eSupplement 1
Study Protocol

2
Strength Training for ARthritis Trial (START)

Principal Investigator
Stephen P. Messier, Ph.D. Health and Exercise Science, Wake Forest University

Co-Investigators
Daniel Beavers, Ph.D. Department of Public Health Sciences, Wake Forest School of Medicine
Kim Bennell, Ph.D. University of Melbourne
Jeff Carr, M.D. Department of Radiology and Radiologic Sciences, Vanderbilt University
Paul DeVita, Ph.D. East Carolina University
Ali Guermazi, M.D. Boston University Medical Center
David Hunter, M.D. University of Sydney
Richard Loeser, M.D. Division of Rheumatology, University of North Carolina at Chapel Hill
Shannon Mihalko, Ph.D. Health and Exercise Science, Wake Forest University
Barbara Nicklas, Ph.D. Center for Aging, Wake Forest School of Medicine
Sandra Schultz, Ph.D. University of North Carolina at Greensboro
Jeff Williamson, M.D. Center for Aging, Wake Forest School of Medicine

Funding
National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
1. Hypotheses & Specific Aims

Primary Hypotheses (1.1 - 1.2) and Aims

**Hypothesis 1.1:** High-intensity (75-90% 1RM), long-term (18-month) strength training in older adults with knee OA will decrease knee pain compared to low-intensity (30-40% 1RM) strength training and attention control groups of equal duration.

**Aim 1.1** To determine if an 18-month, high-intensity strength-training intervention significantly decreases pain relative to low-intensity strength training and attention control groups.

**Hypothesis 1.2:** High-intensity, long-term strength training in older adults with knee OA will decrease knee-joint compressive forces during walking, a mechanism that affects the OA disease pathway, compared to low-intensity strength training and attention-control groups of equal duration.

**Aim 1.2** To compare the effects of 18 months of high-intensity strength training, low-intensity strength training, and attention control on knee-joint compressive forces during walking.

Secondary Hypotheses (2-5) and Aims

**Hypothesis 2.1:** *Short-Term Effect (6 months).* The initial clinical and mechanistic effects of high-intensity strength training will be superior to those of low-intensity strength training and an attention control.

**Aim 2.1.** To determine if 6 months of high-intensity strength training reduces pain, improves function, and mobility, and reduces knee-joint compressive loads significantly more than 6 months of low-intensity strength training or an attention control.

**Hypothesis 2.2:** *Long-Term Effect (18 months).* Combining longer duration (18 months) with high-intensity strength training will improve function and mobility (important clinical outcomes in addition to pain), and additional measures of knee-joint loads significantly more than low-intensity strength training or attention control.

**Aim 2.2.** To compare the clinical (function, mobility) and mechanistic (knee adductor moment, knee AP shear force) effects of the 3 interventions at 18-month follow-up.

**Aim 2.3.** To compare the effects of the 3 interventions on clinical (e.g., pain, function, mobility) and mechanistic (knee joint loads) outcomes at 6- and 18-month follow-up.

**Hypothesis 3:** High-intensity, long-term strength training will improve structural outcomes more than low-intensity strength training or attention control in older adults with knee OA.

**Aim 3.1.** To compare the effects of 18-month, high-intensity strength training, low-intensity strength training, and attention control on OA progression by changes in x-ray (e.g., joint space width).

**Hypothesis 4:** High-intensity, long-term strength training will result in improved thigh composition (increased muscle and decreased fat volume), and improved muscle function compared to low-intensity strength-training and attention control.

**Aim 4.1.** To compare the effects on thigh muscle and fat volume of 18-month, high-intensity strength training, low-intensity strength training, and an attention control.

**Aim 4.2.** To compare the effects of the interventions on components of muscle function, including hip abductor and quadriceps strength, and muscle power.

**Hypothesis 5:** High-intensity, long-term strength training will result in lower systemic concentrations of inflammatory and OA biomarkers than low-intensity strength training or attention control.

**Aim 5.1.** To compare the effects of the interventions on inflammation markers (IL-6, TNFα, sTNFR1, leptin) and OA biomarkers (serum PIIANP, COMP, urinary levels of CTX-II).
2. Research Strategy

2.a. Significance
By 2030, an estimated 67 million American adults will report physician-diagnosed arthritis—a 40% increase in 25 years (44). Osteoarthritis (OA) is the most common form and the leading cause of disability among older adults; its prevalence in the US is estimated at 27 million. Knee OA accounts for a significant portion of this disability, and is largely due to factors that alter knee-joint loading. Results from this project will significantly inform future management of patients suffering from knee OA with enormous public health implications.

Muscle loss and fat gain contribute to the disability, pain, and morbidity associated with knee OA (54), and thigh muscle weakness is an independent, modifiable risk factor (48; 117). While treatment guidelines recommend strengthening to combat sarcopenia in knee OA patients (2; 133), the intensities or loads (defined as percent of one repetition maximum, or %1RM) used in previous studies were below those recommended by the American College of Sports Medicine (62) (60-80% 1RM). Further, they were generally short, between 6 and 24 weeks (35; 57; 61; 75; 85; 90; 97; 117); effect sizes were low-to-modest, progression could not be detected, and they provided little lasting clinical benefit. Indeed, short-term exercise benefits are gone as soon as 6 months without exercise (10; 98; 123) but sustained 2 years after long-term supervised treatment ends (98). Few have studied the effectiveness of more intense strength training due to the unsubstantiated belief that it might exacerbate OA symptoms. Preliminary studies, including our pilot study (c.2.1), indicate that high-intensity strength training is safe and well tolerated by healthy older adults (14; 38) and knee OA patients (61).

Obesity is also a major risk factor for knee OA, but not all patients may benefit from or be able to achieve and sustain significant weight loss. Intensive strength training can change thigh composition in older adults and has shown promise in treating the underlying biomechanical (knee-joint loading) and inflammatory disease pathways. Studies in healthy older adults associate it with increased fat-free thigh mass and quadriceps cross-sectional area and decreased percent body fat and thigh subcutaneous fat with minimal alteration in total body weight (63; 86; 113; 114). Sipla and Suominen (114) and Ferri et al. (38) noted increased quadriceps cross-sectional area and lean cross-sectional area, total muscle lean tissue, and mean calf muscle attenuation (less intramuscular fat) in 16-18 weeks. Treuth et al. (121-123; c.2.5) found significant increases in thigh muscle mass and decreased thigh fat mass after 16 weeks of high-intensity strength training in older men and women.

![Figure 1. Hypothesized pathways mediating high-intensity strength-training outcomes at 6 and 18 mos.](image-url)
Our preliminary studies significantly associate reduced thigh, but not trunk, fat with less pain and greater physical function and quadriceps strength. After adjusting for sex and walk speed, knee-joint loads were 5-9 times larger for thigh fat than abdominal fat (c.2.6). High-intensity strength-training also reduced IL-18, a pro-inflammatory cytokine, in HIV-infected patients (71) and IL-6 and CRP levels in older adults with chronic kidney disease compared to controls (15). We must now gather clinical and mechanistic evidence that improved thigh muscle quality has long-term protective effects on joint mechanics, inflammation, and structural progression in knee OA. We expect initial improvements in thigh muscle function, pain, and knee-joint loading with high-intensity strength training after 6 months; 18 months will determine, for the first time, if further changes in thigh muscle function and composition significantly reduce knee-joint forces and inflammatory cytokines resulting in attenuated OA disease progression and a greater decrease in pain (Figure 1).

We will conduct the first long-term clinical trial comparing the efficacy of high- (75-90% 1RM) to low-intensity (30-40%1RM) strength-training and attention-control interventions in older adults with knee OA. It will identify the mechanisms responsible for any changes in structural joint damage, pain, and function and any clinically meaningful benefits. Given the prevalence of OA, the detrimental effects of sarcopenia and obesity, the difficulty many older adults have in losing weight and keeping it off, and the safety and widespread availability of the intervention, this trial has immediate, potentially transformative clinical impact.

2.b. Innovation
The US health system is predicated on providing acute, episodic care that cannot address the altered patterns of chronic disease now facing the public (118). OA is a common chronic disease needing long-term treatments with minimal adverse side effects. Short-term pain trials cannot assess long-term needs. Building on the results of our ongoing intensive weight-loss trial, IDEA, our studies may shape long-term treatment of the major chronic health problem of the future. The proposed study is uniquely designed and equipped to identify a practical, efficacious, nonpharmacologic therapy capable of slowing OA disease progression.

