CommunityRx-Dementia (CRx-D)
NCT04146545

Overall Statistical Analysis Plan (SAP): 12/14/2018

Aim 1 Statistical Analysis Plan: 3/17/2025

Aim 2 Statistical Analysis Plan: 1/14/2025

Overall Statistical Analysis Plan

NIH funding will support a mixed methods study including a randomized controlled trial and qualitative research to evaluate the impact of CRx-D. This research will fill important gaps in knowledge about how to best leverage clinical care visits to identify, support and improve outcomes for ADRD caregivers and PWD. For each aim, we provide exemplary hypotheses.

Aim 1: Among ADRD caregivers who screen positive for one or more unmet HRSNs, we will evaluate the longitudinal effects of CRx-D versus usual care on caregiver self-efficacy (1o outcome) and 2o outcomes including: psychosocial (unmet needs, social isolation, well-being, burden, depression, stress), behavioral (resource use), and health and healthcare utilization. We will also examine effects of the intervention on care recipient well-being, resource needs and use, health and health care utilization (reported by caregiver as the proxy).

> Hypothesis 1a. The intervention group will report higher levels of self-efficacy at 12 months compared to controls.
>
> Hypothesis 1b. Caregivers in the intervention group and their care recipients will have fewer unmet needs and higher rates of community resource use at 12 months than controls and their care recipients.
>
> Hypothesis 1c. Caregivers and PWD in the intervention group will have better health and well-being at 12 months than controls.

RCT Strategy, Timeline, Methods, and Key Measures: A consecutive sample of caregivers seeking care at UCMC will be enrolled Y1Q3-Y2. Eligible subjects will include patients who live in the target region, identify as an ADRD caregiver and agree to receive text messages from the research team (>98% of people screened for CRx-2 could receive texts). Recruitment will occur in three outpatient sites: primary care, gynecology and Senior Care – South Shore (SCSS). Researchers will consecutively approach patients in the waiting room, screen willing patients for eligibility and enroll eligible patients.

An estimated 6% of all US adults today are ADRD caregivers (and the number is growing). Two thirds of ADRD caregivers are ≥ 50 years. AA/B people are twice as likely as whites to have dementia and are more likely to be home-dwelling. Approximately 40K patients who live in target region will be seen in the 3 recruitment sites during a 12 month period (~87.5K visits/year); 60% of these patients are ≥50 years and 75% are AA/B. Given the race and age composition of the target patient population, we estimate that 6-9% of patients approached for enrollment will be ADRD caregivers.

We will need to approach between 7K-10.5K patients living in the target region to recruit 364 ADRD caregivers (344 for the RCT + 20 for pretesting). This conservative assumption derives from this calculation: 10,550 patients * 0.06 ADRD caregiver rate * 0.71 willing to be screened (rate from CRx-2 trial) * 0.81 cooperate (rate from CRx-2 trial). In CRx-2, we approached 3,100 UCMC patients over 8 months (387/month) to enroll 411 people into a trial with similar burden (length and number of surveys). For this study, we will need to approach 387-583 patients/month and have staffed and budgeted accordingly. To account for unforeseen challenges in recruiting, we have allotted 24 months for enrollment. If enrollment does not meet expected targets during the first quarter, we

will expand recruitment to the ER. This site sees 32-40K patients/yr from the target region and was a successful recruitment site for CRx-1 and -2. In addition, SCSS sees more than 250 PWD/yr from the target region. We can also recruit ADRD caregivers from SCSS. In two recent ADRD caregiver studies, we easily met recruitment goals at SCSS with high cooperation rates (>90%). From this group, we will recruit a caregiver advisory board using a community-engaged research approach that we created together with community members and have applied in several prior successful studies.

Preliminary data indicate that 52% of caregivers will have ≥1 unmet HRSNs. To ensure equal numbers between groups, we will use stratified randomization to allocate subjects to usual care or CRx-D by number of unmet HRSNs (0 or ≥1). Randomization will occur immediately after enrollment using the method of permuted blocks and the uniform random number function in Stata (v15). A stratum will be closed for enrollment once its accrual goal has been reached.

(1) Usual Care Arm: the usual care arm includes typical clinic visit procedures, including, but not limited to: meeting with a Patient Service Representative, receipt of a printed after visit summary (AVS), notification of any financial obligations and scheduling any future visits as appropriate. Usual care may also include information about community resources from the patient's healthcare team.

(2) CommunityRx-Dementia (CRx-D): In this arm of the study, caregivers will receive usual care in addition to CRx-D. CRx-D will be initiated by the Community Resource Specialist (a researcher) at the baseline (index) primary care visit (academic primary care, geriatrics, ob/gyn clinics). Caregivers assigned to CRx-D and will receive: (a) a brief, structured educational intervention about common resource needs among ADRD caregivers (e.g., "It is recommended that we talk with caregivers of people with dementia about community resources that may be helpful. Many caregivers benefit from resources like support groups or respite care. It is also common for caregivers to need help with things like food or rent and utility support."); (b) delivery and review of a personalized resource "prescription" (HealtheRx-D) for vetted resources near the caregiver's preferred location to address HRSNs (e.g. food, housing, interpersonal safety and social support) and ADRD caregiver needs (e.g. ADRD education, support groups, respite care, advance care planning); (c) demonstration of use of an online Community Resource Finder that participants can use to find and share community resources (for HRSNs, caregiving and any self-care needs), give feedback and request additional resource information; and (d) coaching on how to activate community resources, including how to reach the Community Resource Specialist.

