

**Reducing Burden on Hispanic Caregivers of Relatives with Dementia -- The Northern
Manhattan Hispanic Caregiver Intervention Effectiveness Study (NHICE)**

José A. Luchsinger, MD, MPH¹, Louis Burgio, PhD², Mary Mittelman, DrPH³, Ilana Dunner,
LMSW⁴, Jed A. Levine, MA⁵, Jian Kong, MS⁶, Stephanie Silver, MPH⁶, Mildred Ramirez, PhD⁶,
Jeanne A. Teresi, PhD, EdD⁶

¹Departments of Medicine and Epidemiology, Columbia University Medical Center, New York, New York

²Burgio Geriatric Consulting, Tuscaloosa, Alabama

³Departments of Psychiatry and Rehabilitation Medicine, New York University School of Medicine, New York, New York

⁴Riverstone Senior Life Services, New York, New York

⁵CaringKind, (formerly known as the Alzheimer's Association, New York City Chapter), New York, New York

⁶Hebrew Home at Riverdale, Research Division, and Weill Cornell Medical College, New York, New York

Original Project Title: Northern-Manhattan Hispanic Caregiver Intervention Effectiveness Study (NHICE)

PCORI ID: CE-1304-7160

HSRProj ID: 20143534

ClinicalTrials.gov ID: NCT02092987

To cite this document, please use: Luchsinger J, Burgio L, Mittelman M, et al. 2019. *Reducing Burden on Hispanic Caregivers of Relatives with Dementia -- The Northern Manhattan Hispanic Caregiver Intervention Effectiveness Study (NHICE)*. Washington, DC: Patient-Centered Outcomes Research Institute (PCORI). <https://doi.org/10.25302/3.2019.CE.13047160>

Table of Contents

ABSTRACT	3
BACKGROUND	5
3.1. Introduction	5
3.2. Overview and aims	7
3.3. Participation of patients and other stakeholders in the design, conduct, and dissemination of the research	7
METHODS	8
4.1. Study design	8
4.2. Study population	8
4.2.1. Recruitment	9
4.2.2. Screening	10
4.2.3. Participant subgroups	10
4.3. Study setting(s)	11
4.4. Interventions	11
4.5. Data collection and sources.....	13
4.5.1. Primary outcome measure—burden.....	14
4.5.2. Primary outcome measure—depression	14
4.5.3. Secondary measures for caregiver burden	14
4.5.4. Secondary measures of depressive symptoms	15
4.5.5. Exploratory outcome measures	15
4.5.6. Other measures.....	16
4.6. Research processes	16
4.6.1. Collection of study assessments.....	16
4.6.2. Intervention procedures	16
4.7. Analytical and statistical approaches burden	17
RESULT	22
5.1. Recruitment and Retention	22
5.2. Sample and group characteristics	23
5.3. Correlations of primary and secondary outcomes	29
5.4. Results of analyses of primary and secondary outcomes	29
5.5. Results for secondary measures of burden and depression.....	38
5.6. Results for exploratory outcomes.....	38
5.7. Subgroup analyses	39
5.7.1. Subgroup analyses for primary outcomes	41
5.7.2. Subgroup analyses for secondary outcomes.....	41
5.7.3. Subgroup analyses for exploratory outcomes	42
5.8. Dose-response analyses outcomes	43
5.9. Sensitivity analyses	43
5.10. Adverse events	43
DISCUSSION	43
6.1. Subpopulation considerations	45
6.2 Limitations	46
6.3 Future research	46
6.4 Generalizability and implementation of study results.....	46
CONCLUSION	47
REFERENCES	48

2. Abstract

Background: Dementia prevalence and dementia caregiving are increasing. Caregiver burden leads to adverse health outcomes. Behavioral interventions designed to improve adverse health outcomes in caregivers are available. Two of the best-known caregiver interventions with demonstrated efficacy are the Resources for Enhancing Caregiver Health (REACH), and the New York University Caregiver Intervention (NYUCI). The effectiveness of these interventions has never been compared. Thus, caregivers of persons with dementia and caregiver service providers do not know if one intervention is better than the other. In addition, a paucity of data exists on the effectiveness of these interventions in urban Hispanics, who have a higher risk of dementia caregiving burden compared with non-Hispanic whites.

Objectives: Our primary objective was to compare the effectiveness of the NYUCI to a short adaptation of REACH (REACH-OUT [Offering Useful Treatment]), in improving depressive symptoms and caregiver burden among unpaid Hispanic caregivers of persons with dementia. Our exploratory objectives were to examine outcomes such as stress and physical health, and examine whether sociodemographic characteristics (eg, family position) predicted the comparative effectiveness of the interventions.

Methods: We conducted a 1:1 randomized trial comparing the NYUCI and REACH-OUT in 221 Hispanic unpaid caregivers of persons with dementia. The interventions lasted 6 months. Participants underwent a baseline evaluation before the interventions, and a follow-up assessment after the completion of interventions. All participants received referrals for social work services also. Our primary outcomes were depressive symptoms, measured with the Geriatric Depression Scale (GDS), and caregiver burden, measured with the Zarit Caregiver Burden Scale (ZCBS). We compared changes in outcomes from baseline to follow-up using mixed models with an intent-to-treat approach.

Results: The mean age of caregivers was 58 years; 82.8% were women, 63.3% were adult children, and 31.7% were spouses; 110 were randomized to the NYUCI and 111 to REACH-OUT. No differences in pertinent characteristics existed at baseline. For the ZCBS, a net reduction (improvement in burden) of about five points (maximum score = 88) occurred in each intervention group ($p = 0.006$). Using the model-based estimates, the effect size for REACH-OUT was a 5.18-point reduction ($p = 0.0006$) and for the NYUCI a 4.58-point reduction ($p = 0.0022$). For the GDS, no significant changes occurred from baseline to follow-up (0.03 points [$p = 0.9529$] for REACH-OUT; 0.19 points [$p = 0.6991$] for NYUCI). For our exploratory stress outcome (Perceived Stress Scale), a reduction of approximately 2 points (maximum score = 40) occurred in each group ($p = 0.0296$). The reduction of the ZCBS within spouse and child subgroups was significant

among spouse caregivers (8.00 points, $p = 0.0047$ for REACH-OUT; 9.44, $p = 0.0010$ for NYUCI), but not among children (3.23, $p = 0.0765$ for REACH-OUT; 2.11, $p = 0.2163$ for NYUCI). The reduction of the ZCBS was also greater among persons aged 58 years and older than among persons less than 58 years old. Most persons younger than 58 years were adult children.

Conclusions: No differences in the effectiveness for REACH-OUT and the NYUCI existed for any outcome. Both REACH-OUT and the NYUCI were effective in reducing burden and stress, but not depressive symptoms. Subgroup analyses suggest that the effect on burden was significant only among spouse caregivers.

Registration: ClinicalTrials.gov NCT02092987

3. Background

3.1. Introduction

In 2013 the Alzheimer's Association estimated that 1 in 8 elders has Alzheimer's dementia and that these individuals are cared for by 15.4 million unpaid caregivers, resulting in \$214 billion in annual costs.¹ The RAND corporation reported that dementia is the most expensive condition to care for, including the costs from unpaid caregiving by relatives.² Caring for persons with dementia is challenging.³⁻⁵ They require intense supervision and care, risking caregivers' psychological, physical,⁶⁻⁸ and financial health.¹ Caregiver stress leads to premature nursing home placement for the person with dementia,^{8,9} but caregivers report emotional and physical stress even after nursing home placement.^{10,11}

Dementia prevalence in Hispanics is several times higher than in non-Hispanic whites (NHWs) nationally (27.9% versus 10.9% in persons aged 75-84 years; 62.9% versus 30.2% in persons 85 years and older)¹ and in New York City.^{12,13} Our project, **Northern-Manhattan Hispanic Caregiver-intervention Effectiveness study (NHICE)**, *focused on Hispanics because they suffer disproportionately from dementia and its related caregiving burden*. Hispanic unpaid caregivers experience more depression and higher burden due to unique cultural characteristics. Most caregivers report no guilt after nursing home placement,¹ but this is less common in Hispanics,^{1,14} who traditionally delegate less care to paid caregivers.¹⁴ A 2011 nationwide *telephone survey* conducted by the Alzheimer's Association highlighted Hispanic caregivers' unique characteristics. They are younger on average than NHW and non-Hispanic black (NHB) caregivers,¹ less likely to be married than NHWs, more likely to have children or grandchildren under age 18 in their household than NHWs and NHBs, more likely to be a primary caregiver than NHWs and Asian-Americans (AAs); more likely to earn < \$50 000 annually than NHWs and AAs, and more likely to need help balancing work and family than NHWs.¹ Small studies have shown Hispanics experience more strain and less social support than other racial/ethnic groups,^{15,16} despite extensive social networks,¹⁷ and less acculturated Hispanic caregivers experience more depression.¹⁸ Hispanics tend to underutilize formal services for the care of persons with dementia, believing that caregiving is a responsibility of the family, a phenomenon that has been named "familism."¹⁹

The two best known interventions for caregivers of persons with dementia, the New York University Caregiver Intervention (NYUCI)²⁰⁻²² and the Resources for Enhancing Caregivers' Health (REACH),²³ have never been compared, and a dearth of data exists on these interventions for Hispanics. The NYUCI is primarily oriented toward improving social networks that support families, while REACH focuses on

building skills related to caregiving. Because the NYUCI has a greater focus on the family than REACH, consistent with the reported familism in Hispanics, we hypothesized that the NYUCI would be more effective in *reducing depressive symptoms and caregiver burden* in Hispanics compared with REACH.

REACH demonstrated efficacy among Cuban-Hispanics in South Florida and Mexican-Hispanics in California in a randomized controlled trial (RCT).²³ The NYUCI demonstrated efficacy in white spouse caregivers, and more recently in white adult children caregivers in Minnesota,²⁴ but did not have evidence of efficacy from an RCT in Hispanics and had scant data on nonspouse caregivers.

The NYUCI was evaluated in a longitudinal RCT over more than two decades that included 406 spouse caregivers (89% white sample).²⁵ The intervention alleviated the deleterious effects of caregiving on the mental and physical health of spouse caregivers, and postponed or prevented nursing home placement of their spouses with Alzheimer's dementia. The intervention's effects on caregiver depression, which appeared after the 6-month intervention period, around the 10th month, were long-lasting and continued through nursing home placement and death of the person with Alzheimer's dementia.^{21,26} The NYUCI is listed and described in detail on the National Registry of Evidence-based Programs and Practices website.

REACH I was a National Institutes of Health-funded cooperative agreement wherein six intervention sites investigated six different intervention packages with a common outcomes assessment battery.²⁷ Spouse and adult child caregivers were trained in behavior management and general problem-solving skills.²⁸ REACH I was effective in reducing behavior problems in care recipients.²⁹ REACH II, designed based on REACH I, was an *intensive* tailored caregiver intervention with multiple components to address multiple needs of caregiver-care receiver dyads. The therapeutic approach was problem solving. Equal numbers of NHBs, NHWs, and Hispanics were recruited. There was significant improvement in a composite measure of quality of life compared with controls in all ethnic groups.²³ The REACH II protocol was too intensive for practical use in the community. Consequently, in 2010 the REACH Offering Useful Treatment (OUT) protocol was developed; this is the protocol implemented in NHICE. REACH-OUT is listed as an evidence-based treatment on the Alliance for Caregiving and Roselyn Carter Center for Caregiving websites.

Despite being the two most widely implemented caregiver interventions, the effectiveness of the NYUCI and REACH-OUT had never been compared before NHICE (RQ-1). NHICE was the first pragmatic randomized clinical trial^{30,31} to compare the effectiveness of the two best-known dementia caregiver interventions.

3.2. Overview and aims

The goal of NHICE was to compare the effectiveness of the NYUCI with REACH-OUT in alleviating depressive symptoms and caregiver burden. The primary aim was to compare the effectiveness of the NYUCI and REACH-OUT for 200 Hispanic caregivers of persons with dementia in New York City in a pragmatic RCT. We compared changes in 2 primary outcomes between baseline and follow-up: depressive symptoms, measured with the Geriatric Depression Scale (GDS), and caregiver burden, measured with the Zarit Caregiver Burden Scale (ZCBS). **We hypothesized** that depressive symptoms and caregiver burden would improve in both intervention groups, but that these outcomes would improve more or deteriorate less in caregivers receiving the NYUCI.

We examined the following **exploratory aims**:

1. Compare the NYUCI and REACH-OUT in changes in caregiver health (measured with the Patient Reported Outcomes Measurement Information System [PROMIS] physical health scale) and stress (measured with the Perceived Stress Scale).
2. Examine the predictors (demographics, family position, acculturation, and social support) of success of the NYUCI and REACH-OUT.

3.3. Participation of patients and other stakeholders in the design, conduct, and dissemination of the research

The scientific stakeholders responsible for the study design were Columbia University Medical Center (CUMC), where the study was based, and our research partner, the research division of the Hebrew Home for the Aged in Riverdale, New York, led by Jeanne Teresi, PhD. Dr. Teresi and her team were primarily responsible for designing the statistical plan, calculated the sample size, and assisted the principal investigator in identifying the primary and secondary outcomes for the study. Other scientific stakeholders included New York University (NYU) and Burgio Geriatric consulting. NYU is the institution where coinvestigator Mary Mittelman, DrPH is based. Dr. Mittelman was the developer of the NYUCI, and she was responsible for the training and supervision of the counselor delivering the NYUCI in NHICE. Burgio Geriatric consulting is the base for coinvestigator Louis Burgio, PhD, who codeveloped REACH and REACH-OUT and conducted the efficacy studies for REACH. Dr. Burgio was responsible for the training and supervision of the counselor delivering REACH-OUT. Our community-based partner, Riverstone

Senior Life Services, the stakeholder representing persons with dementia, their informal caregivers, and providers of social work services, was responsible for the design and delivery of the social work support services that all participants received. Riverstone helped with outreach and recruitment. The New York City chapter of the Alzheimer's Association, now called CaringKind, was the other major stakeholder representing persons with dementia, their caregivers, and service providers. CaringKind's main role was promoting the study citywide and helping with recruitment.

