

A Comparison of Treatment Methods for Patients Following
Total Knee Replacement

University of Pittsburgh

PCORI Award Number: CER – 1310-06994

Study Principal Investigator: Dr. Sara Piva

Trial Registration: [ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT02237911) Identifier NCT02237911

08/31/2014

PRINCIPAL INVESTIGATOR (LAST, FIRST, MIDDLE): Piva, Sara, R.

A Comparison of Treatment Methods for Patients Following Total Knee Replacement.

Study Protocol

Design Overview - This comparative effectiveness study is designed to combine patient-centered research questions with rigorous research methods that minimize bias and balance internal and external validity. The pragmatic randomized clinical trial was selected to compare the treatments in ways in which they are delivered in routine clinical practice, so improving generalizability when informing what interventions work best. This will be a three group, single blind, pragmatic randomized clinical trial to answer the question: “In patients at later stages post TKR, given their personal characteristics, what are the comparative effectiveness and harms of the interventions available to them, among individual outpatient rehabilitative exercise, community-based group exercise classes, and usual medical care, especially with regard to physical function and activity?” This study focuses on: (1) outcomes of physical function and activity, both genuinely important to patients, (2) identifying predictors of functional recovery, and (3) potential harms of comparators. The design is depicted in the Figure above. In short, patients post-TKR will be informed about the study by their health care provider or by public announcement. Screening for eligibility will be done in two phases; over the telephone using a questionnaire, and in-person history and physical exam. Following pre-intervention evaluation, subjects will be randomized to one of the 3 groups:

1) clinic-based individual outpatient rehabilitative exercise; 2) community-based group exercise classes; or 3) usual medical care. The usual medical care group will continue their usual care whereas the other two groups will exercise during 12

Table 1- Time points and outcome measures of the study	Baseline	Month 3 and 6	Call Month 1.5 & 4.5
Aim 1			
Physical Function (WOMAC-PF & performance tests)	x	x	
Physical Activity (SWA & CHAMPS)			
Aim 2			
Demographics and Biomedical	x		
Medication	x		
Comorbidities	x		
Psychosocial (fear, anxiety, self-efficacy, depression, coping)	x	x	
Physical Impairments (pain, range of motion, muscle strength)	x	x	
Exploratory Aim			
Adverse Events		x	x
Attrition		x	
Adherence with intervention	monitored each visit		
Co-Interventions		x	x

weeks. Endpoint measures will be assessed prior to and after intervention (3 months), and 6 months after randomization. After the 6 months follow-up, subjects in the usual medical care group will be randomized to either clinic-based individual outpatient rehabilitation exercise or community-based group exercise.

Randomization- Patients will be randomized with a probability 2:1 of receiving one of the exercise interventions as compared to usual medical care. The unbalanced design was chosen because we expect a larger functional recovery in both exercise groups as compared to the usual care group requiring higher sample size in these two arms to be able to detect smaller differences between them compared to larger differences from the usual care arm. The study coordinator will carry out the randomization through a web-based computer system at the end of baseline visit, thereby preserving allocation concealment. We

PRINCIPAL INVESTIGATOR (LAST, FIRST, MIDDLE): Piva, Sara, R.

will use an adaptive randomization to minimize imbalances in important prognostic variables at baseline including gender, age, BMI, physical function, and knee range of motion. The minimal sufficient balance algorithm⁵² will be used as this method accommodates both categorical and continuous variables. Our data center staff have experience creating this type of randomization algorithm within the framework of the clinical trial database, necessary because the allocation is assigned based on the instantaneous imbalances instead of being generated as a fixed list prior to the beginning of the trial. Dr. Moore will monitor randomization and imbalances monthly.

OUTCOMES- The outcomes of this study are entirely patient-centered and selected directly from patient input during structured interviews. They include a broad range of measures as described in Table 1.

