

***Agency for
Healthcare Research
and Quality***

***“Design and Evaluation
of Three Administration
on Aging (AoA) Programs:
Chronic Disease Self-
Management Program
Evaluation Design”***

***Final Evaluation Design
Report***

Contract #: HHSA290200710071T

May 16, 2011

Daver Kahvecioglu
Terry Moore
Marios Michaelides
Sarah Ruiz
Rosanna Bertrand

***IMPAQ International, LLC
& Abt Associates, Inc***

Attention: William Spector
AHRQ/CDOM
540 Gaither Road
Rockville, MD 20850
William.Spector@ahrq.hhs.gov

TABLE OF CONTENTS

TABLE OF CONTENTS.....	i
EXECUTIVE SUMMARY	iii
CHAPTER 1: INTRODUCTION	1
A. Description of the Project Task.....	1
B. Description of Efforts to Date	1
CHAPTER 2: BACKGROUND	2
A. Description of the Stanford University CDSMP	2
B. Other CDSMP Models	3
C. AoA's Role in Funding CDSMPs.....	3
CHAPTER 3: INPUTS TO DESIGNING THE CDSMP EVALUATION	5
A. Review of CDSMP Studies, Gray Literature, and Data Sources	5
A.1 Results of Past CDSMP Evaluations	5
A.2 Strategies Used to Evaluate CDSMP Programs	6
A.3 Outcomes of Interest.....	9
A.4 Dual Eligibles and CDSMP	12
B. Conference Calls with CDSMP Representatives	13
C. Data from the CDSMP Technical Assistance Contractor	17
C.1 Evaluation Universe.....	18
C.2 Challenges of Recruiting Adequate Sample Size	18
D. Key Findings from Technical Expert Panel (TEP).....	19
E. Further Exploration Needed	20
E.1 Measuring Program Differences.....	22
CHAPTER 4: EVALUATION DESIGN – OUTCOME EVALUATION.....	24
A. Research Questions	25
B. Evaluation Methodologies	25
C. Required Sample Sizes	26
D. Site Selection and Recruitment.....	29
E. Baseline Data Collection and Participant Tracking System	31
F. Random Assignment	32
G. Follow-Up Data Collection	34
G.1 Survey Procedures.....	34

G.2	Survey Instrument Development	35
G.3	Administrative Data Review	35
G.4	IRB and OMB Clearance.....	36
H.	Data Analyses	37
H.1	Descriptive Analyses of Evaluation Sample.....	37
H.2	Impact Analyses Using Means Differences.....	38
H.3	Impact Analyses Using Regression Models	38
H.4	Sensitivity Analyses.....	39
CHAPTER 5:	TIMELINE	41
References	42
Appendix A:	Estimated Costs for Evaluation Design.....	45
Appendix B:	Final Literature Review with Appendices A-D	
Appendix C:	Stanford Patient Education Chronic Disease Sample Questionnaire	
Appendix D:	CDSMP Site Interview Summary	
Appendix E:	Technical Expert Panel Meeting Summary	

EXECUTIVE SUMMARY

The goal of this project is to develop a research design that estimates the effect of Chronic Disease Self-Management Programs (CDSMP), as administered in Administration on Aging (AoA)-funded settings, on outcomes of participants among AoA's target population, i.e., individuals aged 60 or older. The 24 states which were original CDSMP grantees in 2006 and 2007 are eligible for the evaluation. Public health care expenditure in the Medicare and Medicaid programs is an outcome of particular interest to AoA, because there has been little research examining this outcome for CDSMP participants. AoA seeks the most rigorous evaluation methodology and is therefore interested in the feasibility of a randomized controlled trial (RCT) evaluation of CDSMPs. AoA is also interested in a rigorous alternative design if there are feasibility challenges with the RCT design. We propose Propensity Score Matching (PSM) design with matched controls from Medicare administrative records. An overview of the proposed design elements is provided below. After the evaluation design is finalized, AoA will procure the services of an independent evaluation contractor to carry out a national evaluation of the CDSMP.