1. We will be the first to combine long duration and high-intensity (75-90% of 1 RM) strength training in a knee OA population to determine if it significantly affects (1) thigh composition, (2) long-term clinical outcomes, such as pain, function, and mobility, and (3) biomechanical and inflammatory pathways to slow progression.
2. The NIH Obesity Task Force prioritized research into the various fat depots that contribute to serious health conditions (88). We will be the first to determine whether changes in thigh composition mediate knee OA progression through effects on knee-joint loads, as our pilot data suggest.
3. Our efficacy study will be the first to determine the proper intensity and effects of long-term strength training on knee OA.
4. Chang et al. (17) observed that greater hip internal abductor moments based on greater strength reduced medial tibiofemoral (TF) OA progression by 50%. We will be the first to use high-intensity strength training targeting hip and knee musculature and predict it will reduce knee-joint loads both by strengthening hip abductors, quadriceps, and hamstrings and changing thigh muscle composition (less fat and more muscle).
5. We will be the first to directly compare the effects of duration in a high-intensity strength-training cohort, but even short-term improvements in pain and function will be clinically important.
6. The dataset generated by this study will advance our novel work in biomechanical modeling of a pathologic population to develop and fine-tune new therapeutic strategies.

2.c. Approach

c.1. Investigators. Our investigative team’s expertise has advanced understanding of nonpharmacologic treatment for knee OA in 3 successful, large-scale, randomized clinical trials (RCT): Fitness Arthritis in Seniors Trial (FAST), Arthritis Diet and Activity Promotion Trial (ADAPT), and the ongoing Intensive Diet and Exercise for Arthritis (IDEA). We routinely exceed our recruitment goals, including minorities; retain more than 80% of randomized participants over the long-term; and develop and implement safe, effective exercise interventions.

Dr. Messier (PI) and co-investigators are well known for their research on the effects of exercise and weight loss on gait, strength, function, and pain in knee OA. Dr. Messier has over 20 years’ experience in NIH-funded OA clinical trials as co-PI of FAST and PI of ADAPT and IDEA. The Arthritis Foundation named the primary outcome paper from ADAPT one of the top 10 advances for 2004.

c.2. Preliminary Studies
c.2.1. Strength Training for ARthritis Trial (START): pilot study. Messier, Legault, Loeser, Mihalko, DeVita, Carr, Nicklas. This study showed that older adults with knee OA can tolerate high-intensity strength training. We randomized 21 aged ≥55 yrs with self-reported knee OA disability into a high (H)-(70-85%) 1RM) or low (L)-intensity (30-40% 1RM) program. Both were 1 hr/d, 3 d/wk for 16 wks. Compliance for the H group was 86% with 2 dropouts (reasons: retirement; medical not associated with intervention) and 82% for the L group with 2 dropouts (reasons: traveling; spouse). Follow-up data showed that the H group had less pain (H:4.8±0.8 vs L:6.4±0.7: ES = 0.71), better function (H:12.5±2.5 vs L:15.7±2.2: ES = 0.45), and faster walk speed (H:1.08±0.04 m/s vs L:1.01±0.04 m/s: ES = 0.58). Weight loss was H: 0.2 kg; L: 0.9 kg. The small sample limited detection of significant between-group differences. Mean differences and effect sizes support our model (Fig. 1): a longer intervention and larger sample should improve the H group’s better clinical outcomes.

c.2.2. A randomized trial comparing aerobic exercise and resistance exercise with a health education program in older adults with knee osteoarthritis (FAST). Messier et al. JAMA 1997; 277:25-31.(35) Retention was 82%. Results showed that long-term, low-intensity strength training for older adults with knee OA was well tolerated, improved pain and function, and should be part of the standard-of-care.

c.2.3. Function, pain, and the modifying effects of regional adiposity and strength in older adults with knee OA. Messier, Legault, Loeser, Mihalko, DeVita, Carr, Nicklas. We recruited 34 older adults with radiographic knee OA to determine the associations among regional adiposity, strength, pain, and function. Lower % lower extremity, but not % trunk, fat was significantly associated with reduced pain (r = 0.55, p= 0.02), greater physical function (r = 0.66, p = 0.003), and higher knee-extensor strength (r = -0.58, p = 0.009).

c.2.4. Intensive Diet and Exercise for Arthritis (IDEA) (1R01AR052528-01). Messier, Legault, Loeser, Mihalko, DeVita, Carr, Nicklas, Hunter, Williamson, Eckstein(82). This 5-year NIAMS-funded study tests the hypothesis that intensive weight loss will reduce inflammation and joint loads enough to alter OA disease progression, either with or without exercise (walking and low-intensity strength training). The 454 overweight and obese (BMI = 27-40 kg/m²), older (age ≥ 55 yrs) participants with TF OA are randomized to one of three 18-month interventions: intensive dietary restriction-plus-exercise; exercise-only control; or intensive dietary restriction-only. The primary aims compare effects on inflammatory biomarkers and knee-joint loads. We randomized 454 to exceed our recruitment goal of 450. Minority recruitment is 19%.

c.2.5. Strength-training effects on total and regional body composition in older men. Nicklas et al. (123). High-intensity strength training in men aged = 60± 4 yrs for 16 weeks significantly (6.6%) increased muscle cross-sectional area and significantly (9.0%) reduced subcutaneous fat in the midthigh without change in total body weight. Thigh muscle and fat mass did not change in the sedentary control group. These findings show, for the first time, that high-intensity strength training can significantly improve thigh composition in older adults.

c.2.6. Relationship between fat depots and knee-joint loads in knee OA. Messier, Legault, DeVita, Loeser, Carr, Nicklas. This preliminary study used CT to determine associations between thigh and fat depots and knee-joint loads in older, overweight and obese adults with knee OA (n = 174). External knee moments and joint forces were calculated using musculoskeletal modeling. After adjusting for gender and walk speed, total fat was significantly associated with peak knee compressive and shear forces and extension moments. β values were driven by strongly significant (p < 0.05) relationships between subcutaneous fat depots and knee-joint forces and were 5-9 times larger for thigh than abdominal fat. These data are the first to associate specific regional fat depots, especially thigh subcutaneous fat, with knee-joint loads.

c.2.7. Arthritis, Diet, and Activity Promotion Trial (ADAPT). Messier et al. ADAPT was a randomized, single-blind, 18-month clinical trial to determine whether long-term exercise (walking plus low-intensity strength training) and dietary weight loss are more effective, separately or combined, than usual care in improving physical function (primary outcome), pain, and mobility in 316 community-dwelling adults age ≥60 yrs with BMI≥28, knee pain, radiographic evidence of knee OA, and self-reported physical disability. They were randomized into healthy lifestyle control (HL), diet (D), exercise (E), and diet-plus-exercise (D+E) groups; 80% completed the study (252). Adherence was: HL, 73%; D, 72%; E, 60%; and D+E, 64%. The database launched 19 publications; 3 are noted. Exercise and dietary weight loss… (2004) Arthritis Rheum. (84) Messier et al.
This primary outcome paper indicated that modest weight loss combined with moderate exercise provided the best overall improvement in function, pain, and mobility with no adverse effects on disease progression.

**Effects of weight loss on knee-joint loads...** *(2005) Arthritis Rheum* *(81)* Messier, DeVita, et al. Results indicate that weight loss reduces stress on the affected knee; support inclusion of knee-joint loads as a primary mechanistic aim; and demonstrate our expertise in modeling to estimate bone-on-bone joint forces.

**Effect of an exercise and dietary intervention on serum biomarkers** *(2008) Osteoarthritis Cartilage,* *(18)* Loeser, Messier, Legault, et al. Intervention groups differed significantly at 18-month follow-up for TGF-β1 (p = 0.02), with lower levels among the diet-only group than the healthy lifestyle group. *Data on HA, COMP, and KS suggest no measurable harm to joint tissues from 18 months of weight-bearing exercise.*

c.2.8. Optimizing Body Composition for Function in Older Adults. Nicklas, Carr, et al. This preliminary study compared the effects of a 4-month weight-loss intervention, alone (WL) and with high-intensity (70% 1RM) resistance training (WL+RT) in 40 obese women and 48 obese men aged 65-79 yrs. Retention was 92%, and compliance to exercise, 84%. Individuals assigned to RT lost 30% less lean mass (WL: -2.5 ± 2.0 kg vs WL+RT: -1.7 ± 2.0 kg), and their leg power was significantly greater (p = 0.0001).

c.2.9. Summary. These studies demonstrate our expertise in designing and implementing RCTs that include high- and low-intensity strength training; achieving outstanding compliance and retention rates; measuring fat depots and muscle mass using CT; calculating knee-joint loads using musculoskeletal modeling; analyzing serum levels of pro-inflammatory biomarkers and serum and urinary levels of OA biomarkers; measuring functional status using a variety of reliable and valid tests and questionnaires; and provide strong evidence in support of our aims.