Physical Function

- For patient-reported function we will use the Western Ontario and McMaster Universities Osteoarthritis Index Physical Function Subscale (WOMAC-PF). The score at 3 months will be the primary end point. This instrument was developed based on patient input and is the instrument of choice to assess outcome post-TKR.⁵³ The WOMAC-PF consists of 17 items related to physical function. Each item is scored on a 5-point Likert-type Scale with descriptors from 0-4 (none, mild, moderate, severe, and extreme difficulty). Scores of each item are summed for a maximum total score on the WOMAC-PF of 68. Higher scores indicate worse functional limitations. Reliability and validity of this instrument have been established.⁵⁴⁻⁵⁶ Prior to choosing the WOMAC-PF as the primary outcome, we received input from patients. We asked for 15 patients post TKR to compare the WOMAC, the Lower Extremity Functional Scale (LEFS) and the Knee Outcome Survey (KOS). The patients observed that 4 activities of the LEFS (running on even ground, running on uneven ground, making sharp turns while running fast, and hopping) were not applicable as they are unusual and contra-indicated motions after TKR. They also noted that the focus on symptoms of pain and knee buckling of the KOS were not as relevant to them as the activities limitations queries in the WOMAC. Patients selected the WOMAC as the most patient-centered of the 3 instruments, but added that we should also test their true ability to perform tasks (see below).
- For performance-based physical function we will measure a battery of 6 tests easily performed in the clinical setting. This selection was based on the interaction with hundreds of patients post-TKR in the last decade of work and during study design. During these interactions, patients expressed (and we observed) significant limitations in activities such as going up and down stairs, walking long distances, walking fast, balancing on one leg, and standing up from a chair. We will assess the performance of patients in all of these activities. During one of the planning meetings, patients also unanimously stated having difficulty getting up from the floor while gardening, cleaning the house and playing with their grandchildren and suggested that we include it as a test, which we did. So, the tests used in this study cover important domains of walking, dynamic and static balance, mobility, muscle strength and power, endurance, and coordination. These tests are reliable, responsive, and discriminate between low to high functional ability in older adults.^{42,57-62} Description of the tests follows: (1) Self-selected gait speed- measured in m/sec while patients walk at their regular pace over 4 meters. (2) Chair rise- seated in a chair (18" height) without armrests with arms crossed over the chest. Patients are timed during 5 repetitions of rising to a full upright position and sitting back down in the chair without assistance. (3) Single leg stance test – we will record the time of balancing on one leg while keeping the hands on the hips. The test lasts up to 60 sec and is stopped if the swing leg touches the floor, support foot moves on the floor, or arms swing away from the hips. (4) Stair ascend/descend test- patients will be timed while climbing up and down a set of 11 stairs (30 cm depth, 17 cm height) using a handrail on the preferred side. (5) Six min walk test- patients are instructed to cover as much distance as possible during 6 min with the opportunity to stop and rest if required. The test is conducted on an

PRINCIPAL INVESTIGATOR (LAST, FIRST, MIDDLE): Piva, Sara, R.

unobstructed, rectangular circuit (marked in meters). (6) Sitting-rising test- it scores the ability of patients to sit and rise from the floor. Assistive devices will not be permitted during these tests.

Physical Activity- Most patients post-TKR do not reach recommended levels of physical activity to improve health and prevent morbidity.³² In the discussions we had with a number of senior citizens at community centers as well as patients from physical therapy clinics, they invariably complained about the way their knees impacted their daily routines, especially with respect to physical activities that involved substantial walking. Physical activity was included in our research design because patients who participated in a prior study stated repeatedly that they considered that exercise at later stage post TKR helped them to increase physical activity and that it is important to them.

- Real-time physical activity will be measured by the SenseWear Minifly (SWM)(Body Media Inc., Pittsburgh PA). The SWM is an innovative technology that provides real-time measures of physical activity in subject's homes or communities during normal activities of daily life. The SWM has good reliability, validity and has been used by our team.^{41,63-66} Subjects will be instructed to wear the SWM on the back of the right arm during 24 hours/7 days (except during shower or water activities). Traditionally, physical activity has been measured at moderate intensities, which is the default for most monitors. We will record moderate intensity activities, but will also capture data on sedentary behavior and physical activity performed at light intensity (e.g., slow walking, light cycling, cooking, and scrubbing). This is highly relevant to patients post-TKR as they perform most of their activities at light intensities. For example, a patient who before intervention walks for an average of 10 min/day could increase walk time to 30 min/day after intervention. Yet, as this patient walks slowly (light intensity), devices that only capture moderate intensities or above would miss this important improvement. The SenseWear technology is unique in its ability to capture this degree of change in light intensity physical activity.
- Self-reported physical activity will be assessed using the Community Healthy Activities Model Program for Seniors questionnaire (CHAMPS). The CHAMPS is a reliable, valid, and responsive instrument.^{67,68} It assesses activities such as hobbies, work- and social-related activities, walking, swimming, dancing; and will complement the information obtained from the SenseWear technology.

