Preventing Alzheimer's Disease With Cognitive Training

NCT03848312

WIRB® Protocol #20182630

4 NOVEMBER 2019

STATISTICAL ANALYSIS PLAN

Power and Sample size

In a prior study, Edwards et al. found a hazard ratio of 0.71 when assessing the effect of SPT on the reduction of dementia risk as compared to the no contact control group and a hazard ratio of 0.75 when compared to memory training (Edwards et al., 2017). These hazard ratios correspond to 29% and 25% risk reduction, respectively. Analyses of ACTIVE data indicate that among those with normal cognitive function at baseline, 15.6% of participants randomized to SPT met either MCI or dementia criteria by three years as compared to 21.4% of controls, representing a 32% risk reduction of MCI or dementia, *OR*=0.68 95%CI=0.49-0.94, *p*=.02. After simulating a reduction in incidence of MCI or dementia of 15%, 20%, and 25% in the intervention group across the study period, we powered our study to detect an effect size of 20% risk reduction of MCI or dementia (i.e., *OR*=0.80). Given the previously observed effect sizes, this is a conservative assumption. We plan for an attrition rate of 20%. The attrition rate of 20% is based on prior SPT studies. As examples, ACTIVE had a 5.5% attrition rate per year of follow-up, and the 3-year assessment of the SKILL study had 22% attrition.

We powered our study based on at least 10.5% of participants in the control arm converting to either MCI or dementia during the 3-year follow-up period, which is equivalent to the 3.5% per year conversion rate to MCI among healthy adults aged 70 to 75 years reported previously (Roberts et al., 2012). The 70- to 75-year age bracket corresponds to the median age group in our older adult samples. In comparison, among those in the ACTIVE study with normal baseline cognition, 17.3% overall met MCI or dementia criteria at three years, rendering our predicted conversion rate conservative.

Power was estimated using conventional a-priori methods. We computed sample size and power using PASS 12 software for comparing two binomial proportions. For aim 1, power calculations indicate that at a .05 level of significance with a 20% attrition rate, the two-group, randomized design will require a total sample size of 7600 participants to ensure at least 82% power to detect a 20% reduction from an estimated 3-year incidence rate of 10.5% for conversion to MCI or dementia. Given the anticipated 20% attrition of the total sample of 7600, and the 10.5% conversion rate, we expect to have a final sample of 6080 participants, of whom about 640 will exhibit MCI or dementia across the 3 years of follow-up under the null hypothesis (i.e., 3.5% per year conversion rate).

For aim 2, we will explore if the extent of the amyloid burden or ApoE4 status modify the effects of SPT in a sample of converters to MCI or dementia. For this exploratory aim, a formal sample size and power calculation is not performed. We will calculate the effect sizes and the 95% confidence intervals for the interactions.

Randomization

For allocation concealment, Dr. Ji implemented the randomization, which is stratified by site. The randomizations are generated by using the randomization module in REDCap. The random assignments are stored on REDCap and are password protected. An unblinded study statistician will be the only person who can access the randomization list. To disseminate the random assignments, Dr. Ji created a link for the Site Adminstrators to access the REDCap randomization system. The Site Administrators are able to retrieve one assignment at a time. It is not be possible for the Site Administrators to predict the next assignment. Investigators are blind to participants' randomized conditions. The study team makes concerted, consistent efforts to maintain blinding. All participants randomized will be included in the analysis data set, irrespective of intervention

adherence. The study team will make every effort to attempt to evaluate individuals randomized at the 3-year follow up, irrespective of adherence.

Data Management and Data Quality Assurance

Please see the **Data Safety and Monitoring Plan** for detail. The Health Informatics Institute (HII) will implement data management and quality assurance. The HII is directed by Co-I Krischer who will supervise a staff of 8 with effort devoted to this project including positions of statistician, software developer, data architect, quality assurance officer, data engineer, systems engineer, data manager, and study monitor. HII will develop, maintain, and generate the study databases. The design will employ .Net as the web interface and Oracle REDCap as the backend. The HII will develop a secure, efficient, centralized, customizable, scalable, and coordinated web-based data management platform for the integration of the data to be collected, including participant characteristics, questionnaires, and specimen tracking. The HII platform will provide web-based data capture complemented by automated quality assurance (QA) measures preserving the utility of project data for downstream data integration and analyses. In addition to the HII Platform real-time tracking of data quality and completeness, the HII will regularly generate reports to allow for remote data auditing and promote data curation and QA. Regular study monitoring and site visits will also be conducted for QA as described in the **Data Safety and Monitoring Plan**.

