and structured interviews).

4.2 SCIENTIFIC RATIONALE FOR STUDY DESIGN

Rationale for cluster randomized trial with delayed intervention control group. Given the amount of interaction that occurs among members of the senior community centers, it is imperative that we conduct a cluster randomized trial and randomize by senior community center to minimize contamination between arms. It is also consistent with intervention delivery as a group and studying organizational level factors. A delayed intervention control group was selected so that all participants eventually receive the potentially beneficial OTM intervention, while a non-intervention control group is maintained in the initial phase. In our previous trial, we discovered that participants were unhappy when they were randomly assigned to the Usual Care control group. As a result, to assess the longer impact of OTM, we will have to compare disparate time points, 24-week change in OTM centers vs 12-week change in control centers. Delaying the control group 24-weeks would have been ideal. Yet, 24-weeks of no intervention may cause greater decline, resulting in a greater effect for OTM. Thus, our approach is more conservative. We elected this approach, because we feel that the benefits for the facility and for participant retention greatly outweigh the minor disadvantage of a lack of a pure reference for this secondary hypothesis.

It is imperative that we randomize at the facility level, based on the need to prevent cross-contamination, and preserve the group-based nature of intervention delivery. Unlike traditional trials in which participants are randomized as they are recruited, randomization by facility requires a different methodology because all sampling units (i.e., participating facilities) for a wave are known and available at the time of randomization for that wave. It further affords the benefit of incorporating the facility characteristics, such as location (urban/rural) and size, into the randomization scheme to ensure a balance between arms. We anticipate that the 44 participating facilities will be largely known at the beginning of the study but acknowledge the possibility of some facilities in later waves changing their commitment to participate. Therefore, we will randomize facilities by 3 waves of 12, 16 and 16 facilities each. Number of facilities for each wave was driven by the need for an even number and the desire to have a slightly smaller number of facilities at start-up in the first wave. For each wave, first we will create several strata of facilities based on county. We are planning to recruit only urban facilities from Allegheny County to increase minority participants; the other counties are mainly suburban and rural. Therefore, our stratification can also be viewed as one that is based on urban/suburban/rural location. Second, within each stratum, using the high-quality pseudo-random deviate generator in SAS®, the study statistician Dr. Perera will randomize facilities to two generically labeled arms (e.g., A and B) in a 1:1 ratio. Third, he will descriptively compare facility characteristics between the two arms. Finally, once a satisfactorily balanced assignment to A and B arms has been achieved, labels A and B will be randomly assigned to immediate OTM and control (delayed OTM) intervention groups. The study statistician will reveal a facility's randomized group assignment to unblinded intervention delivery personnel only after all baseline assessments are complete at the facility. We are also cognizant of the unlikely but potential need to add more sites if a site changes its commitment to participate, and/or if recruitment of participants per site is less than anticipated. In such an event, we will recruit suburban sites from Allegheny County, as well as additional sites from other counties, and increase the number of facilities randomized in waves 1 and/or 2. Reaching the participant recruitment target with a greater number of facilities has a beneficial rather than an adverse effect on statistical power.

4.3 JUSTIFICATION FOR INTERVENTION

The OTM exercise program is a group-based program led by a certified OTM instructor. The frequency and duration of OTM is 50 minutes, twice weekly for 12 weeks for a total of 24 classes. The frequency and duration of OTM classes is based on our previous research which demonstrated the efficacy of the intervention. The OTM exercise program is based on principles of motor learning that enhance smooth and automatic movement control. The program contains a warm-up (5 minutes), stepping patterns (15 minutes), walking patterns (15 minutes), strengthening exercises (10 minutes), and cool-down exercises (5 minutes). The warm-up and cool down contain gentle range of motion exercises and stretches for the lower extremities and trunk. Considered essential components, the stepping and walking patterns are progressively more difficult, goal-oriented patterns that promote the timing and coordination of stepping integrated with the phases of the gait cycle. Conceptually, the exercise is intended to achieve its effects by shifting the center of pressure posterolateral and then forward, encouraging hip extension prior to stepping, loading the trailing limb, coordinating activation of the abductors of the soon-to-be swing leg with adductors of the stance limb, and shifting the center of pressure in medial stance to unload the stepping limb. Progression is based on separately increasing the speed, amplitude, or accuracy of performance prior to undertaking a more complex task. The strengthening exercises are conducted sitting and standing, targeting the lower extremity muscles. In our preliminary work, we have conducted the OTM program in over 20 different locations with older adults with varying levels of ability. We were able to successfully individualize the program, in that participants reported feeling both challenged and safe. The optimal dose of OTM is unknown. In the current trial we will record modifications to mode and frequency of delivery of OTM and examine whether these modifications impact effectiveness.

4.4 END-OF-STUDY DEFINITION

The total number of visits for this exercise study is 28 visits in about 24 weeks. The timing of the sessions will vary depending on which exercise session the community-based organization is assigned to. If the organization is assigned to the first session the visits will include: 1 visit for testing prior to starting the exercise program, 24 visits for the exercise program, 1 visit for testing after the exercise program, 1 visit for the focus group, and 1 final testing session 12 weeks after the conclusion of the exercise program. If the organization is assigned to the second session the visits will include: 2 visits for testing prior to starting the exercise program, 24 visits for the exercise program, 1 visit for testing after the exercise program, and 1 visit for the focus group.

A participant is considered to have completed the study if he or she has completed the baseline assessment, up to 24 intervention sessions, the focus group, and the 12-week and 24-week follow-up assessments. The end of the study is defined as completion of the last 12-week or 24-week follow-up assessment shown in the Schedule of Activities (SoA), **Section 1.3** and all of the focus groups and interviews.

5 STUDY POPULATION

The target population of our study is community-dwelling older adults.

5.1 INCLUSION CRITERIA

To be eligible to participate in this study, an individual must meet all the following criteria:

1. Be 60 years of age or older
2. Attend a senior community center
3. Can ambulate household distances independently
4. Can participate in group exercise programs

5.2 EXCLUSION CRITERIA

An individual who meets any of the following criteria will be excluded from participation in this study:

1. Impaired cognition, defined as inability to understand the informed consent process.
2. Plans to leave the area for an extended period.
3. Any acute or unstable illness or medical condition.

5.3 LIFESTYLE CONSIDERATIONS

N/A

5.4 SCREEN FAILURES

Screen failures are defined as participants who consent to participate in this study but are not subsequently assigned to the study intervention or entered in the study. Individuals who do not meet the criteria for participation in this trial (screen failure) because of meeting one or more exclusion criteria that are likely to change over time may be rescreened. Examples include the successful treatment of an acute medical conditions or the stabilization of an illness or chronic condition. Rescreened participants will be assigned the same participant number as for the initial screening.

5.5 STRATEGIES FOR RECRUITMENT AND RETENTION

We plan to enroll 10-15 participants from each center and conduct multiple sections of the class if the enrollment is over the maximum class size of 10. Recruitment strategies will be like those used in our PCORI trial, wherein we were able to successfully recruit the needed number of participants (424 randomized from 32 sites in 2.5 years). We will make a brief presentation at the senior community center to introduce and explain the study. We will also involve the management and frontline staff of the centers to promote the study through posted calendars, newsletters, and town meetings. Potential participants will be asked to give their contact information to the study staff so that they can be contacted and scheduled. After obtaining written informed consent, a screening exam will take place to identify potential exclusion issues. Potential participants who meet all inclusion/exclusion criteria will undergo baseline testing. A standard demographics questionnaire will be used to determine age, gender, race, marital status and work history, to adequately describe the research participants.

There are no sex/gender, race or ethnicity restrictions for participation in this study. Given that walking difficulty is more common in women than men and that women are more likely to participate in group exercise classes, we expect to enroll slightly more women than men into study. Based on census data and our past group exercise study we anticipate 70-75% of the participants to be female. Minorities will be represented in proportion to the ethnic and racial character of the senior community centers in which the study will be conducted. We anticipate at least 75% of the participants to be white. We have specifically selected diverse counties and targeted centers within those counties that are in underserved communities. In Allegheny County, which includes the City of Pittsburgh, we are targeting senior centers in some of the poorer areas of the inner city. Fayette County was selected because it is more rural, has a high level of impoverishment, and a low health ranking. In some counties where the Area Agency on Aging (AAA) plays a significant role in the programming of the senior community centers (i.e., Washington, Fayette, Greene, and Westmoreland Counties), we will work with the Agency administrator to select centers which best represent the county and facilitate diversity. We will work closely with each of the centers to target a representative sample of center participants. Based on our prior study conducted primarily in Allegheny County, individuals who participated in OTM at the senior community centers were 71% white, 78% female, and 72% had less than college education.