Potential Predictors of Treatment Response- Variables at the baseline that we believe may be potential predictors of treatment response, in addition to physical function (described above) include:

- Demographics and Biomedical Characteristics such as age, gender, race, education, BMI, self-rated health (excellent, good, fair, poor, or bad), discharge placement, number of prior rehabilitation sessions, TKR technique, and surgeon experience will be collected.
- Medication prescribed and over-the-counter used for pain will be recorded in the medication form. We will record the current and highest dosage of any pain medication used during the last month.
- Comorbidity data will be gathered using the Cumulative Illness Rating Scale.⁶⁹ Subjects will be asked to indicate conditions with which they have been formally diagnosed by a physician in the past.
- Psychosocial Measures will include factors that our and other groups have demonstrated to be associated with physical function in knee OA and include:^{48,70-72} 1) Fear; 2) Anxiety; 3) Self-Efficacy; 4) Depression; 5) Coping. We hypothesize that patients with higher fear, anxiety, depression and catastrophizing coping along with low self-efficacy who participated in the community-based group exercise classes will be able to improve physical function, whereas the ones who participated in the individual outpatient exercise will not. Fear-avoidance beliefs will be measured by the Tampa Scale for Kinesiophobia.⁷³ Anxiety will be measured using the Beck Anxiety Index.⁷⁴ Self-efficacy will be measured by the Arthritis Self Efficacy Scale.⁷⁵ Depression will be assessed by the Center for Epidemiologic Studies Short Depression Scale.⁷⁶ Coping will be measured by the Coping Strategy Questionnaire.^{77,78}

PRINCIPAL INVESTIGATOR (LAST, FIRST, MIDDLE): Piva, Sara, R.

- **Knee Impairments** – We hypothesize that patients with increased knee pain, decreased knee range of motion, and lower extremity muscle strength who participated in the outpatient rehabilitation will be able to improve physical function whereas the ones who participated in the community-based group exercise classes will not. Knee pain in the surgical and non-surgical knee will also be measured using an 11-point pain scale. Knee range of motion will be measured by a standard goniometer. We will measure the strength of the muscle groups that have been related to the outcome of TKR^{45,79} including knee extension strength and hip abduction strength measured using an isokinetic dynamometer (Biodex System 4 Pro, Shirley, NY) as we described before.⁴⁵

Capturing Potential Harm of Interventions- Growing evidence suggests that regular physical activity does not have deleterious effects on the survival of the prosthesis.^{80,81} The exercises we propose are not expected to cause any harm because they exert relatively low forces over the prosthesis.⁸²⁻⁸⁵ Regardless of the anticipated safety of the exercises, we will ascertain that interventions offer no harm by closely monitoring adverse events:

- **Adverse Events-** We will capture information on (1) changes in knee pain, swelling, and stiffness; (2) difficulty to bear weight on the surgical leg; (3) falls; (4) if subject has been hospitalized or disabled (> ½ day in bed or required to cut back on routine activities), and; (5) if subject had a TKR on the other knee. For information on falls we will ask subjects if they have had any falls using a 2-3 minute questionnaire developed at the Pittsburgh Pepper Center. One-week test-retest reliability (Kappa) was established as 0.89 in a sub-sample of 43 participants enrolled in a Pepper Center observational cohort study of older adults (Perera S. unpublished data). Falls will be defined as unintentionally coming to rest at a lower position. In this study serious adverse events (SAEs) are defined as hospitalization, death, or permanent disability. Adverse Events (AEs) are defined as exercise-related discomforts that remains for 3 days or longer (e.g., muscle and joint soreness/pain), minor injuries (e.g., strains, sprains), and non-injurious falls. Transient side effects are defined as complaints of increased pain, stiffness, or muscle weakness for periods of 2 days or shorter. We will compare the measures of harm such as adverse events and attrition by treatment interventions. If they differ between groups, they will be used as a covariate in the final outcome analysis.

