STATISTICAL ANALYSIS PLAN

4/21/2025

Sponsor Name: NIH/National Institute on Aging (NIA)
Sponsor Type: Federal - NIH
Sponsor Role: Funding
Grant/Contract Number: R01 AG061973-01
Funding Through: Federal-NIH
Award Number: R01 AG061973-01
Grant Title: GamePlan4Care: Web-based Delivery System for REACH II

ClinicalTrials.gov Identifier: NCT04540198

Contents

1   Overall Summary from ClinicalTrials.Gov

Daily care and supervision of a person living with dementia (PWD) has been defined as "intense caregiving" and is associated with significant daily burdens and an overall threat to the caregiver's quality of life. Despite evidence suggesting that negative consequences can be remediated with community-based supports, those services remain elusive to caregivers due to the systemic challenges of turning interventions into services. This study is based on a practical approach of applying technology to an existing evidence-based intervention, Resources for Enhancing Alzheimer's Caregiver Health II (REACH II), refined with real-world user feedback and rigorously tested to create an online family caregiver support system that has the potential of both scalability and sustainability. This two-group randomized controlled trial will compare the relative impact of GP4C to an education-based online site called Resources4Care (R4C) on a wide range of family caregiver outcomes. Family caregivers will be randomized to one of the two conditions and will complete an assessment battery at baseline and at the 6-month follow-up.

2   Specific Aims

Family care of an older adult is a valued tradition of society and has become an essential element of the US healthcare system with 83% of long term care provided to older adults coming from family members or other unpaid helpers[1]. As the population of older adults grows, almost doubling in size from 2012 to 2040[2], so too will the expectations placed on American families. This is especially true of families who support community living for individuals living with Alzheimer's disease and related dementias (ADRD), which has been defined as "intense caregiving" and is associated with significant daily burdens and an overall threat to the caregiver's quality of life[3]. A robust body of ADRD research demonstrates that interventions aimed at supporting and educating caregivers can significantly improve the quality of care delivery and improve the well-being and quality of life for both caregivers and persons with dementia (PWD)[4]. Despite the fact that over 200 interventions have been found to be effective for dementia caregivers in randomized clinical trials (results confirmed in 7 meta-analyses and 17 systematic reviews), only six have undergone translation efforts resulting in publication[5]. The lack of credible translation of evidence-based interventions into widely available community-based services means that the clear majority of the current 16.1 million family caregivers do not have access to evidence informed long term care supports and services. Thus, *family caregivers remain at risk of the known negative consequences of caregiving despite evidence suggesting that those negative consequence can be remediated*. There is an immediate and critical need to address this issue which is of significant societal impact.

This study is based on a practical approach of applying technology to an existing evidence-based intervention, Resources for Enhancing Alzheimer's Caregiver Health II (REACH II), refined with real-world user feedback and rigorously tested with the goal of creating an online family caregiver support system that has the potential of both scalability and sustainability. The feasibility of our proposition has been established by the creation of the proof-of-concept system ***GamePlan4Care (GP4C)***, demonstrating that modern internet technology can be used to automate key components of REACH II delivery. Such automated REACH II components include active engagement of caregivers in skill-building to address diverse challenges in caregiving and flexible tailoring of the intervention based on participant needs. ***Therefore, GamePlan4Care is positioned for a Phase 1 study to incorporate the full breadth of REACH II material and optimize the user experience for efficacy testing in a randomization clinical trial that engages BSWH clinics and community-based organizations (CBOs) where family caregivers already seek services.***

Two specific staged AIMS are proposed in this study.

AIM 1. **Advance the current proof-of-concept GP4C into a viable delivery system for the REACH II intervention.** Further development will include uploading the full breadth of REACH II education and skill-building materials, usability testing and updates to the user interface/user experience (UI/UX), creating a portal of delivery tools used by Dementia Care Specialist (DCS; i.e., interventionist), and establishing a comparable education-focused online system (Resources4Care; R4C) to be used in comparison to the multicomponent, skills-training GP4C system. Development activities will be conducted by the Baylor Scott & White Health (BSWH) Office of Digital Health with engagement of community-based organizations, family caregivers, and external advisors, with oversight by the GP4C research team.