Methodology

Experimental Option (RCT): After reviewing a large number of successfully implemented randomized controlled trials, we recommend this approach for the CDSMP evaluation. After applicants give informed consent, they will be randomly assigned to either a treatment or a control group, with equal probability. This method is the gold standard in program evaluation since it is the only method that can produce a comparison group that is similar to the treatment group on observed *and unobserved* characteristics thanks to the randomization. Due to ethical concerns, the control group will be put on a 6-month wait list to receive the program rather than being denied.

At the time of application to the CDSMP, sites will collect from applicants paper-and-pencil baseline data, and a 6-month follow-up phone survey will be conducted by the evaluation contractor. In parallel, Medicare administrative data spanning the 12 months prior to and the 12 months after the CDSMP workshop will be obtained from CMS. This procedure allows estimation of impacts on self-reported and Medicare outcomes at 6 months with the highest rigor. A 12-month follow-up phone survey (again, conducted by the evaluation contractor) is also worth considering because a longer period may be necessary to observe the impacts of CDSMP. However, the 12-month impacts cannot be estimated with as much rigor as the 6-month impacts, because the control group will be allowed to take the CDSMP workshops after 6 months. 12-month impact estimates will be biased downwards. Given the lack of a clean control group, we also recommend a pre-post test to estimate the 12-month impacts. This approach has been adopted frequently in the CDSMP evaluation literature.

In conclusion, the proposed RCT design is the most rigorous design given the ethical limitation of waitlisting the control group for six months only rather a year or two years.

Quasi-Experimental Option (PSM): We are proposing PSM as the next best option after RCT. When RCT is not feasible, PSM is an ideal alternative in situations where there is a rich set of characteristics available for a large population pool from which matches to the treatment group can be selected. Treatment group is the group of people who take the CDSMP workshops. Availability of rich administrative records for Medicare beneficiaries residing in service areas of the CDSMP host sites makes it feasible to implement a rigorous PSM. This method basically provides a mechanism to select a comparison group (administrative control group) by simultaneously matching on a multitude of characteristics of the treatment group including recent Medicare utilization and expenditures, health status, and demographics. Unlike the RCT, however, PSM cannot guarantee matching on unobserved characteristics. The PSM results are valid to the extent that these unobserved characteristics are not important in determining outcomes of interest or that they are captured by the extensive set of observed characteristics.

Similar to the RCT option, Medicare administrative data spanning the 12 months prior to and the 12 months after the CSMP workshop will be obtained for the treatment and comparison groups. An advantage over the RCT option, this option allows rigorous estimation of CDSMP impact on Medicare utilization and expenditures not only at 6 months, but also at 12 months since the comparison group will be contamination-free beyond 6 months. However, unlike the RCT option, the baseline, 6-month, and 12-month follow up surveys will only be administered to the treatment group as it will be very difficult to reach out to an adequate number of administrative controls. Due to lack of survey data for the comparison group, impact on self-reported 6-month outcomes cannot be estimated as rigorously as in the RCT option. Instead, a pre-post test is going to be utilized.

Outcomes and Survey Instruments

Self-reported outcomes include health status, health behavior, self-efficacy, quality of life, and cognitive symptom management. For comparability of results with earlier studies, we recommend that these self-reported outcomes be measured with the survey instruments developed by Stanford University. Claim-based outcomes include Medicare utilization and expenditures (number of hospitalizations, length of hospital stay, number of emergency room visits), and Medicare expenditures in various settings (inpatient, outpatient, physician, home health, skilled nursing facility, durable medical equipment, prescription drug/Part D). Even though CDSMP has the potential to improve outcomes for Medicaid beneficiaries (i.e., the dual-eligible population), it does not seem feasible to include Medicaid outcomes in the study for a number of reasons, including (1) a long lag time in receiving data and (2) inconsistencies across states that make the data difficult and costly to analyze. Although we had discussions with the sites and the technical assistance contractor, we were unable to determine the degree to which AoA grantees are specifically targeting dual-eligible beneficiaries. Furthermore, there were no data available to our team to assess whether dual eligibles constitute a reasonably large group among CDSMP participants. Nevertheless, we recommend keeping the dual eligibles in the study and evaluating the impact of CDSMP on their self-reported and Medicare outcomes if there happens to be an adequate number of them participating in the study. Furthermore,

assuming an adequately large number of dual eligibles, it may also be possible to assess the impact of CDSMP on some components of Medicaid outcomes such co-pays paid by Medicaid for dual eligibles.