All caregivers will complete an interviewer-administered baseline survey. Those randomized to the intervention arm will complete the survey prior to receipt of the CRx-D intervention. In addition to the printed "HealtheRx-D" delivered at the time of the baseline visit, the HealtheRx will be texted to the participant and, if the participant wishes, emailed within 24 hours of the baseline visit. The initial encounter will be followed by a series of timed text messages from the Community Resource Specialist offering ongoing community resource linkage support. At any time, a participant can reply "stop" to stop receiving text messages from the CRx-D Community Resource Specialist. Table 2 describes the frequency and content of automated text messages that will be sent to caregivers randomized to CRx-D. The CRx-D Community Resource Specialist, K. Paradise, is trained in NowPow's system and will use its "nudge" technology to send tracked text messages at

the intervals outlined below. The content of these messages is based on a text messaging experiment conducted during the CMMI HCIA CRx study. The timing of these messages is informed by the Critical Time Intervention model. Participants will be informed at the time of enrollment and in text message communications that the Community Resource Specialist will respond to text messages within 24 hours during regular work days and within 48 hours of weekends or holidays. For all participants, phone numbers will be verified via text message at the time of 7 day, 30 day, 90 day and 12 month surveys and via text message each month in between follow-up surveys. We will collect a back-up phone, mailing address and email contacts for participants (this contact info is typically documented in the electronic medical record). To maximize retention, a brief check-in will be conducted at 180 days.

The baseline survey will be administered following randomization by a trained interviewer. All caregivers will receive usual care. Participants will be blinded to the arm to which they are randomized. For caregivers randomized to the intervention arm, CRx-D will be delivered by the interviewer. Follow-up phone surveys will be conducted using computer-assisted personal interviewing (CAPI) at 7, 30, 90 days and 12 months post-enrollment. To maximize retention, a brief check-in will be conducted at 180 days and we will use text messaging and phone calls monthly between surveys to verify contact information and to schedule and remind participants about upcoming surveys. Intervention group participants can also text, email or call the CRx-D resource specialist for resource support throughout the follow-up period.

We will conduct two waves of baseline and 1 week follow-up pre-testing (Y1Q3), each with 5 ADRD caregivers with 1 or more unmet HRSNs and 5 ADRD caregivers with 0 unmet HRSNs (20 total with the two waves) to ensure survey quality and minimize risk of measurement and CAPI error. Table 1 summarizes example measures, sources and planned time points for collection. Dynamic factors will be assessed at appropriate intervals (e.g., during the last 30 days or in the interval since the last survey).

The primary outcome for Aim 1 will be assessed using the 4-item self-efficacy sub-domain from the 2015 Caregiver Dementia Care and Self-Efficacy Survey developed by Jennings et al (NIA-5P30AG028748, CMS-1C1-12-0001). We selected this relatively new measure because it is, to our knowledge, the only available measure validated specifically for ADRD caregivers that assesses self-efficacy for accessing services.

Per Jennings' protocol, responses to each of the 4 items in the self-efficacy domain will be scored from 1 to 5 (1=strongly disagree to 5=strongly agree) and averaged. We will also assess caregiver self-efficacy for self-care using our measure from the CRx-2 trial. This measure derives from Bandura's Self-Efficacy theory: "How confident are you in your ability to find resources in your community that help you manage your [your care recipient's] health?" (5-point Likert scale "not at all confident" to "completely confident"). The pragmatic CRx-2 intervention was delivered, via the routine workflow, to adult patients by clinic staff or a nurse at discharge from an outpatient visit with no specific counseling. The proposed CRx-D intervention is more intensive (e.g. delivered by the research team with a brief education component, ongoing support, caregiver access to the Community Resource Finder). We therefore hypothesize that CRx-D may also promote general self-efficacy (a broader concept than Jennings' and the CRx-2 measure) and will explore this hypothesis by administering the Generalized Self-Efficacy Scale (range 10-40) at baseline and at follow-up.

Secondary and exploratory outcomes will also be measured longitudinally. These outcomes include:

Psychosocial Outcomes HRSNs will be assessed at baseline and follow-up using the CMS tool. The "core" tool includes 10 validated items to assess 5 needs domains: food, housing, interpersonal safety, transportation, and utility support. We will also include the domains "family and community support" (2 items) and "financial strain" (1 item). Each domain will be scored per CMS instructions. For follow-up surveys after 1 week, we will assess needs occurring in the interval since the last interview and, when a need is identified, during the past 30 days specifically. In addition to assessing unmet HRSNs, we will assess unmet caregiving-specific resource needs, using a tool we created from prior studies and our own qualitative work.

Caregiver well-being will be assessed using PROMIS Neuro-QoL ShortForm v1.0 – Positive Affect and Well-Being. Well-being of the PWD will be assessed by caregiver proxy using the Alzheimer's Disease Related Quality of Life (ADRQL) measure. Caregiver burden will be assessed using the Dementia Caregiver Burden Scale, a validated measure created by Peipert et al. Depression will be assessed using the CES-D, the most widely used measure of depression in research, and perceived stress will be assessed using the validated and widely used Perceived Stress Scale. Activation is a domain of interest for our exploratory work. We will measure caregiver activation for self-care using the validated Patient Activation Measure. We will also measure caregiver activation for PWD's care using the validated Managing Your Loved One's Health (MYLOH) measure developed by consultant Dr. Borson and colleagues.