Participants will be paid a total of $90 if they complete all parts of this study. If for any reason participants complete part but not all of the study, the terms of this payment will be as follows: $20 for completing each of the 3 testing sessions and $30 for participating in the focus group. Payment will be made to participants at the completion of the testing session using a University of Pittsburgh “Vincent” credit/debit card. Community-based organization representatives will be compensated $40 for completion of study questionnaires/interviews. OTM instructors will be compensated $60 for completion of exercise class logs and study questionnaires/interviews.

6 STUDY INTERVENTION(S) OR EXPERIMENTAL MANIPULATION(S)

6.1 STUDY INTERVENTION(S) OR EXPERIMENTAL MANIPULATION(S) ADMINISTRATION

6.1.1 STUDY INTERVENTION OR EXPERIMENTAL MANIPULATION DESCRIPTION

All study participants will receive the intervention. Those participating at facilities randomized to receive the program immediately will begin at week 1 and those who participate at facilities randomized to receive the delayed intervention will begin after week 12. Our aims include investigation of effectiveness of OTM when implemented in real-world conditions using the targeted clinical endpoints of change in gait speed and mobility; and exploration of the extent to which certified instructors deliver the program as intended and how those measures of fidelity impact gait speed and mobility.

OTM is a group exercise program to improve walking in older adults by targeting the timing and coordination of walking.[8] The program challenges the brain to match the timing and sequences of movements with posture to improve the smoothness and efficiency of walking. To achieve this ‘better walking,’ OTM includes a variety of progressing stepping and walking patterns that are based on the biomechanics of gait. OTM classes are designed to meet for 50 minutes, twice a week for 12 weeks. The program includes a warm-up, progressive walking and stepping patterns, lower extremity strengthening patterns focused on the muscles important for walking, and a flexibility/cool-down component.

6.1.2 ADMINISTRATION AND/OR DOSING

Dose will be quantified as the frequency, and duration of exercise classes (50-minute classes, twice weekly for 12 weeks is ideal) and by the attendance of participants. Each community center is required to document when exercise classes are offered. We will obtain this information directly from the community center records. This will give a measure of the dose of OTM made available (offered) to the participants and will be recorded as number of classes (0-24), number of weeks (0-12) and total minutes of exercise (0-1200). We will also measure dose at the participant level (i.e., how many classes each participant attends, 0-24). The OTM instructor will keep a class log of each session to record attendance.

6.2 FIDELITY

6.2.1 INTERVENTIONIST TRAINING AND TRACKING

The OTM exercise program will be delivered by certified OTM instructors. Individuals must be a licensed health professional (physical therapist, physical therapy assistant, occupational therapist, occupational therapy assistant, etc.) or a certified fitness instructor, and have current CPR certification to become an OTM instructor. The OTM instructor training consists of an 8-hour course which includes lecture and hands-on laboratory components. All instructors must pass a written and practical examination to become certified. During the training, instructors are taught how to identify participants who are having difficulty and how to modify the program to fit their needs. In addition, all participants will be screened using the PAR-Q prior to starting the exercise program and each OTM class will be limited to 10 participants to maintain a reasonable and safe instructor to participant ratio.

We use a modified version of the Conceptual Framework for Implementation Fidelity[20], a well-accepted model, to guide our investigation. The framework depicts the vital elements of intervention fidelity and their relationship to one another. Intervention fidelity can be measured by adherence to content and dose and competence in the quality of the intervention. Carroll et al propose that many of these elements interact and may influence the association between intervention, fidelity, and outcomes.[20] For our investigation, we will quantify intervention fidelity by measuring instructor adherence (to content and dose of OTM) and competence (quality of delivery). Adaptations or modifications to the evidence-based program occur during the implementation process to address differences between the context in which the intervention was originally designed and tested and the one into which it is ultimately implemented.[21] Some adaptations are planned/purposeful and occur with the intention to retain fidelity where others are unplanned and in reaction to a specific circumstance. While some modifications may facilitate implementation and sustainability, they can also erode treatment integrity. Therefore, is it critical to monitor both fidelity and adaptations to gain a more complete understanding of the implementation process and the levels of fidelity necessary to promote desired outcomes. We will record, define, and classify adaptations to OTM to inform future implementation efforts.

6.3 MEASURES TO MINIMIZE BIAS: RANDOMIZATION AND BLINDING

It is imperative that we randomize at the facility level, based on the need to prevent cross-contamination and preserve the group-based nature of intervention delivery. Unlike traditional trials in which participants are randomized as they are recruited, randomization by facility requires a different

methodology because all sampling units (i.e., participating facilities) for a wave are known and available at the time of randomization for that wave. It further affords the benefit of incorporating the facility characteristics, such as location (urban/rural) and size, into the randomization scheme to ensure a balance between arms. We anticipate that the 44 participating facilities will be largely known at the beginning of the study but acknowledge the possibility of some facilities in later waves changing their commitment to participate. Therefore, we will randomize facilities by 3 waves of 12, 16 and 16 facilities each. Number of facilities for each wave was driven by the need for an even number and the desire to have a slightly smaller number of facilities at start-up in the first wave. For each wave, first we will create several strata of facilities based on county. We are planning to recruit only urban facilities from Allegheny County to increase minority participants; the other counties are mainly suburban and rural. Therefore, our stratification can also be viewed as one that is based on urban/suburban/rural location. Second, within each stratum, using the high-quality pseudo-random deviate generator in SAS®, the study statistician Dr. Perera will randomize facilities to two generically labeled arms (e.g., A and B) in a 1:1 ratio. Third, he will descriptively compare facility characteristics between the two arms. Finally, once a satisfactorily balanced assignment to A and B arms has been achieved, labels A and B will be randomly assigned to immediate OTM and control (delayed OTM) intervention groups. The study statistician will reveal a facility's randomized group assignment to unblinded intervention delivery personnel only after all baseline assessments are complete at the facility. We are also cognizant of the unlikely but potential need to add more sites if a site changes its commitment to participate, and/or if recruitment of participants per site is less than anticipated. In such an event, we will recruit suburban sites from Allegheny County, as well as additional sites from other counties, and increase the number of facilities randomized in waves 1 and/or 2. Reaching the participant recruitment target with a greater number of facilities has a beneficial rather than an adverse effect on statistical power.

Given the nature of the intervention being tested in this study and the use of wait-list control sites, it will not be possible to blind the participants or the instructor providing the interventions. To minimize "contamination" of the interventions, facilities will be randomized to the intervention or to the wait-list, instead of individual participants. Study staff involved in randomization or fidelity assessments will not perform follow-up data collection. Study staff performing the follow-up assessments will be blinded to treatment assignments and will be asked not to seek this information when in contact with the participants. Participants will be asked not to discuss their treatment experience during study visits. Only a small number of staff members and the safety monitor will see group-specific study results.

If "unblinding" occurs, a protocol deviation form will be completed to document and record the following information:

- The ID of the unblinded facility,
- The date of unblinding,
- The study staff person responsible for unblinding,
- A list of the person(s) who have been unblinded.

6.4 STUDY INTERVENTION/EXPERIMENTAL MANIPULATION ADHERENCE

We have identified the essential components of OTM. Inclusion of essential components will be determined from exercise class logs. OTM instructors will be required to keep a log of time spent on each program component in each class, for each of the 24 classes. We will review the logs and determine if the essential components (warm-up, walking, stepping patterns, strengthening exercises and cool-down) are included in each of the classes.

Each community center is required to document when exercise classes are offered. We will obtain this information directly from the community center records. This will give a measure of the dose of OTM made available (offered) to the participants and will be recorded as number of classes (0-24), number of weeks (0-12) and total minutes of exercise (0-1200). We will also measure dose at the participant level (i.e. how many classes each participant attends, 0-24). The OTM instructor will keep a class log of each session to record attendance>

6.5 CONCOMITANT THERAPY

N/A

6.5.1 RESCUE THERAPY

N/A

7 STUDY INTERVENTION/EXPERIMENTAL MANIPULATION DISCONTINUATION AND PARTICIPANT DISCONTINUATION/WITHDRAWAL

7.1 DISCONTINUATION OF STUDY INTERVENTION/EXPERIMENTAL MANIPULATION

The participants' participation in this research study is completely voluntary. The participant may withdraw, at any time, their consent for participation in this research study. Any identifiable research information recorded for, or resulting from, their participation in this research study prior to the date that they formally withdrew their consent may continue to be used and disclosed by the investigators. To formally withdraw consent for participation in this research study the participant should provide a written and dated notice of this decision to the principal investigator of this research study. If the participant withdraws from the intervention, study staff will ask permission to continue to follow the participant for follow-up assessment. If participation is discontinued for medical reasons and the participant is unable to complete the performance-based testing, all attempts will be made to obtain the self-reported outcomes.

7.2 PARTICIPANT DISCONTINUATION/WITHDRAWAL FROM THE STUDY

Participation in the research study will be discontinued if a participant's walking ability worsens; for example, if a subject who has osteoarthritis of the knees exhibits a significant increase in pain that begins

to negatively affect their walking ability. A participant can also be removed from the study if any new medical condition, injury or illness is discovered during the course of treatment that would render the participant ineligible based on the exclusion criteria.

