Study: High Intensity Walking to Reduce Frailty
Principal Investigator: Margaret K. Danilovich, PT, DPT, PhD
Statisticians: Lauren Christine Balmert, PhD ; Chen Yeh, MS
Date Updated: 8/26/2021

STATISTICAL ANALYSIS OUTLINE

1. ANALYSIS OBJECTIVES

1.1. To test the effect of walking intensity on both frailty category (i.e., frail, prefrail, and nonfrail) and frailty score as a continuous outcome on the SHARE-FI (frailty instrument). Hypothesis: The HIW group will have: (1A) a greater percentage of participants who improve one frailty category and (1B) larger SHARE-FI continuous score changes at 4 months in comparison to the CSW group

1.2. To test the effect of walking intensity on mobility, physical function, balance, and total PA. Hypothesis: HIW participants will have: (2A) improved mobility as measured by gait speed, Timed Up and Go, and 6-Minute Walk Test, (2B) improved physical function on the Short Physical Performance Battery, (2C) improved balance as measured by the Berg Balance Scale and Falls Efficacy Scale and (2D) greater total PA measured during a 1-week ambulatory monitoring period at 4 months in comparison to the HIW group.

2. STUDY DESIGN/SAMPLE

2.1. Study Type

We propose a cluster randomized controlled trial to evaluate the impact of HIW on frailty in older adults. We will randomize 200 participants from 10 retirement communities (20 per site) to a 1:1 supervised, 4-month overground HIW (70-80% HRmax) or self-selected, casual speed walking intervention (CSW). The CSW intervention is based upon the NIA Go4Life recommendations. The 4-month intervention will consist of two phases: (1) a 1-month ramp-up phase in which HIW participants will gradually increase intensity levels to reach 70-80% HRmax and (2) a 3-month intervention phase whereby HIW participants will walk at 70-80% of HRmax. CSW participants will walk at the self-selected, casual speed for all 4 months. Primary and secondary outcomes will be measured at baseline and 4 months. Our overall hypothesis is that HIW will reduce or reverse frailty in comparison to CSW at 4 months, as well as improve overall mobility, balance, physical functioning, and daily PA.

2.2. Sample Size

We will study 10 retirement community sites with 20 participants per site for a total of 200 participants. In our previous work with a similar population of older adults, overall attrition was 17% at 3 months. We now assume 20% attrition since this study will follow participants out to 4 months. Power calculations conservatively assume 20% attrition at both the site and participant level; with 20% attrition at the site level (one site per arm dropout) and 20% participant attrition at remaining sites (16 participants per site), we will have at least 128 participants for whom we will have complete baseline and follow-up data required to compute outcomes on frailty and physical functioning measures at four months.

We anticipate the CSW intervention arm to exhibit only small improvements in frailty category; we hypothesize that less than 5% of participants will improve according to our definition based on SHARE-FI category. Since power considerations rely heavily on the control proportion of success, we allowed improvement to range from 1%-5% for all calculations; we further included a calculation assuming 10% improvement to explore the robustness of our sample size in the event of incorrect assumptions. We anticipate an approximate 25%-50% improvement in the HIW arm, corresponding to a difference on the order of 20% to 45% between arms. With limited knowledge on intra-cluster correlation (ICC), power considerations allowed for varying levels of correlation within cluster. Since we expect between 20% and 45% absolute difference increase in frailty improvement in the HIW arm when compared to the CSW arm, we anticipate adequate power to detect these differences in frailty under almost all foreseeable scenarios with regard to ICC and control proportions. Even if our assumption regarding control proportion was incorrect such that 10% of those in the CSW arm improve, the worst case scenario (ICC = 0.05) will still allow for 90% power to detect a difference of at least 27% across the two arms. Our minimum expected sample size of 64 participants per arm will also allow us to detect the

differences in SHARE-FI, gait speed, 6 minute walk test, Berg, SPPB, Falls Efficacy Scale, and TUG at 16 weeks as illustrated in the table below.