Behavioral Outcomes As the field and related technology advance, "closed loop" data documenting community resource use will emerge. These data require that both the referring institution (e.g. the primary care clinic) and the receiving organization (e.g. ADRD support group) are using the same technology platform. If closed loop data become available, we will use those. For now, we plan to assess self-reported resource use (by the caregiver and by the care recipient by proxy) using items developed for CRx-1 and 2 to measure contacting, going to and sharing information about resources. Resource use will also be assessed using qualitative data from text messages between the Community Resource Specialist and the caregiver. Data generated by caregivers' use of the web-based Community Resource Finder will be used to assess resource searches, selection and sharing as well as unfulfilled requests (searches that resulted in no available resources). Text message data and Community Resource Finder data will be collected via use of the NowPow system and transferred regularly to the research team.

Health and Healthcare Outcomes Caregiver physical health will be assessed using the Medical Outcomes Study SF-12 and EMR data. Physical health of the PWD will be assessed by proxy using a single item global measure of physical health. Healthcare utilization will be assessed using the DEXCOM, a self-report tool developed by Co-PI E. Huang et al and used in CRx-2. We will explore outcomes such as days to, number and duration of ER/hospital visits, and patterns of 911 and ambulance use to test the hypothesis that supporting caregivers with community resources is associated with less need for emergency healthcare services.

Analysis and Sample Size Justification: To determine the impact of CRx-D on the highest risk group, Aim 1 analyses will focus on 172 caregivers with at least 1 HRSN. Descriptive statistics will be used to summarize, overall and by study arm, sociodemographic characteristics and primary

and secondary outcomes at each measured time point. The mean, standard deviation, median, and inter-quartile range will be generated for continuous variables; frequency counts and percentages will be generated for categorical variables.

The main analyses will follow the principle of intent-to-treat, and include all 172 ADRD caregivers with at least 1 unmet HRSN in the study arm to which they were allocated. We aim to minimize missing data by employing interviewer-administered surveys and communicating to participants the protections of a certificate of confidentiality. Using these strategies, item non-response ranged from 0-3% in the CRx-2 trial. To avoid potential bias due to missing data, analyses will be conducted using multiple imputation with the chained equations method or inverse probability weighting to account for dropout and/or item non-response. In addition, characteristics of caregivers who complete each follow-up survey will be compared to caregivers who do not.

To fully leverage the longitudinal data, generalized linear mixed-effects models (GLMM) will be fit with particular interest in the time (7, 30, 90 days and 12 months post-enrollment) by study arm interaction; baseline values will be included as a covariate. Based on this model, appropriate contrasts to test the intervention effect at specific time points (e.g., 12 months) will be constructed. These contrasts will permit examination of early versus late intervention effects, valuable information for future iterations of this and other ADRD caregiver self- and family management interventions. GLMM can be used for continuous (e.g., mean caregiver self-efficacy for providing care score, patient satisfaction score, stigma score), ordinal (e.g., physical health of PWD by proxy) or binary (e.g., resource use, resource sharing) outcomes. Models will include covariates moderately associated with the outcome to remove differences attributable to baseline covariates and increase precision of the estimate of the treatment affect. Latent growth curve modeling is a viable alternative to GLMM and will be explored when sufficient variability in the distributions of patterns of outcomes by time points exist (e.g., with the binary outcomes there will be a limited number of potential patterns).

Potential moderators of the intervention effect (e.g., race, sex) will also be assessed by inclusion of a moderator by study arm interaction term. Although in the CRx-2 trial we saw no gender differences in outcomes, special consideration will be given to sex (as measured by self-reported gender) as a biological variable (moderator) and variation in intervention effects will be examined in accordance with NIH policy. In addition, based on preliminary evidence of similarity across race and ethnic groups from the CRx-2 trial, we hypothesize that the intervention effect will be similar and will analyze for dramatic departures from this outcome (e.g., whether the intervention has no or negative effect among non-Hispanic white caregivers versus a strong positive effect among AA/B caregivers) through effect modification analyses such as testing of the race by intervention interaction. Also, we will use mediation analyses to identify factors that mediate the relationship between the study arm and the outcome of interest. Improved understanding of moderating and mediating factors will facilitate future improvements to the intervention. For example, if the relationship between intervention effect and caregiver burden at time t is mediated by self-efficacy at time t-1, as posited in our conceptual framework, future improvements might focus the intervention's messaging more explicitly on fostering caregiver self-efficacy.

Transition models will be used to further study the impact of CRx-D on state changes in other outcomes, including unmet needs, well-being and resource use. This model will be of the general

form: logit P(Yit = 1 | Yit-1 = yit-1) = X + yit-1 + yit-1*X where X is the intervention group variable and Yit is a binary outcome variable for subject i at time t. This approach will allow us to determine the frequency of transitioning into and out of a state (e.g., having any unmet HRSNs to having no unmet HRSNs) and how the intervention impacts on these transitions. This model can be extended to permit ordinal outcomes (e.g., self-efficacy for finding services to help care for the PWD). We will assume a first-order Markov model but check the sensitivity of the model to this assumption and will use robust standard errors.

We will also use survival analysis methods and methods for analyzing count data (e.g., Poisson or negative binomial regression) to explore acute and emergency healthcare utilization. For example, time from enrollment to first home EMS or ER visit will be calculated for each subject and those who never have one of these events will be censored at the end of study follow-up. Stratified Cox proportional hazards regression models will be fit with intervention group as the independent variable and clinical enrollment site as a stratification factor, due to the potential for different underlying rates based on clinical site. The proportional hazards assumption will be checked using log-log plots of survival and Schoenfeld residuals.