If a participant is identified by the research staff to have a temporary medical condition that would impact their participation in the research study (i.e., any of the exclusion criteria), the participant will be required to obtain medical clearance for participation in exercise from their primary care physician prior to returning to the exercise intervention.

Any identifiable research information recorded for or resulting from participation in this research study prior to the date that the subject formally withdrew their consent will continue to be used and disclosed by the researchers for the purposes described above.

7.3 LOST TO FOLLOW-UP

A participant will be considered lost to follow-up if he or she fails to return for three scheduled visits and study staff are unable to contact the participant after at least 3 attempts.

The following actions must be taken if a participant fails to return to the clinic for a required study visit:

- The site will attempt to contact the participant, reschedule the missed visit within 1 week, counsel the participant on the importance of maintaining the assigned visit schedule and ascertain if the participant wishes to and/or should continue in the study
- Before a participant is deemed lost to follow-up, the investigator or designee will make every effort to regain contact with the participant (where possible, 3 telephone calls and, if necessary, a letter to the participant's last known mailing address or local equivalent methods). These contact attempts will be documented in the participant's study file.
- Should the participant continue to be unreachable, he or she will be considered to have withdrawn from the study with a primary reason of lost to follow-up

8 STUDY ASSESSMENTS AND PROCEDURES

8.1 ENDPOINT AND OTHER NON-SAFETY ASSESSMENTS

Screening for study participation includes:

- Determining the individual's age to be at least 60 years by self-report of birthdate,
- Observing the ability to ambulate household distances independently,
- Asking whether individual can participate in group exercise classes,
- Asking whether individual plans to leave the area for an extended period during the intervention,
- Completion of the Physical Activity Readiness Questionnaire (PAR-Q).

Outcomes will be assessed in both groups (OTM and OTM delayed) at baseline (week 0), week 12, and week 24 at the community center sites by research staff masked to group assignment.

Mobility. The primary outcome is gait speed, a performance-based measure. Gait speed is assessed using a standardized protocol. After explanation and demonstration, the participant completes 2 walks on a 6-meter course at their usual walking speed. The subject is timed during the central 4 meters of the walk. The initial and final meters are to allow for acceleration and deceleration. Gait speed will be averaged over the 2 walks.

The Patient-Reported Outcomes Measurement Information System (PROMIS) measure is the self-report measure of mobility (PROMIS-Mobility). The 15 item PROMIS Mobility scale will be used to assess self-reported perceptions of mobility; a single Mobility score is obtained from the short form, with higher scores indicating greater mobility.

Satisfaction. Participant satisfaction will be assessed on the last day of the exercise program using a survey that includes five items measured on a Likert scale and three yes/no questions. The Likert items assess degree of satisfaction with various components of the exercise program (i.e., the exercises, instructor, music, space, and the overall program). The yes/no questions assess if the program met the participants' expectations, if they feel they benefited from the program, and if they would recommend the program to others.

Factors impacting fidelity (quantitative approach).

Facilitation strategies. To support OTM instructors, we will encourage them to contact us through the website, email, or phone if they have any questions or concerns about delivering the OTM program. The OTM support staff will provide feedback and suggestions for class delivery and will help the instructor actively problem-solve through the issue. In most instances, verbal consultation is adequate. If the problem or issue persists, the OTM instructor can request an in-person visit/consultation with an OTM support staff member. We will track the frequency at which the instructors utilize the support system as a measure of facilitation strategies.

Responsiveness of participant, instructor, and senior center management. Responsiveness by those delivering and receiving the intervention can influence intervention fidelity. Factors may include a participant's acceptance of the responsibilities associated with the intervention, the instructor's beliefs about the intervention, and participants' and instructors' beliefs about the usefulness of OTM. We will capture responsiveness using three reliable and validated measures: the Acceptability of Intervention Measure (AIM), the Intervention Appropriateness Measure (IAM) and the Feasibility of Intervention Measure (FIM).

Acceptability of Intervention Measure (AIM). The AIM will be used to measure acceptability of OTM among participants, instructors and senior center managers. The AIM contains 4 questions on a 5-point Likert scale that ask if the intervention meets their approval, is appealing, liked and welcomed. Scores range from 4 to 20 with higher scores representing greater acceptability. As the acceptability may change over time, we will administer the AIM to participants, instructors, and senior center managers before and after the intervention.

Intervention Appropriateness Measure (IAM). We will use the IAM to measure the perceived fit, relevance or compatibility of the intervention for a given participant, instructor, and senior center manager. The IAM contains 4 questions on a 5-point Likert scale that ask if the intervention seems fitting, suitable, applicable and a good match. Scores range from 4 to 20 with higher scores representing greater perceived fit. As an individual's perception of the appropriateness of the intervention may change over time, we will administer the IAM before and after the 12-week OTM intervention.

Feasibility of Intervention Measure (FIM). We will use the FIM to measure the extent to which OTM can be successfully used or carried out within a given agency or setting (i.e. feasibility). The FIM contains 4 questions scored on a 5-point Likert scale that ask if OTM seems implementable, possible, doable and easy to use. Scores range from 4 to 20 with higher scores representing greater feasibility. We will administer the FIM to center managers and instructors prior to the initiation of the OTM intervention.
Recruitment. We will obtain demographic characteristics such as gender and race of all members of the senior centers from the center manager.

Context. To identify items to include in the context questionnaire, we used the Consolidated Framework for Implementation Research (CFIR) to guide our selection of items representing the inner setting (i.e. implementing organization). Items represent the constructs of structural characteristics, networks and communications, culture, implementation climate and readiness for implementation. Some specific items on the questionnaire include financial resources, prior experience offering evidence-based programs, compatibility of OTM with other organized programs and mission of the center, and amount and type of programming offered at the center. The lead manager or individual responsible for evidence-based programming at the center will complete the questionnaire prior to the initiation of the OTM intervention.

Factors impacting fidelity (qualitative approach).
We will use qualitative methods, focus groups and semi-structured interviews, to better understand the factors and how or why they may impact fidelity and effectiveness. Specifically, we will delve into the intervention complexity from the instructor perspective; the acceptability of the intervention by the manager, instructor, and participant; the appropriateness of the intervention as viewed by managers, OTM instructors, and participants; and the feasibility of the intervention as viewed by the managers and the OTM instructors. The qualitative data will build on our quantitative data and will be used to help explain the results of the quantitative findings (i.e. provide depth of understanding). The University Center for Social and Urban Research (UCSUR) will conduct the focus groups and semi-structured interviews.

Focus groups: We will conduct focus groups of the participants at each of the sites. To encourage participation, focus groups will occur at the senior center, and the participants will be compensated for their participation. These focus groups will occur after the post-intervention quantitative data collection. We will work with the University Center for Social and Urban Research (UCSUR) Survey Research Program and the Qualitative Data Analysis Program to create the scripts for the focus groups. Using the responses from the AIM and IAM questionnaires, we will develop questions that expand on the participants' questionnaire responses. For example, a question on the AIM is "OTM is appealing to me" and participants can respond: "completely disagree," "disagree," "neutral," "agree," or "completely agree." If the participants select "disagree" with this statement, we will ask them why OTM is not appealing, delving into specific reasons.

Semi-structured interviews: UCSUR will conduct semi-structured interviews with all the OTM instructors and a representative from the senior centers (i.e. manager). The interviews will be conducted in person or over the phone and are anticipated to take approximately 30 minutes. Instructors and senior center staff will be compensated for their time to encourage interview completion. All interviews will occur after completion of the questionnaires (i.e. quantitative data collection). Again, we will use questionnaire responses (i.e. AIM, IAM, and FIM) to develop interview scripts, probing for clarity and elaboration on their questionnaire responses. For example, if an instructor or manager states that OTM is not a good match for the participants (IAM question #4), we will ask them to explain why they feel this way and if they can provide a rationale. Likewise, if they state that OTM does not seem possible (FIM question #2)

we would ask them to explain why it is not possible and ask for potential suggestions. Further, we will ask the OTM instructors about their feelings on OTM intervention complexity, including questions on instructor support materials, such as the training manual.

8.2 SAFETY ASSESSMENTS

N/A

8.3 ADVERSE EVENTS AND SERIOUS ADVERSE EVENTS

8.3.1 DEFINITION OF ADVERSE EVENTS

Expected adverse events (AEs) will be captured through interviews at 12 and 24 weeks, based on the Health Status Update questionnaire. Reportable events are those events that have potential implications for participant safety and require individual reporting. Reportable AEs are defined as serious adverse events (SAEs), AEs that have potential implications for participant safety, unexpected AEs, and injury that occurs while a participant is under the supervision of study related personnel. These RAEs will require individual event reporting as described below. The timely and complete account of RAEs will be a critical requirement for the protection of human subjects in this trial.

