Care Ecosystem: Navigating Patients and Families Through Stages of Care
NCT02213458
August 6, 2018

Care Ecosystem Trial Statistical Analysis Plan

Primary Outcome

The primary outcome is patient quality of life as measured by the Quality of Life – Alzheimer's Disease (QoL-AD)[1], a 13-item measure using a 1-4 ordinal scale for each item. The item scores are summed for a total score ranging from 13-52. This will be measured at baseline and 12 months.

Secondary Outcomes

Secondary patient outcomes will all focus on healthcare utilization and will be caregiver reported number of visits to the emergency department, number of hospitalizations, and utilization of an ambulance. Questions are adapted from the Health and Retirement Survey[2] and will be measured at baseline and 12-months. At baseline, utilization for the past 12 months will be queried for a stable estimate, and frequencies will be divided by two. On the 12-month survey, caregivers will be asked these questions about the prior 6 months. The period between baseline and 6-months post-randomization will not be included, because during this time the care plan and implementation will be in process for PWD randomized to the Care Ecosystem. Caregiver depression will be measured using the Patient Health Questionnaire 9-item version (PHQ-9).[3] The total score range is 0-27 with greater scores indicating more depressive symptoms. Caregiver burden will be measured using the Zarit Burden Interview (Zarit-12).[4] This 12-item measure has a 1-4 ordinal scale that is summed for a total score from 0-48 with scores of 17 or greater representing a high level of burden. Caregiver self-efficacy will be measured using a 4-item measure to assess perceived knowledge around where to get needed services and support, and confidence around managing future caregiving challenges.[5]

Caregiver measures will be measured at baseline, 6 months, and 12 months, starting earlier in accordance with the care model timeline: care will be first focused on immediate caregiver needs, and

the care plan for the patient will be developed over a series of monthly phone calls. The primary comparison for all outcomes will be at 12 months compared to baseline by group. All outcome measures will be collected via the phone with the caregiver by a research coordinator blind to treatment group assignment.

Statistical Methods

Design: Patient-caregiver dyads will be randomized into two groups: Survey of Care (SoC) and Care Ecosystem (CE) within each of two sites (California and Nebraska or Iowa). Subjects will be randomized 2:1 CE:SoC. To increase the likelihood of similar types of patients with dementia (PWDs) in the CE and SoC groups, we will stratify the sample using randomization blocks by dementia severity (mild, moderate, and advanced) and referral site for the California cohort, and dementia severity and geographic locale (urban / suburban and rural / frontier) for the Nebraska cohort.

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, medians, standard deviations and ranges. Categorical data will be summarized by number and percent. Bivariate analyses comparing groups and time points will be summarized using t-tests and one-way analysis of variance for continuous variables that are not extremely skewed and by Wilcoxon rank sum and the Mann-Whitney U test for variables requiring non-parametric statistics. Categorical analyses will use Chi-squared tests of independence, Fisher exact tests, and Z-tests of proportions.

To evaluate the treatment effects, we will use linear mixed effects models. Baseline dementia severity, as measured by the Quick Dementia Rating Scale (QDRS),[a,6] will be included as a covariate in

[a] The original plan was to use the Functional Assessment Staging Tool as the measure of dementia severity. The QDRS was published shortly after study initiation and was felt to be a more appropriate measure for the st

all treatment effect analyses. The other predictors will be treatment group, and time (baseline, 6-months, and 12- months for caregiver outcomes; and baseline and 12-months for patient outcomes). To evaluate for treatment effects, an interaction term between treatment group and time will be included to examine treatment group-related differences in the change from baseline to later assessments for each outcome. The random effect will be included to capture differences in implementation of the interventions within patients over time and the hierarchical nature of the trial. Although this study involves 7 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 and supports the primary outcome result. 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 priori expectations, then we will note that the result(s) with $p<0.05$ lack biological plausibility and could be due to chance, despite meeting the conventional cutoff for statistical significance.[7-11]

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) based on regressing on the remaining item scores that had been answered. If the results differ for the two approaches, then we will report on the differences.

REFERENCES

1. Logsdon RG, Gibbons LE, McCurry SM, Teri L. Assessing quality of life in older adults with cognitive impairment. Psychosomatic medicine 2002;64:510-9.

2. Hurd MD, Martorell P, Delavande A, Mullen KJ, Langa KM. Monetary costs of dementia in the United States. N Engl J Med 2013;368:1326-34.

3. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. Journal of general internal medicine 2001;16:606-13.

4. Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. Gerontologist 1980;20:649-55.

5. Merrilees JJ, Bernstein A, Dulaney S, et al. The Care Ecosystem: Promoting self-efficacy among dementia family caregivers. Dementia 2018:1471301218814121.

6. Galvin JE. The Quick Dementia Rating System (QDRS): A Rapid Dementia Staging Tool. Alzheimer's & dementia (Amsterdam, Netherlands) 2015;1:249-59.

7. Feise RJ. Do multiple outcome measures require p-value adjustment? BMC medical research methodology 2002;2:8.

8. Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology (Cambridge, Mass) 1990;1:43-6.

9. Savitz DA, Olshan AF. Multiple comparisons and related issues in the interpretation of epidemiologic data. American journal of epidemiology 1995;142:904-8.

10. Perneger TV. What's wrong with Bonferroni adjustments. BMJ (Clinical research ed) 1998;316:1236-8.

11. Bacchetti P. Peer review of statistics in medical research: the other problem. BMJ (Clinical research ed) 2002;324:1271-3.