For the primary outcome (self-efficacy for providing care at 12 months) a sample of 86 caregivers/arm allows for detection of an intervention group difference of 0.43 points (sample size calculation using a group allocation of 1:1; power=0.8; α=0.05, mean=2.8 and SD=1). With this sample size, power is sufficient to see at least a half a standard deviation (SD) difference in caregiver self-efficacy between intervention and control groups; half a standard deviation is considered a minimally important difference. Assuming 25% attrition, we will still be able to detect a difference equal to half an SD. Attrition in the pragmatic CRx-2 trial was 20%; this was an even lower-intensity intervention with no ongoing resource support nor touchpoints between follow-up surveys. In higher-intensity caregiver support studies, attrition rates range from 4% (MIND Study at 18 months) to 10% (REACH-II Study at 6 months). We will follow all caregivers through 12 months even if they no longer care for the PWD.

Aim 2. Among all ADRD caregivers (those who do and do not screen positive for unmet HRSNs), evaluate acceptability of the intervention as well as the effects of CRx-D plus usual care versus usual care alone on the health care experience, including satisfaction with care (1o outcome), experiences of stigma during clinical care and likelihood of sharing community resource information with others.

> Hypothesis 2a. Among intervention group participants, the CRx-D intervention will be acceptable to caregivers with and without unmet HRSNs.
>
> Hypothesis 2b. There will be no evidence, comparing the intervention group to controls, that CRx-D decreases caregiver satisfaction with care or promotes caregiver stigma.
>
> Hypothesis 2c. Among ADRD caregivers with no unmet HRSNs at baseline, those randomized to the CRx-D intervention will demonstrate a lower likelihood of developing an unmet HRSN. In addition, they will exhibit a higher likelihood of sharing resource information with others compared to those randomized to usual care.

Strategy, Timeline, Methods, and Key Measures: Aim 2 will utilize RCT data (collected as described in Aim 1) from ADRD caregivers with and without unmet HRSNs. For Aim 2, the primary outcome of interest is satisfaction with care. Satisfaction will be assessed 7 days post-discharge from the caregiver's primary care visit using the Patient Satisfaction Questionnaire 18-item short form (PSQ-18). The PSQ-18 is widely used in research and shows internal consistency with the long form of the PSQ, an 80-item survey developed in 1976. The 18 items are grouped into 7 domains; each domain is scored separately (there is no overall score). The general satisfaction sub-scale of the PSQ (2/18 questions) is the most relevant PSQ domain for our study.

To assess stigma, we will use the Discrimination in Medical Settings Scale (DMS). The DMS was adapted from the Everyday Discrimination Scale for medical settings and has been tested for reliability and validity among African Americans. It includes 7 items, each scored on a 1 to 5 scale to generate a mean overall score.

Using the same transition model approach described for Aim 1, we will study the impact of the intervention on state changes in unmet needs specifically among caregivers without unmet needs at baseline. This approach will fill a gap in knowledge about the joint incidence of unmet HRSNs and ADRD caregiving needs over time.

Lastly, among those assigned to the intervention group, we will examine caregivers' propensity for sharing information about resources, comparing a) those with and without unmet HRSNs at baseline and b) among those without baseline unmet HRSNs, by study arm. Based on Maslow's Theory of Human Motivation, we hypothesize higher caregiving self-efficacy and well-being rates at 12 months among caregivers who shared resource information with others and that the effect will be highest for caregivers with unmet baseline HRSNs.

Analysis and Sample Size Justification Physician concern about negatively impacting patient satisfaction or causing stigma are barriers to intervening to address unmet needs. To address these, we hypothesize that satisfaction with care will not be significantly diminished among caregivers randomized to CRx-D compared to those randomized to usual care. Aim 2 analyses will include all caregivers enrolled in the trial (N=344).

We will use the analytic approach described for Aim 1. Here, we discuss analytic issues unique to Aim 2. To facilitate ease of interpretation of the primary outcome, we will linearly rescale the composite score for the PSQ-18 subdomain of general satisfaction (questions 3 and 17). Five possible responses for each of the two items are scored 1 to 5; with a higher score equating to higher satisfaction. Scores will be transformed using the formula: $y= 100 * (x-a)/(b-a)$ where y=the transformed score, x=the original score, a=the minimum possible score and b=the maximum possible score. Transformed items will be averaged to generate the composite score (range 0-100). Of particular interest for the non-inferiority hypothesis (H2b) will be the confidence interval for the difference (usual care - CRx-D) and whether or not the upper limit of the confidence interval exceeds the non-inferiority margin. Of note, superiority of the intervention with respect to stigma, unmet needs and other outcomes will also be tested as this can be done without penalty. Caregiver satisfaction will be measured at 7 days following the baseline outpatient visit; no longitudinal analyses will be conducted for this outcome. Typically, analyses of patient satisfaction data adjust for case-mix (age, education and general health status); this will be feasible, but given the randomized design and single-site analysis, adjustment should not be necessary. In addition,

to determine whether there are substantial differences in the intervention's effect on ADRD caregivers with unmet HRSNs compared to ADRD caregivers without unmet HRSNs, the intervention by unmet HRSNs interaction will be tested. As in Aim 1, we will examine for differential intervention effects on primary and secondary outcomes among racial and ethnic groups, and by sex, using effect modification analyses. Some analyses will be restricted to the group without HRSNs (N=172) to provide a more complete assessment of the impact of the intervention on this group (e.g., incidence of unmet HRSNs over time).