The SO will review and approve study-defined expected AEs and will be involved in the regular monitoring of the RAE reporting system. Reporting of expected AEs to the SO, the NIA, and the IRB will be through the database. RAEs will be monitored and tracked separately and reviewed by the SO and the NIA as outlined below.

All unanticipated problems involving risks to study participants or others will be reported to the IRB, the NIA, and the Safety Officer within 48 hours unless they are also SAEs which will be reported within 24 hours of learning of the event. Reports of unanticipated problems will be forwarded to the OHRP within two weeks of the event. All unanticipated problems will be reviewed by the PI and the study team and a corrective plan to prevent additional unanticipated problems will be developed and implemented.

Classifying Adverse Events
An AE is any unfavorable or unintended medical occurrence in a human study participant that has taken place during the course of a research project, including any abnormal sign, symptom, or disease, whether or not related to participation in the research.

For the purposes of this study, any event that meets the criteria for an SAE, is unexpected, or results in injury to the participant while he/she is under the supervision of study related personnel will be classified as an RAE. Adequate review, assessment, and monitoring of RAEs require they be classified as to severity, expectedness, and potential relatedness to the study intervention.

8.3.2 DEFINITION OF SERIOUS ADVERSE EVENTS

Serious adverse event (SAEs) - SAEs will be defined as any adverse event that results in death, is life threatening, or places the participant at immediate risk of death from the event as it occurred, requires or prolongs hospitalization, causes persistent or significant disability or incapacity, results in congenital abnormalities or birth defects, or is another condition which investigators judge to represent significant hazards.

8.3.3 CLASSIFICATION OF AN ADVERSE EVENT

The following subsections will include a discussion of how AEs will be classified.

8.3.3.1 SEVERITY OF EVENT

An AE is any unfavorable or unintended medical occurrence in a human study participant that has taken place during the course of a research project, including any abnormal sign, symptom, or disease, whether or not related to participation in the research.

For the purposes of this study, any event that meets the criteria for an SAE, is unexpected, or results in injury to the participant while he/she is under the supervision of study related personnel will be classified as an RAE. Adequate review, assessment, and monitoring of RAEs require they be classified as to severity, expectedness, and potential relatedness to the study intervention.

Severity
The following guidelines will be used to determine level of severity:
Mild: Awareness of signs and symptoms, but easily tolerated and causing no loss of time from normal activities. No specific medical attention is required.
Moderate: Discomfort enough to cause a low level of inconvenience or concern to the participant and may interfere with daily activities. Symptoms may require minimal, local or noninvasive medical intervention.
Severe: Events interrupt the participant's normal daily activities and are usually incapacitating. Significant symptoms may require hospitalization or invasive medical intervention.
Life-threatening/Disabling: Events that may involve acute, life-threatening metabolic or cardiovascular complications (such as circulatory failure, hemorrhage, sepsis) or life –threatening physiological consequences. Intensive care or emergent invasive procedure is required.
Death: Causing death.

Severity is not synonymous with seriousness. A severe headache is not necessarily a RAE. However, mild chest pain may result in a day's hospitalization and thus would be classified as a RAE.

8.3.3.2 RELATIONSHIP TO STUDY INTERVENTION/EXPERIMENTAL MANIPULATION

Relatedness
The PI in consultation with the Co-Investigators and the independent SO monitor will determine the degree to which RAEs are related to study procedures using the criteria below.

Definitely related: The adverse event is clearly related to the investigational procedure – i.e., an event that follows a reasonable temporal sequence from administration of the study intervention, follows a known or expected pattern of response to the study intervention, that is confirmed by improvement on stopping and reappearance of the event on repeated exposure, and that could not be reasonably explained by the known characteristics of the participant's clinical state.

Possibly related: An adverse event that follows a reasonable temporal sequence from administration of the study intervention of that follows a known or expected pattern of response to the study intervention, but that could readily have been produced by a number of other factors.

Unrelated: The adverse event is clearly not related to the investigational procedure (i.e. another cause of the event is most plausible; and/or a clinically plausible temporal sequence is inconsistent with the onset of the event and the study intervention and/or a causal relationship is considered biologically implausible).

8.3.3.3 EXPECTEDNESS

AEs will be assigned as to whether they were expected or unexpected based on current knowledge. Categories are defined as follows:

Expectedness
Expected: An AE that is anticipated on the basis of prior experience with the intervention under investigation; an event that can be attributed to the underlying condition of the participant being studied; or an event that can be attributed to the patient population being studied (see section--- Expected AEs). Expected adverse events will be captured through interviews at 12 and 24 weeks, based on the Health Status Update questionnaire.

Unexpected: An unexpected AE is defined as medical events that occur during study participation, but do not commonly occur in the study population and which are not listed in the informed consent document or study protocol.

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

The occurrence of an adverse event (AE) or serious adverse event (SAE) may come to the attention of study personnel during study visits and interviews of a study participant presenting for a study visit.

All AEs, not otherwise precluded per the protocol, will be captured on the appropriate case report form (CRF). Information to be collected includes event description, time of onset, clinician's assessment of severity, relationship to study procedures (assessed only by those with the training and authority to make a diagnosis), and time of resolution/stabilization of the event. All AEs occurring while on study will be documented appropriately regardless of relationship. All AEs will be followed to adequate resolution.

Any medical or psychiatric condition that is present at the time that the participant is screened will be considered as baseline and not reported as an AE. However, if the study participant's condition deteriorates at any time during the study, it will be recorded as an AE.

Changes in the severity of an AE will be documented to allow an assessment of the duration of the event at each level of severity to be performed. Documentation of onset and duration of each episode will be maintained for AEs characterized as intermittent.

The Study Coordinator will record events with start dates occurring any time after informed consent is obtained until 7 (for non-serious AEs) or 30 days (for SAEs) after the last day of study participation. At each study visit, the investigator will inquire about the occurrence of AE/SAEs since the last visit. Events will be followed for outcome information until resolution or stabilization.

8.3.5 ADVERSE EVENT REPORTING

The study PI has primary responsibility for the safety of participants as it relates to the study protocol. The study coordinator will be responsible for reviewing adverse events and assuring accurate and timely reporting of the adverse events. The co-investigators will review, evaluate, and classify adverse events and provide follow-up for events until they are resolved. The SO will be responsible for monitoring study data for evidence of adverse effects attributable to participation in the study. The PI will be responsible for reporting study-defined AEs and SAEs to the University of Pittsburgh institutional review board (IRB) according to their timeline and format. Any deaths will be reported to the NIA Program Officer and the Safety Officer within 24 hours of knowledge of the event.

Unexpected Adverse Event. An unexpected AE may be witnessed by a member of the research team or a staff member may be told about the occurrence or an unexpected event that may meet the criteria for reporting. Unexpected AEs that have a potential relationship to study procedures and activities will be reported to the University of Pittsburgh IRB and the SO within 24 hours of learning of the event.

Adverse Events Related or Possibly Related to the Research Intervention. An AE which in the opinion of the principal investigator, the incident, experience or outcome more likely than not was caused by the procedures involved in the research will be reported to the IRB and the SO within10 working days of the investigator learning of the event. If the AE is both serious (SAE) and unexpected it will be reported to the IRB, NIA Program Officer, and the Safety Officer within 48 hours of the study's knowledge of the SAE.

Reporting Expected AEs and RAEs to the SO
The SO will review tabulated data on non-serious and expected AEs on a semi-annual basis and monitor for adverse event rates out of proportion to those expected. The PI will forward the individual Reportable Adverse Event Records, including a narrative for each event, as well as a table showing all RAEs to the SO and NIA semi-annually. The SO will review all RAEs that are temporally related to the interventions in aggregate form at its scheduled meetings. All RAEs that relate to hazards of the study interventions or are cause for urgent concern will be reported to the SO, the NIA and the University of Pittsburgh IRB immediately after recognition of their importance.

8.3.6 SERIOUS ADVERSE EVENT REPORTING

Serious Adverse Events. The occurrence of SAEs will be obtained on an Adverse Event form during the 12 and 24-week testing sessions and by any study staff member who learns of a serious event. All study staff will be trained to recognize when there is a SAE and to follow appropriate reporting procedures as they become aware of a reportable event.

All adverse events that are both serious (SAEs) and unexpected will be reported to the University of Pittsburgh IRB, the SO, and NIA Program Officer within 48 hours of the study's knowledge of the event. The expedited report of the SAE will be submitted by telephone, fax, or email. Recognizing that the information available during this 48-hour period may not be sufficient to permit accurate completion of the required adverse event reporting forms, a detailed written SAE report will be completed as soon as the information is available. Follow up information can be requested by the SO or the NIA Program Officer or its representative. The summary of all other SAEs will be reported to the NIA Program Officer and the Safety Officer, quarterly, unless otherwise requested by the Safety Officer.

