STATISTICAL ANALYSIS PLAN
FOR

Computerized Cognitive Training to Improve Cognition in Diabetic Elderly Veterans

(CCT_DEV)

Protocol ID: IIR 11-285

NCT01736124

August 5, 2020

1. Introduction

Diabetes mellitus (DM) has consistently been associated with increased risk for cognitive decline, mild cognitive impairment, and dementia in the elderly. Even minor cognitive impairments in nondemented individuals dramatically affect disease self-management. This, in turn, is associated with poor glycemic and blood pressure control in diabetes, which by themselves increase the risk of dementia, provoking a reinforcing cycle of disease. Thus, it is imperative to find interventions to delay or prevent cognitive compromise in diabetic patients, that can be relatively easily and rapidly implemented, and that are not cost prohibitive. This is especially true in the VA, in view of the high incidence of both diabetes and dementia in the growing population of elderly Veterans.

Epidemiologic evidence suggests modifiable life-style factors, including cognitive activity, may prevent or delay the onset of cognitive decline. Computerized cognitive training (CCT) is an intervention that has shown promising results in the improvement of cognitive functioning, more consistently in non-demented elderly, with additional benefits from booster training sessions. To date, studies of CCT have typically only examined cognitive outcomes, and only shortly after the intervention. The proposed CCT program, Cognifit, is designed to improve cognition of elderly persons by targeting their weak cognitive functions, using a personally tailored training plan. The study will provide an evaluation of the effects of CCT on DM self-management behavior and clinical outcomes, in addition to cognition. Participants for this randomized clinical trial are recruited from the James J. Peters Veterans Affairs Medical Center (VAMC) in the Bronx, NY and the Ann Arbor VAMC (MI).

2. Trial Population

Non-demented DM elderly veterans from the James J. Peters (Bronx, NY) and Ann Arbor (MI) VAMCs have been recruited for this study. Some eligibility criteria are mentioned for screening during the recruitment process.

Inclusion criteria are:

- 55 years old or above
- A diagnosis of type 2 diabetes
- Access to computer and internet
- Identification and provision of contact information of an informant
- Self-management score 18 (of a possible 20) or below on the Diabetes Self-Management Questionnaire (DSMQ)

Exclusion Criteria are:

- Dementia (diagnosis or CDR ≥ 1) or prescribed AD medications
- Major medical, psychiatric, or neurological conditions that affect cognitive performance; MMSE <25

- Severe impairment of vision, hearing or fine motor control necessary for computer operation

Eligibility criteria are assessed after potential participants are enrolled by providing signed informed consent at baseline (Visit 1). Following baseline assessment and determination of eligibility, participants are randomized to the CCT intervention, Cognifit, or the active control task. The CCT intervention has games that adapt to the changing performance level of the participant; the active control alternative, has the same games except that they do not adapt to performance, remaining instead at baseline levels. At Visit 2 are introduced to their respective program and are asked perform it 3 days per week (every other day), for 20 minutes, for 24 total sessions. Four months after the intervention, subjects receive a 1-week booster training. Visits 3, 5, and 6 are outcome assessments at the immediate finish of the intervention, six months post-intervention, and 12 months post-intervention, respectively (Figure 1).