The proposed sample size of 344 is sufficient based on a non-inferiority margin of 7.6 points in patient satisfaction (primary outcome) assuming group allocation 1:1, power=0.8, one-sided α=0.025, standard deviation=25. The non-inferiority margin of 7.6 points is less than half of the standard deviation for mean transformed general satisfaction scores in a population-based sample (e.g., male [mean=65, SD=23], female [mean=64, SD=24]); half a standard deviation is considered a minimally important difference.

Aim 1 Statistical Analysis Plan

Table of Contents

Note: Sections 1-5 are copied verbatim from the CommunityRx-Dementia R01 application (submitted 12/14/2018; R01AG064949). Any notes or points of clarification are denoted in brackets.

Data were analyzed by co-authors K. Wroblewski, MS and Suyeon Lee, PhD.

Introduction

The purpose of Aim 1 is to study the effect of a screening-based approach to delivery of the CRx-D intervention on caregiver self-efficacy [primary outcome] and more distal caregiver and PWD health and healthcare outcomes.

Measures Used

Caregiver Self-efficacy [Aim 1]
The primary outcome for Aim 1 will be assessed using the 4-item self-efficacy sub-domain from the 2015 Caregiver Dementia Care and Self-Efficacy Survey developed by Jennings et al for the NIA- and CMS-funded Alzheimer's and Dementia Care Program.[1] We selected this measure because it is, to our knowledge, the only available measure validated specifically for ADRD caregivers that assesses self-efficacy for accessing services. Per Jennings' protocol, responses to each of the 4 items in the self-efficacy domain will be scored from 1=strongly disagree to 5=strongly agree and averaged. We will also assess caregiver self-efficacy for self-care using our measure from the CRx-2 trial. This measure derives from Bandura's Self-Efficacy theory: "How confident are you in your ability to find resources in your community that help you manage your [your care recipient's] health?" (5-point Likert scale "not at all confident" to "completely confident").[2]

Psychosocial Outcomes
HRSNs will be assessed at baseline and follow-up using the CMS tool.[3] The "core" tool includes 10 validated items to assess 5 needs domains: food, housing, interpersonal safety, transportation, and utility support. We will also include the domains "family and community support" (including social isolation) (2 items) and "financial strain" (1 item). Scoring will be per CMS instructions. For follow-up surveys after 1 week, we will assess needs occurring in the interval since the last interview and, when a need is identified, during the past 30 days. In addition to assessing unmet HRSNs, we will assess unmet caregiving-specific resource needs using a measure building on prior studies and our own qualitative work. Caregiver well-being will be assessed using PROMIS Neuro-QoL ShortForm v1.0 – Positive Affect and Well-Being.[4] Well-being of the PWD will be assessed by caregiver proxy using the Alzheimer's Disease Related Quality of Life (ADRQL) measure.[5] Caregiver burden will be assessed using the Dementia Caregiver Burden Scale, a validated measure created by Peipert et al.[6] Depression will be assessed using the CES-D,[7] the most widely used measure of depression in research, and perceived stress will be assessed using the validated and widely used Perceived Stress Scale.[8] Activation is a domain of interest for our exploratory work. We will measure caregiver activation for self-care using the validated Patient Activation Measure.[9] We will also measure caregiver activation for PWD's care using the validated Managing Your Loved One's Health (MYLOH) measure developed by consultant Dr. Borson and colleagues.[10,11]

Behavioral Outcomes
We plan to assess self-reported resource use (by the caregiver and by the care recipient by proxy) using items developed for CRx-1 and 2 to measure contacting, going to and sharing information about resources. Resource use will also be assessed using qualitative data from text messages between the community resource specialist and the caregiver. Data generated by caregivers' use of the web-based Community Resource Finder will be used to assess resource searches, selection and sharing as well as unfulfilled requests (searches that resulted in no available resources). Text message data and Community Resource Finder data will be collected via use of the NowPow system and transferred regularly to the research team.

Health and Healthcare Outcomes

Caregiver physical health and quality of life will be assessed using the Medical Outcomes Study SF-12 and EMR data.[12] Physical health of the PWD will be assessed by proxy using a single item global measure of physical health. Healthcare utilization will be assessed using the DEXCOM,[13] a self-report tool developed by Co-PI E. Huang and used in CRx-2.

Sample Size

Aim 1 [primary] analyses will include ADRD caregivers who screen positive for one or more unmet HRSNs (N=172). For the primary outcome (self-efficacy for providing care at 12 months) a sample of 86 caregivers/arm allows for detection of an intervention group difference of 0.43 points (sample size calculation using a group allocation of 1:1; power=0.8; α=0.05, mean=2.8 and SD=1). With this sample size, power is sufficient to see at least a half a standard deviation (SD) difference in caregiver self-efficacy between intervention and control groups; half a standard deviation is considered a minimally important difference.[14] Assuming 25% attrition, we will still be able to detect a difference equal to half an SD. Attrition in the pragmatic CRx-2 trial was 20%; this was an even lower-intensity intervention with no ongoing resource support nor touchpoints between follow-up surveys.

Analysis Plan

To determine the impact of CRx-D on the highest risk group, Aim 1 analyses will focus on 172 caregivers with at least 1 HRSN. Descriptive statistics will be used to summarize, overall and by study arm, sociodemographic characteristics and primary and secondary outcomes at each measured time point. The mean, standard deviation, median, and inter-quartile range will be generated for continuous variables; frequency counts and percentages will be generated for categorical variables.