8.3.7 REPORTING EVENTS TO PARTICIPANTS

N/A

8.3.8 EVENTS OF SPECIAL INTEREST

N/A

8.3.9 REPORTING OF PREGNANCY

N/A

8.4 UNANTICIPATED PROBLEMS

8.4.1 DEFINITION OF UNANTICIPATED PROBLEMS

This protocol uses the definition of Unanticipated Problems as defined by the Office for Human Research Protections (OHRP). OHRP considers unanticipated problems involving risks to participants or others to include, in general, any incident, experience, or outcome that meets **all** of the following criteria:

- Unexpected in terms of nature, severity, or frequency given (a) the research procedures that are described in the protocol-related documents, such as the Institutional Review Board (IRB)-approved research protocol and informed consent document; and (b) the characteristics of the participant population being studied;

- Related or possibly related to participation in the research ("possibly related" means there is a reasonable possibility that the incident, experience, or outcome may have been caused by the procedures involved in the research); and
- 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 PROBLEMS REPORTING

All unanticipated problems involving risks to study participants or others will be reported to the IRB, the NIA, and the Safety Officer within 48 hours unless they are also SAEs which will be reported within 24 hours of learning of the event. Reports of unanticipated problems will be forwarded to the OHRP within two weeks of the event. All unanticipated problems will be reviewed by the PI and the study team and a corrective plan to prevent additional unanticipated problems will be developed and implemented.

8.4.3 REPORTING UNANTICIPATED PROBLEMS TO PARTICIPANTS

N/A

9 STATISTICAL CONSIDERATIONS

9.1 STATISTICAL HYPOTHESES

- Primary Endpoint(s):

We hypothesize that, compared to a randomized control group followed for 12 weeks, the OTM group will have greater gains in gait speed (primary outcome) and self-reported mobility at 12 and 24 weeks of follow-up.

- Secondary Endpoint(s):

We hypothesize that organizational, instructor, and participant level factors will impact fidelity. For example, instructors who utilize greater facilitation strategies will have higher fidelity, poor organizational acceptance of the program will be related to lower fidelity, and fewer organizational resources will lead to lower fidelity.

We hypothesize that higher fidelity will be associated with greater mobility gains, and there may be critical fidelity thresholds.

9.2 SAMPLE SIZE DETERMINATION

Prior data used for the sample size justification include a between-subject standard deviation of 0.13 m/s for pre- to post-intervention gait speed change from the OTM PCORI trial[25], a conservative 80% participant

retention rate to account for the pragmatic nature of the study (compared to over 90% observed under more controlled conditions), and an average cluster size of 13.14 with a coefficient of variation of 0.5529 and an over-ground gait speed intra-cluster correlation of 0.053 in a subset of same type of sites from the OTM trial.[25] The sample size inflation factor (design effect) based on this information is 1.86.[26,27] We used published methodologies[28-31] implemented in commercially available PASS 2012® (Number Cruncher Statistical Systems, Kaysville, Utah) for sample size computations. We will be able to detect statistical significance of a difference as small as 0.05 m/s in pre- to post-intervention gait speed change between interventions with 80% statistical power in a two-tailed test conducted at α=0.05 with 251+251=502 participants or 22+22=44 sites including 4 contingency sites. A change of 0.05 m/s used for computations is considered small but meaningful,[5] and if the true effect is 0.06 m/s, we would have 92% power. We actually observed a much greater adjusted intervention effect of 0.09 m/s in community centers in our OTM PCORI trial.[25] Therefore, we are confident that our sample size affords ample statistical power and sensitivity to detect sufficiently small magnitudes of intervention effect in Aim 1.

Aim 2 involves estimation of metrics of intervention fidelity and potential impacting factors. We will be able to estimate with 95% confidence within-intervention arm rates of any participant-level dichotomized metric (such as percent with at least 80% class attendance) with a margin of error no greater than ±9.6% and ±6.8%, respectively, combined and separate arm analyses.[32,33] In Aims 2 and 3, we will be able to detect within- and combined-arm correlations representing influencing factors-fidelity associations (Aim 2) and fidelity-effectiveness associations (Aim 3) as small as 0.27 and 0.19 with 80% power in two-tailed tests conducted at α=0.05.[31,34,35] In the unlikely event that we have to increase the number of facilities due to lower than expected enrollment per facility, it will have a beneficial rather than an adverse effect on the above computations of statistical power.

9.3 POPULATIONS FOR ANALYSES

The older adult target population is adults 60 years of age or older, who are able to walk household distances without the help of another person, do not have plans to leave the area for an extended period of time, do not have any acute or unstable illness or medical condition, and can participate in a group exercise program

Community-based organization staff - will be the manager or director of the organization and the individual hired by the organization to deliver the On the Move exercise program (i.e. OTM instructor).

9.4 STATISTICAL ANALYSES

9.4.1 GENERAL APPROACH

We will perform all main analyses based on intention-to-treat philosophy following the a priori plans outlined below. All statistical analyses will be performed and/or overseen by the study statistician, Dr. Perera, using SAS® version 9 (SAS Institute, Inc., Cary, North Carolina). Participant flow will be summarized using a CONSORT extension diagram for cluster randomized studies. Data will be summarized by treatment group and time point, as well as baseline to follow-up change using appropriate descriptive statistics. First, the baseline participant characteristics will be compared between the two arms. Any significant differences will be noted and accounted for as additional covariates in the sensitivity analyses.

9.4.2 ANALYSIS OF THE PRIMARY ENDPOINT(S)

We will fit a series of linear mixed models,[36] using the SAS® MIXED procedure with baseline to follow-up changes in continuous outcomes as the dependent variable; intervention arm (immediate OTM/delayed OTM control), follow-up time point (12/24 weeks) and their interaction term as the fixed effects of interest; the baseline value of the continuous outcome as a fixed effect covariate; a facility random effect to account for clustering of participants within facilities; and an exchangeable working correlation structure to account for correlation between multiple post-intervention measurements over time from the same participant. We will appropriately construct means contrasts to estimate active vs control intervention effects difference to test specific hypotheses. The magnitude and statistical significance at α=0.05 of the active vs control difference at 12 weeks for gait speed will serve as the formal test of the primary hypothesis. If significant, we will define a meaningful responder as having at least a 0.05 m/s improvement in gait speed over 12 weeks, compute percentage of responders in both arms, and conclude a number-needed-to-treat (NNT) as the reciprocal of the difference in percentages. We will employ the same analytic strategy described above for assessing 12-week intervention effect on our secondary outcome, PROMIS-mobility. While post-intervention assessments of the delayed intervention groups are not typically used in analyses to preserve the randomized nature of the study design, we will conduct a sensitivity analysis with that data via an additional means contrast which compares the average improvement of 0-12 weeks in immediate OTM and 12-24 weeks in delayed intervention group (both due to OTM) with the 0-12 week changes in the delayed intervention (due to control condition).

We will employ two strategies for evaluating the longer term 24-week changes due to intervention. Unlike during the first 12-week period, there will be no 24-week control group for comparison, due to the 12-week delayed intervention delivery in the control group. As such, first, using the first 12 weeks of the control group as a conservative (because they are likely to have a greater decline over 24 weeks if not for the delayed intervention) reference, we will construct an additional means contrast to compare 24-week OTM changes to 12-week control group changes. Second, we will assess longitudinal changes (during 0-12, 12-24, and 0-24 week periods) in the immediate OTM arm. Means contrasts restricted to the immediate OTM arm will be constructed to estimate changes over 0-12 weeks (anticipated to be highly significant), 0-24 weeks (anticipated to be also significant), and 12-24 weeks (anticipated to be not significant).

9.4.3 ANALYSIS OF THE SECONDARY ENDPOINT(S)

Aim 2 quantitative analysis. General analytic strategy will focus on estimation supplemented by testing exploratory hypotheses comparing intervention fidelity across factors of interest. We will compute appropriate descriptive statistics (means, standard deviations, medians, interquartile ranges, rates, proportions) based on continuous/categorical nature of the metrics and construct 95% confidence intervals. We will consider dichotomizing metrics known to have heavily skewed distributions (e.g. class attendance) or sparse categories prior to analysis. We will compare fidelity metrics across

immediate/delayed nature of the delivery of intervention and organizational-, instructor- and participant-level factors of interest such as organization location, instructor experience, participant baseline characteristics and expectations. We will make comparisons using some of the same methods described above under the baseline comparisons section.

To identify the impact of organizational, instructor and participant level factors on intervention fidelity we will focus on correlating factors potentially impacting fidelity (e.g., AIM, IAM and FIM at pre-, post- and pre- to post-intervention change) to intervention fidelity metrics. We will combine data from both 0-12 weeks in the immediate and 12-24 weeks from delayed groups to make maximal use of the available data. Instead of simple correlations, we will use techniques that account for clustering of classes and participants by facility. Specifically, we will fit a series of linear mixed or GEE models, as appropriate, using intervention fidelity metrics as dependent variables; each of the factors potentially impacting fidelity as the independent variable of interest; and a facility random effect or an exchangeable correlation structure to account for clustering.

