STATISTICAL ANALYSIS PLAN

PROTOCOL TITLE: PHYSICAL TELEREHABILITATION IN VETERANS WITH MULTIPLE SCLEROSIS

NCT02346734

APPROVAL DATE: 7/27/15

Statistical Analysis Plan

Preliminary analysis will be descriptive in nature. Outliers will be identified and verified for accuracy. Distributions of the outcome variables will be examined for symmetry, and transformations will be considered for seriously skewed variables. Bivariate associations will be examined using cross-tabulations or scatter-plots. We will use chi-square tests for assessing crude associations between two categorical variables, t tests for comparing two group means, and correlation coefficients for crude relationships between two continuous variables. Specifically, we will check whether the two study groups are comparable in terms of baseline covariates by t tests or chi-square tests. We will compute means and standard deviations (or medians and inter-quartile ranges if distributions are skewed) for continuous measurements at baseline, 3 months, and 6 months.

Data analysis plan for the hypothesis related to our primary aim: The physical telerehabilitation will result in improved timed 25-foot walk and Berg Balance Scale Score at 6 months as compared to a control group. We will compute the changes in timed 25-foot walk and Berg Balance Scale Score between 6 months and baseline for each subject. We will examine the distribution of the changes and transformations will be applied if needed. Since we use randomization within each EDSS stratum, we will use ANOVA with the treatment group, EDSS group and EDSS by treatment interaction terms to compare the average changes in the two groups to test the above hypothesis, as well as examine the changes within each EDSS stratum, although our study is not powered to detect the stratum-specific changes. We will use intention-to-treat (ITT) analysis and include all subjects in the study group they are randomized to, regardless of compliance.

In the following secondary analyses, we will include the baseline status of the outcomes in thecorresponding models. Secondary analysis for the hypothesis related to our primary aim: As a secondary analysis, we will also perform explanatory analysis that examines the effect of the intervention after adjusting for potential confounders and effect modifiers. We will build an initial multivariate linear regression model for the change in timed 25-foot walk that includes treatment indicator and all baseline covariates (socio-demographic variables, MS profile variables and psychosocial status variables) that are significantly different between the two groups. Insignificant covariates will be removed from the model. However, they will be further assessed to see whether they are confounders by examining the change in the regression coefficient of the treatment indicator. We will add confounders back to the model. Biological plausible interactions will be assessed in the regression model and we will keep interactions that are biological interpretable and significant in the final regression model.

Data analysis plan for hypothesis related to the secondary aim on socio-behavior outcomes at 6 months: We will compute the changes in the corresponding scores between 6 months and baseline in MS self-efficacy, exercise adherence, patient-provider communication, social support, quality of life, depression and satisfaction with care. We will perform the analysis similarly as for the timed 25 foot walk. The baseline status of the outcomes will be included in the model.

Data analysis for hypothesis related to the secondary aim on impairment, activity, and participation (including EDSS, MS Functional Composite, MS Impact Scale, Modified Fatigue Impact Scale, MS Walking Scale, Modified Ashworth Scale) at 6 months will be carried out similarly as for the timed 25 foot walk. The baseline status of the outcomes will be included in the model. Data analysis for Aim 4: Evaluate primary and secondary aims in specific subgroups stratified by gender, race, and EDSS, and explore possible interactions. If data allows, we will carry out multiple regression analyses to examine whether gender, race exercise adherence, and categorized EDSS are effect modifiers for the association of the intervention and the primary outcomes (the changes in timed 25 foot walk and Berg Balance Scale Score at 6 months). For example, if an interaction of gender by intervention indicator is significant, we will evaluate the primary outcomes and secondary outcomes by gender.

Data analysis for Aim 5: Compare the longitudinal trajectories of primary and secondary outcomes in the intervention and control groups. We will use linear mixed models for longitudinal data to study the trajectories of primary and secondary outcomes in the intervention and control groups. Specifically, we will use the repeated measures of an outcome variable (such as timed 25 foot walk) at baseline, 3 months, and 6 months simultaneously in a linear mixed model. We will include the intervention indicator, an indicator for higher EDSS score (4-6.5), a categorical time variable with 3 levels, the interaction terms by intervention and time, the baseline measurement of this outcome (in this example, it is timed 25 foot walk) as fixed effect covariates in the model. A subject level random effect will be

included in the model to handle the correlations among repeated measurement within subjects.

Strategies for handling Missing data: We will take proactive steps to minimize withdraw from the trial. We will report the percentage of subjects with complete data by intervention group. For those lost to follow up, we will compare them with those who have completed the trial to see whether there are differences in social-demographic and MS status variables. The linear mixed models can use all available partial data from each subject, even with data missing at some time points. The results are unbiased if missing at random (MAR) is a reasonable assumption. If MAR assumption is not reasonable for our data, we will perform sensitivity analyses to assess biases due to loss to follow up