The main analyses will follow the principle of intent-to-treat,[15] and include all 172 ADRD caregivers with at least 1 unmet HRSN in the study arm to which they were allocated.

To fully leverage the longitudinal data, generalized linear mixed-effects models (GLMM) will be fit with particular interest in the time (7, 30, 90 days and 12 months post-enrollment) by study arm interaction; baseline values will be included as a covariate. Based on this model, appropriate contrasts to test the intervention effect at specific time points (e.g., 12 months) will be constructed. These contrasts will permit examination of early versus late intervention effects, valuable information for future iterations of this and other ADRD caregiver self- and family management interventions. GLMM can be used for continuous (e.g., mean caregiver self-efficacy for providing care score, patient satisfaction score, stigma score), ordinal (e.g., physical health of PWD by proxy) or binary (e.g., resource use, resource sharing) outcomes. Models will include covariates moderately associated with the outcome to remove differences attributable to baseline covariates and increase precision of the estimate of the treatment affect. Latent growth curve modeling is a viable alternative to GLMM and will be explored when sufficient variability in the distributions of patterns of outcomes by time points exist (e.g., with the binary outcomes there will be a limited number of potential patterns).

Potential moderators of the intervention effect (e.g., race, sex) will also be assessed by inclusion of a moderator by study arm interaction term. Although in the CRx-2 trial we saw no gender differences in outcomes, special consideration will be given to sex (as measured by self-reported gender) as a biological variable (moderator) and variation in intervention effects will be examined in accordance with NIH policy.[16] In addition, based on preliminary evidence of similarity across race and ethnic groups from the CRx-2 trial, we hypothesize that the intervention effect will be similar and will

analyze for dramatic departures from this outcome (e.g., whether the intervention has no or negative effect among non-Hispanic white caregivers versus a strong positive effect among AA/B caregivers) through effect modification analyses (e.g. testing race by intervention interaction). Also, we will use mediation analyses to identify factors that mediate the relationship between the study arm and the outcome of interest. Improved understanding of moderating and mediating factors will facilitate future improvements to the intervention. For example, if the relationship between intervention effect and caregiver burden at time t is mediated by self-efficacy at time t-1, as posited in our conceptual framework, future improvements might focus the intervention's messaging more explicitly on fostering caregiver self-efficacy.

Transition models will be used to further study the impact of CRx-D on state changes in other outcomes, including unmet needs, well-being and resource use. This model will be of the general form: logit $P(Y_{it} = 1 \mid Y_{it-1} = y_{it-1}) = X + y_{it-1} + y_{it-1}*X$ where X is the intervention group variable and Y_{it} is a binary outcome variable for subject i at time t. This approach will allow us to determine the frequency of transitioning into and out of a state (e.g., having any unmet HRSNs to having no unmet HRSNs) and how the intervention impacts on these transitions. This model can be extended to permit ordinal outcomes (e.g., self-efficacy for finding services to help care for the PWD). We will assume a first-order Markov model but check the sensitivity of the model to this assumption and will use robust standard errors.

We will also use survival analysis methods and methods for analyzing count data (e.g., Poisson or negative binomial regression) to explore acute and emergency healthcare utilization. For example, time from enrollment to first home EMS or ER visit will be calculated for each subject and those who never have one of these events will be censored at the end of study follow-up. Stratified Cox proportional hazards regression models will be fit with intervention group as the independent variable and clinical enrollment site as a stratification factor, due to the potential for different underlying rates based on clinical site. The proportional hazards assumption will be checked using log-log plots of survival and Schoenfeld residuals.

Handling of Missing Data

We aim to minimize missing data by employing interviewer-administered surveys and communicating to participants the protections of a Certificate of Confidentiality. Using these strategies, item non-response ranged from 0-3% in the CRx-2 trial. To avoid potential bias due to missing data, analyses will be conducted using multiple imputation with the chained equations method or inverse probability weighting to account for dropout and/or item non-response. In addition, characteristics of caregivers who complete each follow-up survey will be compared to caregivers who do not. We will follow all caregivers through 12 months even if they no longer care for the PWD.

Additional Information, Clarification, Changes

Aim 1:

For each Aim, there is a single primary outcome, and no multiplicity adjustment will be made to the analyses of the secondary outcomes. Analyses will also be performed in the overall group (irrespective of baseline HRSNs) and among those without baseline HRSNs.

These analyses aim to fill a gap in knowledge about the effectiveness of a universal intervention approach. The treatment group by baseline HRSN status interaction will also be tested.

Since a majority of participants were female and African American or Black, formal interaction testing was not performed.

Models will also be fit with the intervention group separated into two subgroups (those with at least 1 adhoc HealtheRx vs not).

Aim 2 Statistical Analysis Plan

Table of Contents

Note: Sections 1-5 are copied verbatim from the CommunityRx-Dementia R01 application (submitted 12/14/2018; R01AG064949). Any notes or points of clarification are denoted in brackets.

Data were analyzed by co-authors K. Wroblewski, MS and Suyeon Lee, PhD.

Introduction

The purpose of Aim 2 is to test the hypothesis that a universal approach to delivery of the CRx-D intervention is acceptable to caregivers and does not diminish caregiver satisfaction with care [primary outcome] or cause caregiver harm relative to usual care.