Data Analysis

Dr. Krischer will oversee the data management team and statisticians who will conduct the interim data analyses for the Data Safety Monitoring Committee (DSMC) meetings and the final data analyses at the end of the study, in consultation with the MPIs and executive committee. We will follow established standards and guidelines on clinical trial data analysis and reporting such as the CONSORT Guideline and the Intent-to-Treat Principle (ITT). A CONSORT diagram will be produced. For baseline analysis, descriptive statistics including means, standard deviations, frequencies, percentages, histograms, boxplots, etc. for the entire sample as well as the intervention and control groups will be generated, compiled, and reviewed.

Data issues such as abnormal values or outliers will be investigated and reconciled by collaborations among the research team. Given the relatively large sample size, we expect the treatment arms to be balanced in their baseline characteristics due to the randomization. However, as a check on the randomization, baseline characteristics will be compared by two sample *t*-tests or nonparametric tests for continuous variables and Chi-Square Test or Fisher's Exact Test for categorical variables. Unbalanced baseline factors will be recorded and included in the outcome analysis as covariates. Sex as a biological variable will also be examined as a covariate. The level of significance for the primary hypotheses will be set at two-tailed .05. In the presence of missing data on the outcome variable, we will follow the Intent-to-Treat (ITT) Principle, as described above. That is, all randomized individuals will be included in the planned analyses to the extent that they have data available, irrespective of adherence to the intervention to which they were assigned. The planned sample size includes provision for those who drop out, thus ensuring the study power. As a check on potential endpoint ascertainment bias, percentages of missing outcome data in the intervention and the control groups will be calculated and compared according to baseline characteristics. And then the comparison of the characteristics of patients with complete data vs. those with missing data will be performed to observe if there are systematic differences. Second, multiple imputation (MI) will be performed to impute the missing values and the previously proposed analyses will be run subsequently. The MI analysis should address the missing at random (MAR) data. In case there is potentially nonignorable missing data such as considerable different attrition rates in the intervention vs. control groups, nonignorable missing data models such as the selection model approach or the

pattern mixture model approach will be performed as sensitivity analysis of our proposed statistical data analysis plan.

Aim 1. **To ascertain the effectiveness of SPT to reduce incidence of MCI or dementia**

For Aim 1, a Chi-square test will be performed to compare whether the proportions of MCI or dementia at the end of the trial are statistically different by randomization. To further address Aim 1, we will estimate the risk of conversion to MCI or dementia as a function of the intervention using binary logistic regression. Results will be presented in the form of odds ratios, which will represent risk of conversion to MCI or dementia. An odds ratio below 1.00 will indicate reduction in risk as a function of assignment to the intervention group. A 95% confidence interval will be used. In addition to the *p* value for statistical significance, point estimates and confidence interval of the differences of the proportions will also be reported to show the direction and magnitude of the effects. If there are unbalanced factors identified in the baseline analyses, then logistic regression analysis will be performed with group assignment as the main predictor and the unbalanced baseline factor(s) as covariates. Adjusted OR and its 95% confidence interval for the group assignment will be estimated to account for the unbalanced covariates. Secondary subgroup analyses based upon site, sex, and age will be explored. As well, a "per protocol" analysis is planned to examine the intervention effect in the subset of participants who were adherent, completing at least 75% of the assigned sessions. Although measuring time to event within Cox proportional hazards regression is an alternative approach, we will not be able to apply Cox regression as exact timing of MCI/dementia incidence will not be known.