We will also monitor attrition, adherence and co-interventions:

- **Attrition** is defined in this study as the number of patients dropping out of the study in each group.
- **Adherence** to intervention will be estimated by the proportion of sessions attended in each group and the proportion of patients missing each session. Adherence to the individualized outpatient exercise will be recorded by the physical therapist. Adherence to the group exercise classes will be obtained from reports generated by the community centers.
- **Co-interventions** will be queried during every assessment. Some subjects could decide to seek additional treatment options while enrolled in the study. We will ask about additional treatment sought (e.g., seeking specialized care), and participation in exercises besides the ones prescribed by the study. If the subject participated in additional exercises, we will record the frequency, duration, and type of the exercise and will analyze these data for their potential effect on the primary outcome.

Interventions – The choice of comparators was based on concerns that patients, physicians, physician assistants, physical therapists, and community groups have about the interventions available to them.

Clinic-based individual outpatient rehabilitative exercise: Results of studies on the effectiveness of outpatient rehabilitative exercise at later stage post TKR appear to support this practice.⁸⁻¹¹ Yet, as the evidence is limited, clinicians only prescribe this intervention to a relatively small number of patients. All patients we interviewed unanimously said that outpatient exercise at later stage post TKR was essential to allow them to function at higher levels. Discussions with physicians indicated that they would feel more comfortable increasing the utilization of exercise at later stage post TKR provided there was stronger evidence about its benefits and risks. The exercise program that we will use in this group has been shown

PRINCIPAL INVESTIGATOR (LAST, FIRST, MIDDLE): Piva, Sara, R.

to be safe and feasible, and combines the best research evidence from ours and other groups of investigators.⁸⁻¹¹ Details of the program have been published (<http://www.ncbi.nlm.nih.gov/pubmed/20378678>). Subjects will participate in 12 supervised sessions of exercise followed by a home exercise program. The 12 sessions will be delivered during 3 months in the following schedule: 2 sessions per week during weeks 1-3; 1 session per week in weeks 4 to 7; and 1 session every 2 weeks after that (in week 9 and 11). The gradual weaning will ensure that subjects learn the exercises, which helps increase adherence with the home exercise program. The home program mimics common rehab practice. Patients are instructed to perform at home the same exercises as in the clinic starting after the 3rd week of supervised intervention in a way that they exercise twice a week (either supervised exercise in the clinic or at home) during the 3-month intervention phase. Each exercise session will last about 60 minutes. Treatment sessions will utilize a pragmatic approach and will include the following components:

- Warm-up with stretching of lower extremity muscles and range of motion exercises;
- Moderate to vigorous intensity strengthening exercises of the major lower extremity muscle groups (knee extensors, knee flexors, hip extensors and hip abductors);
- Moderate intensity aerobic training using a treadmill or exercise bicycle;
- Functional activities such as getting up from and sitting down in a chair, squatting, walking in place, kneeling, stair climbing and dancing;
- Agility and balance exercises.

To facilitate adaptation, exercises will initiate at low intensity and will be gradually increased to the target level as tolerated. Exercises will progress only if subjects do not experience increased pain, effusion, or decreased range of knee motion. All components of the exercise program will be used with each subject because they target impairments commonly present post-TKA (muscle weakness, deconditioning, functional limitations, and imbalance). However, the selection of what exercises should be emphasized and their progression will be individualized. For example, during the balance component we may use tandem walking in a patient with difficulty to walk on a straight line. We initiate the tandem walking in a way that is easy for the patient (i.e., little trial variability, stable surface, without manipulation of objects) and progress to more complex activity (i.e., varying speeds and direction of walking, tandem walk over unstable surface such as foam, and manipulate objects such as a carrying a book). We will progress as the subject demonstrates improved skill at the task.

All treatment sessions in this arm of the study will be performed at the Physical Therapy Clinical and Translational Research Center (PT-CTRC) which is an outpatient facility located on the campus of the University of Pittsburgh and the supervised treatments will be provided by physical therapists. The PT-CTRC is located in a building that is on several major mass transit routes, has handicapped accessibility, and has a large covered parking lot.

Community-based group exercise: When we initially designed this study, we had chosen only two comparators: (1) individualized outpatient rehabilitation; and (2) usual medical care. After interviewing numerous patients post-TKR, we modified the research design and included community-based exercise as a comparator. Patients indicated that clinic-based physical therapy was expensive and that they also experienced difficulty getting convenient appointments. Many indicated that they had tried community-based exercise as an alternative means to further improve their surgical outcomes, and that they derived great benefit from it. The inclusion of this comparator is solely based on stakeholders input, as we could not find any peer-reviewed publications that included a community-based exercise arm post TKR.