H₁: *Within 12 months of usability testing with 32 family caregivers, the GP4C research team will reach consensus with the BSWH Office of Digit Health and an external Advisory Committee that GP4C is a viable platform for online delivery of REACH II with* R4C *as an appropriate comparison condition.*

AIM 2. **Compare the relative impact of GP4C and R4C, both of which include access to a Dementia Care Specialist, on a wide range of family caregiver outcomes.** Family caregivers will be randomized to one of the two conditions and will complete an assessment battery at baseline and at the 6-month follow-up. Primary outcome measures will include burden, depression, social support, and caregiver's health and interaction with health care providers. Secondary outcomes will include caregiver stress, positive aspects of caregiving (PAC), cognitive change in care-recipient. The CBOs will designate a "GP4C Champion" to promote the enrollment of 240 family caregivers who have computer access and report regular use of the internet. Use of GP4C & R4C and contact with DCS will be recorded by the system as implementation data.

H₁: Family caregivers randomized to the GP4C condition will report greater improvement on primary and secondary outcomes as compared to family caregivers in the R4C condition.

GP4C will fill a critical gap in caregiver support services as an evidence-based, internet-enabled system capable of providing immediate, tailored education and skills training to caregivers who can access live support from a DCS via phone or web-based video. Compared to existing evidence-based interventions, CBOs will face fewer barriers to scaling and sustaining GP4C, a critical achievement in efforts to serve the growing number of family caregivers who face the challenges of dementia caregivers.

This Statistical Analysis Plan is limited to AIM 2,

a Phase 1 study of two dementia caregiver interventions.

3   Summary of the Two Study Groups

- Experimental: GamePlan4Care (GP4C)
    - o Participants in this arm will have access to full functionality and content of the online system GamePlan4Care (GP4C) including educational resources, skills training, and support tailored to their unique caregiving needs. Additional individualized feedback will be automatically generated based on responses to online questions and will include links to relevant site educational/skill-building content. Participants will be assigned a Dementia Care Specialist who will facilitate caregiver interactions with the online material and provide skills training via telephone or web-video conference. Study participants assigned to GP4C will receive 9 automated emails and 4 phone calls over a 6-month period.

- Active Comparator: Resources4Care (R4C)
    - o Participants in this arm will receive access to Resources4Care (R4C), a feature-limited version GamePlan4Care system. R4C will serve as an online hub for articles and videos about Alzheimer's disease and dementia. Educational topics will include information on: 1) Alzheimer's Disease & Dementia, 2) Caregiving, 3) Caregiver Stress and 4) Home Safety. R4C will present a page on each topic with active links to two additional online sources on the same topic. Study participants assigned to R4C will receive two emails from their DCS encouraging the caregiver to review specific education materials. Each email will be followed by brief "check-in" calls (15-min each) at three months and five months after randomization.

4   Entry Criteria

4.1   Inclusion Criteria

- Must be age 18 years or older
- Providing at least 8 hours of weekly care and/or supervision (on average) for a friend or family member with a self-reported diagnosis of Alzheimer's disease or a related dementia. Family will be subjectively determined by the caregiver to enable a broader definition of a" "family" member often found in minority communities (e.g., a person not related by blood but who serves in the role of an "aunt" or "grandchild").
- The family member, named as the care recipient (CR) in this proposal, must be diagnosed with AD/ADRD (self-report from the caregiver accepted) and is experiencing signs of dementia as verified by the family caregiver on the AD8 informant interview. A score of 2 or greater is the inclusion criteria.
- Must demonstrate access to a home computer with internet access to research staff and report using the computer to access the internet at least three times per week, on average.
- English-speaking caregivers
- Must reside within the recruitment area (Target counties within Texas: Bastrop, Bell, Blanco, Burnet, Caldwell, Coryell, Fayette, Hamilton, Hays, Lampasas, Lee, Llano, Milam, Mills, San Saba, Travis, Williamson)

4.2   Exclusion Criteria

- Current participation in another caregiving evidence-based program
- Previous participation in usability testing for current system development

4.3   Sample Size Considerations

The primary endpoint in this power analysis is the Zarit's caregiving burden score. Assuming 20% attrition rate at the 6-month follow-up, we plan to initially enroll totally 240 family caregivers who will be equally randomized into two will be equally randomized into two arms (120 per arm), and we expect a minimum of 100 caregivers per arm to complete the 6-month assessment. The sample size of 100 in each arm can guarantee 80% power at significance level 0.05 to detect the effect size of 0.28 standard deviation (SD) difference between baseline and 6 months follow-up in a two-sided paired t-test. (See Table 4). Nicholas et al.[6] reported the effect size of 0.33SD of Zarit's score from the REACH VA intervention (also between baseline and 6-month follow-up). While GP4C will use a different delivery strategy than used in the REACH II and REACH VA trials, core therapeutic materials and techniques of the intervention are maintained. Thus, we expect GP4C to have a similar impact on caregiver outcomes as observed in prior REACH II trial. Thus, we conclude that enrolling 120 caregivers per arm (100 available at 6 months) will achieve sufficient statistical power to detect the effectiveness of the proposed interventions, even if the GP4C full functionality intervention is slightly less robust than the original REACH interventions.

