Statistical Analysis Plan

Improving Neurological Health in Aging via Neuroplasticity-based Computerized Exercises (INHANCE)

Protocol Number: PSC-0903-19

NCT04149457

Document Date: February 22, 2023

The data analysis plan defines a primary ITT population, a single primary outcome measure, a set of exploratory outcome measures, a single primary evaluation time point, an exploratory evaluation time point, a primary statistical analysis methodology, a criterion for statistical significance, and guidance for the interpretation of results.

The primary ITT population is defined as all participants who were randomized and completed the first training session. This includes all randomized participants except those who drop or withdraw post randomization and pretraining.

The primary outcome measure assesses forebrain FEOBV binding in the anterior cingulate cortex at baseline and posttest between the intervention and active control groups. The exploratory outcome measures are FEOBV binding across additional ROIs (posterior anterior cingulate, primary auditory cortex, primary sensorimotor cortex, parietal lobe, frontal lobe, occipital lobe, temporal lobe, global cortex, hippocampus, parahippocampal gyrus, putamen, caudate, striatum, nucleus basalis of Meynert), NIH EXAMINER composite score, Double Decision and Freeze Frame train-to-task assessment scores, heart rate variability, pupillometry, and MRI volumetrics. The 3-month follow-up assessment is an exploratory timepoint.

The primary statistical analysis methodology is a linear mixed model approach. We will first compare intervention and active control groups in the ITT population at baseline using t tests or chi-square tests to determine if differences in baseline variables remain after the randomization process. Variables that show a significant α ($P<.05$) will be noted and used as covariates in the model. We will examine the data for each outcome measure using a linear mixed effects model with treatment group and time as fixed factors, and covariates as necessary from the baseline analysis. Missing data will be accounted for using iterative full-information maximum likelihood estimation. Within-group effects for each time point (posttraining, follow-up) will be calculated using data from each group. Between-group effects for each time point will be calculated in the same way, adding an interaction term (training group × time) to estimate the effect of cognitive training on outcome measure change. P values and effect sizes will be reported.

To evaluate the efficacy of BrainHQ to upregulate acetylcholine, we will conduct the analysis based on the baseline and posttest data. Finding a significant α of $P<.05$ on the FEOBV primary outcome measure will support the statement that the intervention increases acetylcholine binding.

To evaluate the endurance of cognitive effects following the completion of training (NIH EXAMINER), we will conduct confirmatory analyses based on the posttest and follow-up assessment data to evaluate the endurance of the change (maintenance, decline, or increase).

Showing maintenance or improvement between posttest and follow-up will support the statement that a brief intensive training epoch maintains (or improves) cognition after training has ceased.

Additional exploratory analyses will use Pearson or Spearman correlations as relevant and will compare FEOBV binding against the NIH EXAMINER composite score, all exploratory ROIs, the train-to-task assessments, heart rate variability, pupillometry, MoCA performance, demographic variables, and the number of levels trained for the ITT. We will also conduct these analyses for those who completed the minimum effective dose of 10 hours of training or achieved a score on the Double Decision posttest assessment of 400 milliseconds or faster. We will compare the NIH EXAMINER composite score with the primary and exploratory ROIs,

train-to-task assessments, heart rate variability, pupillometry, MRI volumetrics, MoCA, and the number of levels trained for the ITT, and for those who completed the 10-hour minimum and achieved 400 milliseconds or faster on the Double Decision posttest assessment. Responder analyses will evaluate key moderators (age, sex, MoCA, baseline cognitive performance on NIH EXAMINER composite, baseline cognitive performance on the train-to-task assessments, and number of levels completed) to assess who responds most to training. No correction for multiple comparisons will be made for exploratory analyses and all trending relationships ($P<.10$) will be reported.