Selection of Sites and Sample Size

Due to the decentralized way in which the CDSMP program is implemented, the evaluation activity will be coordinated at the host site level (rather than the state level). March 2011 data from the CDSMP technical assistance contractor indicate that approximately 30,000 people attended the workshop in the prior 12 months. We estimate that evaluation sample frame is 14,500 of these in about 500 host sites. The frame consists of individuals 60 years of age or older, taking the generic CDSMP workshop in the 24 states that are the original grantees. A key finding is that most of the host sites are very small (average is 30 attendees per site) and a few large sites account for many of the workshop participants.

Based on the statistical power calculations we conducted, we recommend a baseline sample of 3,000. In order to recruit this many study participants we estimate that at least 20 of the largest sites will need to be recruited for the study along with a long intake period of 12 months. In addition, based on our analysis of the technical assistance contractor data and calls with sites, we strongly recommend that AoA provides additional funds to the evaluation sites in order for them to sustain and double the number of participants. The biggest challenge for the RCT option is being able to recruit the target sample size. To ensure national representation, up to 30 other sites may be probabilistically selected for a total of up to 50 sites. However, given the potential challenges of recruiting enough participants, it is advisable to include as many large sites as possible.

There are many uncertainties around the circumstances influencing the ability to recruit adequate sample size. It is necessary that the evaluator reassess these circumstances. Since the PSM option significantly reduces the recruitment burden, it may end up being the feasible design option based on this future assessment.

Our power calculations also indicate that an unrealistically large sample size (50,000) would be required to detect statistically significant CDSMP effects on Medicare expenditures. This is due to very high variability in these expenditures. The estimated impact will be the best estimate available, but it will be imprecise.

Evaluation Grants from AoA:

Given the absence of a centralized, robust, and streamlined registration system, and because local staff are overstretched, we strongly recommend that AoA provides funds to help the selected sites develop a uniform registration system. To alleviate concerns about loss of clients to randomization, we also strongly recommend that AoA provides additional funds to help selected sites with marketing, advertising, and recruiting with the intent of doubling the customers.

Evaluation Timeline and Cost:

Both the RCT and PSM evaluations are expected to take four years and approximately \$2.6 million to complete with 50 evaluation sites. For 20 evaluation sites, estimated costs for RCT and PSM are \$2.2 million and \$2.1 million, respectively. Details are provided in Appendix A.

Combination Option (First RCT, then PSM, if necessary):

Considering the desire to conduct an RCT and the uncertainties around its feasibility, an optimal evaluation strategy would be to start with the RCT implementation and to switch to PSM, if necessary. At the midpoint of the 12-month intake period, the evaluator would assess the number of study recruits and will make a determination on the likelihood of reaching the target sample size by the end of the intake period. If the assessment is negative, the evaluator would switch to the PSM design. This switch should not cause any significant delays or significant additional evaluation costs if planned carefully before the evaluation is implemented.

CHAPTER 1: INTRODUCTION

A. Description of the Project Task

Chronic Disease Self-Management Programs (CDSMPs) are designed to empower adults with chronic disorders to better self-manage their conditions and improve their physical and mental health. The best known and most thoroughly studied CDSMP was developed by the Division of Family and Community Medicine, Stanford University School of Medicine, with funding from the Agency for Healthcare Research and Quality (AHRQ) and the State of California Tobacco-Related Diseases office. This lay-led program emphasizes the individual's role in managing illness and building self-efficacy so that he or she can be successful in adopting healthy behaviors. The Administration on Aging (AoA) funds the CDSMP and similar programs through grants to State Units on Aging and Public Health Departments, which disburse funding to local Area Agencies on Aging (AAAs) in many states.