Aim 2 qualitative analysis. Thematic analysis[37,38] will be used to develop a list of codes that identify the major emergent conceptual categories in the data. Initially, two coders will read through a subset of the focus group and interview transcripts to identify the major themes. The code lists will then be consolidated into a single codebook by the project manager, to then be approved by the principal investigator and research team. The two coders will then independently apply these codes to all of the focus group and interview transcripts using ATLAS.ti 8, a qualitative data analysis software program that enables multi-coder projects. Krippendorff's Alpha will be used to measure inter-coder reliability, and disagreements in applications of the codes will be adjudicated jointly by the principal investigator and project manager. Based on the results of the adjudication, the project manager will remove invalid code applications from the ATLAS.ti data files and consolidate the two sets of codings into a single set of code applications for the focus group and interview transcripts. To understand the themes discussed in the focus groups and interviews, ATLAS.ti 8 will then be used to generate a list of quotations for each individual code.

Integration of quantitative and qualitative. We will sequentially collect quantitative then qualitative data to provide depth of understanding of the potential impactors of intervention fidelity. We will use joint displays to integrate the quantitative and qualitative data.[39] For example, to better understand the feasibility of delivering OTM we will construct a table that links OTM instructors with high and low scores on the Feasibility of Intervention Measure to specific qualitative quotes gathered from those instructors.

Aim 3 analysis. We will employ a similar strategy to Aim 2 but use pre- to post-intervention change in mobility measures as the dependent variable; intervention fidelity metrics as independent variables of interest; and pre-intervention value of outcome as a covariate. Fidelity metrics will be used in their native continuous (e.g., number of classes attended) or categorical (e.g., adherent vs nonadherent to content) scales in the above models to establish the associations with greatest statistical power. In addition, with alternative dichotomized operational definitions of continuous fidelity metrics using varying cut-points, we will explore whether there are any thresholds that elicit a large difference in

mobility gains, which may indicate a level of fidelity above which effectiveness encounters diminishing returns (e.g., <80% vs ≥80% attendance).

9.4.4 SAFETY ANALYSES

N/A

9.4.5 BASELINE DESCRIPTIVE STATISTICS

Due to the large trial size and randomization scheme balanced with respect to location of facility, it is unlikely that facility and baseline participant characteristics will be significantly different between the arms. However, we will control for any that are found to be significantly different by including them as additional covariates in the sensitivity analyses. In such an event, we will not alter the primary analytic strategy to preserve its a priori nature and predictability. To compare facility characteristics, we will use independent samples *t*-, Wilcoxon rank sum, chi-square and Fisher's exact tests, as appropriate, based on distributional properties and continuous/categorical nature of the characteristic being compared. To compare continuous baseline participant characteristics, we will fit a series of linear mixed models[36] with each one as the dependent variable, intervention arm as a fixed effect, and a site random effect to account for clustering of participants within site. To compare categorical participant characteristics, we will first dichotomize them and fit generalized estimating equations (GEE) models[40,41] with a binomial distribution, logit link function, intervention arm as the independent factor, and an exchangeable working correlation structure to account for clustering.

9.4.6 PLANNED INTERIM ANALYSES

We do not plan to perform any interim analyses for efficacy. We will follow the recommendations of the Safety Officer (SO) to establish the specific rules for interim analyses for safety, either before commencement or in the very early stages of the trial. Specifically, we a priori plan to use the proportion of completed person-time of follow-up as the information statistic; test-statistic for comparison of proportions of (serious) adverse events to define the Brownian motion process; and an alpha-spending function discussed in Lan and DeMets (1983) to define a safety boundary. We will defer to the recommendation of the SO regarding specific time points for interim safety analyses.

9.4.7 SUB-GROUP ANALYSES

N/A

9.4.8 TABULATION OF INDIVIDUAL PARTICIPANT DATA

Individual participant data will not be listed by measure and time point.

9.4.9 EXPLORATORY ANALYSES

We will perform sensitivity/exploratory analyses to examine the robustness of our results against various assumptions and extend findings. We will compute descriptive statistics for metrics of uptake and maintenance beyond the active phase in all facilities. We will perform a series of exploratory subgroup analyses by stratifying the above analyses by various organizational-, instructor- and participant-level characteristics, and their results will be clearly identified as exploratory and interpreted with a hypothesis generating philosophy. We will consider gender as one such stratification criterion. If analyses of residuals from the linear mixed models show violations of statistical assumptions, we will consider fitting models after Box-Cox transforming the continuous variables to make them approximately conform to a normal distribution.[42] As mentioned, we will perform a secondary 12-week efficacy analysis additionally including the 24-week data from the delayed intervention arm. We will also repeat fitting the models after additionally controlling for any participant characteristics found to be significantly different between the intervention arms or deemed important (such as sex). We will also perform a set of sensitivity analyses ignoring missing data and with last-value-carried-forward imputation, although they are generally not recommended.

10 SUPPORTING DOCUMENTATION AND OPERATIONAL CONSIDERATIONS

10.1 REGULATORY, ETHICAL, AND STUDY OVERSIGHT CONSIDERATIONS

10.1.1 INFORMED CONSENT PROCESS

10.1.1.1 CONSENT/ASSENT AND OTHER INFORMATIONAL DOCUMENTS PROVIDED TO PARTICIPANTS

Written informed consent will be obtained prior to conducting the screening exam. One of the study investigators will explain the study to the research subject and the research subject will be given the consent form to read. The investigator will then answer any questions that the subject has about the study. The investigator will ask the subject several questions about the study to make sure they understand the study procedures. Finally, the subject will be asked to sign the consent form.

10.1.1.2 CONSENT PROCEDURES AND DOCUMENTATION

Informed consent will be obtained prior to any data collection. At the initial visit, a study investigator will describe the study to the participant. The participant will also be given the consent document to read. The investigator will answer any questions that the participant might have. The investigator will then ask the participant basic questions about the study (such as time commitment, risks involved, etc.) to determine if the participant understands the study. Once the participant understands the study, written informed consent will be obtained and the investigator will document so in the participant's research record. Individuals who cannot understand what is being asked of them (i.e., the study procedures) will be excluded from the study.

10.1.2 STUDY DISCONTINUATION AND CLOSURE

Participation in the research study will be discontinued if a participant's walking ability worsens; for example, if a subject who has osteoarthritis of the knees exhibits a significant increase in pain that begins to negatively affect their walking ability. A participant can also be removed from the study if any new medical condition, injury or illness is discovered during the course of treatment that would render the participant ineligible based on the exclusion criteria.

If a participant is identified by the research staff to have a temporary medical condition that would impact their participation in the research study (i.e., any of the exclusion criteria), the participant will be required to obtain medical clearance for participation in exercise from their primary care physician prior to returning to the exercise intervention.

10.1.3 CONFIDENTIALITY AND PRIVACY

Research assessments and interventions are conducted onsite at the community-based organizations. Participants will be screened and tested individually to protect their privacy. Participants are informed via the consent process that the exercise class is conducted in a group of 10 participants. Although not completely private, this level of exposure to others during exercise is similar to what one might experience at a community-based fitness class.

The Principal Investigator (Dr. Jennifer Brach) with the assistance of the Co-Investigators will provide ongoing review of the accrued research data to ensure the safety of research participants and the provisions for protecting the privacy of participants and the confidentiality of research data.

Participant’s confidentiality will be protected in the data collection process. All personnel involved with the research have read and signed a Confidentiality statement, and approval is being obtained from the University of Pittsburgh Biomedical Institutional Review Board. Consent forms and data collection forms that identify the participant by name will be stored in a locked cabinet. All computers are password protected. If the data are used in scholarly presentations or journal articles, the investigators will protect the anonymity of individual participants and will report only aggregate data (eg group means) where appropriate. The Principal Investigator will review data confidentiality processes monthly or as indicated with the project staff (Study coordinator and Physical Therapists). The Investigators are all certified in Research Practice Fundamentals, Human Subjects Research Module.

10.1.4 FUTURE USE OF STORED SPECIMENS AND DATA

We recognize with the critical importance of sharing data from this study; thus, our data sharing practices will adhere to the following principles:

1. Data sets will be available at the time the primary paper is accepted for publication or at 2 years after the last participant is followed whichever comes first. The identity of human subjects will be

protected. In all instances, data sets will be stripped of all patient identifiers including the date that human subjects were entered into the study.
2. Data will be available by secure transfer to investigators who complete a data sharing agreement for data from this project.
3. The sharable data will be archived at the study's home site and will be available for sharing by submitting a request to the PI. Once a request is approved by the PI and the requester agrees to the conditions for data use, the data will be transferred.