Table 4. Sample size and power by effect size

Effect size of 0.28 standard deviation (SD) difference		Effect size of 0.33 standard deviation (SD) difference	
Sample Size	Power	Sample Size	Power
70	.646	70	.777
80	.706	80	.830
90	.757	90	.872
100	.800	100	.905
110	.834	110	.929
120	.868	120	.948
130	.893	130	.962
140	.914	140	.972
150	.931	150	.980

5   Outcome Measures

Change from Baseline Caregiving Burden at 6 months

A 12-item version of the Zarit Caregiver Burden Interview[7]. The Zarit Caregiver Burden Interview (ZBI) is an assessment tool for evaluating caregiver burden. The ZBI consists of 12 items representing a statement related to some aspect of perceived burden. Respondents (i.e., caregivers) rate each item ranging 0 (=never) to 4 (=nearly always). Total ZBI score is the summation of 12 items ranging from 0 to 48. Higher scores indicate greater burden.

[Time Frame: Baseline and six months]

Change from Baseline Depression at 6 months

Levels of depressed symptoms:

A 10-item version of the Center for Epidemiological Studies Depression (CESD) Scale[8,9] is an assessment tool for evaluating depression. Each item represents a statement for which respondents indicate how often in the past week they have felt that way from 0 (= rarely or none of the time) to 3 (= most or almost all of the time). Total score is the summation of 10 items ranging from 0 to 30. Higher scores indicate higher levels of depressed symptoms.

[Time Frame: Baseline and six months]

Change from Baseline Self-rated health at 6 months

Assessment of perception of general health status:

A single item of self-rated health[10] assess caregiver's perception on his/her own health from 0 (=excellent) to 4 (=poor).

[Time Frame: Baseline and six months]

Change from Baseline Social Support at 6 months

Assessment of availability of support and satisfaction with support from others:

Two constructs from Social Provision Scale[11] (reliable alliance and guidance) are used to evaluate social support. Two constructs assess availability of support and satisfaction with support from others. Four items for each construct (total of 8 items) describe a statement for which respondents agree to what extent their relationships are with other people from 1 (= strongly disagree) to 4 (=strongly agree). A total score is the summation of 8 items ranging from 8 to 32. Higher scores indicate higher levels of social support.

[Time Frame: Baseline and six months]

Change from Baseline Interaction with Health Care Providers at 6 months

Assessment of the frequency and experiences of medical care management and coordination for care-recipients:

Nine items from the National Study of Caregiving (NSOC)[12,13] ask about experiences and frequency of caregiver's interactions with care-recipients' health care providers. Each item represents a statement for which respondents indicate how often in the last 6 months they interacted with health care providers/workers.

[Time Frame: Baseline and six months]

Change from Baseline Caregiver Stress at 6 months

Assessment of levels of stress:

The Perceived Stress Scale (PSS)[14] is an assessment tool for evaluating stress level. The PSS consists of 10 stress items. Each item represents how often in the last month respondents felt that way from 0 (=never) to 4 (=very often). A total score is the summation of 10 items ranging from 0 to 40. Higher scores indicate higher levels of stress.

[Time Frame: Baseline and six months]

Change from Baseline Reported Positive Aspects of Caregiving at 6 months

Assessment of favorable aspects of caregiving experiences:

An 11-item of Positive Aspect of caregiving (PAC)[15] is an assessment tool for evaluating favorable aspects of caregiving experiences. Each item represents a statement of mental or affective state related to caregiving experiences. Respondents rate to what extent each statement corresponds to their mental or affective state ranging from 0 (=disagree a lot) to 4 (=agree a lot). A total score is the summation of 11 items ranging from 0 to 44. Higher scores indicate higher levels of positive experiences of caregiving.

[Time Frame: Baseline and six months]

Change from Baseline Neuropsychiatric Symptoms in care-recipient/corresponding caregiver distress at 6 months

Presence and severity of neuropsychiatric symptoms and levels of relevant caregiving distress:

Neuropsychiatric Inventory Questionnaire-Q (NPI-Q)[16], a 12-item self-administered questionnaire (NPI-Q), completed by the caregivers about care-recipients for whom they care, is a tool to assess the presence and severity of 12 Neuropsychiatric Symptoms (NPS) in patients with dementia, as well as the caregiver's corresponding distress. Each item asks the presence (1=yes, 0=no), severity (1=mild, 2=moderate, 3=severe), and caregiving distress (0=not at all; 5=extreme or very severe).