Measures Used

Satisfaction with Care

Satisfaction will be assessed 7 days post-discharge from the caregiver's ambulatory visit using the Patient Satisfaction Questionnaire 18-item short form (PSQ-18).[17] The PSQ-18 is widely used in research and shows internal consistency with the PSQ long form, an 80-item survey developed in 1976. The 18 items are grouped into 7 domains; each domain is scored separately (there is no overall score). The general satisfaction sub-scale of the PSQ (2/18 questions) is the most relevant PSQ domain for our study.

Experiences of Discrimination

To assess stigma, we will use the Discrimination in Medical Settings Scale (DMS) [and it will be assessed on day 7].[18] The DMS was adapted from the Everyday Discrimination Scale for medical settings and has been tested for reliability and validity among African Americans and used in studies of Korean American and American Indian women. It includes 7 items, each scored on a 1 to 5 scale to generate an overall score [range 7-35, higher scores indicating more frequent discrimination].

Sample Size

Aim 2 analyses will include all caregivers enrolled in the trial (N=344). The non-inferiority margin of 7.6 points is less than half of the standard deviation for mean transformed general satisfaction scores in a population-based sample (e.g., male [mean=65, SD=23], female [mean=64, SD=24]); half a standard deviation is considered a minimally important difference. [Calculations assumed power=0.8, one-sided alpha=0.025, and standard deviation=25]

Analysis Plan

***Aim* 1 Analytic Plan [Aim 1 analytic plan is included here because it was referenced in Aim 2 due to overlap)**

To determine the impact of CRx-D on the highest risk group, Aim 1 analyses will focus on 172 caregivers with at least 1 HRSN. Descriptive statistics will be used to summarize, overall and by study arm, sociodemographic characteristics and primary and secondary outcomes at each measured time point. The mean, standard deviation, median, and inter-quartile range will be generated for continuous variables; frequency counts and percentages will be generated for categorical variables.

The main analyses will follow the principle of intent-to-treat, and include all 172 ADRD caregivers with at least 1 unmet HRSN in the study arm to which they were allocated.

To fully leverage the longitudinal data, generalized linear mixed-effects models (GLMM) will be fit with particular interest in the time (7, 30, 90 days and 12 months post-enrollment) by study arm interaction; baseline values will be included as a covariate. Based on this model, appropriate

contrasts to test the intervention effect at specific time points (e.g., 12 months) will be constructed. These contrasts will permit examination of early versus late intervention effects, valuable information for future iterations of this and other ADRD caregiver self- and family management interventions. GLMM can be used for continuous (e.g., mean caregiver self-efficacy for providing care score, patient satisfaction score, stigma score), ordinal (e.g., physical health of PWD by proxy) or binary (e.g., resource use, resource sharing) outcomes. Models will include covariates moderately associated with the outcome to remove differences attributable to baseline covariates and increase precision of the estimate of the treatment affect. Latent growth curve modeling is a viable alternative to GLMM and will be explored when sufficient variability in the distributions of patterns of outcomes by time points exist (e.g., with the binary outcomes there will be a limited number of potential patterns).

Potential moderators of the intervention effect (e.g., race, sex) will also be assessed by inclusion of a moderator by study arm interaction term. Although in the CRx-2 trial we saw no gender differences in outcomes, special consideration will be given to sex (as measured by self-reported gender) as a biological variable (moderator) and variation in intervention effects will be examined in accordance with NIH policy.[16] In addition, based on preliminary evidence of similarity across race and ethnic groups from the CRx-2 trial, we hypothesize that the intervention effect will be similar and will analyze for dramatic departures from this outcome (e.g., whether the intervention has no or negative effect among non-Hispanic white caregivers versus a strong positive effect among AA/B caregivers) through effect modification analyses (e.g. testing race by intervention interaction). Also, we will use mediation analyses to identify factors that mediate the relationship between the study arm and the outcome of interest. Improved understanding of moderating and mediating factors will facilitate future improvements to the intervention. For example, if the relationship between intervention effect and caregiver burden at time t is mediated by self-efficacy at time t-1, as posited in our conceptual framework, future improvements might focus the intervention's messaging more explicitly on fostering caregiver self-efficacy.

Aim 2

We will use the analytic approach described for Aim 1 [above]. Here, we discuss analytic issues unique to Aim 2. To facilitate ease of interpretation of the primary outcome, we will linearly rescale the composite score for the PSQ-18 subdomain of general satisfaction (questions 3 and 17). Five possible responses for each of the two items are scored 1 to 5; with a higher score equating to higher satisfaction. Scores will be transformed using the formula: y= 100 * (x-a)/(b-a) where y=the transformed score, x=the original score, a=the minimum possible score and b=the maximum possible score. Transformed items will be averaged to generate the composite score (range 0-100; [higher scores indicating greater satisfaction]). Of particular interest for the non-inferiority hypothesis will be the [95%] confidence interval for the difference (usual care - CRx-D) and whether or not the upper limit of the confidence interval exceeds the non-inferiority margin. Of note, superiority of the intervention with respect to [satisfaction or] stigma will also be tested as this can be done without penalty. Typically, analyses of patient satisfaction data adjust for case-mix (age, education and general health status); this will be feasible, but given the randomized design and single-site analysis, adjustment should not be necessary. In addition, to determine whether there are substantial differences in the intervention's effect on ADRD caregivers with unmet HRSNs compared to ADRD caregivers without unmet HRSNs, the intervention by unmet HRSNS interaction will be tested. As in Aim 1, we will examine for differential intervention effects on primary and secondary outcomes among racial and ethnic groups, and by sex, using effect modification

analyses. Some analyses will be restricted to the group without HRSNs (N=172) to provide a more complete assessment of the impact of the intervention on this group (e.g., incidence of unmet HRSNs over time).