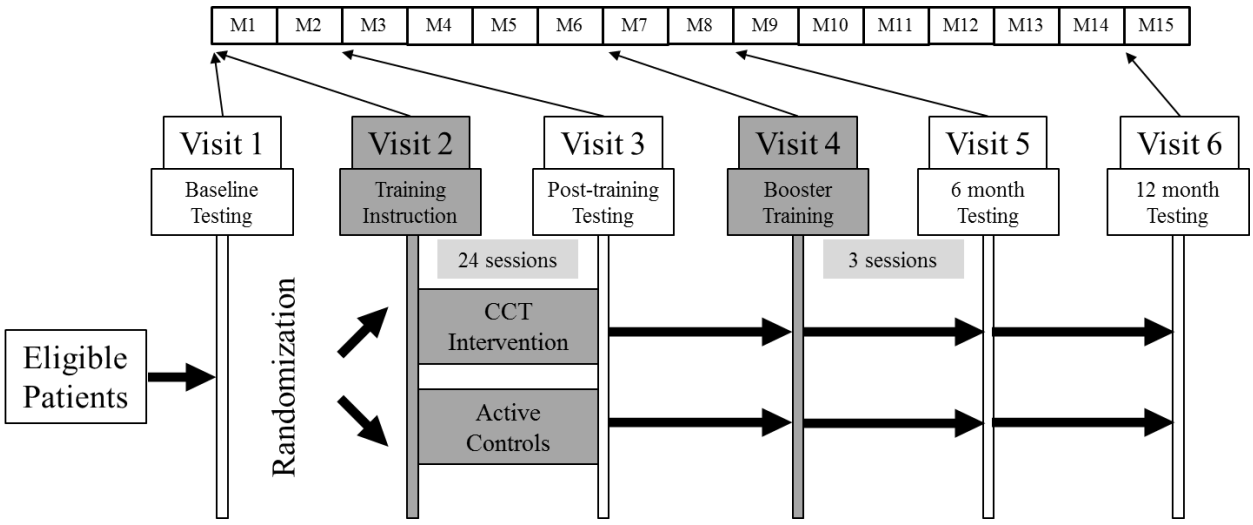

Figure 1. Timeline of visits from the initial assessments (Month 1; M1) to the final assessments 14 months later.

One of the challenges of internet-based behavior change programs is keeping participants interested and engaged for a sufficient duration, which can be important to ensure a therapeutic benefit. To strengthen participation rates and adherence, staff members call participants once a week, to ask if they are performing the program and if they have any technical difficulties. Nevertheless, we acknowledge that despite these efforts, attrition of subjects occur. Drop out are recorded and Cognifit maintains records, available to this study, of the date and time of each session for each participant in each group.

3. Primary Analysis Objective

The primary outcome measure of this randomized clinical trial (RCT) is the change in DSMQ from baseline (Visit 1) to 6 months after the intervention (Visit 5). The primary analysis is the intent-to-treat (ITT; all enrolled and eligible participants who are randomized) comparison of the two arms of the trial.

4. Analysis Sets/Populations/Subgroups

Intent-to-treat (ITT)

- Differences between the two arms will be assessed among participants with a Clinical Dementia Rating (CDR) score of zero (i.e., excluding those with uncertain dementia [CDR = 0.5]).
- Differences between the two arms will be assessed among participants with and without at least 12 years of education.
- Differences between the two arms will be assessed by sex.

For each seven secondary outcomes, the primary analysis will like the primary analysis for the DMSQ.

For each primary and secondary outcome, a secondary analysis will include the primary covariates to reduce variation attributable to subject differences.

Analyses by CCT Performance

- Group differences between the two arms will be assessed among participants who complete at least 18 of the 24 CCT training sessions as well as the follow-up visit being examined (i.e. Visit 3, 5, or 6).
- Total time spent training (across all sessions by the participant) are data calculated by Cognifit and available to the investigators. This will be used to explore whether training time modifies group differences.
- CCT performance scores are calculated at each session by Cognifit using assessment measures that remain constant regardless of any changes for the participant. For those with at least 18 of 24 trainings sessions, a linear slope will be calculated to explore whether this modifies group differences.

5. Endpoints and Covariates

Endpoints

The primary endpoint is the change from baseline DSMQ to Visit 5.