Patients randomized to the community-based group exercise group will attend exercise classes at local community senior centers at the same frequency/duration as the clinic-based exercise group; 2 times per

PRINCIPAL INVESTIGATOR (LAST, FIRST, MIDDLE): Piva, Sara, R.

week during 3 months. The exercise classes offered at these community centers each last approximately 60 minutes, which matches the time of the clinic-based exercise sessions. The supervised group exercise classes for older adults that the subjects will participate consists of a variety of exercises designed to increase general muscular strength, low impact cardiovascular exercise, range of movement, and activity for daily living skills. The participants in these classes are monitored carefully for any signs of physical discomfort and are allowed to rest whenever necessary. No specific body region is targeted with these exercise classes. The goal is to provide participants with strength, balance, and generalized fitness training for the entire body. Some of the specific exercises included in these classes are: partial squats, leg and knee extension/flexion, elastic tubing for strength training of the upper arm and chest muscles, coordination drills with a gym ball such as bouncing, throwing and catching, and cardiovascular exercise using treadmill, bikes or aerobic series on the floor. The classes are taught by physical fitness instructors and are designed specifically for older adults. The instructors are well trained in emergency procedures and are certified in first aid and CPR. The instructors will be briefed on the research design and special needs of patients post TKR. All of the instructors will meet with the principal investigator quarterly for a review of the research protocols.

We have the support of two of the largest community centers in Pittsburgh that provide these classes and other services to older adults, the Squirrel Hill Jewish Community Center (JCC) and the Vintage Center. These centers will serve as the location sites for this arm of the study. The directors of these centers have agreed to allow us to randomize subjects in this arm of the study to participation in group exercise programs at their facilities. Subjects randomized to this group will be given temporary membership to the Vintage and JCC (see letters of support). These centers are located in two distinct urban settings; the Vintage Center is located in a predominantly low income and African-American neighborhood and the JCC is located in a largely middle-upper income and European-Caucasian neighborhood. Research subjects will be recruited from both of these areas of Pittsburgh and those randomized to this group be given the choice of community center at which they would like to participate in the group exercise classes. Both locations have adequate off-street and metered-street parking, and are located on major mass transit routes.

The partnership with the leaders of the community centers is not new to this study. These collaborations have been established prior to planning the proposed study. We are currently implementing another study on patients with lumbar spinal stenosis in which patients also participate in community based exercise (funded by PCORI). We believe the community centers are valuable and powerful resources that are underutilized by the health system. Although some health insurance companies have begun to cover the costs of group exercise programs such as the *Silver & Fit* and *Silver Sneakers* for their Medicare Advantage subscribers, few older adults take advantage of this benefit. Many older patients with arthritis and other chronic health conditions could benefit to a large degree from community based exercise programs along with other activities offered by these centers. To that end, community based exercise program should be considered a “treatment” and patients should be prescribed this intervention in the same way as they are prescribed any other intervention. This would represent a paradigm shift from more conventional treatment approaches. Therefore, more evidence is necessary to inform payers and policy makers about the potential value of this novel “treatment”. Evidence from this proposed study will complement other research initiatives currently being implemented by our group that may ultimately help to increase community participation and improve the health of patients with arthritis.

Usual medical care: Our goal is to match our study conditions as closely as possible to those experienced in clinical settings, so for patients randomized to the usual medical care group, no attempt will be made to interfere with the usual care received from the doctor or independently sought by the subjects. There is some debate about what constitutes usual medical care at later stage post TKR. Our definition of usual medical care was carefully informed from a literature review and extensive discussions with patients, surgeons, family practitioners and rehabilitation clinicians. Our literature review demonstrated that the

PRINCIPAL INVESTIGATOR (LAST, FIRST, MIDDLE): Piva, Sara, R.

majority of patients are discharged from physical therapy within 2 months post TKR.^{1-4,86,87} Hence, as this study will recruit patients who are at least two months post TKR, these patients will likely no longer be participating in any structured exercise program at the time of study enrollment. To avoid a potential confounder effect, we will exclude patients that are currently participating in structured exercises at the time of study enrollment. Our preliminary studies demonstrated that this exclusion criterion will only minimally affect recruitment. If patients in this group decide by themselves to participate in exercise AFTER they are enrolled into the study, information about frequency, duration, and content of exercises will be recorded and analyzed for their potential effect on the outcomes.