### 3. Research Design and Methods

**3.a. Study Design.** We will randomize 372 adults age ≥50 yrs with self-reported disability due to knee OA into one of 3 groups: high-intensity strength training, low-intensity strength training, and attention control. We now propose an 18-month, high-intensity strength-training intervention for older adults with knee OA, focused on improving thigh composition (more muscle and less fat). The study duration will be 20 months including testing visits. The primary clinical aim is to compare the interventions’ effects on knee pain, and the primary mechanistic aim is to compare their effects on knee-joint compressive forces during walking, a mechanism that affects the OA disease pathway. Secondary aims will compare intervention effects on additional clinical measures of disease severity (e.g., function, mobility); thigh muscle and fat volume, measured by CT; components of thigh muscle function, including hip abductor strength and quadriceps strength, and power disease progression, additional measures of knee-joint loading; and inflammatory and OA biomarkers.

**3.b. Measures**

b.1. **Western Ontario McMasters Universities Osteoarthritis Index (WOMAC).** We will measure self-reported physical function and pain (primary clinical outcome) using the LK version of WOMAC *(9)*. The pain index assesses participants’ pain on the same scale, ranging from 0 (none) to 4 (extreme). The pain subscale consists of 5 items and total scores can range from 0-20, with larger scores indicating greater pain. Participants must score 4 or greater in order to be included in the study. The LK function subscale asks participants to indicate on a scale from 0 (none) to 4 (extreme) the degree of difficulty experienced in the last 48 hours due to knee OA. Individual scores for the 17 items are totaled to generate a summary score that could range from 0-68, with higher scores indicating poorer function. This instrument has been validated and recommended by the Osteoarthritis Research Society as the health status measure of choice in older adults with knee OA.

b.2. **Gait.** *The primary mechanistic outcome is maximal knee compressive force; secondary outcomes include external knee adduction moment and AP shear force* *(5; 6; 80; 81; 108).* A 25-reflective marker set, 6-camera Motion Analysis System (60 Hz), and 6-channel force plate (AMTI, 480 Hz) will obtain 3D kinematic and kinetic gait data. The former will be processed using EVaRT 4.4 software and a Butterworth low-pass filter (6 Hz cutoff). For each participant, 3 of 6 successful trials will be analyzed; i.e., within ±3.5% of the participant’s freely chosen speed, and the entire foot must contact the force plate in a visually normal stride. Orthotrak 6.0 β4 gait analysis software will generate hip, knee, and ankle data. Smoothed coordinate data, ground reaction,
and gravitational and inertial forces will inform an inverse dynamics model to calculate 3D moments and forces at hip, knee, and ankle joints. These moments and forces will be used in the knee model developed by DeVita et al. (24) for use in knee OA subjects (80; 81). Isolated activation of the quadriceps muscles is insufficient to balance the external adductor moment (106); co-contraction of hamstrings and tension in lateral soft tissue (IT Band, lateral collateral ligament) are required to maintain frontal plane equilibrium. Shelburne et al. (111) found that much of the internal abductor moment that counteracts the external adductor moment comes from the quadriceps and gastrocnemius, which are included in our model. Our test-retest reliability intraclass correlations (ICC) for 21 knee OA patients with mean age 65.7 yrs (SD = 5.8) were r = 0.86 for external peak knee flexor moment, r = 0.94 for external peak adductor moment, and r = 0.95 for peak knee compressive force (82). Lab technicians will be blinded to treatment group.

b.3. Screening. The Eligibility Questionnaire will address joint pain, physical function, activity level, co-morbid diseases, willingness to participate for 18 months, height and weight (to determine BMI), caregiver status, status of significant others, and distance of home from the center (within 50-mile radius).

b.4. Medical History. The medical history and self-administered comorbidities questionnaires will obtain information regarding past and current health status. The medications questionnaire, adapted from the ARIC study and widely used in field research and our studies, is designed to obtain information about all prescription and over-the-counter medicines and supplements used during the 2 weeks prior to interview.

b.5. Physical Exam. A study physician or physician assistant will perform a routine medical exam which will include a review of medical history and current medications and blood pressure.

b.6. Cognitive Functioning. The Montreal Cognitive Assessment (MOCA) will measure cognitive functioning. At baseline; a score of <20 will be evaluated by the study physician, who will determine eligibility. The Digit Symbol Substitution Test (DSST) measures motor persistence, sustained attention, response speed and visual motor coordination (104).

b.7. Measures of Quality of Life. The SF-36 (128) is the most widely used and carefully validated measure of HRQL. It yields two broad summary scores: physical health and mental health. Mood is indicative of quality of life. Depression will be measured using the Center for Epidemiologic Studies Depression Scale (94), which is designed to measure depressive symptoms experienced during the previous week. Persons scoring >17 on the CES-D depression scale will be evaluated by the study physician, who will determine eligibility. The FAST-23 will be used to measure physical disability. A self-efficacy for adherence measure (76) will be used to assess beliefs in one’s ability (confidence) to continue exercising at various intensities and frequencies. The Positive and Negative Affect (PANAS) measures both positive and negative affect, leading to more insightful outlooks regarding participants’ feeling states. This scale consists of 20 items that reflect the intensity of how the participant “feels” right now (130). Among the various components of subjective well-being, the satisfaction with life scale is narrowly focused to assess global life satisfaction. Confidence may be an important predictor of exercise behavior. Locus of control indicates whether or not an individual attributes the cause of certain obstacles to oneself or to one’s surroundings. Measuring locus of control may help to discern why or why not an individual is adhering to an exercise program. The ABC questionnaire will also be used to measure one’s confidence. The Physical Self Perception Profile (PSPP) is a 30-item instrument used to assess self-esteem relative to several domains of physical functioning in a hierarchical, multidimensional fashion (39).

b.8. Mobility. Our measure will be 6-min walk distance. Participants are told to walk as far as possible in 6 minutes on an established course. No personal timing devices are permitted, and participants are not provided feedback during the test. Results are significantly correlated to treadmill time and symptom-limited maximal oxygen consumption (r = 0.52 and r = 0.53, respectively) and have a 3-month test-retest reliability of 0.86 (94).

b.9. Body Composition by DXA. Dual-energy x-ray absorptiometry (DXA), using a fan-beam scanner, Delphi A™ (Hologic, Waltham, MA), will measure whole body changes in total fat and lean mass (85; 123; 124). Percent coefficients of variation (%CV) are 1.2% for whole body FM; 0.5% for whole body LM; 0.9% for whole body BMD; 1.2% for PA spine BMD; and 0.9% for total hip BMD.
b.10. Body weight, height, hip/waist circumference, BMI. Body weight, height, and hip and waist circumference will be obtained at baseline and follow up visits using standard techniques. The range of BMIs for inclusion is 20 to 44.9 kg/m².

b.11. Thigh Composition by Computed Tomography. Bilateral volumetric measures of thigh adipose tissue and skeletal muscle will use a standard CT protocol. Participants will be placed supine on the CT couch with their legs held in a neutral position by dedicated Velcro straps. A calibration phantom (Image Analysis, Columbia, KY) with known CT densities will be placed in the scan fields of view (FOV) posterior to the legs. A toponogram of the femur will be obtained to measure its length from the superior aspect of the head to the inferior aspect of the medial condyle. This length will be trisection, and the proximal and mid third junction determined. The scan’s start location will be 25 mm above and the end 25 mm below this point, providing a total 50 mm of axial plane images, using 120 KV, 150 mAs, full reconstructions, standard and bone reconstruction kernels, 1.25 mm slice thickness, and display FOV (dfov) of 50 cm and 35 cm (both thighs) and 20 cm (center, right and left leg) to ensure complete coverage of the subcutaneous tissue, even in the largest participants. The high-resolution 20-cm dfov target each femur for potential future analysis of cortical bone structure.

CT image analysis software and protocol: CT images will be analyzed (reader masked to treatment group) using NIH’s Center for Information Technology Medical Image Processing, Analysis, and Visualization (MIPAV) software with custom subroutines developed by Dr. Carr’s group. This approach enhances workflow, consistency, and accuracy (79). We have extensive experience in analyzing thigh CT images for adipose tissue (total, subcutaneous, intermuscular) and muscle (volume and quality based on CT attenuation) and have demonstrated very high precision and reproducibility for changes in phenotypes.

b.12. Blood and Urine Sample Collection. Blood samples (~50 ml per visit) for assessing biomarkers will be collected via venipuncture at a specific time in the morning at least 2 hours after rising and a 12-hour fast at baseline and at 6- and 18-month assessment visits. Urine samples (second am void, 20 ml per visit) will be collected in 250 ml specimen cups by each participant for analysis of new and emerging OA biomarkers. Serum, plasma, and urine will be aliquoted and stored at -80°C until analysis.