Seven secondary endpoints, also change from baseline to Visit 5 (except for medication adherence), are:

- Memory – averaged from neuropsychological tests standardized using baseline scores:
    - Logical Memory - Immediate Recall
    - Logical Memory - Delayed Recall
    - CERAD Word List Memory – Immediate Recall

CERAD Word List Memory – Delayed Recall
CERAD Word List Memory –Recognition

- Executive function/attention – averaged from neuropsychological tests standardized using baseline scores:
  - Cancellation Test - Letter
  - Cancellation Test – Shape
  - Trail Making Test – A
  - Trail Making Test – B
  - Digit Span - Forward
  - Digit Span - Backward
- Prospective memory
  - Rivermead time-based subtest (Appointment)
- Medication adherence – by change in renewals and refills of diabetes medications between one year before Visit 1 and one year after Visit 3
- Hemoglobin A1c (HbA1c)
- Diastolic blood pressure – measured twice while sitting and at rest with a five minute interval
- Systolic blood pressure – measured twice while sitting and at rest with a five minute interval

Primary covariates

- age
- sex
- years of education
- baseline MMSE

Secondary covariates

- site – Bronx, NY VAMC versus Ann Arbor, MI VAMC
- ethnicity – Non-Hispanic European Ancestry vs. Other
- Cognitive Self-Report Questionnaire
- Geriatric Depression Scale
- Short – Test of Functional Health Literacy in Adults
- Total Training Sessions

Calculating time

Dates of Visits 1 through 6 are recorded and individual CCT training dates are maintained by the Cognifit and available for use. As indicated in Figure 1 (above), pre-specified intervals are targeted for each visit, but actual visits vary due to scheduling

contingencies with participants and staff. Thus, intervals between assessments are determined from days since Visit 1 (baseline assessment**).**

6. Handling of Missing Values and Other Data Conventions

Missing data will be accounted for using iterative full information maximum likelihood estimation of the relevant model parameters.

A separate three-step procedure will be used for the sixteen neuropsychological test scores at each visit. The first two steps substitute values for missing data and the third step creates derived scores for use in analyses.

The first step is limiting the participants for whom their neuropsychological data at that visit will be analyzed. Although most participants have data on all the variables, others have some missing data. A participant's data for a visit will be analyzed only if there are at most two missing variables of the 16. A participant must have data analyzed at Visit 1, but may have data analyzed after missing other visits.

The second step is creation of a set of complete data for a visit, by substituting values for any missing data. For all neuropsychological tests except Trails A and Trails B, higher scores indicate good cognition. These two tests are multiplied by minus one, reversing them to make them agree with the other tests. For each variable, using the mean and standard deviation (SD) of the non-missing data, z-scores are calculated by z = (value – mean) / SD for participants with data on that variable. Participants with missing data have no z-score; the "mean substitution" strategy would substitute the mean of the z-scores, which is zero by definition. An improved strategy substitutes the mean of the participant's z-scores for the non-missing variables in that cognitive domain (in addition to memory and executive functions/attention, the remaining 5 neuropsychological tests are in a language domain that is not an outcome measure of this RCT). After this substitution, each participant has score, which is transformed back to the original units by new value = score * SD + mean. This returns the z-scores for non-missing observations to their original values, and provides an estimate for a participant with missing data based on the participant's results in the cognitive domain of the variable. The result is a set of new values with no missing data, for every participant with data analyzed for that visit. Note that the means and standard deviations for the new values differ from those of the original incomplete data.

The third step is to create derived scores for each visit from the new values data of that visit, and the mean and SD of the Visit 1 new values. For every visit, derived score = (new value - Visit 1 new value mean) / Visit 1 new value SD. For Visit 1, the derived score is the z-score, but this is not the case for the other visits, since their derived score is not based on the mean and SD of the new values for that visit. By calculating derived scores this same way at every visit, the participant's change score between a pair of visits is the same as the change for the neuropsychological test, after transformation in the same way as both derived scores for that test. At every visit, derived domain scores for memory and executive functions/attention are sums of the relevant derived scores. By calculating domain scores this same way at every visit, domain change scores between a pair of visits are the weighted sums of the respective change scores for the

neuropsychological tests, with weights based on the transformations of the derived scores.