In the proposed trial, those participants who show evidence of cognitive decline will receive detailed neuropsychological and clinical examinations and diagnoses. We acknowledge that by using the MoCA as a screening test and the diagnosis of MCI or dementia as a gold standard, then testing only the individuals who exhibit cognitive decline may lead to verification bias. To explore this issue further, we will estimate the proportion of those who screen positive (i.e., MoCA score below 26) who do not meet the primary end point of MCI/dementia diagnosis, arriving at an estimate of false positives. Stratified random sampling will be used to identify a cohort of 100 cognitively-normal participants (stratified by site) who screen negative (i.e., a MoCA score of 26 or higher) at follow-up to arrive at an estimate of false negatives. Given that the established sensitivity of the MoCA is 90% for MCI and 100% for dementia, we expect to find about 10 cases of MCI and no cases of dementia in this cognitively-normal subsample (Nasreddine et al., 2005). Thus, we expect a false negative rate of ~10%. Using PASS software, we computed the comparison cohort sample size using an exact Clopper-Pearson 95% confidence interval for a rate of 10%. Results indicate that with a subsample of n=100, we can detect a false negative rate ranging from 5.0-17.4%. We will apply the statistical method for correcting verification bias as outlined in Zhou, McClish and Obuchowski (2011). To minimize systematic difference in those who screened positive for cognitive impairment on MoCA and the subsample of 100 screening negative, we will use binary logistic regression to compute the propensity scores for the MoCA positive and the MoCA negative groups based on nearest neighbor matching while adjusting age, sex, and education. The propensity score method will be carried out in the SAS procedure PSMATCH. Once the propensity scores are estimated, we will stratify the MoCA positives and the MoCA negatives into five subgroups based on the quintiles of the estimated propensity scores. The goal is to gain insights in factors that might be contributing to the treatment effect comparisons evaluating the MoCA as a tool to identify those in need of more comprehensive clinical evaluation.

We acknowledge that death represents a competing risk in the proposed analyses. Therefore, we will apply the Semi-Markov modeling approach developed by Kryscio et al. (2013) to adjust for mortality in our data analysis of MCI/dementia conversion. Of course, another issue is attrition. We will examine time to attrition in survival analysis. To fully explore the impact of competing risks, we will

apply joint modeling of survival (mortality risk) and longitudinal data in modeling cognitive status (normal vs. MCI/dementia) over time (See Arbeev, Akushevich, Kulminski, Ukraintseva, & Yashin, 2014).

Exploratory Aim 2. To explore if amyloid burden or ApoE4 status modify the effects of SPT

For Aim 2 a logistic regression model will be fitted using clinical diagnosis of cognitive status (MCI/dementia vs. normal) as the primary dependent variable, randomization to the intervention vs. control arm (SPT vs. computer games) as the independent variable, and interaction terms between randomized arm and ApoE 4, and between randomized arm and amyloid burden, to test for moderation effects. A statistically significant interaction ($p<.05$) will be used as criteria for determining moderation effects of ApoE4 or amyloid burden. For significant interactions, odds ratios and 95% CIs in strata defined by the moderator (ApoE4 positive vs. negative; high vs. low amyloid burden based on SUVR cutpoint of 1.25) will be reported separately to characterize moderation effects (Bullilch et al., 2017).

Interim Analysis

Interim analyses will include monitoring of adverse events and study progress (e.g., enrollment goals, sample diversity, protocol deviations) which will be conducted and presented to the DSMC on an annual basis. Monitoring of early futility and analysis of efficacy on the primary endpoint will be conducted when 50% of the target study population has been followed for the planned three years and will be reviewed for assessment of effectiveness.

The Lan-DeMets (1983) spending function with an O'Brien-Fleming boundary will be used to protect the type I error probability from early and multiple testing and to assess the significance of the interim results that emerge during the trial (Cox, 1972; Diggle, Liang, & Zeger, 1994; Simon, Wittes, & Ellenberg, 1985). The spending function that approximates the O'Brien-Fleming boundaries is:

$$\alpha_1(t^*) = 2 - 2\Phi\left[\frac{Z_{\alpha/2}}{\sqrt{t^*}}\right]$$

where t^* is the information fraction ($0 < t^* \leq 1$), α_1 is the $\alpha\text{-}level$ of the interim (one-sided) test and α is the type I error for the comparison of the treatment arms based upon the observed rates of conversion to MCI or dementia. The monitoring plan will allow for early termination based on the treatment effect on incidence rates using the ANCOVA model that incorporates pre-specified covariates.