Inflammatory Markers. IL-6, TNFα, sTNFR1, and leptin were chosen for their known implication in OA (31; 65; 72; 90; 91; 101). They have been shown to change with 1 year of moderate strength training (91). IL-6 is our primary inflammation measure. All inflammatory marker assays will be performed in the WFUHS Cytokine Core Laboratory under Dr. Nicklas. All samples will be measured in duplicate, using the average for analyses. Commercially available (R&D Systems, Minneapolis, MN) enzyme-linked immunosorbent assay (ELISA) kits will be used: high-sensitivity Quantikine® for IL-6. In our laboratory, inter- and intra-assay coefficients of variation (CV) for IL-6 are 5.4% and 3.5%, respectively; for TNFα, 11.8% and 6.2%, respectively; and under 5% for the soluble receptor assays.

OA Biomarkers. Serum samples will be assayed for COMP using ELISA (AnaMar Medical, Uppsala, SW). We will measure serum levels of the N-propeptide of type IIA procollagen to determine type II collagen synthesis and urine levels of a C-terminal crosslinking telopeptide of type II collagen (CTX-II or CartiLaps®) to determine degradation (40) using ELISA (Nordic Biosciences, Herlev, DK). Urine CTX-II results will be corrected for creatinine levels measured by a standard colorimetric method. We will also store aliquots to test for promising new inflammatory and OA biomarkers that may become available during the study.


Strength. Knee flexion/extension concentric and eccentric strength will be assessed at baseline, FU6, FU12, and FU18 using a Kin-Com 125E isokinetic dynamometer set to 30 deg/sec. Strength test-retest reliability for 5 men and 5 women tested twice in our lab, 7-10 days apart, had an ICC between 0.75 and 0.93 (83). Since we suggest that intensive strength training can reduce knee-joint loads by counterbalancing the external knee-adductor moment with strong hip abductors, hip-abductor isometric strength will be measured using a handheld dynamometer (Lafayette Instruments, Lafayette, IN) with a load-cell system to measure static force, as described by Dierks et al. (25). Participants lie on their sides on a padded treatment table with a pillow between their legs to slightly abduct the hip and the dynamometer force pad 5 cm proximal to the lateral knee-joint line. A strap attached to the table secures the dynamometer to the lower extremity and eliminates any influence of tester strength. Jaramillo et al. (59) reported a test-retest ICC (2.1) of 0.95 for this test. We are using it in our pilot study (2.c.2.1). We will alternate the first joint (hip vs. knee) tested between participants.
Power. The Nottingham power rig will be used to measure bilateral leg extensor power because it correlates well with such functional measures as chair-rise, stair-climbing, and walking speed in elderly subjects (7). This measurement is safe and acceptable for all age groups (7; 8).

b.14. X-ray. Bilateral anteroposterior (AP) weight-bearing knee x-rays using a positioning device and the modified Lyon-Schuss technique (76) will be used to identify TF OA and skyline views to identify PF OA. The former will be repeated at FU18 to assess changes in joint-space width (JSW) and Kellgren-Lawrence (K-L) score. At baseline the AP x-ray will be taken at 3 different beam angles (5, 10, and 15 degrees) to ensure that the optimal beam angle is used for each participant. The angle that is chosen as the optimal beam angle will be used for the follow up x-ray. We will exclude people with severe PF OA (JSN = 3 on OARSI scale) and control for severity (none-to-moderate) in statistical analyses. We define medial TF disease based on our previous definitions (132), definite radiographic OA (KL=2-3), plus at least grade-1 medial JS narrowing (0-3 scale) using the OARSI atlas. Because our goal is to enter subjects with a varus alignment, subjects with a valgus orientation greater than 2° (outward) on the short AP films will be excluded prior to obtaining full-length x-rays for measurement of the mechanical axis (115).

To assess alignment, a full-length AP radiograph of each lower extremity will be obtained at baseline with subjects positioned following Sharma et al. (109). Mechanical alignment is the measure of the angle formed by the intersection of the lines connecting the centers of the femoral head and intercondylar notch and the centers of the ankle talus and tibial spines, with varus knee angles >0° inward, and valgus angles >0° outward. Subjects must have knee angles between -2° and 10° to be eligible.

Disease progression will be defined as change in x-ray medial TF JSW. A physician, masked to treatment group, will measure JSW. Severity of TF OA will also be measured using the K-L scale.

b.15. Physical Activity. Physical Activity Scale for the Elderly (129), has proven reliable in many of our clinical trials, including a group of 254 men and women aged ≥65 yr.

b.16. Adverse Events. Participants will be asked if any adverse event has occurred prior to each intervention session. Participants will also be asked to recall adverse events at the 6, 12, and 18 month testing visits. Events related and unrelated to the study will be collected.

3.c. Study Population. Participants will be ambulatory, community-dwelling men and women age ≥50 years with: (1) mild-to-moderate radiographic medial TF OA (KL = 2-3); (2) knee varus malalignment (varus angle ≥2 degrees and ≤10 degrees); (3) BMI ≥20 kg/m² and ≤44.9 kg/m²; and (4) no participation in formal strength training for more than 30 min/week in the past 6 months. We exclude people with BMI >44.9 kg/m² because of difficulty in using CT and DXA equipment and lower exercise compliance (32; 58) and <20 kg/m² because of limited thigh fat. We include only people with moderate varus malalignment and medial knee OA but not predominant lateral compartment or severe patellofemoral (PF) compartment disease because (1) the medial compartment is the most common disease site, and (2) medial progression is strongly associated with moderate varus malalignment (17; 36; 131), independent of BMI (87). People with extreme malalignment (>10 degrees varus) might experience greater progression in a strengthening program (68; 104). Medial bone-marrow lesions are seen mostly in patients with varus limbs, who are most likely to progress medially (37). This approach will engage an enriched cohort of progressors to better determine our intervention’s ability to slow the disease (61). All participants may maintain their medications, including NSAIDs. If pain decreases, they may reduce them with their physician’s consent. Medication use will be recorded at baseline and 6-, 12-, and 18-month follow-up testing. Exclusion criteria are listed in below:

Exclusion Criteria:

1. Significant co-morbid disease that would threaten safety or impair ability to participate in interventions or testing (Method: Medical history; medications physical exam; telephone pre-screen)
   a. Symptomatic or severe coronary artery disease; peripheral vascular disease
b. Severe HTN
c. Active cancer other than skin cancer
d. Anemia
e. Dementia
f. Liver disease
g. COPD
h. Inability to walk without an assistive device
i. Blindness
j. Type 1 diabetes; type 2 diabetes and on thiazolidinedione agents,
k. Other types of arthritis
l. Severe Hip OA

(2) Previous knee injury/trauma, (Method: telephone pre-screen)
(3) Severe tibiofemoral OA (KL = 4), (Method: AP and skyline knee x-ray)
(4) No definite medial tibiofemoral OA (KL = 0, 1), (Method: AP and skyline knee x-ray)
(5) OA disease location and alignment restrictions: predominant knee OA other than medial tibiofemoral,
Severe patellofemoral OA (JSN = 3 using OARSI atlas), OA Lateral tibiofemoral OA > medial tibiofemoral OA, (Method: Knee AP and skyline view x-rays, lower extremity long x-ray)
(6) Valgus alignment > 2° or extreme varus alignment (> 10°), (Method: lower extremity long x-ray)
(7) Severe or low obesity, BMI < 20 or > 4.49 kg/m² (Method: Ht/Wt)
(8) Not having knee pain: (Method: ≤ 3 on the pain scale, WOMAC and Telephone Screen)
(9) Excess alcohol use, ≥ 21 drinks per week, (Method: Med History and Telephone Screen Questionnaire)
(10) Inability to finish 18-month study or unlikely to be compliant (Method: Telephone Screen, Screening Interviews)
   a. Lives > 30 minutes from site
   b. Planning to leave area ≥ 1 month during the next 18 months
   c. Unwilling to attend a minimum of 3 sessions/week for 18 months
   d. Unwilling to discontinue pain medication use for 3 days prior to testing visit
(11) Conditions that prohibit CT
   a. Severe claustrophobia (Method: Med History)
(12) Significant cognitive impairment, diagnosis of dementia (Method: Montreal Cognitive Assessment score < 20 will be evaluated by the study physician, who will determine eligibility)
(13) Age, age < 50 (Method: Telephone Screen & Demographics Forms)
(14) Knee or Hip replacement (Method: Telephone Screen & Med History Forms)
(15) Knee surgery in past 6 months or plans to have surgery in the next 18 months (Method: Telephone Screen & Med History Forms)
(16) Knee injection in past 3 months plans to have injection in the next 18 months (Method: Telephone Screen & Med History Forms)
(17) Current participation in strength training program. (Method: Telephone screen)
(18) Current difficulty with daily activities (Method: Telephone Screen)
(19) Inability to perform training protocol. (Method: Telephone Screen)
   a. Inability to walk short distances
   b. Human assistance required to complete ADLs
(20) Currently participating in another intervention study (Method: Telephone screen)