We formed a caregiver advisory committee comprising 10 participants in the study who were fluent in English and Spanish. This advisory committee was consulted on the dissemination plan and provided feedback on what could be improved in terms of support services for caregivers. Although caregivers were not directly involved in the design of the study, we held 3 focus groups with caregivers in January 2010 that informed the development of the project; in particular, the need to include social services referral for all participants. The participants in these focus groups informed us that, given the prevalent low socioeconomic status of the community surrounding CUMC, social work service referral was necessary in order for caregiver interventions to be effective.

4. Methods

4.1. Study design

NHICE was a Pragmatic RCT^{30,31} designed to compare the effectiveness of the NYUCI and REACH-OUT. We took into consideration the PCORI draft methodology report in choosing the study design.³² The possibilities included randomized trials and nonrandomized studies. Because a comparison of the NYUCI and REACH-OUT had never been conducted, we decided to conduct an RCT. We considered this trial as pragmatic because we did not have the strict inclusion and exclusion criteria of the efficacy studies of the NYUCI and REACH, and because we conducted the study in a community setting, partnering with stakeholders in a way that resembled how the interventions might be implemented in real life.

4.2. Study population

In general, the sampling frame was the population of adult Hispanic unpaid relative caregivers of persons with dementia living in New York City. Most participants were from the community of Northern Manhattan, where CUMC and Riverstone Senior Life Services are located. We considered for inclusion all

relative informal caregivers, including divorced spouses and nonmarried partners of the person with dementia. Paid caregivers, such as home aides, were not eligible to participate. Our previous experience taught us that this sample is the most representative of Hispanic caregivers in New York City.³³ Using their networks of caregivers, our community partners—Riverstone and CaringKind—provided assistance with outreach and recruitment activities.

4.2.1. Recruitment

We had learned from previous experience that we needed to engage multiple sources of recruitment continually, with an emphasis on sources where persons with dementia or their caregivers (who were the target participants of our study) seek services. The study team held weekly meetings on Tuesdays, and at each meeting reviewed the progress in recruitment and had a checklist of recruitment sources summarized here:

- Clinical practices in which persons with dementia and/or their caregivers were followed for medical care at CUMC/New York Presbyterian Hospital. We promoted the study among primary care clinicians, nurses, and social workers. These providers referred potential participants. The principal investigator confirmed the diagnosis of dementia. The team then contacted the caregiver.
- Promotion of the study among caregivers serviced by the Alzheimer’s Association, New York City Chapter (now renamed CaringKind), one of our stakeholder partners
- Recruitment of caregiver clients from Riverstone Senior Life Services, our other stakeholder partner
- Outreach to agencies that provide services to caregivers in New York City

The following were our sources of actual recruitment:

- Alzheimer’s Association New York City Chapter (75 participants)
- Alianza Dominicana (1 participant)
- Ambulatory Care Network of New York Presbyterian Hospital (41 participants)
- Memory clinic at CUMC (10 participants)
- Mount Sinai Hospital Center Alzheimer Disease Assistance Center (20 participants)
- Riverstone Senior Life Services (24 participants)
- Talks at caregiver events (40 participants)

4.2.2. Screening

Once the name of a potential participant with contact information was received, the dedicated research assistant in the study conducted a screening call. If the participant was deemed eligible, the participant was scheduled for the baseline exam.

Inclusion criteria were as follows:

- Both care recipient and caregiver self-identified as Hispanic
- Both care recipient and caregiver between the ages of 18 and 90
- Caregiver related to persons receiving care either as a spouse (including common law partners) or a blood or in-law relative
- Care recipient diagnosed with dementia and reported at least 1 memory/cognition and 1 daily functioning symptom in our screening questionnaire
- Caregiver physically able to provide care
- Caregiver had no diagnosis of major psychiatric disorder other than depression
- Caregiver had no depression with psychotic features or suicidal ideation or attempts in the past five years
- Caregiver expected to live in New York City for the study duration
- Caregiver had least one relative or close friend living in the New York City Metropolitan Area (New York, New Jersey, Connecticut). The rationale for this inclusion criterion is that the NYUCI requires the participation of relatives or friends of the primary caregiver in the intervention.

Exclusion criteria were as follows:

- Not Hispanic
- Caregiver not a relative as defined in the inclusion criteria
- Paid caregivers, such as home aides
- Care recipient without dementia
- Caregiver had a major psychiatric disorder other than depression, such as schizophrenia
- Caregiver had depression with psychotic features or suicidal ideation in the past five years

4.2.3. Participant subgroups

We learned from anecdotal experience³³ that the following characteristics formed relevant subgroups to our project:

- Family position of caregiver (spouse versus adult children). This is important because adult children are more likely to have competing burdens from work and family.
- Age (older versus younger). This is important because older caregivers may have more disability, although they may be retired and have more time for the intervention and for caregiving itself. We expected that age would be highly correlated with family position; that is, that spouses would more likely be older, and adult children more likely be younger.
- Acculturation (more acculturated versus less acculturated). We speculated that this was important because more acculturated persons might do better with the intervention that was more skills-based (REACH-OUT), and less acculturated persons might do better with the NYUCI, which focuses more on enhancing family social support.
- Satisfaction with social support (more versus less satisfaction with social support). We thought that since the NYUCI targets building social support in the circle of family and friends, it would be more effective among those with lower social support.

4.3. Study setting(s)

The study was based at CUMC. Counselors for REACH-OUT and the NYUCI, the study coordinator who collected outcomes data, and the principal investigator were all based at the Division of General Medicine at CUMC. The interventions and data collection were conducted at CUMC or at a place of the participant's preference, including his or her home. At CUMC, the study procedures were conducted at the outpatient facility of the Columbia Clinical and Translational Science Award, called the Irving Institute for Clinical Translational Research, or at its community facility, called the Columbia Community Partnership for Health. The data collected were transmitted to the research division of the Hebrew Home for the Aged in Riverdale, where the data coordinating center (DCC) was located. The DCC was responsible for randomization to the study interventions, data management, and the analyses presented in this report. The New York City Chapter of the Alzheimer's Association, renamed CaringKind in 2016, assisted with study dissemination and with participant recruitment. Riverstone provided social work services for study participants and provided help with outreach and recruitment.

4.4. Interventions

Most dementia caregiving research, including the NYUCI, REACH, and its translation REACH-OUT, has been guided by Pearlin's stress process model (SPM).³⁴ The SPM proposes that caregiving context variables affect each part of the stress process and can have implications for the types of stressors facing

caregivers. It posits that the perception or appraisal of those stressors and social support are mediators that can affect outcomes such as caregiver depression. These mediators can be affected in positive ways by both REACH-OUT and NYUCI.

NYUCI and REACH-OUT have important similarities: (1) They are based on the same conceptual model, (2) they are designed to ameliorate caregiver depression and subjective stress/burden, (3) they involve direct interaction between the counselor and the caregiver/family, (4) the interventions focus on similar skills/topics during the intervention, and (5) both allow tailoring of the intervention to the needs of the caregiver/family. However, they differ on two characteristics that prior research has found to be potent factors in interventions: therapeutic technique and mode of delivery. The table below compares the NYUCI and REACH-OUT. The duration and “dosing” of both interventions in NHICE were similar. Thus, we could compare the effects of mode of delivery independent of the duration and intensity of the interventions. Counselors with similar qualifications who were fluent in English and Spanish delivered the interventions.

Comparison of REACH-OUT and NYUCI Interventions

	REACH-OUT	NYUCI
Mode of Delivery	<ul style="list-style-type: none"> • Sessions in the caregiver’s home • One-on-one interaction between counselor and caregiver • Family involvement encouraged but not required 	<ul style="list-style-type: none"> • Sessions in caregiver home or other mutually convenient place • One-on-one interaction between counselor and caregiver • Family involvement required • Support group participation strongly advised
Caregiver Skills/Counseling Topics	<ul style="list-style-type: none"> • Caregiver self-care and health behaviors • Emotional well-being • Enhancing social support • Environmental safety • Behavior management • Education about dementia and effects on the caregiver • Resource referrals 	<ul style="list-style-type: none"> • Caregiver self-care and health behaviors • Enhancing social support with emphasis on family support • Environmental safety • Behavior management • Education about dementia and effects on the caregiver and family

Comparison of REACH-OUT and NYUCI Interventions

	REACH-OUT	NYUCI
		<ul style="list-style-type: none"> • Resource referrals
Therapeutic Techniques	<ul style="list-style-type: none"> • Skills training through use of formal problem solving • Use of written action plans, homework assignments, and counselor tracking of progress toward goals 	<ul style="list-style-type: none"> • Interpersonal counseling with caregiver and family counseling required
Degree of Tailoring	<ul style="list-style-type: none"> • Caregivers taught full skill set; however, dosage of specific skills based on standardized risk assessment • Ad hoc access to counselor via phone to discuss action plans 	<ul style="list-style-type: none"> • Ad hoc counseling available as needed • Counseling is stage-related • Topics and styles of counseling tailored to the caregiver, person with dementia, and family

4.5. Data collection and sources

Rationale for the primary, secondary, and exploratory outcome measures. The primary constructs that we sought to examine as outcomes were caregiver burden and depressive symptoms, because these were the main constructs examined in previous studies of REACH and the NYUCI. We chose a primary measure for each of these constructs in order to be parsimonious and to avoid chance findings. We chose the ZCBS as the measure of caregiver burden because this measure showed improvement in Hispanic communities in South Florida in the original REACH studies.^{23,36,37} For depressive symptoms, we chose the Geriatric Depression Scale (GDS). Although numerous measures of depression exist, the selection of the GDS as the primary measure of depression was based on the use of this measure in studies of the NYUCI, and findings from the literature related to differential item functioning (DIF) in measures of depression. Although there are few studies of DIF among Hispanics, a review³⁸ of DIF in depression measures showed that more common depression measures were biased for ethnically diverse groups. The GDS has fewer such items and fewer items with somatic content. The sample size and power considerations were made based on the ZCBS and the GDS. Thus, the conclusions of the study depend on the results for these primary outcomes. However, we opted to include secondary

outcome measures examining the same constructs in order to complement the primary outcomes. The rationale for including these outcomes is that they have been used in other caregiver studies, or are scales that are increasingly used. We strived to include measures from the Patient Reported Outcomes Measurement Information System (PROMIS). We included exploratory outcomes measures addressing other constructs, such as stress and physical function, because they might be important for caregivers in addition to the primary outcomes.

Described below are the reliability and unidimensionality estimates for the scales used in NHICE. Ordinal alpha and McDonald's omega from a single factor model were used to estimate reliability, and the explained common variance was used as an estimate of essential unidimensionality of the measures. The estimates of the psychometric properties of the measures for this sample were good to excellent.

4.5.1. Primary outcome measure—burden

Zarit Caregiver Burden Scale (ZCBS).³⁹ The ZCBS consists of 22 items, such as “Do you feel embarrassed by his/her behavior?” Items were rated on a 5-point Likert scale ranging from “never” to “nearly always.” The ZCBS score increases with higher caregiver burden. The ordinal alpha coefficient for this sample was estimated at 0.934, the explained common variance was 33.305, and the McDonald's omega was 0.936.

4.5.2. Primary outcome measure—depression

Geriatric Depression Scale (GDS).⁴⁰ The GDS consists of 30 yes/no items measuring depression, such as “Do you frequently feel like crying?” Items were rated in the deviant direction, and then summed to create the scale, with a higher score indicating greater depression. The alpha coefficient for this sample was 0.957, the explained common variance was 65.723, and the McDonald's omega was 0.959.

4.5.3. Secondary measures for caregiver burden

Lawton Caregiver Appraisal Scale: Burden subscale.⁴¹ The Lawton Caregiver Appraisal Burden subscale has nine items, such as “Do you feel that because of the time you spend with your Elder that you don't have enough time for yourself?” The five-point Likert scale responses ranged from “never” to “nearly always.” The Lawton scale score increases with higher caregiver burden. The alpha coefficient for this sample for the burden subscale was 0.914, the explained common variance was 83.700, and the McDonald's omega was 0.914.

Montgomery Caregiving Objective Burden Scale⁴²: The Montgomery Caregiving Objective Burden Scale consists of six items, such as whether caregiving responsibilities have “decreased time you have to yourself.” Items were rated on a five-point Likert scale ranging from “not at all” to “a great deal.” The Montgomery scale score increases with higher caregiver burden. The alpha coefficient for this sample was 0.965, the explained common variance was 96.409, and the McDonald’s omega was 0.966.

4.5.4. Secondary measures of depressive symptoms

PROMIS Depression Short Form^{43,44}: The PROMIS Depression Short Form contains eight items measured on a five-point scale; responses range from “never” to “always” over the past seven days. Items include “I felt worthless” and “I felt hopeless.” The PROMIS depression score increases with higher depressive symptoms. The alpha coefficient for this sample was 0.957, the explained common variance was 80.132, and the McDonald’s omega was 0.957.

Patient Health Questionnaire (PHQ-9)⁴⁵: The PHQ-9 assesses 9 depressive symptoms associated with diagnosis of depressive disorders. Responses range from “not at all” to “nearly every day” over the past two 2 weeks. The PHQ-9 score increases with higher depressive symptoms. The alpha coefficient for this sample was 0.874, the explained common variance was 39.963, and the McDonald’s omega was 0.876.

4.5.5. Exploratory outcome measures

Perceived Stress Scale⁴⁶: The Perceived Stress Scale contains 10 items to appraise the stressfulness of life situations over the past month; responses are “never,” “almost never,” “sometimes,” “fairly often,” and “very often.” Items include “In the last month, how often have you felt that you could not cope with all the things you had to do?” The Perceived Stress Scale score increases with higher reported stress. The alpha coefficient for this sample was 0.905, the explained common variance was 67.330, and the McDonald’s omega was 0.907.