Power to Detect Meaningful Clinical Differences Between Treatment Arms in Secondary Outcomes

Measure	Difference	SDs	Power to Detect Meaningful Difference under Same Assumptions
SHARE-FI	1.4 points	0.18	>95% power
Gait speed	0.10 m/sec	0.16	Variable, depending on ICC (61%-83%)
Six-minute walk test	50 meters	17	>95% power
Berg	5 points	1.4	>95% power
SPPB	0.5 points	0.3	>95% power
Falls Efficacy Scale	10 points	3.3	>95% power
TUG	3.4 seconds	1.38	>95% power

3. OUTCOMES

Primary

- Frailty (classified as non-frail, pre-frail, or frail based on SHARE-FI) – a dichotomous outcome will be considered, indicating whether an individual moves up at least one "category" of frailty

Secondary

- SHARE-FI (continuous measure)
- Gait speed
  - This speed is measured in meters per second. The distance for tests is 6 meters (gait_distance) and completion time is measured in seconds over two trials (gait_self_1), (gait_self_2). Create a new variable that is the average of (gait_distance)/(gait_self_1), (gait_distance)/(gait_self_2)
- Endurance (6 minute walk test)
  - six_min_dist
- Overall mobility (TUG)
  - tug_time_score
- SPPB
  - create a new variable that is the sum of (sbs_a), (sts_b), (ts_c), (gest_score), (rcs_score)
- Berg Balance Scale
  - create a new variable that is the sum of (berg_sit_stand), (berg_stand_unsp), (berg_sit_unsp), (berg_stand_sit), (berg_transfer), (berg_eyes_close), (berg_feet_tgt), (berg_reach), (berg_retrieve), (berg_turn_behind), (berg_turn360), (berg_alter_ft), (berg_1ft_frnt), (berg_1_leg)
- Falls Efficacy Scale
  - *sum of (fes_bath), (fes_reach), (fes_walk), (fes_prep), (fes_bed), (fes_answr), (fes_chair), (fes_dress), (fes_groom), (fes_toilet)*
- Physical Activity

4. STATISTICAL ANALYSIS PLAN

4.1. *GENERAL CONSIDERATIONS*

We will use descriptive statistics (mean±standard deviation for continuous measures and frequency/percentages for categorical measures) to summarize all baseline data in aggregate as well as across study arms. Analyses will assume parametric modeling assumptions are appropriate (e.g., normality for linear models); in cases of violation of these assumptions, we will explore nonparametric and/or variable transformations. Analyses will be based on the intent-to-treat (ITT) principle, and sensitivity analyses will examine an as-treated dataset according to a set of pre-defined criteria for adherence. Longitudinal analyses will further include fixed time effects and random participant effects to further account for within-participant

correlation across time. All quantitative statistical analyses will be performed using SAS software version 9.4 (The SAS Institute Inc, Cary, NC). Residual diagnostics will assess model assumptions and, as appropriate, transformation of variables, non-parametric methods, and/or inclusion of higher-order terms in analytic models may be explored (e.g., interaction, quadratic, cubic, etc.). Analyses will assume a 5% level of significance. If primary outcome analyses are significant at the 5% level, we will employ Hochberg's procedure to correct for multiple hypothesis tests for all secondary analyses due to the potentially large number of statistical tests. According to this process, we will order all secondary p-values from smallest to largest. We will reject all m hypotheses if $p(m)<0.05/m$. Otherwise, we will reject all (m-1) hypotheses if $p(m-1)<0.05/(m-1)$.

4.2. *ANALYTIC STRATEGY*

4.2.1. *Primary Analyses*

Primary outcome for analyses is the 4 month classification according to SHARE-FI criteria: whether a participant improved at least one category of frailty classification. This dichotomous outcome will be analyzed via a generalized linear mixed model (GLMM) with fixed arm, baseline SHARE-FI score effects, and random cluster effects (i.e., random site effects). This will allow for distinction between within and between cluster variation and more efficient analyses that will account for intra-cluster correlation (ICC).

4.2.2. *Secondary Analyses*

For each dependent measure, the difference between baseline and 4 month scores will be calculated. Generalized linear mixed models will be performed with fixed arm, baseline score effects, and random cluster effects (i.e., random site effects). This will allow for distinction between within and between cluster variation and more efficient analyses that will account for intra-cluster correlation (ICC).