Handling of Missing Data

We aim to minimize missing data by employing interviewer-administered surveys and communicating to participants the protections of a Certificate of Confidentiality. Using these strategies, item non-response ranged from 0-3% in the CRx-2 trial. To avoid potential bias due to missing data, analyses will be conducted using multiple imputation with the chained equations method or inverse probability weighting to account for dropout and/or item non-response. In addition, characteristics of caregivers who complete each follow-up survey will be compared to caregivers who do not.

Additional Information, Clarification, Changes

September 2022

For each Aim, there is a single primary outcome, and no multiplicity adjustment will be made to the analyses of the secondary outcomes.

Aim 2:
Since the non-inferiority margin of 7.6 for satisfaction with care is a relative effect size of 0.3 standard deviations, this relative effect size will be used as the non-inferiority margin for the secondary outcome of discrimination.

More formally, non-inferiority testing will use linear regression: $Y = \beta0 + \beta1*T + \varepsilon$, where Y is the outcome (satisfaction with care or discrimination) and T is the treatment group indicator. The 95% CI for β1 will be of particular interest.

Since a majority of participants were female and African American or Black, formal interaction testing was not performed. However, a sensitivity analysis was performed that restricted analyses to females.

An additional sensitivity analysis was performed that restricted the analysis to those who self-completed the 1-week survey online (vs. via phone interview).

REFERENCES

1. Jennings LA, Reuben DB, Evertson LC, Serrano KS, Ercoli L, Grill J, Chodosh J, Tan Z, Wenger NS. Unmet needs of caregivers of individuals referred to a dementia care program. J Am Geriatr Soc. 2015 Feb;63(2):282–289. PMCID: PMC4332558

2. Bandura A. Guide for constructing self-efficacy scales. Self-Efficacy Beliefs of Adolescents [Internet]. Information Age Publishing; 2006. p. 307–337. Available from: https://www.uky.edu/~eushe2/Bandura/BanduraGuide2006.pdf

3. Centers for Medicare & Medicaid Services. The AHC Health-Related Social Needs Screening Tool [Internet]. [cited 2023 Jun 1]. Available from: https://innovation.cms.gov/files/worksheets/ahcm-screeningtool.pdf

4. Salsman JM, Victorson D, Choi SW, Peterman AH, Heinemann AW, Nowinski C, Cella D. Development and validation of the positive affect and well-being scale for the neurology quality of life (Neuro-QOL) measurement system. Qual Life Res. 2013 Nov;22(9):2569–2580. PMCID: PMC3855608

5. Kasper JD, Black BS, Shore AD, Rabins PV. Evaluation of the Validity and Reliability of the Alzheimer's Disease-Related Quality of Life (ADRQL) Assessment Instrument. Alzheimer Dis Assoc Disord. 2009;23(3):275–284. PMCID: PMC3086660

6. Peipert JD, Jennings LA, Hays RD, Wenger NS, Keeler E, Reuben DB. A Composite Measure of Caregiver Burden in Dementia: The Dementia Burden Scale-Caregiver. J Am Geriatr Soc. 2018 Sep;66(9):1785–1789. PMID: 30094817

7. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. Applied Psychological Measurement. 1977 Jun 1;1(3):385–401.

8. Cohen S. Perceived Stress Scale [Internet]. Mind Garden, Inc.; 1994 [cited 2016 Feb 29]. Available from: http://www.mindgarden.com/documents/PerceivedStressScale.pdf

9. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): Conceptualizing and Measuring Activation in Patients and Consumers. Health Serv Res. 2004 Aug;39(4 Pt 1):1005–1026. PMCID: PMC1361049

10. Borson S, Mobley P, Fernstrom K, Bingham P, Sadak T, Britt HR. Measuring caregiver activation to identify coaching and support needs: Extending MYLOH to advanced chronic illness. PLoS ONE. 2018;13(10):e0205153. PMCID: PMC6181336

11. Sadak T, Wright J, Borson S. Managing Your Loved One's Health: Development of a New Care Management Measure for Dementia Family Caregivers. J Appl Gerontol. 2018 May;37(5):620–643. PMID: 27384048

12. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992 Jun;30(6):473–483. PMID: 1593914

13. Wan W, Skandari MR, Minc A, Nathan AG, Winn A, Zarei P, O'Grady M, Huang ES. Cost-effectiveness of Continuous Glucose Monitoring for Adults With Type 1 Diabetes Compared With Self-Monitoring of Blood Glucose: The DIAMOND Randomized Trial. Diabetes Care. 2018;41(6):1227–1234. PMCID: PMC5961392

14. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care. 2003 May;41(5):582–592. PMID: 12719681

15. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ. 1999 Sep 11;319(7211):670–674. PMCID: PMC28218

16. ORWH. NIH Policy on Sex as a Biological Variable | ORWH [Internet]. [cited 2017 May 31]. Available from: https://orwh.od.nih.gov/research/sex-gender/nih-policy/

17. Marshall GN, Hays RD. The Patient Satisfaction Questionnaire Short-Form (PSQ-18) [Internet]. RAND. 1994 [cited 2018 Dec 13]. Available from: https://www.rand.org/content/dam/rand/pubs/papers/2006/P7865.pdf

18. Peek ME, Nunez-Smith M, Drum M, Lewis TT. Adapting the everyday discrimination scale to medical settings: Reliability and validity testing in a sample of African American patients. Ethn Dis. 2011;21(4):502–509. PMCID: PMC3350778