PROMIS Physical Function Short Form^{47,48}: The PROMIS Physical Function Short Form contains eight items measured on a five-point scale. Four items range from “without any difficulty” to “unable to do” (eg, “Are you able to go for a walk of at least 15 minutes?”) and four range from “not at all” to “cannot do” (eg, “Does your health now limit you in lifting or carrying groceries?”). The alpha coefficient for this sample was 0.970, the explained common variance was 89.228, and the McDonald’s omega was 0.970.

4.5.6. Other measures

We measured acculturation with the Bi-dimensional Acculturation Scale.⁴⁹ The Bi-dimensional Acculturation Scale has two domains (Hispanic and non-Hispanic) with four scales under each domain (language use, linguistic proficiency, electronic media, and total). Total domain scores were dichotomized to indicate a low or high level of adherence to the specific domain. Biculturalism was defined as having a high level of adherence in both the Hispanic and non-Hispanic domains. Cronbach alpha for the domains and subscales ranged between 0.91 and 0.98, and McDonald's omega ranged between 0.91 and 0.98. Satisfaction with social support was measured using the Stokes Social Network Scale.⁵⁰ No reliability estimates are available for this measure.

4.6. Research processes (including patient and stakeholder engagement, if applicable)

4.6.1. Collection of study assessments

There were two assessments, baseline and follow-up. The baseline examination was scheduled after screening and conducted within one month before commencement of interventions. The follow-up examination was scheduled within one month after the completion of the six-month intervention. A dedicated coordinator who was independent of the study counselors and was blinded to the interventions collected the data. The coordinator collected the data using a computer-assisted personal interview system, which had prepopulated rules and ranges that minimized errors and missing data. The DCC conducted audits of a random sample of 10% of all assessments on a monthly basis in order to ensure the quality of data collection. Any issues were addressed with the coordinator in real time via email or telephone and were further discussed in monthly meetings that included the principal investigator, the coordinator, and the personnel of the DCC. These meetings were held the first Friday of every month.

4.6.2. Intervention procedures

The study had two arms, the NYUCI and REACH-OUT. Both have six visits of similar duration in six months, in addition to ad hoc contacts in those six months. An important premise of NHICE was that the two interventions had similar "doses" and only the mode of delivery was different. In addition to the interventions, all participants were referred to a dedicated case worker at Riverstone.

Summary of the NYUCI. The *first component* consisted of two individual and four family counseling sessions that included relatives or close friends of the caregiver. The content of these sessions was determined by the needs of the caregiver and participating family members (eg, learning techniques for management of troublesome patient behavior, and promoting communication among family members). These sessions lasted between one and 1.5 hours. The *second component* of the intervention was participation in a support group to provide the caregiver with continuous emotional support and education. The third component of the treatment was “ad hoc” counseling—the continuous availability of the interventionist to help deal with crises and with the changing nature and severity of symptoms of dementia. The emergence of new psychiatric and behavioral problems of patients, which are generally more stressful than the need for assistance with activities of daily living or physical limitations, could precipitate ad hoc calls. A bilingual (English, Spanish) social worker with experience in dementia and caregiver issues delivered the NYUCI.

Summary of REACH-OUT. All aspects of REACH-OUT involved problem-solving techniques and the development of written action plans. The goal of REACH-OUT was to engage the caregiver in joint problem solving, creating a written action plan targeting specific caregiving problems. Problem-solving strategies were used to generate relevant information about the “target problem” and the overall caregiving situation, with emphasis on the *context* in which the problem occurred. Problem solving was the guiding strategy for working with caregivers to develop and modify action plans over the intervention period. The basic steps of problem solving were (1) define the problem; (2) set goals; (3) brainstorm with caregiver and list possible solutions on paper; (4) select solutions; (5) develop an action plan based on these solutions; and (6) implement the action plan, track progress, and make adjustments as needed. The first step in REACH-OUT is a risk appraisal to determine how much emphasis to place on each of the intervention components. Thus, the intervention was standardized with respect to the components available but varied with respect to the dosing or depth of treatment delivered for each component. The tailoring of the intervention was guided by the findings of the risk appraisal.

Six specific skill sets were taught to caregivers: (1) education about dementia, caregiving, and stress; (2) helping the caregiver stay healthy; (3) keeping the home safe for the dementia patient; (4) teaching the caregiver how to maintain emotional wellbeing through relaxation; (5) behavior management; and (6) enhancing social support. The REACH-OUT intervention was delivered over six months. There were six in-home sessions, approximately one to 1.5 hours in length, similar to the NYUCI. Under exceptional conditions, up to two in-home sessions could be substituted by telephone sessions lasting from 30

minutes to 90 minutes. As part of the REACH-OUT intervention, caregivers received a notebook with educational information about dementia, self-care, safety, and other relevant caregiver issues. This notebook included copies of action plans. The Caregiver Notebook served as a tool for organizing intervention materials and a resource guide for the caregiver to use during and after the intervention. The interventionist reviewed specific issues related to dementia, caregiving, and stress using education material provided in the Caregiver Notebook and provided individualized action plans to address targeted problems such as care recipient behaviors, caregiver health behaviors, and social support. The interventionist taught caregivers how to enhance their emotional well-being, by teaching relaxation techniques and, if indicated, Pleasant Events Training. The REACH-OUT counselor had similar qualifications as the NYUCI counselor.

Section 4.4. provides a comparison of the REACH-OUT and NYUCI interventions. While both interventions are delivered by counselors with similar qualifications, have the same number of scheduled visits, and involve education about dementia and dementia caregiving, the most salient difference is that REACH-OUT emphasizes systematic skill-building while the NYUCI emphasizes improvement of social support through counseling and participation of relatives and friends in the intervention.

Justification for duration, treatment dose, and adherence. We tested both the NYUCI²⁶ and REACH²³ as six-month interventions in efficacy studies. In the case of the NYUCI, the efficacy studies included an ad hoc intervention period after the original 6 months. The adaptations of the NYUCI and REACH implemented in NHICE lasted six months. Both the NYUCI and REACH-OUT lasted six months and have an equal number of study visits. Thus, it seemed reasonable to postulate that the dose of both interventions was comparable. However, we measured intervention fidelity and ad hoc contact with the therapists for each intervention during the 6-month intervention period in order to ensure that both interventions were comparable.

Social work services. The interventionist in each arm offered all participants referral for social work services with a case worker at Riverstone. If participants accepted, a case worker at Riverstone Senior Life Services conducted an intake evaluation and customized social work services for each individual. The case worker kept a tracking file of all services provided using a datasheet.

4.7. Analytical and statistical approaches

We performed randomization to either of the 2 arms using a 1:1 allocation at the level of the individual. No clustering or other design effects required modeling. No data transformations were required. A biostatistician at the DCC performed randomization on a weekly basis. We performed a recruitment algorithm using a Statistical Analysis Software brand (SAS[®]) macro after subjects completed the baseline interview. We used a random number from 0 to 1 to determine intervention assignment. We set the standard cut score at 0.5 for the first n subjects, eg, five. We assigned those who received a random number from 0 to 0.5 to REACH-OUT and those with a random number greater than 0.5 to the NYUCI. Before randomization, we estimated the number of subjects randomized to each arm. If more than the n subjects were randomized initially, the cut score for the next subject was equal to the ratio of the experimental group ($n1$) to the subjects already randomized (m) for that group ($n1/m$). For example, because group A provided eight subjects for randomization and the n is set to 5, the first five subjects (n) were randomized to the standard cut score 0.5 (about half to the NYUCI and half to REACH-OUT). The sixth subject's randomization cut score equaled the number of subjects in the experimental group ($n1$) divided by the total number of subjects randomized within that group (in this case, the denominator m is 5). The seventh subject's cut score was adjusted according to the previous six subjects, and so on. The data coordinating center notified the pertinent intervention counselor of the intervention assignment after the participant had been consented and the baseline assessment had been completed. A coordinator independent of the counselor and blinded to the intervention assignment conducted the baseline assessment.

We conducted the primary, exploratory, and subgroup analyses using repeated measures mixed models to examine differences between the groups. We used an unstructured covariance structure in all analyses. Group and residual variances did not need to be modeled. For continuous outcomes, we used PROC Mixed from SAS[®].

Analysis for primary aim 1: Our primary approach to analyses was intent-to-treat (ITT). The primary analyses used mixed random effects models, and a FIML approach, with sensitivity analyses using generalized estimating equations (GEE). We modeled the change from pretreatment to posttreatment values of continuous outcomes as functions of baseline values, intervention, and the interaction of baseline and intervention. We used a general longitudinal mixed effects model, using SAS PROC MIXED, to allow for the possible group heterogeneity in residual variances and serial correlations that required

modeling to satisfy model assumptions and improve model fit. There could have been violations of the more rigid assumptions involved in analysis of covariance (ANCOVA), such as homoscedasticity, so that modeling the group heterogeneity in residual variances would be necessary. Based on prior analytic experience with the outcome variables, we did not expect that transformations would be necessary. Prior to analyses, we examined baseline values of all variables from each arm. We planned examination of baseline differences on key variables between completers and those lost to follow-up to inform the nature of the missing data. The ITT analyses performed using SAS PROC MIXED permit all individuals with at least one observation to be included.

Depending on the observed correlation between the dependent variables, we performed multivariate analysis of variance (MANOVA) in sensitivity analyses, and multivariate ANCOVA (MANCOVA) in subgroup analyses. A significant interaction term for one of the groups indicated that the effect of one of the treatments is different for ZCBS and GDS; in that case we estimated two treatment effects for each outcome. If the interaction was not significant we fit a model with only main effects for the outcomes and treatment and estimated the treatment effect common for the ZCBS and GDS. In addition to significance testing, we estimated the treatment effects with 95% CI. Bartlett's test of sphericity informed about the degree of intercorrelation among the outcome measures in order to determine suitability of the basic MANOVA model. We selected the final covariate using collinearity diagnostics and examination of correlations.

Heterogeneity of Treatment Effect (HTE). As specified in the methodology report to PCORI, because there is an interest in examining subgroups to help determine for which individuals interventions may be effective, we performed descriptive HTE by adding the potential effect modifiers into the MANCOVA model as interaction terms.

We used time of assessment (baseline versus follow-up) as the indicator for study wave interval. However, we used actual time (in months) in sensitivity analyses. We performed sensitivity analyses using an ANCOVA approach, in which the follow-up value was the outcome, adjusted by the baseline and group status variables.

We conducted dose-response analyses using repeated measures mixed models using PROC Mixed to examine the effect of the number of visits on the outcomes. We used an unstructured covariance structure in all analyses.

4.7.1 Original power calculations

We performed power calculations on an intent-to-treat basis. We specified two primary outcomes and set the p-value at 0.05 for a two-tailed test. We based the sample size calculations on the number of subjects needed to provide adequate power to test the primary hypothesis related to group differences in caregiver depressive symptoms and burden at six-month follow-up. The primary power calculations assumed separate analyses of burden and depression measures and scenarios in which loss to follow-up as large as 20% occurred.

Effect sizes (Cohen's *d*) for depression and burden: Studies of caregivers have used different depression measures; for example, the original REACH study used the Center for Epidemiologic Studies Depression Scale, whereas the original NYUCI study used the GDS. Both studies used the ZCBS. For the ZCBS, we used the estimates from REACH because they included a sample of Hispanic caregivers; for the depression measure we used estimates from the GDS provided from a pilot study.³⁵ Based on these studies, we used the following data for estimation: The baseline ZCBS standard deviations in the pilot study were 9.46 and 10.86 in the treatment and control groups, respectively. In the original Hispanic REACH sample, the estimates of the parameters were as follows: $\mu T1$ (treatment group) = 16.9 (9.6); $\mu T2$ (6-month follow-up) = 14.9 (9.1); $\mu C1$ (control group) = 17.4 (9.9); $\mu C2$ = 15.9 (9.9). In the pilot study, the estimated standard deviation for the GDS was 7.47.

We could not find a robust reference for what constitutes clinically significant changes in the GDS and ZCBS. Thus, we chose to express effect sizes in Cohen's *d*, which is a statistical way to express effect sizes. A Cohen's *d* of 0.2 is considered a small effect size, 0.5 a moderate effect size, and 0.8 a large effect size. It is usually accepted that moderate and large effect sizes are more likely to be clinically significant—that is, to be important for study participants—than small effect sizes, but this is relative.

Assuming power of 0.80, with 100 per group (200 total) we originally projected that we would be able to detect a moderate effect size (Cohen's *d* = 0.4)—equivalent to about 4.21 ZCBS (score range 0-88) and 3.21 GDS (score range 0-30) points, and equivalent to a 3- to 4-point endpoint mean difference between groups on the GDS and ZCBS. We examined sample size requirements for the detection of other endpoint differences: 4.0, 4.5, and 5 points on the ZCBS and 3.0, 3.5, and 4 points on the GDS. Sample sizes to detect this range of effects are: 111, 88, and 71 on the ZCBS and 115, 84, and 64 for the GDS, respectively.

In summary, Cohen's d ranged from 0.33 to 0.50 or between 3 and 5 points on the ZCBS and 2.5 and 4 points on the GDS endpoint means, roughly equivalent to a 0.3 to 0.5 SD endpoint difference in means, a moderate effect size for the primary analyses, assuming 100 per group.

In the current study, the observed standard deviations for the ZCBS at baseline and follow-up were about 18 for REACH-OUT and 17 and 16 for the NYUCI, larger than those observed in the original studies.

4.7.2. Sensitivity power calculations

We performed sensitivity power calculations for examining a MANCOVA model with covariates. Conservatively, under the assumptions specified above regarding attrition and inclusion of potential effect modifiers, 80 subjects per group would provide power ≥ 0.80 to detect a 4- to 5-unit differential change in depression and burden. This was based on testing the Time X Group interaction, entering subgroups (eg, spouse, child), and the interaction of group by these terms in a MANCOVA, allowing for heterogeneous variances and modeling correlations between measures and over time. Even if the pooled variance was higher than assumed, medium effect sizes would still be detectable. Thus, we estimated that 80 to 100 subjects per arm would provide sufficient power to detect the hypothesized difference between the two study arms.