Because the human subjects included in the project could potentially be identified, given the relatively limited geographic distribution of these subjects, investigators requesting data will be required to sign a data sharing agreement. For external requests, the agreement will be executed through the University's Office of Research, and for internal requests, through email. The data sharing agreement will include the following:

1. The data set may only be used for research, analysis, and statistical reporting.
2. The data set must be kept in a secured environment that allows only authorized users to have access to the data.
3. The data will not be disclosed in a fashion that would make it possible to seek the identity of the human subjects. The data will not be released when the number of observations is less or equal to ten.
4. The individual signing the data sharing agreement will not attempt to link, and will prohibit others from attempting to link, the data to any individual human subjects.
5. The individual signing the data sharing agreement will not attempt to learn, and will prohibit others from learning, the identity of any persons included in the data set.
6. Destruction of the data set after completion of the proposed projects and any required holding period.

We will use NIH guidelines for deidentification under the privacy rule: HIPAA Privacy Rule and Its Impacts on Research (nih.gov).

10.1.5 KEY ROLES AND STUDY GOVERNANCE

Principal Investigator	**Medical Monitor or Independent Safety Monitor**
Jennifer Brach, PhD, PT,FAPTA, Professor	*Lena Makaroun, MD, MS*
University of Pittsburgh	*VA Pittsburgh Healthcare System*
100 Technology Drive, Office 504, Pittsburgh, PA 15219	University Drive C (151C) Building 30 Pittsburgh, PA 15240
412-383-6533	*(412) 360-2253*
jbrach@pitt.edu	*lkm35@pitt.edu*

10.1.6 SAFETY OVERSIGHT

The Principal Investigator (PI) will be responsible for ensuring participants' safety on a daily basis. The Safety Officer (SO) will act in an advisory capacity to the NIA Program staff and the Principal Investigator to monitor participant safety, evaluate the progress of the study, to review procedures for maintaining the confidentiality of data, the quality of data collection, management, and analyses.

The responsibilities of the SO include:

- Review the research protocol, informed consent documents and plans for data safety and monitoring.
- Recommend subject recruitment be initiated after receipt of a satisfactory protocol.
- Evaluate the progress of the trial, including periodic assessments of data quality and timeliness, recruitment, accrual and retention, participant risk versus benefit, performance of the trial sites, and other factors that can affect study outcome.
- Consider factors external to the study when relevant information becomes available, such as scientific or therapeutic developments that may have an impact on the safety of the participants or the ethics of the trial.
- Review study performance, make recommendations and assist in the resolution of problems reported by the Principal Investigator.
- Protect the safety of the study participants.
- Report to NIA on the safety and progress of the trial.
- Make recommendations to the NIA and the Principal Investigator concerning continuation, termination or other modifications of the trial based on the observed beneficial or adverse effects of the treatment under study.
- If appropriate, review interim analyses in accordance with stopping rules, which are clearly defined in advance of data analysis and have the approval of the SO.
- Ensure the confidentiality of the study data and the results of monitoring.
- Assist the NIA by commenting on any problems with study conduct, enrollment, and sample size and/or data collection.

10.1.7 CLINICAL MONITORING

N/A

10.1.8 QUALITY ASSURANCE AND QUALITY CONTROL

Intervention fidelity. Quality will be determined through direct observation of video recordings of the instructor led exercise classes. We will video record two exercise classes of the 24 classes offered (one from the first 6 weeks and one from the second 6 weeks). We will use video recordings of the classes instead of live observation for two reasons: video recording will likely decrease reactivity effects (i.e. instructors performing differently because they are anxious of being watched by an observer) and will permit the reviewers to watch the videos repeatedly and to stop, rewind, and review material multiple times if needed.51 It also makes it more feasible for our content experts to review the recordings when their schedule permits. The content experts will use a standardized checklist to document if the instructor is delivering the program with adequate skill. The checklist includes 20 items such as appropriately

demonstrating exercises, modifying the exercises as needed, and progressing exercises appropriately. Competence will be quantified by the total number of items present on the checklist; scores will range from 0-20. We will ensure the reliability of the content checklist by having two content experts independently review and code a subset of the videos, with any disagreements resolved via discussion or a third reviewer. We will use the videos to also assess the accuracy of the instructor logs by documenting whether the essential program components were completed.

10.1.9 DATA HANDLING AND RECORD KEEPING

10.1.9.1 DATA COLLECTION AND MANAGEMENT RESPONSIBILITIES

Data will be managed by the University of Pittsburgh Physical Therapy Data Center (PTDC) staff. The PTDC will be responsible for creating the web-based application for data entry in collaboration with Drs. Brach and Perera. The Research Electronic Data Capture (REDCap) system, hosted at the University of Pittsburgh, will be used as the web-based data collection and management system for this study. REDCap is a secure, web-based application designed with the flexibility to support data capture for a variety of research projects. Case Report Forms (CRFs) will serve as the basis for the structure of the REDCap database with the data entry screens being as visually similar to the CRFs as much as possible. Blank CRFs, Study Protocol, and the Manual of Procedures (MOP) will be available in the document sharing application in the REDCap system. The PTDC will also be responsible for monitoring data quality as the study is ongoing and merging any data not directly entered into a combined database for the study. All statistical analyses will be performed and/or overseen by the study statistician, Dr. Perera.

10.1.9.2 STUDY RECORDS RETENTION

Data will be stored separately from informed consent data in different database tables. Access to the informed consent data will be restricted using role-based security. Each research data table will only contain the unique study ID assigned to each participant so the participant is not directly identifiable. Information will be securely captured in the password protected project and data will be stored behind the University of Pittsburgh firewall (Pitt CSSD, Network Operations Center).

All paper records related to involvement in this research study will be stored locally in a locked file cabinet. The subject will be indicated by a study ID and the information linking these numbers with the subject's identity will be kept separate from the research records at the study site. All paper research records will be destroyed when such is approved by the sponsor of this study or, as per University policy, at 7 years following study completion, whichever is later.

10.1.10 PROTOCOL DEVIATIONS

This protocol defines a protocol deviation as any noncompliance with the clinical trial protocol, International Council on Harmonisation Good Clinical Practice (ICH GCP), or Manual of Procedures (MOP) requirements. The noncompliance may be either on the part of the participant, the investigator, or the

study site staff. As a result of deviations, corrective actions will be developed by the site and implemented promptly.

These practices are consistent with ICH GCP:

- Section 4.5 Compliance with Protocol, subsections 4.5.1, 4.5.2, and 4.5.3
- Section 5.1 Quality Assurance and Quality Control, subsection 5.1.1
- Section 5.20 Noncompliance, subsections 5.20.1, and 5.20.2.

It will be the responsibility of the site investigator to use continuous vigilance to identify and report deviations within seven working days of identification of the protocol deviation, or within seven working days of the scheduled protocol-required activity. Protocol deviations will be sent to the reviewing Institutional Review Board (IRB) per their policies. The PI will be responsible for knowing and adhering to the reviewing IRB requirements. Further details about the handling of protocol deviations will be included in the MOP.

10.1.11 PUBLICATION AND DATA SHARING POLICY

This study will be conducted in accordance with the following publication and data sharing policies and regulations:

National Institutes of Health (NIH) Public Access Policy, which ensures that the public has access to the published results of NIH funded research. It requires scientists to submit final peer-reviewed journal manuscripts that arise from NIH funds to the digital archive PubMed Central upon acceptance for publication.

This study will comply with the NIH Data Sharing Policy and Policy on the Dissemination of NIH-Funded Clinical Trial Information and the Clinical Trials Registration and Results Information Submission rule. As such, this trial will be registered at ClinicalTrials.gov, and results information from this trial will be submitted to ClinicalTrials.gov. In addition, every attempt will be made to publish results in peer-reviewed journals. Data from this study may be requested from other researchers 2 years after the completion of the primary endpoint by contacting Jennifer Brach (jbrach@pitt.edu). Considerations for ensuring confidentiality of these shared data are described in Section 10.1.3.

10.1.12 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 trial will be disclosed and managed. Furthermore, persons who have a perceived conflict of interest will be required to have such conflicts managed in a way that is appropriate to their participation in the design and conduct of this trial. The study leadership in conjunction with the NIA has established policies

and procedures for all study group members to disclose all conflicts of interest and will establish a mechanism for the management of all reported dualities of interest.