3.d. Randomization. A stratified block randomization with block size unknown to investigators and staff will ensure equal accrual to each study arm. Prestratification will balance pretrial BMI values (25.0-29.9, 30.0-34.9, 35.0-40.0 kg/m²) and gender, which could predict intervention effect and associations between secondary outcome variables. A computer program will randomize participants into the 3 groups, verify eligibility, and provide identification number and intervention assignment. This system worked very successfully in IDEA.
3.e. Interventions. Both strength-training interventions consist of 5-min warm-up, 40-min training, and 15-min cool-down. Blood pressure will be measured prior to each exercise session. A post exercise blood pressure will also be taken. The 60-min sessions will be conducted 3 times/wk for 18 months at the Clinical Research Center. The first two will introduce participants to proper techniques, and at the third, 1-repetition max (1RM) tests will determine the starting resistance for each exercise. Intensity (load) is defined as %1RM (62). Each exercise will be performed on a Nautilus resistance-training machine with 60-90 s of rest between sets; 1RM is defined as the maximum weight one can lift in a single repetition. Participants will keep a session log of each exercise, its weight setting, and number of sets and repetitions achieved. 

Valuable feedback from the pilot study enabled us to optimize each exercise for this disabled population. Although our hypotheses focus on the lower extremity, experience indicates that participants want a well-rounded program. Thus, for both groups the program will include 6 lower body exercises with each leg exercised separately: hip abduction and adduction; leg curl, extension, and press; and seated calf; and 4 upper body and core exercises: compound row, vertical chest, lower back, and abdomen (Appendix B). We use Nautilus machines based on time, safety, and availability, but results will be generalizable to any strength-training method. Participants who plan absences of ≥2 sessions will use Thera-Bands in a home-based program (Appendix C). Upon their return, interventionists will determine the progression needed to reach prior intensity. Previous strength training trials with older adults, including our work (2.c.2.1, 2.c.2.5), predict small fluctuations in body weight (< 1 kg) as muscle mass increases and fat mass decreases. Interventionists will be alert to any substantial change (≥ 2 kg) and, if necessary, the participant will be referred to the medical director.

e.1. High-intensity intervention (H). The H group will perform 3 sets of each exercise at 75-90% of 1RM, within the intensity range necessary to maximize muscular hypertrophy (64). Each block will have the following structure and be repeated with training loads recalibrated to each new 1RM:

| Weeks 1-2 | 3 sets by 8 reps. | Intensity: 75% of 1RM |
| Weeks 3-4 | 3 sets by 8 reps. | Intensity: 80% of 1RM |
| Weeks 5-6 | 3 sets by 6 reps. | Intensity: 85% of 1RM |
| Weeks 7-8 | 3 sets by 4 reps. | Intensity: 90% of 1RM |
| Week 9. | Taper. | Alternate exercises and 1RM testing |

Based on our pilot study, most participants will have no difficulty progressing at 2-wk intervals, but variation is inevitable, and interventionists will rate perceived exertion (RPE) at completion of each workout. On a 10-point Borg category ratio-RPE scale, the H group should be working between 5 (hard)-8 (very hard), and the L group between 2 (easy)-4 (somewhat hard) (23). At the end of each block, we add taper periods—2 days (Monday, Friday) of alternate exercises, separated by a 1RM testing day (Wednesday)—because they have been shown to increase performance in older women (112). Table 2 equates the workloads for both groups.

e.2. Low-intensity intervention (L). The L group will perform 3 sets of 15 repetitions at 30-40% of 1RM using the exercises described above. Each 8-week block will have the following structure:

| Weeks 1-2 | 3 sets by 15 reps. | Intensity: 30% of 1RM. |
| Weeks 3-4 | 3 sets by 15 reps. | Intensity: 35% of 1RM. |
| Weeks 4-6 | 3 sets by 15 reps. | Intensity: 40% of 1RM. |
| Weeks 7-8 | 3 sets by 15 reps. | Intensity: 35% of 1RM. |
| Week 9. | Taper week. | Alternate exercises and 1RM testing |

Repeat weeks 1-8 with training loads recalibrated to each new 1RM

| Table 2. Sample workloads and total volume for high- and low-intensity interventions, assuming 1RM = 100 lbs. Total volume = total repetitions × intensity × resistance (assume 100 lbs). |
|---|---|---|
| Intervention | Sets/Repetitions/Intensity | Volume |
| **Low Intensity** | | |
| Weeks 1-2 | 3 sets of 15 reps at 30% 1-RM | 45 reps * 30 lbs = 1350 lbs * 2 wks = 2700 lbs |
| Weeks 3-4 | 3 sets of 15 reps at 35% 1-RM | 45 reps * 35 lbs = 1575 lbs * 2 wks = 3150 lbs |
| Weeks 5-6 | 3 sets of 15 reps at 40% 1-RM | 45 reps * 40 lbs = 1800 lbs * 2 wks = 3600 lbs |
| Weeks 7-8 | 3 sets of 15 reps at 35% 1-RM | 45 reps * 35 lbs = 1575 lbs * 2 wks = 3150 lbs |
Weeks 1 thru 8 | Total volume = 12600 lbs  
---|---  
High Intensity |  
Weeks 1-2 | 3 sets of 8 reps at 75% 1-RM | 24 reps*75 lbs = 1800 lbs*2wks = 3600 lbs  
Weeks 3-4 | 3 sets of 8 reps at 80% 1-RM | 24 reps*80 lbs = 1920 lbs*2wks = 3840 lbs  
Weeks 5-6 | 3 sets of 6 reps at 85% 1-RM | 18 reps*85 lbs = 1530 lbs*2wks = 3060 lbs  
Weeks 7-8 | 3 sets of 4 reps at 90% 1-RM | 12 reps*90 lbs = 1080 lbs*2 wks =2160 lbs  
Weeks 1 thru 8 | Total volume = 12600 lbs  
Low/High Ratio = 1.0

**e.3. Attention-control group.** The control is modeled after ADAPT’s usual-care comparison group (84), providing attention, social interaction, and health education. Participants attend 60-min organized workshops 2 times/month for the first 6 months and then 1 time/month for the remaining year. The control group will meet at the Winston-Salem Senior Center. This arm aims to control for attention from study staff and general levels of participant time; to encourage recruitment, adherence and benefit; and not to influence the primary outcomes directly: no evidence suggests that health education alone will affect pain or knee-joint loads during walking. Over the 18 months, interactive presentations will cover such topics as foot care, nutrition, managing medication, and sleep practices, and experts will give wide-ranging lectures. An experiential component will encourage participants to seek more information about their health and related practices. They will complete homework, review topics, and engage in small group discussions to increase their involvement in this study arm. Each workshop will end with seated stretching to enhance adherence and increase perceived benefit without directly affecting the knees or study outcomes. Prior studies suggest older adults are less likely to participate if they think any treatment group does not provide personal benefit.

**e.4. Techniques to Improve Adherence and Retention.** Time-intensive behavioral studies require significant commitment (22; 84; 125). START’s design evolved from social cognitive theory (SCT), group dynamics, and over 20 years’ experience in RCTs: our 18-month trials FAST and ADAPT had over 80% retention and 64-70% adherence. We conservatively estimate 80% retention and 65% adherence rates over this intervention; adherence will be calculated by dividing the total number of sessions completed by the number scheduled.