Volunteers assigned to the usual medical care arm will be randomized to an exercise arm after completing the 6-month control period. This will be done based on the request from the patients who helped to design this study. The patients said that they would be less likely to participate in the trial if they would not receive some clear benefit from it. Being able to receive exercise upon completion of the 6-month control period was considered by the patients a good alternative to equalize the potential benefits from study participation. Thus, we made the usual care group an effective wait list control group. This approach is further justified for the following reasons: 1) it is expected to enhance compliance in this group; 2) it addresses the ethical concern of asking volunteers to join an exercise research study and then asking them not to exercise; and 3) data from this group can be included in the analysis of Aim 2 and the exercise data can be included in the analysis of attrition. We believe one of the strengths of the proposed study is the attempt to compare the 3 interventions in the manner they occur in the clinical settings to accurately guide decision-making. After careful consideration of the comparison groups, we also decided against a control group in which we would match the level of attention (hour-per-hour) between the 3 intervention groups because that would deviate from our pragmatic approach to evaluate the interventions as they occur in the real world.

ANALYTIC METHODS

Sample Size

Calculations of sample size were based on the primary endpoint of WOMAC-PF at 3 months. We propose to recruit a total of 240 subjects (96 in each exercise arm and 48 in the usual care group) to allow approximately 86 subjects in each exercise arm and 43 in the usual care group available for a complete case analysis at 3 months (assuming 10% attrition at 3 months). We propose a larger number of subjects in both exercises groups (outpatient rehabilitative exercise and community-based group exercise) because we expect a moderate effect size between these two interventions; whereas we propose a smaller number of subjects in the usual care group because we expect a large effect size between the usual medical care group and either of the exercise groups. Thus, with an alpha level of 0.05, 2 tails test, a sample size of 172 (n=86 in each exercise group) will provide 81% power to detect a difference as small as 3.3-point difference between the two exercise groups in WOMAC-PF (SD of 7.7 from our pilot work).¹¹ This would correspond to a difference in less than 2 activities (out of the 17 queried in the WOMAC-PF) having improved from “about moderate difficulty” to “no difficulty”. The sample size of 43 in the usual medical care group will provide 80% power to detect a difference of 5.2-point difference in WOMAC-PF between the usual medical care group and any exercise group. Power analysis was conducted in NCSS/PASS. Moreover, this sample size will provide adequate power (above 80%) to test the hypothesis related to the secondary outcomes (performance-based physical function and physical activity). For example, we will be able to detect differences between the two exercise groups as small as 0.7 sec in the stair climbing test, 0.4 in the chair-rise test, 0.02 m/sec in walking speed, 1.4 sec in the single leg stance, and 35 min/day in light intensity physical activity (based on SD of our own preliminary data). The differences we will be able to detect between the usual care and any intervention groups are as small as 1.01 sec in the stair climbing test, 0.54 in the chair-rise test, 0.03 m/sec in walking speed, 1.6 sec in the

PRINCIPAL INVESTIGATOR (LAST, FIRST, MIDDLE): Piva, Sara, R.

single leg stance, and 49 min/day in light intensity physical activity. These differences are smaller than the known minimal detectable changes for these outcomes.⁸⁸⁻⁹⁰

For Aim 2, we plan to assess predictors of response based on a minimum change score of 20% in BOTH the WOMAC-PF and the composite score of functional performance at 3 months, thus yielding a binary outcome. For this aim, we will have a combination of those initially randomized to one of the exercise arms and those in the usual care group who were later randomized to the exercise arms for a total of approximately 200. If the expected response rate ranges between 50% and 60%, we would be able to detect an odds ratio of 2.2 to 2.4 with 80% power assuming a binary predictor with 50% split in the sample. In addition, with approximately 100 responders, we will be able to accommodate 9 predictors in the multiple logistic regression based on rule of thumb guidelines for sample size in logistic regression ($\min(\text{response/nonresponse}/(\#\text{predictors}+1)) \geq 10$).

Data Analysis

Dr. Moore will oversee data analyses. We will evaluate the statistical properties of baseline and follow-up outcome measures, including potential outliers, normality and missing data. Measures of central tendency (means, medians, other percentiles) and dispersion (standard deviations, ranges) will be computed for continuous variables, whereas frequency distributions will be calculated for categorical data. Distributions of baseline characteristics will be compared between groups to assess effectiveness of randomization, and statistical or clinical differences will be adjusted in the analyses. Data transformations may be applied if needed and guided by clinical meaningfulness.