5. Results

5.1. Recruitment and Retention

Recruitment began on March 1, 2014 and was completed on February 29, 2016. The total duration of the recruitment period was 24 months, as originally planned. Our original recruitment target was 200 and our target date to end recruitment was February 29, 2016. We achieved our recruitment target in December 2015. However, with encouragement from the PCORI project officer and the study data safety and monitoring board, we continued recruitment until the target date, and reached a sample size of 221. We completed follow-up data collection on September 30, 2016. Nearly 60% of all potential participants were enrolled. The follow-up assessment was completed by 95.5% of participants (210 out of 221). This very high completion proportion minimized the problem of missing data. See the Consort Statement in Figure 1 for more details on recruitment and retention.

5.2. Sample and group characteristics

As shown in Table 1, about one-third of the caregivers described their Hispanic subgroup as Dominican, 13% as Puerto Rican, and about one-fourth as “other Hispanic” group. About half of the caregivers were married. The average age of the caregiver was 58.24 years (SD = 11.24). The average education of the caregivers was 12.82 (SD = 3.61) years. About 40% were employed full or part time. Most (63%) caregivers were adult children, and about one-third were spouses. Most caregivers (73%) lived with the care recipient with dementia. The average age of the care receiver with dementia was 79.60 years (SD = 9.18), most care receivers (62%) were women, and 76% of care receivers had a diagnosis of Alzheimer’s disease. Only 10 out of 221 persons (4.5%) failed to complete the follow-up assessments. Caregivers who did not complete follow-up had, on average, lower years of education completed (10.00 versus 12.96 years; $p = 0.013$) and higher satisfaction with social support network (5.70 versus 5.07; $p = 0.042$) as compared with participants who completed the follow-up assessment. No statistically significant differences in other variables existed.

Figure 1. Consort Statement Figure
NHICE Consort Statement

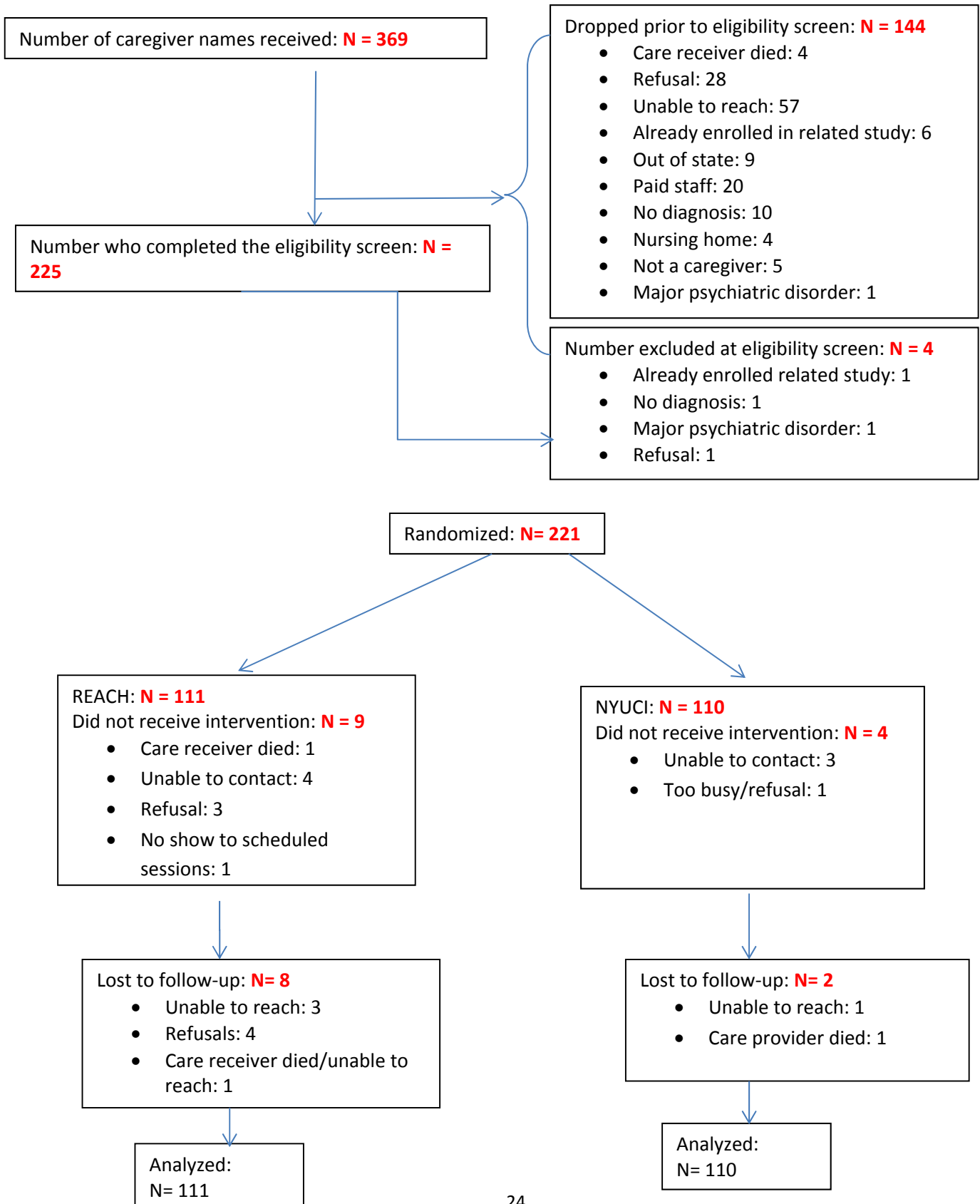


Table 1. Baseline characteristics for the care recipient and caregiver for the total study sample, REACH-OUT arm, and the NYUCI arm. Results for continuous outcomes are shown as means and (standard deviations). Results for categorical outcomes are shown as numbers and percentages.

		Total (n = 221)		REACH-OUT (N = 111)		NYUCI (N = 110)	
Care Recipient							
Age (mean, SD)		79.60	(9.18)	79.44	(8.25)	79.76	(10.07)
Female (N, %)		137	62.0	67	60.4	70	63.6
Hispanic (N, %)		219	99.1	111	100.0	108	98.2
Type of dementia (N, %)	Alzheimer's disease	167	75.6	86	77.5	81	73.6
	Vascular dementia	16	7.2	8	7.2	8	7.3
	Mixed dementia	3	1.4	1	.9	2	1.8
	Other	28	12.7	11	9.9	17	15.5
	Don't know	1	.5	1	.9	0	.0
	Data missing	6	2.7	4	3.6	2	1.8
Caregiver							
Relationship to care receiver (N, %)	Spouse	70	31.7	36	32.4	34	30.9
	Child	140	63.3	66	59.5	74	67.3
	Sibling	1	.5	1	.9	0	.0
	Niece/nephew	3	1.4	3	2.7	0	.0
	Granddaughter	4	1.8	3	2.7	1	.9
	Daughter-in-law	2	.9	1	.9	1	.9
	Cousin	1	.5	1	.9	0	.0
Caregiver lives in same household as care receiver (N, %)		161	72.9	83	74.8	78	70.9
Age (mean, SD)		58.24	(11.24)	58.33	(11.26)	58.15	(11.28)
Female (N, %)		183	82.8	91	82.0	92	83.6
Language prefer to be interviewed in (N, %)	English	81	36.7	37	33.3	44	40.0
	Spanish	140	63.3	74	66.7	66	60.0
Born in the United States (N, %)		54	24.4	20	18.0	34	30.9

		Total (n = 221)		REACH-OUT (N = 111)		NYUCI (N = 110)	
Race (N, %)	White	41	18.6	18	16.2	23	20.9
	Black	6	2.7	2	1.8	4	3.6
	Other	174	78.7	91	81.9	83	84.5
Hispanic subgroups (per Census categories) (N, %)	Mexican	3	1.4	3	2.7	0	.0
	Puerto Rican	28	12.7	17	15.3	11	10.0
	Cuban	4	1.8	3	2.7	1	.9
	Dominican	72	32.6	33	29.7	39	35.5
	Costa Rican	1	.5	0	.0	1	.9
	Guatemalan	5	2.3	5	4.5	0	.0
	Honduran	4	1.8	3	2.7	1	.9
	Panamanian	4	1.8	2	1.8	2	1.8
	Other Central American	2	.9	2	1.8	0	.0
	Chilean	1	.5	1	.9	0	.0
	Colombian	14	6.3	9	8.1	5	4.5
	Ecuadorian	21	9.5	8	7.2	13	11.8
	Peruvian	2	.9	1	.9	1	.9
	Other	60	27.1	24	21.6	36	32.7
	Marital status (N, %)	Single/never married	52	23.5	24	21.6	28
Living with significant other		3	1.4	2	1.8	1	.9
Married		110	49.8	55	49.5	55	50.0
Separated		14	6.3	10	9.0	4	3.6
Divorced		34	15.4	14	12.6	20	18.2
Widowed		4	1.8	3	2.7	1	.9
Data missing		4	1.8	3	2.7	1	.9
Religious preference (N, %)	Catholicism	146	66.1	78	70.3	68	61.8
	Protestantism	5	2.3	3	2.7	2	1.8
	Baptist	2	.9	0	.0	2	1.8

		Total (n = 221)		REACH-OUT (N = 111)		NYUCI (N = 110)	
	Other Christian	33	14.9	13	11.7	20	18.2
	Buddhism	3	1.4	2	1.8	1	.9
	Atheism/Agnosticism	1	.5	0	.0	1	.9
	Other	10	4.5	6	5.4	4	3.6
	None	11	5.0	2	1.8	9	8.2
	Refusal	1	.5	1	.9	0	.0
	Data missing	9	4.1	6	5.4	3	2.7
Education (mean, SD)		12.82	(3.61)	12.90	(3.48)	12.73	(3.74)
Highest education degree (N, %)	Technical degree/diploma	1	.5	0	.0	1	.9
	High school diploma/GED*	72	32.6	35	31.5	37	33.6
	Associate's degree	37	16.7	19	17.1	18	16.4
	College diploma	43	19.5	23	20.7	20	18.2
	MA/MS/other master's	10	4.5	4	3.6	6	5.5
	Doctorate (PhD/MD/JD/Other)	2	.9	1	.9	1	.9
	Other non-US degree	29	13.1	14	12.6	15	13.6
	None	18	8.1	9	8.1	9	8.2
Data missing	9	4.1	6	5.4	3	2.7	
Employment status (N, %)	Full-time	64	29.0	31	27.9	33	30.0
	Part-time	22	10.0	10	9.0	12	10.9
	Homemaker	17	7.7	8	7.2	9	8.2
	Retiree	61	27.6	31	27.9	30	27.3
	Not working < 6 months	9	4.1	5	4.5	4	3.6
	Not working > 6 months	23	10.4	11	9.9	12	10.9

		Total (n = 221)		REACH-OUT (N = 111)		NYUCI (N = 110)	
	Student	2	.9	1	.9	1	.9
	Disabled/unable to work	19	8.6	11	9.9	8	7.3
	Data missing	4	1.8	3	2.7	1	.9
Average total household monthly income (mean, SD)		2677.72	(3644.72)	2575.78	(2068.50)	2771.34	(4,654.57)
People supported by this income (mean, SD)		2.62	(1.26)	2.67	(1.16)	2.58	(1.35)
Private insurance	Yes	95	43.0	44	39.6	51	46.4
Medicare	Yes	63	28.5	31	27.9	32	29.1
Medicaid	Yes	112	50.7	57	51.4	55	50.0
Taking psychiatric medications (N, %)		20	9.0	13	11.7	7	6.4
Number of children (mean, SD)		2.14	(1.47)	2.32	(1.47)	1.97	(1.46)
Number of sisters (mean, SD)		2.25	(2.20)	2.28	(2.08)	2.22	(2.32)
Number of brothers (mean, SD)		1.93	(1.96)	2.09	(2.15)	1.78	(1.74)
Number living in household (mean, SD)		2.88	(1.47)	3.13	(1.58)	2.62	(1.32)
Number < 18 years old (mean, SD)		.33	(.76)	.40	(.87)	.26	(.64)
Persons between 18 and 64 years old (mean, SD)		1.56	(1.18)	1.72	(1.18)	1.40	(1.16)
Persons 65 years or older (mean, SD)		1.03	(.77)	1.05	(.77)	1.00	(.77)
Satisfaction with social support network (mean, SD)		5.10	(1.06)	5.15	(.98)	5.05	(1.14)
Number of people in social support network (mean, SD)		3.61	(2.59)	3.54	(2.60)	3.68	(2.59)
Currently smoke cigarettes (N, %)	Never smoked	203	91.9	103	92.8	100	90.9
	No	5	2.3	2	1.8	3	2.7
	Yes	13	5.9	6	5.4	7	6.4
Alcoholic beverages (N, %)	Never drank	106	48.0	54	48.6	52	47.3
	Previously drank	17	7.7	6	5.4	11	10.0
	Currently drinks	98	44.3	51	45.9	47	42.7

*GED: General educational development test

5.3. Correlations of primary and secondary outcomes

We examined whether the primary and secondary outcomes within a construct were correlated. The bivariate correlations of the burden outcomes at baseline were ZCBS and Lawton, 0.905; ZCBS and Montgomery, 0.582; and Lawton and Montgomery, 0.673. At follow-up, the bivariate correlations were ZCBS and Lawton, 0.894; ZCBS and Montgomery, 0.570; and Lawton and Montgomery, 0.659.

Bivariate correlations of the depression outcomes at baseline were GDS and PROMIS depression, 0.772; GDS and PHQ-9, 0.680; and PROMIS depression and PHQ-9, 0.707. At follow-up, the correlations were GDS and PROMIS depression, 0.816; GDS and PHQ-9, 0.774; and PROMIS depression and PHQ-9, 0.774.