START interventionists will be trained by Drs. Mihalko and Bennell in standardized behavioral techniques developed in a SCT framework. They include frequent contact during the intervention; positive feedback; incentives to reach attendance and performance goals; establishing personal commitment to the project; promoting a sense of community via study logo, cards, and newsletters; and targeted mechanisms for behavioral adherence, including self-efficacy, outcome expectations, and self-regulatory skills. From the outset, the importance of regular attendance will be emphasized. Adherence data will be reviewed regularly to identify any participants who need additional reminders and/or counseling. Our toolbox approach, guided by algorithms of common strategies and decision-making processes, tailors the intervention to each participant’s needs. It identifies a problem, finds a solution, and tests it for a specific period. If the problem is resolved, the strategy is continued until behavior change is consistent. If not, the strategy is terminated, and a new option selected and tested for a specific period. Some options can be used in the group setting, while others require one-on-one interaction via telephone or meeting. For example, if a participant misses 2 consecutive sessions and has no contact with the interventionist, a phone session will be scheduled. The interventionist will assess participant study goals, time management, care-giving concerns, and feelings of connectedness to group objectives. Together, participant and interventionist will develop a specific plan. Collectively, these strategies will increase social cognitive mechanisms for regular participation and enhanced adherence in all groups.

3.f. Testing Procedures & Visits

**f.1. Screening and follow-up visits.** After expressing an interest in participating and passing the initial eligibility criteria during the pre-screening visit (PSV), each enrollee will undergo baseline testing. The enrollee will attend 2 screening visits (SV1 & SV2) to establish eligibility. Informed consent will be completed at SV1. Participants will return for a third visit (randomization visit, RV) to have the CT scans, blood draw, and gait analysis. After completing the randomization visit eligible participants will be randomized into one of the 3 intervention groups. We will randomize up to 385 persons into the study. In order be randomized into the
study each participant has to pass the screening at each screening appointment as well as the randomization visit. The randomized number has been changed so that if we meet our original goal of 372 randomized participants, those participants that have started the screening process will be allowed to continue with their screening and will be randomized if they meet the study eligibility criteria. Participants will return for 6, 12, and 18 month testing. Table 3 shows the data collected at each visit.

f.1.1. Prescreening visit (PSV). Individuals who contact our recruitment office in response to advertising will be asked a series of brief questions that focus on major eligibility criteria. A screening visit appointment will be made for participants who meet these criteria. A medical history form and a medication form will be mailed to the participants for them to complete.

f.1.2. Screening Visit One (SV1) Individuals will go to Worrell Professional Building at Wake Forest University. SV1 includes an explanation of the study and obtaining informed consent. Other assessments include medical history and medication use (previously mailed), height and weight (to calculate and verify BMI). The MOCA and CES-D will be administered. A score of < 20 on the MOCA will be evaluated by the study physician, who will determine eligibility; persons scoring >17 on the CES-D depression scale will be evaluated by the study physician/PA, who will determine eligibility. The DSST will also be administered. Those not excluded will have an interview with a staff member to assure their ability and willingness to comply. All subjects who have met the eligibility criteria will go to the Geriatric Clinical Research Unit (CRU) at Wake Forest University Baptist for a physical exam (including vitals). Afterwards bilateral standing (semi-flexed view), sunrise knee view, and full length x-rays will be collected. X-rays will be administered at Outpatient Radiology at Wake Forest University Baptist Medical Center. This visit will last approximately 3 – 3.5 hours.

f.1.3. Screening Visit Two (SV2). Individuals will go to Worrell Professional Building at Wake Forest to have a gait analysis, strength assessment, and to complete the FAST-23, Satisfaction with Life (SWL), Positive and Negative Affect Scale (PANAS), ABC, Comorbidities, and the Physical Self-Perception Profile (PSPP) questionnaires. The 6 minute walk test will also be performed. Participants will also complete the SPPB. Afterwards the participant will have a DXA scan. This 2nd screening visit will last approximately 2.5 - 3 hours.

f.1.4. Randomization Visit 1 (RV1) Participants will be asked to fast for 12 hours prior to this visit. Participants will go to the Geriatric CRU at WFUBMC for a blood draw. A snack will be given once the blood draw is complete. Urine samples (20 ml per visit) will be collected in 250 ml specimen cups by each participant for analysis of new and emerging OA biomarkers. In the event of acute respiratory, urinary tract, or other infection, we will postpone blood and urine sampling for 1–2 weeks after recovery from symptoms. Afterwards a power test will be performed in the Sticht Center. Finally, the participant will have a thigh CT scan performed at Wake Forest Baptist Outpatient Imaging. This visit will last approximately 1.5 – 2 hours.

f.1.5. 6-month Follow-up Data Collection Visit (FU6) Six month follow up testing will be split into three visits. Participants will go to Worrell Professional Building at Wake Forest University for Visit 1 (FU6V1). Tests include WOMAC, PASE, CES-D, SF-36, Medical History, Medication Form, and adverse event recall. This visit will last approximately 1 hour. A 2nd visit (FU6V2) will be scheduled at Worrell Professional Building to perform the strength tests, the FAST-23, PANAS, SWL, PSPP, Comorbidities self-administered questionnaire, the ABC, the SPPB, 6 minute walk test, height, weight, and the gait analysis. This visit will last approximately 2 – 2.5 hours. Visit 3 (FU6V3) will be scheduled at the Geriatric CRU. Participants will be asked to fast for 12 hours prior to this visit. Participants will go to Geriatric CRU main for a blood draw (50ml will be drawn). A snack will be given once the blood draw is complete. Urine samples (20 ml per visit) will be collected in 250 ml specimen cups by each participant for analysis. Afterwards the participants will complete the power test. The third visit (FU6V3) will last approximately 1 hour.

f.1.6. 12-month Follow-up Data Collection Visit (FU12) Twelve month follow up testing will consist of 2 visits. Participants will go to Worrell Professional Building at Wake Forest University for Visit 1 (FU12V1). Tests include WOMAC, MOCA, PASE, CES-D, SF-36, DSST, FAST-23, PANAS, SWL, Comorbidities self-administered questionnaire, PSPP, ABC, Medical History, Medication Form, adverse event recall. This visit
will last approximately 1 hour. The following tests will be done during the 2nd 12 Month visit (FU12V2): height, weight, 6 minute walk, and the strength tests. This visit will last approximately 1.5 hours.

f.1.7. 18-month Follow-up Data Collection Visit (FU18) At the completion of the 18-month intervention a complete battery of tests performed at baseline will be administered (minus the demographics and eligibility questionnaires and the brief physical exam). Participants will also be asked if any adverse events have occurred since their last testing visit that has not been reported. Participants will complete testing over a series of 3 testing sessions (see table 3). Testing will be performed at Worrell Professional Building and Wake Forest University Baptist Medical Center (Geriatric CRU, Outpatient Radiology, and the Wake Forest Baptist Imaging). At the interview they will be provided with a record of their own data. Data that are not available at this time will be given to the participants when completed.

Table 3

<table>
<thead>
<tr>
<th>Questionnaires</th>
<th>PSV</th>
<th>SV1</th>
<th>SV2</th>
<th>RV</th>
<th>FU6 Visit 1</th>
<th>FU6 Visit 2</th>
<th>FU6 Visit 3</th>
<th>FU12 Visit 1</th>
<th>FU12 Visit 2</th>
<th>FU18 Visit 1</th>
<th>FU18 Visit 2</th>
<th>FU18 Visit 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>eligibility questionnaire</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>informed consent</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>medical history</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medication Form</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographics</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WOMAC</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PASE scale</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOCA</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CES-D</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-36 (general health, quality life)</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Falls</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self Efficacy</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSST</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comorbidities</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SWL</td>
<td>X</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABC</td>
<td>X</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PANAS</td>
<td>X</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAST-23</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSPP</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Physical Performance Tests/Physical Exam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height</td>
</tr>
<tr>
<td>Weight</td>
</tr>
<tr>
<td>Waist/hip/thigh circumference</td>
</tr>
<tr>
<td>Brief physical exam</td>
</tr>
<tr>
<td>6-min walk</td>
</tr>
<tr>
<td>SPPB</td>
</tr>
<tr>
<td>Muscle Function Test: Power</td>
</tr>
<tr>
<td>Muscle function tests: Strength</td>
</tr>
<tr>
<td>Grip Strength</td>
</tr>
<tr>
<td>Gait Analysis</td>
</tr>
<tr>
<td>Knee A-P x-ray</td>
</tr>
<tr>
<td>Knee x-ray skyline view</td>
</tr>
<tr>
<td>Full length lower extremity x-ray</td>
</tr>
<tr>
<td>DXA</td>
</tr>
<tr>
<td>CT scans: thigh</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inflammatory Markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomarkers: blood</td>
</tr>
<tr>
<td>Biomarkers: urine</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adverse Events*</td>
</tr>
</tbody>
</table>

*Participants will also be asked about adverse events prior to each intervention session.