Hypothesis 1.1 - *Subjects in Groups 1 and 2 will demonstrate better physical function as compared to Group 3.*

The primary outcome of this aim is the WOMAC-PF at 3 months.⁹¹ This analysis will use an intention-to-treat approach. We will analyze this aim using contrasts from a linear mixed models analysis for 3 and 6 month function controlling for baseline function and the randomization covariates (age, gender, BMI, physical function, ROM). We will first explore the intervention by time interaction, and then proceed to a main effects model with only group and time. Our primary interest is the 3 month comparison between the clinic-based individual outpatient exercise and the community-based exercise groups. The linear mixed models allow us to maximize the number of individuals used for the analyses as a person can contribute information at both time points, or just one time point. This analysis is advantageous to conducting a simple baseline adjusted ANCOVA at 3 months because persons missing 3 month data would not be included. The linear mixed model “borrows” information pertaining to the relationship between the 3 and 6 month outcomes such that persons missing either (but not both) can still be used in the analyses. To test if the improvements in outcomes are sustained, we will use contrasts from the linear mixed model at 6 months. For missing 3 and 6 month outcomes see missing data section.

For the battery of performance-based tests we will perform the analyses as described above, one for each measure. For these secondary outcomes we will use Hochberg’s step-up procedure to control the experiment-wise Type I error rate ($\alpha=0.05$),⁹² which otherwise would be inflated due to the multiple endpoints. Hochberg’s procedure is more powerful than Bonferroni adjustment and performs well when the number of endpoints is small and correlated with small to mid-size correlations. For this outcome we will also perform analysis using a composite score formed with unit-weighted z scores of constituent tests to provide a more stable measure of the subjects’ underlying functional performance.⁹³

Hypothesis 1.2- *Subjects in Groups 1 and 2 will demonstrate increased physical activity as compared to Group 3.*

Adjusted analysis for the outcome of physical activity will parallel the analyses described above for performance-based physical function. The outcomes of this hypothesis are real-time physical activity data captured by the SWM and self-reported physical activity from the CHAMPS questionnaire

PRINCIPAL INVESTIGATOR (LAST, FIRST, MIDDLE): Piva, Sara, R.

Hypothesis 2- *A group of baseline biomedical and psychosocial measures will associate with treatment response.*

Each subject will be classified as a responder or non-responder based on a minimum change score of 20% in BOTH the WOMAC-PF and the composite score of functional performance at 3 months, thus yielding a binary outcome. Baseline variables will be summarized separately for responders and non-responders. Unadjusted odds ratios will be estimated using univariate logistic regression. To consolidate potential predictors, we will test for colinearity among baseline variables that are associated with response. Baseline measures associated with response at the $p < 0.15$ level in unadjusted models will be added to multivariable logistic regression models to assess predictors of treatment response. We will limit the number of predictors going into any one model to no more than one predictor per 10 responses (or 10 non-responses, whichever is less); if more variables are significant, the model will be limited to the most significant variables, after adjusting for those deemed a-priori to be clinically significant. We expect different predictors for each group. For example, patients with worse physical function and impairments (e.g., limited range of motion and muscle weakness) will do better with individual outpatient exercise whereas those with heightened psychosocial factors (e.g., anxiety and depressive symptoms) will respond better to group exercise in the community.^{48,70-72}

Exploratory Hypothesis- *Adherence and co-interventions will be similar in all groups. The attrition rate and adverse events- mainly number of falls- will be lower in Groups 1 and 2 compared to Group 3.*

We will calculate dropout rates as proportions of subjects randomized and as a cumulative probability of remaining in the study using survival analysis techniques such as the product-limit estimator. This statistics can be estimated at various times following randomization and take into account when dropouts occur. Descriptive statistics will be used for reporting and evaluating implementation of the exercise protocols including the proportion in attendance for each session and the average number of sessions attended by group. The two intervention groups require the same number of exercise sessions (2 per week) for 3 months. To assess the impact of non-adherence, we could conduct a per-protocol analysis for the treatment effect but this would likely result in a biased estimate of the treatment effect due to selection bias of those who are more likely to adhere not representing a true random sample of trial participants. We propose to explore using instrumental variable methodology to estimate the efficacy for our interventions in the presence of non-adherence.⁹⁴ Instrumental variables have been mainly used in the econometrics⁹⁵ but have been proposed as a useful tool in estimating dose-response effects in psychological treatments where participants are expected to attend multiple sessions as part of the intervention protocol.⁹⁶ We propose to use the two-stage IV methods which can be easily implemented in Stata software using simple linear structural models for the effect of sessions attended on the primary outcome of function. We will also calculate the 6-month incidence (and 95% CI) of individual adverse events by organ system and relatedness to the study for each group. We will estimate the incidence of adverse events with specific focus on those deemed definitely, probably, or possibly related to interventions. For adverse events, clinical judgments will be considered more important than statistical testing.

