# In-Home Technology for Caregivers of People With Dementia and Mild Cognitive Impairment: Spanish Language Homes

Unique Protocol ID: 1247267250000-2
Secondary IDs: 2SB1AG059458-04A1
Identifiers: NCT05159596
Document Date: July 12, 2023

# Statistical Analysis Plan

## Primary Outcomes

All caregivers–regardless of randomized study condition–complete questionnaires on three separate occasions during this study: baseline, 3-months, 6-months. The four outcome measures include changes over the 6-month study period for the following scales: Zarit Burden Interview-Short Form, a questionnaire measuring caregiver burden (Zarit, Reever, & Bach-Peterson, 1980). 12 items are rated on 0-4 scale. Range: 0-48. No subscales. Higher scores represent worse outcomes; Center for Epidemiological Studies Depression Scale (CES-D), a questionnaire measuring depression (Radloff, 1977). 20 items are rated on a 0-3 scale and summed (range = 0-60). There are no subscales. Higher scores represent worse outcomes. The clinical cut-off is usually set at a score of 16; Beck Anxiety Inventory (BAI), a questionnaire measuring anxiety (Beck, Epstein, Brown, & Steer, 1988). 20 items are rated on a 0-3 scale and summed (range= 0-60). Higher scores indicate worse outcomes. There are no subscales. A score greater than 36 is considered to be clinically significant.; and Satisfaction with Life Scale, a questionnaire measuring overall life satisfaction and well-being (Diener, Emmons, Larsen, & Griffin, 1985). 5 items scored on a 1-7 scale and summed (Range = 5-35). Lower scores indicate worse outcomes. A score of 20 is considered neutral with higher scores considered increasingly more satisfied and lower scores considered increasingly more dissatisfied.

## Statistical Methods

*Design:* Caregivers will be randomized into two groups: Active In-Home Technology System and Limited In-Home Technology System. Subjects will be randomized 1:1. Randomization of participants in each group will be determined by People Power employees. Analyses of caregiver health and well-being will use 2 X 3 (Condition [active treatment vs control] X Time [time 1 vs time 2 vs time 3]) mixed model analysis of variance. In all analyses, significant main effects and interactions will be decomposed using a priori or post hoc comparisons depending on whether they are hypothesized or not. Bonferroni corrections will be used to control for Type I error resulting from multiple comparisons. Statistical Power. Tests of hypothesized Group differences (active treatment vs control condition) for caregiver health and well-being will have .80 power to detect a medium effect size (one-tailed, alpha = .05). Using repeated measures analysis of variance, tests of the interaction of Condition (2 groups) X Time (3 assessments) will have power of .99 to detect a medium effect size (correlation between repeated measures r = .5, alpha = .05).

*Analysis:* Descriptive statistics will be used to summarize all the clinical characteristics and demographic outcomes. These will be summarized overall, by treatment group, by assessment time, and by treatment group and time. Continuous variables will be summarized using means, standard deviations, and ranges. Categorical data will be summarized by number and percent. To evaluate the treatment effects, we will use repeated measures ANOVAs. The other predictors will be treatment group, and time (baseline, 3-months, and 6-months for caregiver outcomes). To evaluate treatment effects, an interaction term between treatment group and time will be included

to examine treatment group-related differences in the assessments for each outcome. Each significant interaction will be followed up with paired t-tests to examine treatment differences at each set of time points. Although this study involves 4 outcomes, we do not plan formal multiple comparison adjustments if the results fit a coherent pattern that is consistent with the context of similar studies. In this case, each result will reinforce the other, rather than detracting from one another, as required by formal multiple comparisons adjustments such as Bonferroni. Conversely, if only one or a very few measures reach statistical significance and their directions and/or magnitudes do not coherently fit with our a prior expectations, then we will note that the result(s) with $p < 0.05$ lack plausibility and could be due to chance, despite meeting the conventional cutoff for statistical significance.

*Missing Data:* If subjects either cannot or refuse to complete surveys we will use a sensitivity analysis approach in dealing with these missing data by performing analysis with 2 versions of scores: 1) with missing item interpolated using mean replacement by group, and 2) only using responses with 0% missing data. If the results differ for the two approaches, then we will report on the differences.

## REFERENCES

Bedard, M., D.W. Molloy, L. Squire, S. Dubois, J.A. Lever, and M. O'Donnell, The Zarit Burden Interview: a new short version and screening version. Gerontologist, 2001. 41(5): p. 652-7.

Radloff, L.S., The CES-D scale: A self-report depression scale for research in the general population. Applied Psychological Measurement, 1977. 1: p. 385-401.

Beck, A.T., N. Epstein, G. Brown, and R.A. Steer, An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol, 1988. 56(6): p. 893-7.

Diener, E., R.A. Emmons, R.J. Larsen, and S. Griffin, The Satisfaction With Life Scale. Journal of Personality Assessment, 1985. 49(1): p. 71-75.