3.g. Statistical Considerations

g.1. Data Management. Data will be collected on hard copy forms and transformed to an electronic database. We will use a web-based management system to assure integrity and validity. Dynamic reports and periodic statistical analyses will monitor quality. A participant-based inventory system will track recruitment, retention,
adherence, and missing data from entry through exit, close-out, and lock-down of final datasets. Our team developed a similar database for the IDEA study.

To assist with this the site has a built in "dashboard" display which displays the current state of any data collection form at a glance using a combination of color coding and prose. In addition, this dashboard features drill down displays into every useful item of form meta data (field names, data types, validation rules, etc.)

The most significant quality control feature of the architecture is that the data collection forms are all processed with a single common data processing code module. This has a number of benefits. Because every form uses this same code, any errors are flushed out early. Since the code is so commonly exercised, any errors can’t really lurk "undiscovered" for long. The module is relatively simple and well documented. From a maintenance perspective, understanding how any single form works means you understand them all. There is no difference between forms from a processing perspective.

Second, the form specific data collection code is completely machine generated for all forms with only very rare exceptions. This means that it is not hand written and as such there is no opportunity to introduce manual errors.

Each person entering outcome data into the website will undergo a training period of entering data. Duplicate data will be compared monthly and errors will be corrected. A person will not be allowed to enter real data until their error rate is 0.5% or less.

g.2. Statistical Analyses will be conducted according to intention-to-treat principles using SAS. Brigham and Women’s hospital will use a validated computer simulation model of knee OA to assess the value of three major domains in pain management in knee OA patients with comorbidities: 1) PA programs as means of non-pharmacologic pain reduction; 2) pain phenotyping as means of optimizing pharmacologic pain management; and 3) weight management in morbidly obese persons as means of pain reduction and improving outcomes of other OA-focused treatments. Analysis will be done on the following data: Intervention group (Exercise, Diet, Diet+ Exercise); Age (yrs); Gender; Race; BMI (kg/m2); Weight (kg); Height (cm); Gait speed and distance; NSAID use; Comorbidity presence; SF-36 (self-efficacy measure); WOMAC (pain & function measure); Joint Space Width; MRI Progressors; KL grade; Bilateral OA presence; PatelloFemoral OA presence; Biomarker data; Cost Effectiveness.

g.2.1. Primary Aim. All primary analyses will be conducted at the 0.0083 two-sided level of significance (2 outcomes, 3 interventions). We will test the effects at 18 months on knee pain and maximal compressive force using repeated measures ANCOVA with time (6 and 18 mos), intensity (high, low, none) and the interaction (64), which adjusts the means at each time point for potential missing data bias. Maximum-likelihood techniques will estimate parameters. Preliminary analyses will be conducted to check the shape of the distributions and variances between groups and as a function of the covariates. Regression diagnostics and residual plots will help to find appropriate transformations, if necessary. Intervention-effect estimates will be adjusted for baseline values, BMI, and gender; analysis will match design, so the variance estimate will not be biased. In subsequent models, we will control for possible confounders, including PFOA severity (none to moderate) and use of medications, such as analgesics, NSAIDs, bisphosphonates, and glucosamine/chondroitin sulfate. Since we are excluding subjects with severe PF OA and medication has only modest efficacy in OA, we do not expect significant confounding by these variables.

g.2.2. Secondary Aims. Standard repeated measures ANCOVA (as noted above) will be used for secondary aims at the 0.05 significance level. Short-term effects (Aim 2.1) will be determined by comparing 6-mo means from primary analysis models. Similar contrasts for means at 18 mo (Aim 2.2) and for high intensity at 18 mo with short- and long-term results of the other interventions will be tested (Aim 2.3). Cartilage thickness (Aim 3.1) is continuous, so repeated measures ANCOVA with statistical tests and estimates like those described above, verifying assumptions and using appropriate transformations, will be used to determine the interventions' effect at 18 mo.
Repeated measures ANCOVA will be used to determine group differences in thigh muscle and fat measures at 18 mo (Aim 4.1). The main analysis will include baseline BMI and gender. We will investigate any differences in muscle or fat measures among the interventions within each subgroup, including ethnicity. All outcomes related to muscle function (Aim 4.2) are continuous and will be analyzed with models similar to those for Aim 4.1.

As inflammatory marker (Aim 5.1) distributions are often skewed, data will be log-transformed before analysis. The effect of the interventions at 18 mo will be determined with repeated measures models and estimates obtained at each FU visit. For interpretation, we will transform log means and SEs back to their original units.

### g.2.3. Missing Data

If missing data are related to outcomes, our results will be slightly biased. Our models will include variables from previous visits determined to predict loss to satisfy Little and Rubin’s (70) conditions for data considered Missing at Random (MAR). If “informative censoring” occurs, we will compare analyses using subjects with complete data, multiple imputations, or explicit modeling of the censoring mechanism (20; 132).

<table>
<thead>
<tr>
<th>Variable</th>
<th>SD</th>
<th>N/group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>n=143</td>
</tr>
<tr>
<td>WOMAC Pain</td>
<td>2.15</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(16.3%)</td>
</tr>
<tr>
<td>Max Compressive Force, N</td>
<td>630.85</td>
<td>294</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(9.0%)</td>
</tr>
</tbody>
</table>

comparisons per outcome) with 80% retention (2-sample t-test, Nquery Advisor). Standard deviations for pain and maximal compressive force were obtained from the START pilot and ADAPT, respectively; with mean differences of 1.12 (18%) and 657 N (20%).

### g.2.4. Sample-size calculations

**Primary Outcomes** (Table 4). A total sample of 372 (124/group) will provide 80% statistical power to detect differences ≥17.6% in pain and ≥9.6% in maximal compressive force at the 2-sided 0.0083 significance level (3 pairwise comparisons per outcome) with 80% retention (2-sample t-test, Nquery Advisor). Standard deviations for pain and maximal compressive force were obtained from the START pilot and ADAPT, respectively; with mean differences of 1.12 (18%) and 657 N (20%).

<table>
<thead>
<tr>
<th>Variable</th>
<th>SD</th>
<th>Detectable difference (% change)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WOMAC Function</td>
<td>8.1</td>
<td>3.7 (17%)</td>
</tr>
<tr>
<td>Mobility: 6-min walk (m)</td>
<td>199.9</td>
<td>92.6 (19%)</td>
</tr>
<tr>
<td>Thigh Muscle Volume (cm³)</td>
<td>577</td>
<td>267 (8%)</td>
</tr>
<tr>
<td>Thigh Subcutaneous Fat, cm³</td>
<td>281.8</td>
<td>130.5 (16%)</td>
</tr>
<tr>
<td>Thigh Muscle Attenuation, HU</td>
<td>4.9</td>
<td>2.3 (6%)</td>
</tr>
<tr>
<td>Quadriceps Strength: Knee extension strength (N)</td>
<td>78.2</td>
<td>36.2 (16%)</td>
</tr>
<tr>
<td>Proprioception: Mean error in average knee flexion (degrees)</td>
<td>1.9</td>
<td>0.9 (18%)</td>
</tr>
<tr>
<td>Max External Adduction Moment</td>
<td>5.6</td>
<td>2.6 (24%)</td>
</tr>
<tr>
<td>Inflammation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ln(IL-6), pg/ml</td>
<td>0.43</td>
<td>0.20 (1.22)*</td>
</tr>
<tr>
<td>Ln(TNF-α), pg/ml</td>
<td>0.52</td>
<td>0.24 (1.27)*</td>
</tr>
<tr>
<td>Ln(stTNFR1), pg/ml</td>
<td>0.19</td>
<td>0.09 (1.09)*</td>
</tr>
<tr>
<td>Ln(Leptin), pg/ml</td>
<td>0.49</td>
<td>0.23 (1.25)*</td>
</tr>
<tr>
<td>OA biomarker: COMP</td>
<td>2.3</td>
<td>1.07 (8%)</td>
</tr>
</tbody>
</table>

* Number in parenthesis indicates the ratio of the 2 means in their original units.

### 3.h. Recruitment

Our team has an excellent record in recruiting participants for knee OA clinical trials. We exceeded our goals in FAST (209 randomized with a goal of 200), ADAPT (316 with a goal of 300), and IDEA (454 with a goal of 450). Yields for FAST and ADAPT were 10% and 14%, respectively. Our yield for IDEA is...
Our study to hand/or race/ethnicity will disproportionately benefit or suffer from trial interventions. Hence, we have designed to suggest no compelling biologic basis or persuasive empirical data indicating that a group defined by gender approximately 70% women, and we anticipate the same ratio of women to men in this trial. Hence, our recruitment efforts in Forsyth County, NC, it is 18.4% African American, 0.9% Hispanic, less than 0.01% Native American, and 0.01% Asian. In Forsyth chronicle, we had success recruiting and enrolling them in FAST, ADAPT, and IDEA. Recruitment projections in FAST (209 randomized with a goal of 200), ADAPT (316 randomized with a goal of 300), and IDEA (454 randomized with a goal of 450).

