Study Title: A Multidomain Intervention Program for Older People with Dementia

NCT#: NCT04948450

Document Date: October 7, 2022

Document Type: Statistical Analysis Plan

Statistical Analysis Plan

Data were analyzed on an intention-to-treat (ITT) basis, wherein participants were analyzed according to their treatment assignment by randomization, regardless of their compliance.

Descriptive statistics, etc.

Categorical variables were presented as percentages, continuous variables were presented as mean ± standard deviation, medians and confidence intervals. Fisher's exact test was used to compare categorical variables between the two groups, while Mann-Whitney test was used to compare continuous variables between the two groups.

Estimation of Treatment Effects

Continuous outcomes

To investigate the treatment effects on continuous outcomes, we utilized Linear Mixed-effects Model (LMM) with random slopes. The treatment effects were estimated by modeling the interaction between time (T: after vs. before intervention) and the treatment arms (A: treatment vs. control). The specification of the models is provided below:

$$Y_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})^* T_{ij} + \beta_2 T_{ij}^* A_i + \varepsilon_{ij}$$

Where:

$Y_{ij}|T, A, b_i \sim N(\mu, \sigma_y^2)$ is the study outcome of i^{th} individual at j^{th} measurement

β_2 is the treatment effect of interest

$b_i = (b_{0i},\ b_{1i})' \sim N(0,\ D)$ are random intercepts and random slopes, respectively, with D unstructured variance-covariance matrix

$\varepsilon_{ij} \sim N(0, \Sigma)$ are the random errors, with Σ compound symmetry variance-covariance matrix

We constrained the baseline values of the outcome to be the same between the two arms in the model by omitting the group main effect (i.e., main effect of A). This procedure is to gain efficiency of the model and is valid under randomization.[1]

Binary outcomes

For binary outcomes, the Generalized Estimating Equation models (GEE) were applied as our causal effects of interest are the population-average effects. The log link function was utilized to obtain the RR estimates. The specification of the models is provided below:

$$Y_{ij} \sim Bernoulli(\pi_{ij})$$

$$\log(\pi_{ij}) = \gamma_0 + \gamma_1 T_{ij} + \gamma_2 T_{ij}^* A_i$$

Where: γ_2 represents the treatment effect of interest. We specified the exchangeable working assumption and obtained the empirical standard errors. The same constraint for baseline values, as mentioned in the LMMs, was applied.

Handling missing values

We applied direct likelihood method, which is valid under LMM, to address missing values for continuous outcome models.

For binary outcomes models, the Multiple Imputation by Chained Equations (MICE) method was employed. This algorithm created multiple imputed datasets based on a set of imputation models, assuming missing at random. We utilized logistic regressions as the imputation models for binary outcomes. The imputed model included the outcome variable and candidate predictors, including age, sex, marital status, education, and BMI. Due to small sample size, we could not impute all outcomes in the same model, therefore, each outcome was imputed separately. In total, 20 imputed datasets were created for each outcome. The estimates were estimated in each imputed dataset and then were pooled to obtain the single estimate based on Rubin´s rules. We also compare the estimates based on multiple imputation to the results of available-case analyses.