[Time Frame: At Baseline]

Program Evaluation of GP4C/R4C

Assessment of attitude toward GP4C/R4C system:

A 30-items of USE questionnaire will evaluate participant's attitude toward the GP4C system. The USE questionnaire[17] includes four areas: assess usefulness (8 items), ease of use (11 items), ease of learning (4 items), and satisfaction (7 items). Respondents rate agreement with the statements, ranging from strongly disagree (=1) to strongly agree (=7). Total scores are the summation of items on each area. Higher scores in each indicate a more useful, easier to use, easier to learn, and more satisfied on their experience.

[Time Frame: At six months only]

A Composite Psychological Distress Score

A composite score of psychological distress will be constructed. Three measures, depression (CES-D), caregiver stress (PSS), and NPI-Q Distress, will be transformed linearly to have a range from 0 to 100 and then combined with the same weight. This replicates the creation of a dementia burden composite score used in the literature[18].

6   Statistical Methods and Analyses

Caregiver participants will be randomized to the two intervention arms (GP4C and R4C) with the ratio of 1:1. Primary and secondary outcomes, including the Zarit's caregiving burden interview, will be assessed at baseline and 6-month follow-up. Covariates in the analysis include demographic characteristics of the caregivers. Common descriptive statistics (mean, standard deviation, median, interquartile range for continuous variable; frequency and percentage for categorical variables) will be calculated for all covariates in each intervention arm and the two arms will be compared at baseline. Goals of our statistical analysis will be to evaluate the effectiveness of each intervention arm and to compare the two arms. For each intervention arm, the difference between 6-month follow-up and baseline will be computed. Paired t-test and Wilcoxon signed-rank test will be used to detect significant changes in outcome variables at 6-month follow-up since the baseline. Medians, means, and 95% confidence intervals of the differences will be calculated using both empirical percentiles and normal approximation. Multiple comparison adjustments, e.g., Holm-Bonferroni method and Hochberg procedure, will be implemented to control the family wise error rate when testing multiple outcome measures in Section 5.

The two intervention arms will be compared at each follow-up time and for the whole trial period. We will compare the two arms at 6-month follow-up using the two-sample t-test and Wilcoxon rank sum test for each of the two outcomes. The Anderson-Darling test will be used to check the normality assumption at baseline and follow-up times for the two outcome variables. If normal assumption is significantly violated then only nonparametric estimates and tests (i.e., median, IQR, empirical confidence intervals, Wilcoxon signed-rank test, and Wilcoxon rank sum test) will be reported, and data transformation, such as log and square root transformations, will be considered to normalize the outcome variables for further model-based testing. A pre-post analysis will be conducted to model the outcomes at 6-month follow-up to test the intervention arm and covariates adjusting the baseline measurements. The two arms will also be compared incorporating observations from baseline and follow-up in longitudinal analysis of covariance (ANCOVA) models or linear mixed-effects models having the follow-up time and adjusting for covariates.

Given the fact that participants will be randomized into the two arms, the outcome variables will likely be comparable between the two arms at baseline. As a result, the interaction effect between the treatment arm and follow-up time (while adjusting for covariates) will be tested in the longitudinal analysis to evaluate the difference between the two intervention arms because a significant interaction indicates the caregiver's outcome has followed significantly different trajectories in the follow-up period depending on the treatment arm. Missing values may be possible because 1) some caregivers may be lost to follow-up, and 2) some caregiving recipients can be deceased prior to the follow-up. Missing patterns will be inspected and multiple imputation methods, including regression imputation, propensity score imputation, and MCMC imputation will be considered based on the missing mechanism[19]. All tests will be two-sided and a p-value less than or equal to 0.05 will be reported as statistically significant. Statistical Analysis Software[20] will be used to conduct the proposed analysis.