5.4. Results of analyses of primary and secondary outcomes

Table 2 shows the observed means and standard deviations or proportions for the primary and secondary outcomes at baseline and follow-up. Both primary analyses evidenced no appreciable differences between the groups at baseline and six-month follow-up for any of the primary, secondary, and exploratory outcomes (see Tables 3 through 5). Although no differences occurred between the study arms for the ZCBS, an appreciable reduction (meaning improvement) occurred in the score from baseline to six months that was statistically significant (see Table 3 and Figure 2). Intervention and subgroup mean differences from baseline to follow-up are found in Tables 7 through 10.

For the ZCBS, the REACH-OUT group started at 40.32 (SD = 18.44) at baseline and decreased to 35.46 (SD = 18.04). The NYUCI group started at 40.95 (SD = 16.67) at baseline and decreased to 36.24 (SD = 15.88). The result was a net reduction of about five points in each group ($p = 0.006$). Using the model-based estimates, the effect size for REACH-OUT was a 5.18-point decrease ($p = 0.0006$) and for the NYUCI a 4.58-point decrease ($p = 0.0022$) (see Table 6).

For the GDS, the REACH-OUT group started at 9.83 (SD = 7.33) at baseline and decreased to 9.80 (SD = 7.25). The NYUCI group started at 9.73 (SD = 6.55) at baseline and decreased to 9.57 (SD = 7.12). The result was a very small nonsignificant net reduction in each group; REACH-OUT: 0.03-point reduction ($p = 0.9529$); and NYUCI: 0.19-point reduction ($p = 0.6991$) (see Figure 3).

Table 2. Values of outcome variable at baseline and six-month follow-up. Means (and standard deviations) are presented for the two treatment arms, REACH-OUT, and the NYUCI.

		REACH-OUT			NYUCI		
		N	Mean	SD	N	Mean	SD
Zarit Caregiver Burden Scale	Baseline	111	40.32	(18.44)	110	40.95	(16.67)
	Six-month follow-up	100	35.46	(18.04)	101	36.24	(15.88)
Lawton Caregiver Appraisal Scale: Burden	Baseline	111	25.36	(9.61)	110	25.75	(8.74)
	Six -month follow-up	100	23.73	(9.26)	101	23.42	(8.72)
Montgomery Caregiving Objective Burden Scale	Baseline	109	17.34	(7.07)	108	18.36	(7.56)
	Six -month follow-up	100	17.26	(6.84)	101	18.00	(7.10)
Geriatric Depression Scale	Baseline	111	9.83	(7.33)	110	9.73	(6.55)
	Six -month follow-up	103	9.80	(7.25)	108	9.57	(7.12)
PROMIS Depression score	Baseline	111	49.59	(9.44)	110	49.57	(9.57)
	Six -month follow-up	103	48.51	(10.28)	108	48.01	(10.32)
PHQ-9 Severity score	Baseline	95	5.03	(4.29)	93	5.82	(4.61)
	Six -month follow-up	103	4.73	(4.49)	107	5.40	(5.40)
Perceived Stress Scale	Baseline	111	16.32	(8.07)	110	17.17	(7.91)
	Six -month follow-up	103	14.95	(8.22)	108	14.98	(8.10)
PROMIS Physical Function score	Baseline	111	28.92	(7.51)	110	30.08	(8.10)
	Six -month follow-up	103	28.68	(7.38)	108	30.10	(8.36)

Table 3. Results of Repeated Measures Mixed Models for the Primary and Secondary Measures of Caregiver Burden, Specifying an Unstructured Covariance Structure

	Primary Outcome			Secondary Outcomes					
	Zarit Caregiver Burden Scale (n = 221)			Lawton Caregiver Appraisal Scale: Burden (n = 221)			Montgomery Objective Burden Scale (n = 221)		
	Estimate	SE	P Value	Estimate	SE	P Value	Estimate	SE	P Value
Intercept	40.3153	1.6692	< 0.0001	25.3604	0.8721	< 0.0001	17.3196	0.6993	< 0.0001
Group	0.6392	2.3659	0.7873	0.3942	1.2362	0.7501	1.1331	0.9912	0.2542
Administration	-5.1788	1.4819	0.0006	-1.7261	0.6977	0.0141	-0.0418	0.6847	0.9514
Group by administration	0.6021	2.0925	0.7738	-0.4822	0.9850	0.6249	-0.1940	0.9671	0.8412

Table 4. Results of Repeated Measures Mixed Models for the Primary and Secondary Measures of Depressive Symptoms, Specifying an Unstructured Covariance Structure

	Primary Outcome			Secondary Outcomes					
	Geriatric Depression Scale (n = 221)			PROMIS Depression (n = 221)			PHQ-9 (n = 220)		
	Estimate	SE	P Value	Estimate	SE	P Value	Estimate	SE	P Value
Intercept	9.8288	0.6598	< 0.0001	49.5910	0.9023	< 0.0001	5.1755	0.4496	< 0.0001
Group	-0.1016	0.9352	0.9136	-0.0219	1.2790	0.9864	0.5974	0.6383	0.3503
Administration	-0.0300	0.5068	0.9529	-1.2051	0.8809	0.1727	-0.5207	0.4615	0.2604
Group by administration	-0.1621	0.7093	0.8194	-0.3739	1.2336	0.7621	0.1710	0.6496	0.7927

Table 5. Results of Repeated Measures Mixed Models for the Exploratory Outcomes of Stress and Physical Function, Specifying an Unstructured Covariance Structure

	Perceived Stress Scale (n = 220)			PROMIS Physical Function (n = 221)		
	Estimate	SE	P Value	Estimate	SE	P Value
Intercept	16.3153	0.7588	< 0.0001	28.9153	0.7414	< 0.0001
Group	0.8584	1.0756	0.4257	1.1601	1.0509	0.2708
Administration	-1.5772	0.7203	0.0296	-0.1541	0.5305	0.7717
Group by administration	-0.6570	1.0093	0.5158	0.2108	0.7424	0.7767

Figure 2. Plot of Change in Means for the Zarit Caregiver Burden Scale From Baseline to Follow-up Assessment

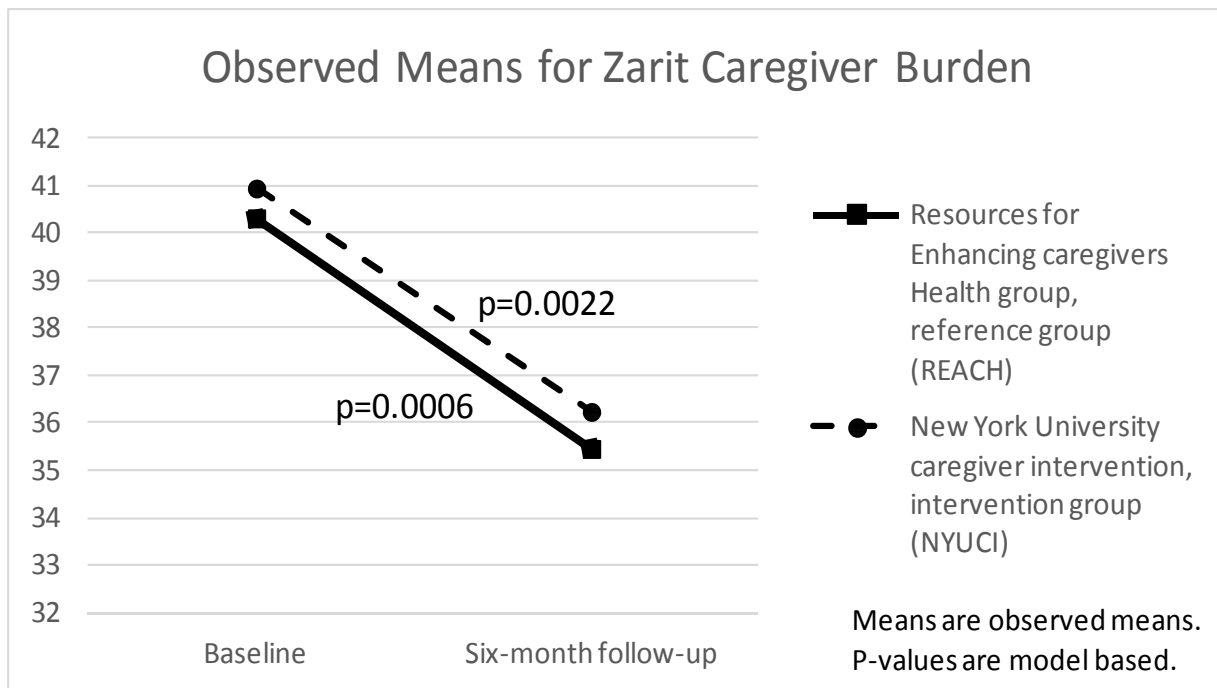


Table 6. Treatment Group Differences from Baseline to Follow-up, Based on Repeated Measures Mixed Models, Specifying an Unstructured Covariance Structure

	Resources for Enhancing Caregivers Health				New York University Caregiver Intervention			
	N	Difference	SE	P Value	N	Difference	SE	P Value
Zarit Caregiver Burden Scale	111	5.1788	1.4819	0.0006	110	4.5767	1.4774	0.0022
Lawton Caregiver Appraisal Scale	111	1.7261	0.6977	0.0141	110	2.2083	0.6952	0.0017
Montgomery Caregiving Objective Burden Scale	111	0.0418	0.6847	0.9514	110	0.2358	0.6829	0.7302
Geriatric Depression Scale	111	0.0300	0.5068	0.9529	110	0.1921	0.4962	0.6991
PROMIS Depression score	111	1.2051	0.8809	0.1727	110	1.5790	0.8636	0.0688
PHQ-9 score	110	0.5207	0.4615	0.2604	110	0.3498	0.4572	0.4451
Perceived Stress Scale	111	1.5772	0.7203	0.0296	110	2.2342	0.7070	0.0018
PROMIS Physical Function	111	0.1541	0.5305	0.7717	110	-0.0567	0.5194	0.9132

Table 7. Analyses stratified by age group showing differences from baseline to follow-up within intervention group, based on repeated measures mixed models, specifying an unstructured covariance structure. Results for primary, secondary, and exploratory outcomes are shown.

	Subgroup (years)	Resources for Enhancing Caregivers				New York University Caregiver			
		Health				Intervention			
		N	Difference	SE	P Value	N	Difference	SE	P Value
Zarit Caregiver Burden Scale	58+	60	6.5869	2.0260	0.0013	52	7.1978	2.1750	0.0011
	≤ 57	51	3.5815	2.1602	0.0987	58	2.2971	2.0025	0.2526
Lawton Caregiver Appraisal Scale: Burden	58+	60	2.7793	0.9557	0.0040	52	2.6503	1.0260	0.0104
	≤ 57	51	0.5225	1.0181	0.6083	58	1.8084	0.9432	0.0565
Montgomery Caregiving Objective Burden Scale	58+	60	-0.3279	0.9384	0.7271	52	-0.5802	1.0086	0.5657
	≤ 57	51	0.4696	1.0030	0.6401	58	0.9136	0.9296	0.3268
Geriatric Depression Scale	58+	60	-0.0710	0.6963	0.9189	52	0.0739	0.7319	0.9196
	≤ 57	51	0.1459	0.7462	0.8451	58	0.2931	0.6808	0.6672
PROMIS Depression score	58+	60	1.1596	1.2103	0.3391	52	1.4841	1.2735	0.2451
	≤ 57	51	1.2530	1.2978	0.3354	58	1.6534	1.1856	0.1645
PHQ-9 score	58+	60	0.1514	0.6287	0.8099	52	0.1667	0.6655	0.8025
	≤ 57	51	0.9669	0.6844	0.1592	58	0.4956	0.6335	0.4349
Perceived Stress Scale	58+	60	1.7157	0.9887	0.0841	52	1.8351	1.0415	0.0795
	≤ 57	51	1.4134	1.0609	0.1841	58	2.5728	0.9705	0.0086
PROMIS Physical Function score	58+	60	0.6531	0.7249	0.3686	52	-0.7759	0.7620	0.3097
	≤ 57	51	-0.4217	0.7769	0.5878	58	0.5466	0.7090	0.4416

Table 8. Analyses stratified by family position of the caregiver (spouse, child) in relation to the care recipient. Differences from baseline to follow-up are shown within intervention group, based on repeated measures mixed models, specifying an unstructured covariance structure. Results are shown for primary, secondary, and exploratory outcomes.

	Subgroup	Resources for Enhancing Caregivers Health				New York University Caregiver Intervention			
		N	Difference	SE	P Value	N	Difference	SE	P Value
Zarit Caregiver Burden Scale	Spouse	36	8.0043	2.7992	0.0047	34	9.4359	2.8102	0.0010
	Child	66	3.2270	1.8112	0.0765	74	2.1096	1.7005	0.2163
Lawton Caregiver Appraisal Scale: Burden	Spouse	36	3.2380	1.1964	0.0074	34	3.9207	1.2009	0.0013
	Child	66	0.7525	0.9001	0.4040	74	1.1428	0.8445	0.1774
Montgomery Caregiving Objective Burden Scale	Spouse	36	-0.9690	1.1954	0.4185	34	-1.1314	1.2052	0.3489
	Child	66	0.7875	0.8921	0.3784	74	0.5129	0.8372	0.5408
Geriatric Depression Scale	Spouse	36	0.3373	0.8728	0.6995	34	-0.2980	0.8871	0.7373
	Child	66	-0.7032	0.6559	0.2849	74	0.2658	0.5969	0.6565
PROMIS Depression score	Spouse	36	2.1732	1.5424	0.1603	34	0.4383	1.5685	0.7802
	Child	66	0.0931	1.1582	0.9360	74	1.9023	1.0557	0.0730
PHQ-9 score	Spouse	36	0.3400	0.7875	0.6663	34	0.2084	0.7930	0.7929
	Child	66	0.7025	0.6143	0.2541	74	0.1563	0.5704	0.7844
Perceived Stress Scale	Spouse	36	1.2535	0.7875	0.1290	34	1.2755	0.7930	0.0048
	Child	66	0.9405	0.6143	0.2082	74	0.8589	0.5704	0.1403
PROMIS Physical Function score	Spouse	36	0.9119	0.9464	0.3364	34	0.3534	0.9618	0.7137
	Child	66	0.0116	0.7112	0.9870	74	-0.2715	0.6472	0.6752

Table 9. Analyses stratified by acculturation status (bicultural, not bicultural) of the caregiver.