<table>
<thead>
<tr>
<th>Inquiries</th>
<th>Eligible after prescreening for SV1</th>
<th>Eligible after SV1 for SV2</th>
<th># Randomized</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,600</td>
<td>780</td>
<td>390</td>
<td>372</td>
</tr>
</tbody>
</table>

Based on these experiences, we plan to use overlapping strategies with a system that provides feedback on each strategy’s effectiveness and cost. We primarily use mass mailings and media (newspaper). Our center has strong ties with local aging service networks and access to senior centers, churches, drug stores, shopping malls, and other sites where older adults gather. Our School of Medicine operates a Best Health store at Hanes Mall in Winston-Salem that provides free health information, and WFU health care professionals regularly present seminars. We have access to a large database of older adults who have signed consent to be contacted about participating in future clinical trials. Online advertising will also be done through the placement of ads on various websites (such as Yahoo). We will also recruit via doctor and participant referrals. A memo containing the study details and copies of our study brochure will be sent to local physician’s offices so they may refer their patients. Current participants in the study may also refer friends and family to the study. Participants may be rewarded prizes/incentives (such as giftcards or items such as mugs, t-shirts, etc. with the study logo) for referrals they may make.

Subjects who phone in response to advertising will be screened by the recruitment coordinator for general eligibility criteria. Subjects who are eligible by initial screening will come to the clinic for 2 screening visits. Informed consent will be obtained on the first face-to-face visit after all procedures and the time commitment the study requires has been explained. Those who meet all eligibility criteria will be randomized at the end of the baseline visits. A copy of the consent form, which outlines the study, will be given to participants, and a copy kept in a locked file.

The recruitment coordinator will work with PI Messier and Williamson, PI of the WFU Pepper Center recruitment core, to plan strategies and activities. Experience has proven that ongoing monitoring of the recruitment process is necessary to achieve study goals. Therefore, we will have bi-weekly meetings to review all recruitment activities, plan new activities, and monitor the number of phone contacts. We anticipate prescreening 2,600 persons to yield 30% for SV1 (first screening visit). Approximately 50% will be eligible after SV1, which includes the major inclusion/exclusion tests. We expect to lose approximately 5% between SV2 and the randomization visits to yield 372 randomized participants.

We will continue our intense recruitment activity in the minority community, advertising in the Winston-Salem Chronicle, a newspaper with a large African-American readership. Although minorities have historically been underrepresented in clinical trials, we had success recruiting and enrolling them in FAST, ADAPT, and IDEA. Minority recruitment in FAST was 26%, ADAPT 24%, and IDEA is 19%. Therefore, our goal of 20% for this study is reasonable and attainable. The percent of minorities in the US population age ≥55 years (2000 census) is 8.7% African American, 5.81% Hispanic, 0.5% Native American, and 2.66% Asian. In Forsyth County, NC, it is 18.4% African American, 0.9% Hispanic, less than 0.01% Native American, and 0.01% Asian. Hence, our recruitment efforts have routinely met or exceeded the percentages both nationally and locally.

Arthritis more commonly afflicts women than men. The FAST, ADAPT, and IDEA cohorts were approximately 70% women, and we anticipate the same ratio of women to men in this trial. A literature review suggests no compelling biologic basis or persuasive empirical data indicating that a group defined by gender and/or race/ethnicity will disproportionately benefit or suffer from trial interventions. Hence, we have designed our study to have sufficient power to address the hypothesis using the entire cohort.
4. Protection of Human Subjects

4.a. Risks

Risks to participants are small. Musculoskeletal injury may occur as the result of the exercise intervention, but during the strength-training portion of a recent study of 316 overweight or obese subjects with knee OA, we had no serious accidents. We will include a blood pressure safety alert trigger for this study. The absolute contraindication to resistance training is set at greater than 180/110 mmHg (participant will not be allowed to exercise), and the relative contraindication at above 160/100 mmHg (participant may be allowed to exercise on a limited basis, using lower weights and/or reps and sets however blood pressure will be monitored throughout the exercise session). DXA tests have very low radiation exposure and do not pose a serious risk. The radiation dose is <1.5 mRem, which is negligible and approximately equal to that obtained during a flight from Los Angeles to Chicago.

Bilateral volumetric measures of thigh adipose tissue and skeletal muscle will use a standard CT protocol. The thigh sequence is centered on the mid thigh as detailed in the attached protocol based on measures from the scout image and is about 33% of the expected exposure of a clinical scan of this region. The average amount of radiation a person will receive from this study is 9.86 mSv. This value is 1.28 times the 7 mSv exposure of residents of Denver. This is a conservative estimate as we are estimating a greater exposure (by including other joint sites in our radiation calculator). Dr. Carr co-chaired the recent AHA Science Advisory committee on radiation exposure with cardiac imaging and is well aware of the complex issues related to individual and population risk with exposure to low levels of ionizing radiation. CT has no reported injuries or deaths related to the imaging procedure. The risk of inducing future malignancies is based on extrapolation of much higher exposures using a conservative linear no-threshold model. The exposure is well below both the 100-mSv low-exposure threshold and the 50-mSv annual exposure for radiation workers. In addition, the small potential increase of cancer risk must be placed in the context of the baseline cancer death rate of ~20% and 40% risk of developing any cancer in the US. In this study of older adults, the risk is comparable to, or less than, other risks encountered in daily life, such as driving or riding in a motor vehicle (41).

4.b. Record-keeping procedures

Records are kept in locked file cabinets, and participant data identified by number only. Data will be used only in aggregate, and no identifying characteristics of individuals will be published or presented. Results of testing are sent to participants' private physicians, if the participants agree. Alert values for all medically useful procedures, such as blood pressure, will be developed, and a system put in place to alert study physicians and participants’ private physicians, depending on urgency. All staff members are taught procedures to deal with medical emergencies, including basic cardiac life support.
Acquisitions being sent to Boston University Medical Center will be uploaded into Efilm Workstation Software (Merge Healthcare, Milwaukee, WI USA) prior to being sent. Data will be returned in a text file and loaded into our secure database.

4.c. Safety Monitoring and Review
NIAMS has assigned a safety officer, with responsibility for reviewing all aspects of the study. The study will be reviewed twice a year via conference call and may convene an executive session at any time. In addition the Pepper Center DSMB will meet twice annually to review study progress and safety.

The safety officer/Pepper Center DSMB will have the following charges:

- To review the entire study protocol, Manual of Procedures, and informed consent and assent forms with regard to recruitment, randomization, intervention, participant safety, data management, plans for auditing participant records, and quality control and analysis plans, and to identify needed modifications. It will then identify the relevant data parameters and the format of the information to be regularly reported.
- To review data relating to efficacy, recruitment, randomization, compliance, retention, protocol adherence, operating procedures, forms completion, intervention effects, gender and minority inclusion, and participant safety over the course of the trial.
- To identify safety problems over the course of the study and to inform the PI via written report; he, in turn, will ensure that appropriate individuals receive the report.
- To identify any need for additional safety data and request them from investigators.
- To propose appropriate analyses and periodically review developing data on safety and endpoints.
- To make recommendations regarding recruitment, intervention effects, retention, compliance, safety, and continuation of the study.
- To send the Program Administrator and the PI written reports of all topics reviewed at each meeting.

Study staff will report non-serious adverse events to project manager and principal investigator within 7 days of notification of the event. The PI and study physician will review non serious adverse events (AE) on a weekly basis. Non serious adverse events will be included in the NIAMS (KAI) safety report and submitted bi-annually to KAI and the Pepper Center DSMB. Serious adverse events (SAEs) will be reported to the PI within 24 hrs whom will notify the NIAMS within 48 hours of notification. Only adverse events classified as serious, unexpected and related to the study will be reported to the IRB.

4.d. Subject Withdrawal
A subject is allowed to withdraw at anytime. If a participant decides to withdraw the reason for withdrawal and date of the withdrawal will be captured on the study status form. If a participant chooses to discontinue intervention, he/she will be encouraged to continue follow-up visits.
5.5.7. Reference List


79. McAuliffe, M. Medical image processing, analysis, and visualization (MIPAV). NIH Center for Information Technology. 4.2.0. 2009. Ref Type: Report