7   References

[1] Friedman EM, Shih RA, Langa KM, Hurd MD. US Prevalence And Predictors Of Informal Caregiving For Dementia. *Health Aff (Millwood)*. Oct 2015;34(10):1637-41. doi:10.1377/hlthaff.2015.0510

[2] Ortman JM, Velkoff VA, Hogan H. *An Aging Nation: The Older Population in the United States*. 2014. P25-1140. May 2014. https://www.census.gov/library/publications/2014/demo/p25-1140.html

[3] Pinquart M, Sorensen S. Differences between caregivers and noncaregivers in psychological health and physical health: a meta-analysis. *Psychol Aging*. Jun 2003;18(2):250-67. doi:10.1037/0882-7974.18.2.250

[4] Sorensen S, Pinquart M, Duberstein P. How effective are interventions with caregivers? An updated meta-analysis. *Gerontologist*. Jun 2002;42(3):356-72. doi:10.1093/geront/42.3.356

[5] Gitlin LN, Marx K, Stanley IH, Hodgson N. Translating Evidence-Based Dementia Caregiving Interventions into Practice: State-of-the-Science and Next Steps. *Gerontologist*. Apr 2015;55(2):210-26. doi:10.1093/geront/gnu123

[6] Nichols LO, Martindale-Adams J, Burns R, Graney MJ, Zuber J. Translation of a dementia caregiver support program in a health care system--REACH VA. *Arch Intern Med*. Feb 28 2011;171(4):353-9. doi:10.1001/archinternmed.2010.548

[7] Bedard M, Molloy DW, Squire L, Dubois S, Lever JA, O'Donnell M. The Zarit Burden Interview: a new short version and screening version. *Gerontologist*. Oct 2001;41(5):652-7. doi:10.1093/geront/41.5.652

[8] Radloff LS. The CES-D scale: A self-report depression scale of research in the general population. *Applied Psychological Measurement*. June 1977 1977;1(3):385-401. doi:10.1177/014662167700100306

[9] Irwin M, Artin KH, Oxman MN. Screening for depression in the older adult: criterion validity of the 10-item Center for Epidemiological Studies Depression Scale (CES-D). *Arch Intern Med*. Aug 9-23 1999;159(15):1701-4. doi:10.1001/archinte.159.15.1701

[10] Idler EL, Hudson SV, Leventhal H. The Meanings of Self-Ratings of Health: A Qualitative and Quantitative Approach. *Research on Aging*. 1999;21(3):458-476. doi:10.1177/0164027599213006

[11] Cutrona CE, Russell DW. The provisions of social relationships and adaptation to stress. In: Jones WH, Perlman D, eds. *Advances in Personal Relationships*. JAI Press; 1987.

[12] Wolff JL, Freedman VA, Mulcahy JF, Kasper JD. Family Caregivers' Experiences With Health Care Workers in the Care of Older Adults With Activity Limitations. *JAMA Netw Open*. Jan 3 2020;3(1):e1919866. doi:10.1001/jamanetworkopen.2019.19866

[13] Freedman VA, Skehan ME, Hu M, Wolff J, Kasper JD. National Study of Caregiving I-III User Guide. Johns Hopkins Bloomberg School of Public Health; 2019. https://urldefense.com/v3/__http:/www.nhats.org__;!!JA_k2roV-A!FU3pD1hB0EGpt5MaAxTK1hA6zV2cSqP_g1oGQVddkPPHVqyUZTiGXjbjHlRxvqC3zSLTLKHI59nVCmtpiM6eg_v8ft765Gkkazk$

[14] Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav*. Dec 1983;24(4):385-96. doi:10.2307/2136404

[15] Tarlow BJ, Wisniewski SR, Belle SH, Rubert M, Ory MG, Gallagher-Thompson D. Positive Aspects of Caregiving:Contributions of the REACH Project to the Development of New Measures for Alzheimer's Caregiving. *Research on Aging*. 2004;26(4):429-453. doi:10.1177/0164027504264493

[16] Kaufer DI, Cummings JL, Ketchel P, Smith V, MacMillan A, Shelley T, et al. Validation of the NPI-Q, a brief clinical form of the Neuropsychiatric Inventory. *J Neuropsychiatry Clin Neurosci*. Spring 2000;12(2):233-9. doi:10.1176/jnp.12.2.233

[17] Lund AM. Measuring Usability with the USE Questionnaire. *STC Usability SIG Newsletter: Usability Interface*. January 2001 2001;8(2):3-6.

[18] Peipert JD, Jennings LA, Hays RD, Wenger NS, Keeler E, Reuben DB. A Composite Measure of Caregiver Burden in Dementia: The Dementia Burden Scale-Caregiver. *J Am Geriatr Soc*. Sep 2018;66(9):1785-1789. doi:10.1111/jgs.15502

[19] Schafer JL, Reid N, Keiding N, Louis T, Tong H, Isham V. Analysis of Incomplete Multivariate Data. Chapman & Hall/CRC; 1997.

[20] SAS Institute Inc. SAS, Version 9.4. Cary, NC 2013.