10.2 ADDITIONAL CONSIDERATIONS

N/A

10.3 ABBREVIATIONS AND SPECIAL TERMS

AE	Adverse Event
ANCOVA	Analysis of Covariance
CFR	Code of Federal Regulations
CLIA	Clinical Laboratory Improvement Amendments
CMP	Clinical Monitoring Plan
COC	Certificate of Confidentiality
CONSORT	Consolidated Standards of Reporting Trials
CRF	Case Report Form
DCC	Data Coordinating Center
DHHS	Department of Health and Human Services
DSMB	Data Safety Monitoring Board
DRE	Disease-Related Event
EC	Ethics Committee
eCRF	Electronic Case Report Forms
FDA	Food and Drug Administration
FDAAA	Food and Drug Administration Amendments Act of 2007
FFR	Federal Financial Report
GCP	Good Clinical Practice
GLP	Good Laboratory Practices
GMP	Good Manufacturing Practices
GWAS	Genome-Wide Association Studies
HIPAA	Health Insurance Portability and Accountability Act
IB	Investigator's Brochure
ICH	International Council on Harmonisation
ICMJE	International Committee of Medical Journal Editors
IDE	Investigational Device Exemption
IND	Investigational New Drug Application
IRB	Institutional Review Board
ISM	Independent Safety Monitor
ITT	Intention-To-Treat
LSMEANS	Least-squares Means
MedDRA	Medical Dictionary for Regulatory Activities
MOP	Manual of Procedures
NCT	National Clinical Trial
NIH	National Institutes of Health
NIH IC	NIH Institute or Center
OHRP	Office for Human Research Protections

PI	Principal Investigator
QA	Quality Assurance
QC	Quality Control
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
SMC	Safety Monitoring Committee
SO	Safety Officer
SOA	Schedule of Activities
SOC	System Organ Class
SOP	Standard Operating Procedure
UP	Unanticipated Problem
US	United States

10.4 PROTOCOL AMENDMENT HISTORY

Version	Date	Description of Change	Brief Rationale
1.0.0	22/10/2021	Not applicable; this is the first version of the protocol	

11 REFERENCES

1. Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. *Implement Sci.* 2017;12(1):108.

2. Onken LS, Carroll KM, Shoham V, Cuthbert BN, Riddle M. Reenvisioning clinical science: unifying the discipline to improve the public health. *Clin Psychol Sci.* 2014;2(1):22-34.

3. Brach JS, Perera S, Gilmore S, et al. Effectiveness of a timing and coordination group exercise program to improve mobility in community-dwelling older adults: a randomized clinical trial. *JAMA Intern Med.* 2017;177(10):1437-1444.

4. Brach J, Perera S, Gilmore S, et al. Stakeholder involvement in the design of a patient-centered comparative effectiveness trial of the "On the MOve" group exercise program in community-dwelling older adults. *Contemp Clin Trials.* 2016;50:135-142.

5. Perera S, Mody S, Woodman R, Studenski S. Meaningful change and responsiveness in common physical performance measures in older adults. *J Am Geriatr Soc.* 2006;54:743-749.

6. Wert D, Perera S, Nutini J, Ricci E, Brach J. Motor Control-based Group Exercise: Can It Be Delivered as Effectively by Lay Leaders? . *Translational J of ACSM.* 2018;3:3.

7. Brach JS, VanSwearingen JM, Gil A, et al. Program to improve mobility in aging (PRIMA) study: Methods and rationale of a task-oriented motor learning exercise program. *Contemp Clin Trials.* 2020;89:105912.

8. Brach J, Francois S, VanSwearingen J, Gilmore S, Perera S, Studenski S. Translation of a motor learning walking rehabilitation program into a group-based exercise program for community-dwelling older adults. *PMR.* 2016;8:520-528.

9. Brach JS, Francois SJ, VanSwearingen JM, Gilmore S, Perera S, Studenski SA. Translation of a motor learning walking rehabilitation program into a group-based exercise program for community-dwelling older adults. *Pm r.* 2016;8(6):520-528.

10. Hoffman J, Ciol M, Huynh M, Chan L. Estimating transitions probabilities in mobility and total costs for Medicare beneficiaries. *Arch Phys Med Rehabil.* 2010;91:1849-1855.

11. Perera S, Patel K, Rosano C, et al. Gait speed predicts incident disability: a pooled analysis. *J Gerontol Med Sci.* 2016;71(1):63-71.

12. Studenski S, Perera S, Patel K, et al. Gait speed and survival in older adults. *JAMA.* 2011;305(1):50-58.

13. Hardy S, Perera S, Roumani Y, Chandler J, Studenski S. Improvement in usual gait speed predicts better survival in older adults. *J Am Geriatr Soc.* 2007;55(11):1727-1734.

14. Shuman V, Coyle PC, Perera S, VanSwearingen JM, Albert SM, Brach JS. Association between improved mobility and distal health outcomes. *The Journals of Gerontology: Series A.* 2020.

15. Coyle PC, Perera S, Albert SM, Freburger JK, VanSwearingen JM, Brach JS. Potential long-term impact of "On The Move" group-exercise program on falls and healthcare utilization in older adults: an exploratory analysis of a randomized controlled trial. *BMC Geriatr.* 2020;20(1):105.

16. Li F, Eckstrom E, Harmer P, Fitzgerald K, Voit J, Cameron KA. Exercise and Fall Prevention: Narrowing the Research-to-Practice Gap and Enhancing Integration of Clinical and Community Practice. *J Am Geriatr Soc.* 2016;64(2):425-431.

17. Beedie C, Mann S, Jimenez A, et al. Death by effectiveness: exercise as medicine caught in the efficacy trap! *Br J Sports Med.* 2016;50(6):323-324.

18. Milat AJ, Bauman AE, Redman S, Curac N. Public health research outputs from efficacy to dissemination: a bibliometric analysis. *BMC Public Health.* 2011;11:934.

19. Yeom HA, Keller C, Fleury J. Interventions for promoting mobility in community-dwelling older adults. *J Am Acad Nurse Pract.* 2009;21(2):95-100.

20. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implement Sci.* 2007;2:40.

21. Wiltsey Stirman S, Baumann AA, Miller CJ. The FRAME: an expanded framework for reporting adaptations and modifications to evidence-based interventions. *Implement Sci.* 2019;14(1):58.

22. Hasson H. Systematic evaluation of implmentation fidelity of complex interventions in health and social care. *Implement Sci.* 2010;5(67):9.

23. Rogers EM. *Diffusion of Innovation.* fifth edition ed. New York, NY: Free Press; 2003.

24. Schoenwald SK, Garland AF, Chapman JE, Frazier SL, Sheidow AJ, Southam-Gerow MA. Toward the effective and efficient measurement of implementation fidelity. *Adm Policy Ment Health.* 2011;38(1):32-43.

25. Brach J, Perera S, Gilmore S, et al. Effectiveness of a " timing and coordination group exercise program to improve mobility in community-dwelling older adults: A randomized clinical trial. *JAMA Intern Med.* 2017;177(10):1437-1444.

26. Eldridge SM, Ashby D, Kerry S. Sample size for cluster randomized trials: effect of coefficient of variation of cluster size and analysis method. *Int J Epidemiol.* 2006;35(5):1292-1300.

27. Kerry SM, Bland JM. Trials which randomize practices II: sample size. *Fam Pract.* 1998;15(1):84-87.

28. Chow S-C, Wang H, Shao J. *Sample Size Calculations in CLinical Research.* CRC Press; 2007.

29. Julious SA. *Sample sizes for clinical trials.* Chapman and Hall/CRC; 2009.

30. Machin D, Campbell M, Fayers P, Pinol A. *Sample SIze Tables for Clinical Studies.* Second edition ed: Wiley; 1997.

31. Zar JH. *Biostatistical Analysis.* second edition ed. Englewood Cliffs, New Jersey: Prentice-Hall; 1984.

32. Fleiss JL, Levin B, Paik MC. *Statistical Methods for Rates and Proportions.* third edition ed. New York, New York: John WIley & Sons; 2003.

33. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med.* 1998;17(8):857-872.

34. Graybill F. *An Introduction to Linear Statistical Models.* New York, New York: McGraw-Hill; 1961.

35. Guenther WC. Desk calculations of probabilities for the distribution of the sample correlation coefficient. *The American Statistician.* 1977;31(1):45-48.

36. Milliken GA, Johnson DE. *Designed Experiments.* Vol 1. New York, NY: Van Nostrand Reinhold; 1992.

37. Boyatzis RE. *Trnsforming qualitative information: thematic analysis and code development.* Thousand Oaks, CA: Sage Publications; 1998.

38. Guest G, MacQueen KM, Namey EE. *Applied thematic analysis.* Thousand Oaks, CA: Sage Publications; 2012.

39. Guetterman TC, JW; Kuckartz, U. Using joint displays and MAXQDA software to represent the results of mixed methods research. In: NMcCrudden M, Schraw, G, & Buckendahl, CW, ed. *Use of visual displays in research and testing: Coding, interpreting, and reporting data.* Charlotte, NC: Information Age Publishing; 2015:145-175.

40. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics.* 1986;42(1):121-130.

41. DIggle PJ, Heagerty P, Liang K, Zeger S. *Analysis of Longitudinal Data.* New York City, NY: Oxford University Press; 2002.

42. Box GP, Cox DR. An analysis of transformations. *Journal of the Royal Statistical Society-Series B.* 1964:211-252.