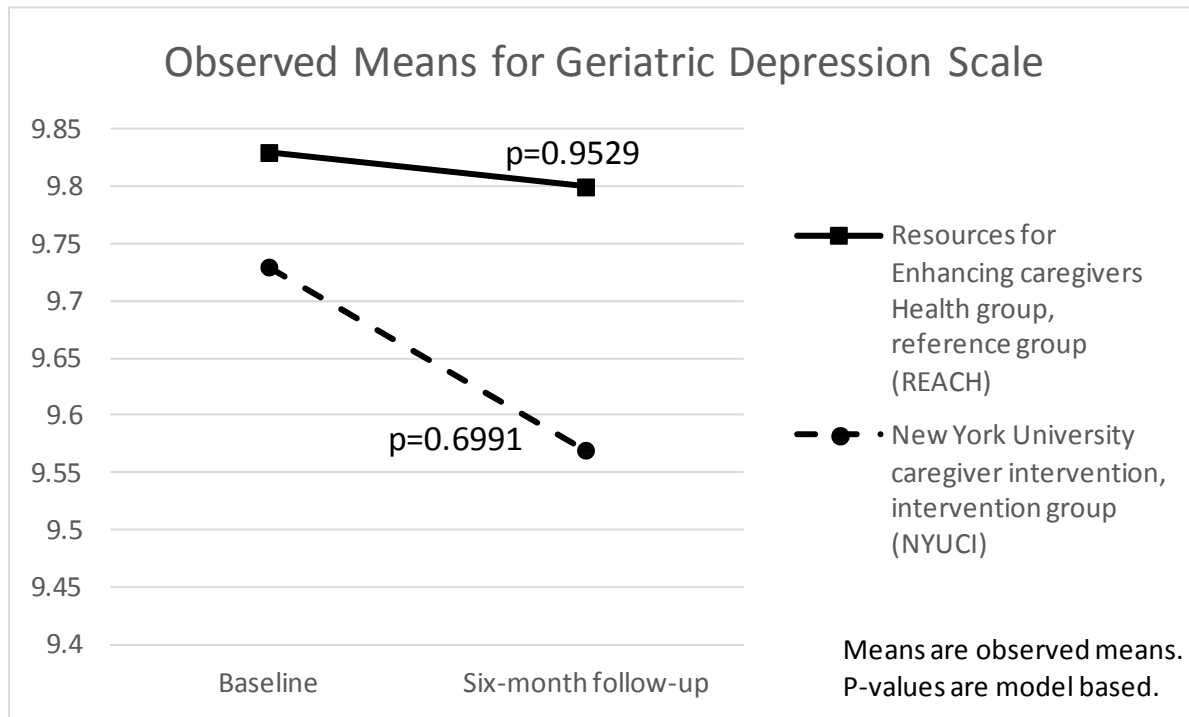
Differences from baseline to follow-up are shown within intervention group, based on repeated measures mixed models, specifying an unstructured covariance structure. Results are shown for primary, secondary, and exploratory outcomes.

	Subgroup	Resources for Enhancing Caregivers Health, Reference Group				New York University Caregiver Intervention, Intervention Group			
		N	Difference	SE	P Value	N	Difference	SE	P Value
		Zarit Caregiver Burden Scale	Bicultural	43	6.9011	2.3806	0.0042	42	5.7400
	Not bicultural	49	4.6463	2.2617	0.0414	47	5.4968	2.2713	0.0165
Lawton Caregiver Appraisal Scale: Burden	Bicultural	42	1.7648	1.1248	0.1184	43	3.6776	1.1516	0.0017
	Not bicultural	49	1.6267	1.0696	0.1301	47	2.0419	1.0730	0.0586
Montgomery Caregiving Objective Burden Scale	Bicultural	42	0.1082	1.1410	0.9246	43	1.3816	1.1660	0.2376
	Not bicultural	49	-0.2097	1.0824	0.8466	47	-0.5472	1.0889	0.6159
Geriatric Depression Scale	Bicultural	43	-0.1308	0.8466	0.8774	42	0.4524	0.8297	0.5863
	Not bicultural	49	0.6916	0.7976	0.3870	47	0.4143	0.7995	0.6050
PROMIS Depression score	Bicultural	43	1.4425	1.4167	0.3099	42	1.4095	1.3926	0.3128
	Not bicultural	49	1.9904	1.3340	0.1374	47	2.4841	1.3395	0.0653
PHQ-9 score	Bicultural	43	0.6790	0.7082	0.3389	42	0.2800	0.6964	0.6881
	Not bicultural	49	0.5939	0.6655	0.3733	47	0.4682	0.6738	0.4880
Perceived Stress Scale	Bicultural	43	1.6528	1.1638	0.1573	42	1.4788	1.1444	0.1979
	Not bicultural	49	3.0537	1.0959	0.0059	47	3.7572	1.1005	0.0008
PROMIS Physical Function score	Bicultural	60	-0.3186	0.8786	0.7173	52	0.5810	0.8603	0.5004
	Not bicultural	51	0.1345	0.8278	0.8711	58	-0.1512	0.8294	0.8555

Table 10. Analyses stratified by satisfaction with social support. Differences from baseline to follow-up are shown within intervention arm, based on repeated measures mixed models, specifying an unstructured covariance structure. Results are shown for the primary, secondary, and exploratory outcomes.

	Subgroup	Resources for Enhancing Caregivers Health				New York University Caregiver Intervention			
		N	Difference	SE	P Value	N	Difference	SE	P Value
Zarit Caregiver Burden Scale	Very satisfied	53	6.1589	2.3827	0.0104	59	5.4400	2.2624	0.0170
	Not very satisfied	50	3.2681	2.2334	0.1448	49	2.3794	2.2490	0.2912
Lawton Caregiver Appraisal Scale: Burden	Very satisfied	53	2.5672	1.1378	0.0250	59	3.1548	1.0745	0.0037
	Not very satisfied	50	0.4380	1.0607	0.6801	49	0.7407	1.0664	0.4880
Montgomery Caregiving Objective Burden Scale	Very satisfied	53	-0.5493	1.0774	0.6107	59	-0.3746	1.0294	0.7163
	Not very satisfied	50	0.3204	1.0213	0.7540	49	0.0956	1.0412	0.9269
Geriatric Depression Scale	Very satisfied	53	-0.5294	0.8227	0.5206	59	0.0310	0.7811	0.9684
	Not very satisfied	50	0.2381	0.7864	0.7623	49	-0.0819	0.7725	0.9157
PROMIS Depression score	Very satisfied	53	-1.4403	1.3484	0.2866	59	0.6205	1.2853	0.6297
	Not very satisfied	50	3.2148	1.3024	0.0143	49	1.2624	1.2861	0.3274
PHQ-9 Severity score	Very satisfied	53	-0.1496	0.7110	0.8336	59	0.3576	0.6849	0.6021
	Not very satisfied	50	0.8997	0.7006	0.2004	49	-0.2194	0.6951	0.7526
Perceived Stress Scale	Very satisfied	53	1.2473	1.1296	0.2707	59	2.5017	1.0839	0.0219
	Not very satisfied	50	1.2608	1.0842	0.2461	49	1.1755	1.0697	0.2730
PROMIS Physical Function score	Very satisfied	53	0.5090	0.8569	0.5531	59	1.3961	0.8125	0.0872
	Not very satisfied	50	-0.3237	0.8152	0.6917	49	-1.5335	0.7989	0.0562

Figure 3. Plot of Change in Means for the Geriatric Depression Scale From Baseline to Follow-up Assessment



5.5. Results for secondary measures of burden and depression

As for the primary burden measure, the ZCBS, no differences existed between the intervention arms in changes in the Lawton and Montgomery scales. Results for the Lawton burden measure (Table 6) showed a significant decline in burden within the two study arms (REACH-OUT, reduction 1.73, $p = 0.0141$; NYUCI, reduction 2.21 points, $p = 0.0017$), similar to the finding for the ZCBS. However, the Montgomery Caregiver Objective Burden Scale did not show such results (REACH-OUT reduction 0.04, $p = 0.95$; NYUCI reduction 0.24, $p = 0.73$). As for the primary depressive symptom scale, the GDS, no differences existed between the intervention arms in the PROMIS depression scale and the PHQ-9. No significant changes existed in the PROMIS depression scale and the PHQ-9 within each intervention arm.

5.6. Results for exploratory outcomes

No significant differences existed between the intervention arms in the exploratory outcomes, the Perceived Stress Scale and the PROMIS Physical Function scale. However, significant decreases occurred in the Perceived Stress Scale for both arms. The REACH-OUT group started at 16.32 (SD = 8.07) at baseline and decreased to 14.95 (SD = 8.22). The NYUCI group started at 17.17 (SD = 7.91) at baseline

and decreased to 14.98 (SD = 8.10). The result was a net decrease of approximately two points in each group ($p = 0.0296$) (see Table 6 and Figure 4).

No appreciable changes occurred in PROMIS Physical Function within each arm (see Table 5 and Figure 5). The REACH-OUT group score was 28.92 (SD = 7.51) at baseline and 28.68 (SD = 7.38) at follow-up. The NYUCI group score was 30.08 (SD = 7.10) at baseline and 30.10 (SD = 8.36) at follow-up.

5.7. Subgroup analyses

We performed subgroup analyses for age (Table 7), family position (Table 8), acculturation (Table 9), and satisfaction with social support (Table 10), for the primary, secondary, and exploratory outcomes. We were unable to evaluate within sex strata due to the low proportion of men (17%). We split the age subgroups at the median (≤ 57 years, 58 years and older). Given that 95% of all caregivers were either the spouse or child of the care recipient, these 2 groups were evaluated. We defined participants as “bicultural” or not using the algorithm associated with the Marín Bicultural Acculturation Scale for Hispanics.⁴⁹ Approximately 17% of the data for the Marín was missing, and we used multiple imputation to replace missing data and assign the acculturation category. Analyses compared those who were considered bicultural against those who were not bicultural. We defined satisfaction with social support groups based on an equal distribution of those who were “very satisfied” with their social support network (49%) and those in the remainder of the categories (“moderately satisfied,” “slightly satisfied,” “slightly dissatisfied,” “moderately dissatisfied,” and “very dissatisfied”), all of whom we defined as “not very satisfied.”

Figure 4. Plot of Change in Means for the Perceived Stress Scale From Baseline to Follow-up Assessment

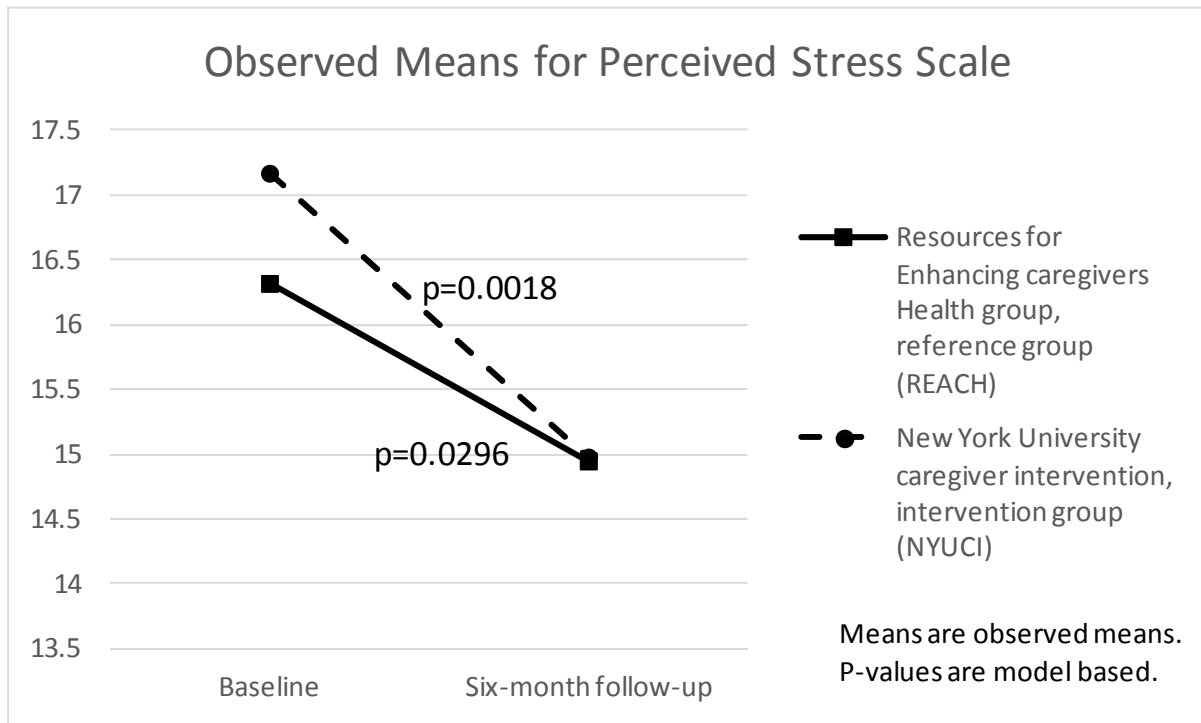
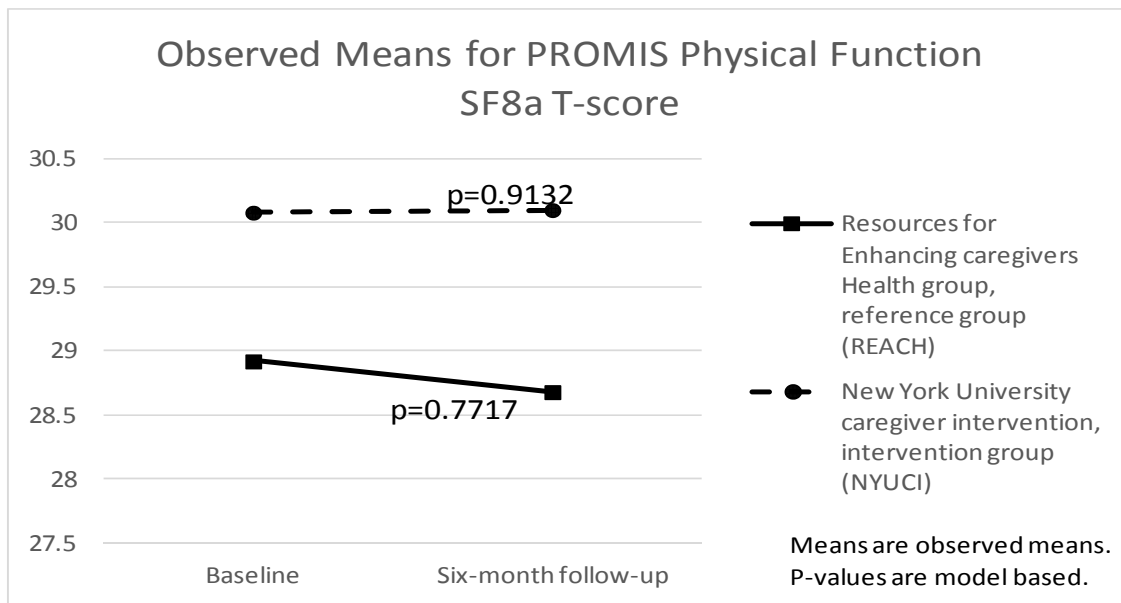