Dropout and missing data analyses

During our pilot study 100% of the patients completed the 2 month follow-up assessment;¹¹ however we are realistic and recognize that such retention rate is very optimistic for this study. Thus, in this study we estimate the attrition to be 10% at the 3 months follow-up. We also estimate the attrition to be 15% at 6 months, based on our pilot work and other studies with similar populations and timeframes. We will compare baseline characteristics between patients with and without the assessment at 3 and 6 months to assess potential biases in the complete case analysis. We will also try to obtain reasons for study drop out to assess the missing data mechanism (missing completely at random, missing at random, non-ignorable missingness). We will use several missing data methods for imputing data and re-analyze using intention to treat (as randomized) to assess the impact of missing data on our conclusions as recommended.⁹⁷ We

PRINCIPAL INVESTIGATOR (LAST, FIRST, MIDDLE): Piva, Sara, R.

will first use multiple imputation (with M=10 imputations) which assumes the data are missing at random. Since the data could be missing not at random, we will use another approach of assigning the lowest observed scores for missing values differentially by treatment group (non-ignorable missingness). The approach assumes the missingness is directly related to the value of the missing data, i.e., the people who are missing data on function have worse function scores (did not come in for assessment because function was worse). Results of all approaches to missing data will be presented in the primary paper for our study. If our significance and interpretation of our treatment effect vary depending on the method of imputation, we will view any conclusions cautiously.

PATIENT POPULATION

Detailed reasoning of inclusion criteria is provided in Human Subjects.

Subjects will be included in the study if they underwent a unilateral TKR 2 to 4 months prior to study, are older than 60 years of age, experience functional limitation in daily activities (score in the WOMAC-PF of at least 9 points), *speaks English sufficient to understand study instructions*, are willing to be randomized to one of the 3 treatment groups, and have medical clearance to participate in the study.

Subjects will be excluded if they:

- Have absolute or relative contraindications to exercise testing as established by the American College of Cardiology/American Heart Association;^{99,100}
- Have a history of uncontrolled cardiovascular disease or hypertension;
- Are unable to walk 50 meters without an assistive device and to comfortably bear weight on the surgical knee;
- Have a history of muscular disease (e.g., muscular dystrophy) or neurological disorder that may affect lower extremity function (e.g., CVA, neuropathy, Parkinson's disease, multiple sclerosis);
- Regularly participate in structured exercise;
- Have a terminal illness;
- Are planning to have another joint replacement during the next 12 months;
- Do not plan to be around during the next 12 months (e.g., plan to travel, relocate to another city, etc).

Inclusiveness of Different Populations- The make-up of the patient population of this study will be representative of the general population of older adults who undergo TKR in the US with respect to age, race, socioeconomic status, and general health. We will be sure to specifically target patients who provide important perspectives, such as patients who are ethnically diverse and have multiple chronic conditions or impaired access to healthcare because of age or disabilities. We will include subjects 60 or older because most patients undergo TKR after 60 years of age. While a higher number of patients are undergoing TKR at a younger age, they are still the minority and would be in better health than the older group. We did not impose an upper age limit for eligibility because many patients over 90 years undergo TKR. We believe that as long as they are cleared to do the surgery and they believe the TKR will be helpful to them, they should be in the study. Last, older patients who undergo TKR also have multiple comorbidities such as obesity, diabetes, high BP and arthritis of other joints^{19,20,101} and they will be included. Enrolling older adults with coexisting comorbidities will improve generalizability and clinical relevance of study results. With respect to race and socioeconomic status, we have been very successful recruiting an adequate proportion of African American and lower income research subjects and, as in the proposed study we will use similar recruitment strategies we expect similar success.