Medication adherence will be measured by the Continuous, Multiple Interval Measure of Medication Gaps (CMG) method. Using the CPRS diabetes medication refill data, we will identify: 1) refills occurring within one year prior to baseline (Visit 1); and, 2) refills occurring within one year after completion of intervention (Visit 3). To calculate pre-intervention and post-intervention continuous medication gap percentages (CMG), cumulative gaps in medication adherence (the sum of the number of days for which each medication was not available, based on refill amounts and dates) will be divided by the total number of days between the first and last medication fill.

For calculation of linear slope of performance scores, a secondary covariate, we will exclude extreme assessment scores (defined here as above or below 2 standard deviations from the mean – calculated excluding the extreme score itself) since such scores are more likely to be spurious than a true reflection of change.

7. Statistical Methodology

Summary of Study Data

Descriptive statistics for all relevant variables will be calculated and examined at baseline and subsequent visits. The memory and executive functions/attention scores will be the derived domain scores after taking account of missing date. Substantially non-normal distributions will be transformed to reduce the influence of extreme values. All continuous variables will be summarized using the following descriptive statistics: n (non-missing sample size), mean, standard deviation, median, quartiles, maximum and minimum. The frequency and percentages (based on the non-missing sample size) of observed levels will be reported for all categorical measures. All summary tables will be structured by group.

Statistical Procedures

The analytic procedure to address the efficacy of CCT will use linear mixed models (LMM) procedures to evaluate change from baseline (Visit 1) to six months post-treatment (Visit 5). Models will include time (coefficient estimate β_1), intervention group (coefficient estimate β_2), and the intervention group by time interaction (coefficient estimate β_3) which indicates whether the change in outcome over time differs between the intervention groups, all as fixed factors, and subjects as a random factor.

The model for the primary outcome (DSMQ) or any of the seven secondary outcome variables (Y) for subject *i* and visit *j* at time t_{ij} is:

$$Y(t_{ij}) = b_0 + b_1 t_{ij} + b_2 \text{Group}_i + b_3 t_{ij}\text{Group}_i + u_i + e_{ij}$$

where bs estimate the fixed effects of group and time, u is the subject random effect, and e is the residual error. If the result of randomization is that there is a significant ($P < 0.05$) discrepancy between the two arms of the RCT on any of the primary covariates,

that covariate will be included in the model. Since this estimate will be valid for the ITT population under the MAR assumption, it will not be biased in favor of groups with greater dropout.

Although hypothesized demonstration of benefit of CCT beyond that of the active control is directional, tests of significance will be a two-sided at the 0.05 level, to test for the possibility that CCT is substantially inferior to the active control.

Similar procedures will be employed to compare the treatment arms on change from baseline to other visits for the all outcomes.

Analysis for medication adherence

In contrast to all other outcome measures, which are assessed at each visit and primarily used to compare Visit 1 to Visit 5, medication adherence is measured only twice the year before Visit 1 and the year after Visit 3. The LMM and model are exactly the same except the two times are Visit 1 and a year after Visit 3 for the particular participant.

Measures to Adjust for Multiplicity, Confounders, Heterogeneity, Etc.

As this study may be considered preliminary to a potential subsequent large-scale multi-site study of CCT throughout the VA, each secondary outcome will be tested at the 0.05 level.

Secondary analyses controlling for primary covariates are described in #4, ITT analysis. With regard to secondary covariates, such as ethnicity, health literacy, or computer experience, if any are unbalanced or correlate with an outcome at $p<0.05$, it will be used as a covariate in a secondary analysis, to reduce within-group variation.

Subject Disposition & CONSORT Diagram

We will report the number of subjects, in each group and in the aggregate, that reached:

- Screen
- Randomization
- Baseline
- At least one follow-up
- Each follow-up

We will also summarize attrition rate and reasons for discontinuation if we have them.

8. Sensitivity Analysis

Not applicable.

9. Interim Analysis

Not applicable.

10. Rationale for Any Deviation from Pre-Specified Analysis Plan

Not applicable

11. QC Plans

All variables are examined for errors.

12. Programming Plan

Data are maintained in Excel files. Statistical analysis will proceed using SAS 9.4, Stata v13.1 and SPSS.

13. Appendices

Not applicable.