Figure 5. Plot of Change in Means for the PROMIS Physical Function Scale From Baseline to Follow-up Assessment



5.7.1. Subgroup analyses for primary outcomes

Examination of the effects of the ZCBS within age subgroups showed that the effects were statistically significant among caregivers 58 years of age and older (6.59 points, $p = 0.0013$ for REACH-OUT; and 7.20, $p = 0.0011$ for NYUCI) but were not statistically significant among younger caregivers (3.58 points, $p = 0.0987$ for REACH; and 2.30, $p = 0.2526$ for NYUCI). Examination of the effects of the ZCBS within spouse and child subgroups showed that the effects were statistically significant among spouse caregivers (8.00 points, $p = 0.0047$ for REACH; and 9.44, $p = 0.0010$ for NYUCI) but results were not statistically significant among children (3.23, $p = 0.0765$ for REACH; and 2.11, $p = 0.2163$ for NYUCI). Age and family position were highly interrelated, with spouses more likely to be older and adult children more likely to be younger. Most spouses (82.9%) were 58 years or older, and most adult children (63.6%) were 57 years old or younger. Examination of the effects of the ZCBS within acculturation status subgroups showed that the reduction in burden was statistically significant both among caregivers who were bicultural (6.90 points, $p = 0.0042$ for REACH-OUT; and 5.74, $p = 0.01795$ for NYUCI) and among those not bicultural (4.65 points, $p = 0.0414$ for REACH; and 5.50, $p = 0.0165$ for NYUCI). Examination of the effects of the ZCBS within subgroups of satisfaction with social support showed that the effects were statistically significant among caregivers who were very satisfied with their social support network (6.16 points, $p = 0.0104$ for REACH; and 5.44, $p = 0.0170$ for NYUCI) and were not statistically significant among those not very satisfied (3.27 points, $p = 0.1448$ for REACH; and 2.38, $p = 0.2912$ for NYUCI). Examination of the GDS within age, relationship, acculturation status, and satisfaction with social support subgroups showed minimal nonsignificant effects within each group.

5.7.2. Subgroup analyses for secondary outcomes

Examination of the effects of the Lawton Caregiver Appraisal Burden measure within age subgroups showed that the effects were statistically significant among caregivers 58 years of age and older (2.78 points, $p = 0.0040$ for REACH; and 2.65, $p = 0.0104$ for NYUCI), but were not statistically significant among younger caregivers (0.52 points, $p = 0.6083$ for REACH; and 1.81, $p = 0.0565$ for NYUCI). Examination of the effects of the Lawton Caregiver Appraisal Burden measure within spouse and child subgroups showed that the effects were statistically significant among spouse caregivers (3.24 points, $p = 0.0074$ for REACH; and 3.92, $p = 0.0013$ for NYUCI) but not statistically significant among children (0.75, $p = 0.4040$ for REACH; and 1.14, $p = 0.1774$ for NYUCI). Examination of the effects of the Lawton Caregiver Appraisal Burden measure within acculturation status subgroups showed that the effects were statistically significant for the NYUCI for caregivers who were bicultural (3.68, $p = 0.0017$), and close to

statistical significance for those who were not (2.04, $p = 0.0586$ for NYUCI), while the effects for REACH were not significant for both those who were bicultural (1.76 points, $p = 0.1184$) and those who were not (1.63 points, $p = 0.1301$). Examination of the effects of the Lawton Caregiver Appraisal Burden measure within satisfaction with social support subgroups showed that the effects were statistically significant for both interventions among caregivers who were very satisfied with their social support network (2.56 points, $p = 0.0250$ for REACH; and 3.15, $p = 0.0037$ for NYUCI) but effects were not statistically significant among those not very satisfied (0.44 points, $p = 0.6801$ for REACH; and 0.74, $p = 0.4880$ for NYUCI).

Examination of the Montgomery Caregiving Objective Burden, PROMIS Depression, and PHQ-9 measures within age, relationship, acculturation status, and satisfaction with social support subgroups showed nonsignificant effects within each group.

5.7.3. Subgroup analyses for exploratory outcomes

Examination of the Perceived Stress Scale within age subgroups showed that effects were not statistically significant for both age subgroups in the REACH intervention group (1.72 points, $p = 0.0841$ for 58 years and older; and 1.41, $p = 0.1841$ for ≤ 57 years). In the NYUCI intervention group, the ≤ 57 years subgroup had statistically significant effects (2.57 points, $p = 0.0086$) while the effects were not significant for the 58 years and older subgroup (1.84, $p = 0.0795$). Examination of the Perceived Stress Scale within spouses showed that the effects were statistically significant for the NYUCI (1.28, $p = 0.0048$) but not REACH (1.25 points, $p = 0.1290$). Among children, the effects were not significant for both REACH (0.94 points, $p = 0.2082$) and the NYUCI (0.86, $p = 0.1403$). Examination of the effects of the Perceived Stress Scale within acculturation status subgroups showed that the effects were statistically significant for both the NYUCI (3.76, $p = 0.0008$) and REACH (3.05 points, $p = 0.0059$) among caregivers who were not bicultural, while they were not significant among those who were bicultural (1.65 points, $p = 0.1573$ for REACH; and 1.48, $p = 0.1979$ for NYUCI). Examination of the Perceived Stress Scale within satisfaction with social support subgroups showed that the effects for the REACH were not statistically significant for both the very satisfied (1.24 points, $p = 0.2707$) and not the very satisfied (1.26, $p = 0.2461$) subgroups. For the NYUCI, the very satisfied subgroup had statistically significant (2.50 points, $p = 0.0219$) effects, while the effects were not statistically significant for the not very satisfied group (1.18, $p = 0.2730$). Examination of the PROMIS Physical Function measure within age subgroups showed nonsignificant effects within each group (Table 7). Examination of the effects of the PROMIS Physical

Function measure within age, relationship, acculturation status, and satisfaction with social support subgroups showed nonsignificant effects within each group.

5.8. Dose-response analyses

Most participants (79.3%) completed four, five, or six intervention counseling sessions (80.1% for REACH-OUT, 78.5% for NYUCI). The REACH-OUT group had an average of 5.02 visits (SD = 1.66), with a median of six visits. The NYUCI group had an average of 4.5 visits (SD = 1.46), with a median of five visits. Thus, little variation existed in the dose-response variable. A significant dose response was not observed. Examination of the dose response with the outcomes yielded no significant effects (P values ranged from 0.132 to 0.884).

In terms of ad hoc calls, the REACH-OUT group had significantly more participants with ad hoc calls ($n = 26, 23.4\%$) than the NYUCI group ($n = 9, 8.2\%$; $p = 0.002$). The difference in ad hoc calls between the groups had no effect on the primary outcomes analyses (ZCBS $p = 0.4968$; GDS $p = 0.4030$; P values are for differences in change in ZCBS and GDS between intervention arms).

We examined the differences in contact with the social worker intervention delivered at Riverstone Senior Life Services between the groups. The REACH-OUT group had significantly more participants with social worker contact outside of the interventions ($n = 48, 43.2\%$) than the NYUCI group ($n = 23, 20.9\%$; $p < 0.001$). However, the number of contacts did not influence the effect of the interventions on the primary outcomes (ZCBS $p = 0.4527$; GDS $p = 0.6706$).

5.9. Sensitivity analyses

The results of the sensitivity analyses were consistent with the primary analyses. Modeling time in months instead of administration yielded similar results. Additionally, the use of an ANCOVA approach yielded similar results. Thus, the results were robust to modeling variation.

5.10. Adverse events

No adverse events related to the study procedures occurred.

6. Discussion

To the best of our knowledge, NHICE was the first study to compare the effectiveness of the best-known caregiver interventions, REACH-OUT and the NYUCI, in any population. In NHICE we focused on

Hispanics because of the lack of data for effectiveness of these interventions in this particular ethnic group, which has a higher prevalence of dementia in New York City compared with whites, and which may experience caregiving in a culturally particular way. The main finding of our study is that both REACH-OUT and the NYUCI were equally effective in reducing the primary outcome of caregiver burden, and were equally ineffective in reducing the other primary outcome of depressive symptoms, at six months. This was contrary to our initial hypothesis that the NYUCI would be more effective than the REACH-OUT intervention.

Despite the differences in the mode of delivery of the NYUCI and REACH-OUT, we found that both were effective in decreasing the primary outcome of caregiver burden, one of the secondary outcomes of burden, and the exploratory outcome of stress. The interventions were not effective in decreasing depressive symptoms. The observations for the primary outcomes were mostly supported by the results for the secondary outcomes. However, although the results for the Lawton caregiver burden measure were similar to the primary burden outcome, the ZCBS, the results for the Montgomery burden scale showed no significant change in the two intervention arms. This observation might be explained by the fact that, in our sample, the ZCBS was highly correlated with the Lawton measure, but less strongly correlated with the Montgomery measure.

The REACH-OUT intervention had significantly more referrals to social work services and more ad hoc calls as compared with the NYUCI. There are multiple potential explanations for this finding. One is that the need for social work services was greater in participants in the REACH-OUT arm. However, examination of baseline variables between the two arms showed no significant differences. Thus, the possibility that participants in the REACH-OUT arm had more social needs and a higher need for ad hoc calls seems unlikely. A potential explanation for the increased number of ad hoc calls for the REACH-OUT is that participants in this arm required further contact beyond the scheduled visits in order to achieve the objectives of that particular intervention as compared with the NYUCI, but we can only speculate on this possibility. Another possibility is that the REACH-OUT interventionist tended to refer more participants to social work services and thus fostered more ad hoc calls. Our analyses suggest that if these tendencies existed, they did not bias the results of the study. Replications of this study should have an independent person conduct the social work referrals blinded to study arm in order to avoid the possibility that the interventionist from a particular arm tended to refer more participants to the case worker. It is unclear at this point if the higher number of ad hoc calls is inherent to the intervention (REACH-OUT versus NYUCI) or the particular interventionist.

It is important to consider whether the social work intervention could have biased the results toward the null (no differences between the arms, or no differences between baseline and follow-up) because it “eclipsed” the effects of the interventions. As discussed above, differential exposure to the social work intervention did not seem to affect the results, but we cannot rule out that the arm that was more exposed to the social work intervention (REACH-OUT) was more affected than the NYUCI by bias toward the null. As explained in section 3.3, the reason for including social work referral for all participants was the perception that social work services were necessary in participants of low socio-economic status for the interventions to be effective. This consideration may not be generalizable to all settings.

6.1. Subpopulation considerations

Our subgroup analyses for the primary outcomes did not favor either intervention. However, our subgroup analyses suggested that both interventions seemed to be particularly more effective in certain subgroups. The strongest subgroup difference was for spouse caregivers, who were more likely to be older, versus adult children caregivers, who were more likely to be younger. We found that spouse caregivers benefited appreciably more from both interventions in the reduction of caregiver burden compared with adult child caregivers. As would be expected given the age differences between spouse and adult child caregivers, we found that older caregivers were more likely to benefit from the interventions. We can only speculate about the potential reasons for this difference. It may be that spouse caregivers were more likely to be retired and to be not taking care of young children, and thus had less stressors than adult children. Adult children may not benefit from the counseling interventions because they are too overwhelmed by their multiple responsibilities, including their jobs and their own children, in addition to the patient with dementia. We found that persons with more satisfaction with social support networks had a greater decrease in burden compared with those with lower satisfaction. This may be due to the fact that both interventions serve to leverage and improve social support, although the NYUCI depends more on this aspect than REACH-OUT, and caregivers with better social support at baseline are better able to take advantage of this aspect of the interventions.

While subgroup analyses for the primary outcomes did not suggest that either intervention was better, some differences were observed for the secondary and exploratory outcomes. When examining the Lawton Caregiver Appraisal Burden measure by acculturation strata, the only statistically significant effect was for the NYUCI in bicultural caregivers. For the examination of the Perceived Stress Scale by age strata, the only significant effect was for the NYUCI in younger caregivers. In analyses for the Perceived Stress Scale by family position, only the effect for the NYUCI was significant among spouses.