The interim analysis will also identify if there is a serious lack of evidence of a treatment effect (i.e. futility analysis). The boundaries are based on the paper by Lachin (2009). The active intervention arm should be "closed" based on the futility of rejecting the null hypothesis at the completion of the trial if the Wald test of the treatment effect coefficient is <=0.1. Lachin showed that a one time use of the boundary for the design parameters above ($\theta \equiv Z_{1-\alpha} + Z_{1-\beta} = 3.00$) raises the type II error minimally. Furthermore, by the laws of probability a single use of each rule will increase the type II error no more than the sum of the increase (i.e., 0.00414+0.00611 = 0.0103).

The application of the results of the interim analysis will be considered in the context of the planned neurological and clinical evaluations. These are relatively expensive and logistically demanding and their contribution to the study objectives might be reconsidered in light of the interim analysis.

References for Statistical Analysis Plan

Arbeev, K. G., Akushevich, I., Kulminski, A. M., Ukraintseva, S. V., & Yashin, A. I. (2014). Joint analyses of longitudinal and time-to-event data in research on aging: Implication for predicting health and survival. Front Public Health, 2:228 PMCID: PMC4222133. *Frontiers in Public Health, 2*(228). doi: 10.3389/fpubh.2014.00228

Bullilch, S., Seibyl, J., Catafau, A. M., Jovalekic, A., Koglin, N., Barthel, H., . . . De Santi, S. (2017). Optimized classification of 18 F-Florbetaben PET scans as positive and negative using an SUVR quantitative approach and comparison to visual assessment. *Neuroimage: Clinical*, 325-332.

Cox, D. R. (1972). Regression model and Life Tables. *Journal of the Royal Statistical Society. Series B (Methodological), 34B*, 187-220.

Diggle, P. J., Liang, K. Y., & Zeger, S. (1994). *Analysis of Longitudinal Data*. Oxford: Clarendon Press.

Edwards, J. D., Clark, D., Xu, H., Guey, L. T., Ross, L. A., & Unverzagt, F. W. (2017). Speed of processing training results in lower risk of dementia. *Alzheimer's & Dementia: Translational Research & Clinical Interventions, 3*, 603-611. doi:https://doi.org/10.1016/j.trci.2017.09.002

Kryscio, R. J., Abner, E. L., Lin, Y., Cooper, G. E., Fardo, D. W., Jicha, G. A., . . . Schmitt, F. A. (2013). Adjusting for mortality when identifying risk factors for transition to MCI and dementia. Journal of Alzheimer's Disease, Jan 1:35(4),823-832. . *Journal of Alzheimer's Disease, 1*(35), 823-832.

Lachin, J. M. (2009). Futility interim monitoring with control of type I and II error probabilities using the interim Z-value or confidence limit. *Clinical Trials, 6*(6), 565-573.

Lan, K. K. G., & DeMets, D. L. (1983). Discrete sequential boundaries for clinical trials. *Biometricka, 70*, 659-663.

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., . . . Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*, 695–699.

Roberts, R. O., Geda, Y. E., Knopman, D. S., Cha, R. H., Pankratz, V. S., Boeve, B. F., . . . Petersen, R. C. (2012). The incidence of MCI differs by subtype and is higher in men The Mayo Clinic Study of Aging. *Neurology*, WNL-0b013e3182452862.

Simon, R., Wittes, R. E., & Ellenberg, S. S. (1985). Randomized phase II clinical trials. *Cancer and Metastasis Reviews, 69*, 1375-1381.

Zhou, X. H., McClish, D. K., & Obuchowski, N. A. (2011). *Statistical Methods in Diagnostic Medicine. Wiley Series in Probability and Statistics.* New York: Wiley Intersciences.