For the Perceived Stress Scale by satisfaction with social support, only the NYUCI had significant effects in the very satisfied groups. While these results suggest that the NYUCI might be more beneficial than REACH-OUT in certain subgroups, they must be interpreted with caution given the multiple analyses for secondary and exploratory outcomes in multiple strata. The results of most subgroup analyses suggest that no differences existed between the interventions by subgroups. The few observed differences must be replicated in larger samples with a priori planned subgroup analyses, to examine whether these differences really exist.

6.2. Limitations

The study was underpowered for analyses of subgroups; sample sizes were sufficient to detect large effects, such as some of those observed in the analyses by family position and age group. Another limitation is the short duration of follow-up. Ideally, the outcomes should be assessed at 12 months and yearly thereafter to examine the “latent” effects of the intervention on the primary outcomes. The null findings for depressive symptoms at 6 months may be explained by this short follow-up, and benefits could appear after 6 months. Another important limitation is that anxiety was not measured. We learned from engaging the study participants that anxiety may be the most important mental health construct to measure and intervene upon.

6.3. Future research

We should conduct studies powered to explore to replicate the findings by subgroup (eg, spouse versus adult caregivers). We should conduct interventions of longer duration, particularly with ad hoc interventions beyond 6 months to 12 months and examine long-term outcomes. We should include anxiety as an outcome, since our study participants reported this outcome as important for them. Finally, our study should be replicated in other ethnic groups (blacks, whites, Asians) and settings (rural versus urban). Since NHiCE showed that adult children and younger subjects seemed to benefit less from the intervention, we should work with these subgroups to examine ways to tailor the interventions in order to improve their effectiveness.

6.4. Generalizability and implementation of study results

Our study demonstrated that both interventions were effective in decreasing caregiver burden and stress among Hispanic dementia caregivers. Thus, there does not seem to be a differential effectiveness of the NYUCI versus REACH-OUT, and the decision of caregivers and service providers to use one versus

the other should be based on personal preference and availability. For example, caregivers who prefer a more systematic approach to tailored skills building may prefer REACH-OUT, and caregivers who prefer involving family members may prefer the NYUCI. Our study focused on urban Hispanics in New York City, and the results of the study may be applicable to only this ethnic group and this particular setting. However, it is important to highlight that persons who were more acculturated to the United States seemed to get similar benefits on burden compared with those who were less acculturated. This could imply that the results of the study are generalizable to the US population at large, and not just Hispanics. Another important issue for implementation may be the need to have social support services for all participants, as was done in NHICE. It is unclear if the interventions as delivered in NHICE could be effective without the possibility of social work services for all participants. However, it is possible that the interventions could be more effective in the absence of social support services, and this is an issue that needs to be addressed by other implementation studies.

7. Conclusions

REACH-OUT and the NYUCI were effective in reducing burden and stress among caregivers of patients with dementia, but they were not effective in reducing depressive symptoms. Neither intervention was better than the other for the primary outcomes. The effectiveness of both interventions on burden seems to be greater among spouse caregivers compared with child caregivers, and among older compared with younger caregivers. We are confident that our results reflect what happens in a 6-month interval but would have liked to assess the effect of the interventions at 12 months and even later, which could have detected effects not seen at 6 months. We would also have liked to formally assess anxiety as an outcome, which we did not include. Finally, the results of our study suggest that spouse and adult child caregivers may be considered separately in terms of study design (eg, power analyses) and interventions.

8. References

1. Thies W, Bleiler L. 2011 Alzheimer's disease facts and figures. *Alzheimers Dement.* 2011; 7(2):208-244.
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(14):1326-1334.
3. Lyketsos CG, Steele C, Baker L, et al. Major and minor depression in Alzheimer's disease: prevalence and impact. *J Neuropsych.* 1997; 9(4):556-561.
4. Lyketsos CG, Sheppard JM, Steinberg M, et al. Neuropsychiatric disturbance in Alzheimer's disease clusters into three groups: the Cache County study.[see comment].2001; 16(11):1043-1053.
5. Lyketsos CG, Lopez O, Jones B, Fitzpatrick AL, Breitner J, DeKosky S. Prevalence of neuropsychiatric symptoms in dementia and mild cognitive impairment: results from the cardiovascular health study. *JAMA.* 2002;288(12):1475-1483.
6. Monin JK, Schulz R. Interpersonal effects of suffering in older adult caregiving relationships. *Psychol Aging.* 2009;24(3):681-695.
7. Yaffe K, Fox P, Newcomer R, et al. Patient and caregiver characteristics and nursing home placement in patients with dementia. *JAMA.* 2002;287(16):2090-2097.
8. Taylor DH Jr, Ezell M, Kuchibhatla M, Ostbye T, Clipp EC. Identifying trajectories of depressive symptoms for women caring for their husbands with dementia. *J Am Geriatr Soc.* 2008;56(2):322-327.
9. Covinsky KE, Newcomer R, Fox P, et al. Patient and caregiver characteristics associated with depression in caregivers of patients with dementia. *J Gen Intern Med.* 2003;18(12):1006-1014.
10. Port CL, Zimmerman S, Williams CS, Dobbs D, Preisser JS, Williams SW. Families filling the gap: comparing family involvement for assisted living and nursing home residents with dementia. *Gerontologist.* 2005;45 Issue Suppl 1(1):87-95.
11. Schulz R, Belle SH, Czaja SJ, McGinnis KA, Stevens A, Zhang S. Long-term care placement of dementia patients and caregiver health and well-being. *JAMA.* 2004;292(8):961-967.
12. Noble JM, Manly JJ, Schupf N, Tang MX, Luchsinger JA. Type 2 diabetes and ethnic disparities in cognitive impairment. *Ethn Dis.* In press.
13. Tang MX, Stern Y, Marder K, et al. The APOE-epsilon4 allele and the risk of Alzheimer disease among African Americans, whites, and Hispanics. *JAMA.* 1998;279(10):751-755.
14. Rosenthal-Gelman C. Learning from recruitment challenges: barriers to diagnosis, treatment, and research participation for Latinos with symptoms of Alzheimer's disease. *J Gerontol Soc Work.* 2010;53:94-113.

15. Cox C, Monk A. Strain among caregivers: comparing the experiences of African American and Hispanic caregivers of Alzheimer's relatives. *Int J Aging Hum Dev.* 1996;43(2):93-105.
16. Valle R, Yamada AM, Barrio C. Ethnic differences in social network help-seeking strategies among Latino and Euro-American dementia caregivers. *Aging Ment Health.* 2004;8(6):535-543.
17. Ranney MJ, Aranda MP. Factors associated with depressive symptoms among Latino family dementia caregivers. *J Ethn Cult Divers Soc Work.* 2001;10(1):1-21.
18. Coon DW, Rubert M, Solano N, et al. Well-being, appraisal, and coping in Latina and Caucasian female dementia caregivers: findings from the REACH study. *Aging Ment Health.* 2004;8(4):330-345.
19. Flores YG, Hinton L, Barker JC, Franz CE, Velasquez A. Beyond familism: ethics of care of Latina caregivers of elderly parents with dementia. *Health Care Women Int.* 2009;30(12):1055-1072.
20. Mittelman MS, Roth DL, Clay OJ, Haley WE. Preserving health of Alzheimer caregivers: impact of a spouse caregiver intervention. *Am J Geriatr Psychiatry.* 2007;15(9):780-789.
21. Mittelman MS, Roth DL, Coon DW, Haley WE. Sustained benefit of supportive intervention for depressive symptoms in caregivers of patients with Alzheimer's disease. *Am J Psychiatry.* 2004;161(5):850-856.
22. Mittelman MS, Brodaty H, Wallen AS, Burns A. A three-country randomized controlled trial of a psychosocial intervention for caregivers combined with pharmacological treatment for patients with Alzheimer disease: effects on caregiver depression. *Am J Geriatr Psychiatry.* 2008;16(11):893-904.
23. Belle SH, Burgio L, Burns R, et al. Enhancing the quality of life of dementia caregivers from different ethnic or racial groups: a randomized, controlled trial. *Ann Intern Med.* 2006;145(10):727-738.
24. Gaugler JE, Reese M, Mittelman MS. Effects of the NYU caregiver intervention-adult child on residential care placement. *Gerontologist.* 2013 Dec;53(6):985-997.
25. Mittelman MS, Roth DL, Haley WE, Zarit SH. Effects of a caregiver intervention on negative caregiver appraisals of behavior problems in patients with Alzheimer's disease: results of a randomized trial. *J Gerontol B Psychol Sci Soc Sci.* 2004;59(1):P27-P34.
26. Mittelman MS, Haley WE, Clay OJ, Roth DL. Improving caregiver well-being delays nursing home placement of patients with Alzheimer disease. *Neurology.* 2006;67(9):1592-1599.
27. Belle SH, Czaja SJ, Schulz R, et al. Using a new taxonomy to combine the uncombinable: integrating results across diverse interventions. *Psychol Aging.* 2003;18(3):396-405.
28. Burgio L, Stevens A, Guy D, Roth DL, Haley WE. Impact of two psychosocial interventions on white and African American family caregivers of individuals with dementia. *Gerontologist.* 2003;43(4):568-579.

29. Bourgeois MS, Burgio LD, Schulz R, Beach S, Palmer B. Modifying repetitive verbalizations of community-dwelling patients with AD. *Gerontologist*. 1997;37(1):30-39.
30. Israel BA, Coombe CM, Cheezum RR, et al. Community-based participatory research: a capacity-building approach for policy advocacy aimed at eliminating health disparities. *Am J Public Health*. 2010 Nov;100(11):2094-2102.
31. Carrasquillo O, Chadiha LA. Development of community-based partnerships in minority aging research. *Ethn Dis*. 2007;17(1)(suppl 1):S3-S5.
32. Committee PM. *Draft Methodology Report: "Our Questions, Our Decisions: Standards for Patient-centered Outcomes Research*. Washington, DC; Patient-Centered Outcomes Research Institute; 2012.
33. Luchsinger JA, Tipiani D, Torres-Patino G, et al. Characteristics and mental health of Hispanic dementia caregivers in New York City. *Am J Alzheimers Dis Other Demen*. 2015 Sep;30(6):584-90.
34. Pearlin LI, Mullan JT, Semple SJ, Skaff MM. Caregiving and the stress process: an overview of concepts and their measures. *Gerontologist*. 1990;30(5):583-594.
35. Luchsinger J, Mittelman M, Mejia M, et al. The Northern Manhattan Caregiver Intervention Project: a randomised trial testing the effectiveness of a dementia caregiver intervention in Hispanics in New York City. *BMJ Open*. 2012;2(5):e001941.
36. Wisniewski SR, Belle SH, Coon DW, et al. The Resources for Enhancing Alzheimer's Caregiver Health (REACH): project design and baseline characteristics. *Psychol Aging*. 2003;18(3):375-384.
37. Amanda FE, Louis DB, Jamie D. Enhancing caregiver health: findings from the Resources for Enhancing Alzheimer's Caregiver Health II Intervention. *J Am Geriatric Soc*. 2010;58(1):30-37.
38. Teresi JA, Ramirez M, Lai JS, Silver S. Occurrences and sources of Differential Item Functioning (DIF) in patient-reported outcome measures: description of DIF methods, and review of measures of depression, quality of life and general health. *Psychol Sci Q*. 2008;50(4):538.
39. Gort AM, Mingot M, Gomez X, et al. Use of the Zarit scale for assessing caregiver burden and collapse in caregiving at home in dementias. *Int J Geriatr Psychiatry*. 2007;22(10):957-962.
40. Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res*. 1982;17(1):37-49.
41. Lawton MP, Moss M, Hoffman C, Perkinson M. Two transitions in daughters' caregiving careers. *Gerontologist*. 2000;40(4):437-448.
42. Savundranayagam MY, Montgomery RJ, Kosloski K. A dimensional analysis of caregiver burden among spouses and adult children. *Gerontologist*. 2011;51(3):321-331.
43. Teresi JA, Ocepek-Welikson K, Kleinman M, et al. Analysis of differential item functioning in the depression item bank from the Patient Reported Outcome Measurement Information System (PROMIS): an item response theory approach. *Psychol Sci Q*. 2009;51(2):148-180.

44. Choi SW, Reise SP, Pilkonis PA, Hays RD, Cella D. Efficiency of static and computer adaptive short forms compared to full-length measures of depressive symptoms. *Qual Life Res.* 2010;19(1):125-136.
45. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med.* 2001;16(9):606-613.
46. Cole SR. Assessment of differential item functioning in the Perceived Stress Scale-10. *J Epidemiol Community Health.* 1999;53(5):319-320.
47. Bruce B, Fries JF, Ambrosini D, et al. Better assessment of physical function: item improvement is neglected but essential. *Arthritis Res Ther.* 2009;11(6):R191.
48. Rose M, Bjorner JB, Becker J, Fries JF, Ware JE. Evaluation of a preliminary physical function item bank supported the expected advantages of the Patient-Reported Outcomes Measurement Information System (PROMIS). *J Clin Epidemiol.* 2008;61(1):17-33.
49. Marin G, Posner SF. The role of gender and acculturation on determining the consumption of alcoholic beverages among Mexican-Americans and Central Americans in the United States. *Int J Addict.* 1995;30(7):779-794.
50. Stokes JP. Toward an understanding of cohesion in personal change groups. *Int J Group Psychother.* 1983;33(4):449-467.

Copyright© 2019. Columbia University Health Sciences. All Rights Reserved.

Disclaimer:

The [views, statements, opinions] presented in this report are solely the responsibility of the author(s) and do not necessarily represent the views of the Patient-Centered Outcomes Research Institute® (PCORI®), its Board of Governors or Methodology Committee.

Acknowledgement:

Research reported in this report was [partially] funded through a Patient-Centered Outcomes Research Institute® (PCORI®) Award (#CE-1304-7160) Further information available at: <https://www.pcori.org/research-results/2013/reducing-burden-hispanic-caregivers-relatives-dementia-northern-manhattan>