The Stanford University CDSMP is an evidence-based disease prevention model designed to help people with chronic diseases better self-manage their conditions, improve their health status, and reduce their need for more costly medical care. The program consists of 2.5-hour workshops once a week for 6 weeks and is generally administered in community settings such as churches, libraries, YW/MCAs, senior centers, public housing projects, community health centers, and cooperative extension programs. The program is also available online in some areas. Because the program is not disease-specific, people with different chronic health problems attend together. The workshops are facilitated by two leaders, who are trained and certified by Stanford University, one or both of whom are non-health professionals or lay people with chronic diseases themselves. Workshop topics include: (1) appropriate exercise for maintaining and improving strength, flexibility, and endurance; (2) appropriate use of medications; (3) communicating effectively with health professionals; (4) nutrition; and (5) techniques to deal with problems such as frustration, fatigue, pain, and isolation.

The goal of the project task is to design an evaluation that examines the effectiveness of the Stanford University CDSMP, as administered by the AoA, in improving participant health status, health behavior, self-efficacy, quality of life (QOL) outcomes, cognitive symptom management, and reducing health care utilization and costs.

B. Description of Efforts to Date

Our evaluation recommendations will be derived from four sources: a literature review of chronic disease self-management programs; a review of CDSMP program materials provided by AoA and the National Council on Aging (NCOA); telephone interviews with staff of CDSMP grantees (State aging and/or public health departments) in Michigan, Illinois, and North Carolina and with staff of host sites within those states; and written and verbal feedback from a three-member technical expert panel (TEP). Findings from each of these approaches will help clarify conceptual and methodological issues with implications for the design and execution of the evaluation.

CHAPTER 2: BACKGROUND

AoA began funding evidence-based programs in 2003, working with AHRQ, the Centers for Disease Control and Prevention (CDC), the Centers for Medicare and Medicaid (CMS), the Health Resources and Services Administration (HRSA), the Aging Network, and other partners. Recent data from the CDSMP technical assistance provider reported that in the last 12 months, 47 states provided workshops to a little over 30,000 participants (NCOA, personal communication, Jan. 27, 2011). Evidence-Based Disease and Disability Prevention (EBDDP) programs are funded through 24 grants around the country. As part of the funding requirements, the grantee must “implement a Stanford University Chronic Disease Self-Management Program (CDSMP), but also gives each state the option to select another evidence-based program which helps reduce chronic disease in its senior population” (see <http://www.healthyagingprograms.org/content.asp?sectionid=32>). Other EBDDP programs cover such topics as physical activity, arthritis, depression, and diabetes self-management. Their cornerstone is that each program is grounded in evidence and uses various education and implementation models to improve the health of older adults (for more information, see http://www.aoa.gov/AoARoot/AoA_Programs/HPW/Evidence_Based/index.aspx).

A. Description of the Stanford University CDSMP

Several key characteristics of the Stanford CDSMP model intervention make it unique compared to other such EBDDP initiatives. These characteristics include, but are not limited to (1) the use of master and lead trainers, and (2) in-person workshops (versus other modes of training). Stanford CDSMPs are counseled to strictly adhere to specific programs and processes in order to attain and maintain fidelity to the program as originally designed. Some of these fidelity requirements, detailed as “must do’s” in the CDSMP Fidelity Toolkit, http://patienteducation.stanford.edu/licensing/Fidelity_ToolKit2010.pdf, are described below.

Master Trainers. Master Trainers train the workshop leaders, work in pairs, and serve as workshop leaders themselves. In addition, Master Trainers:

- Attend 4.5-day Master Training.
- Facilitate one 4-day Leader training within a year of completing Master Training.
- Lead a full 4-day Leader Training at least once a year to remain certified.

Lead Trainers. The workshop leaders, also referred to as lay leaders, are the people who facilitate the CDSMP workshops. They work in pairs, and commit to the following requirements:

- Attend four 6-hour days of training over 2 weeks and complete two practice teachings during training.
- Commit to facilitating at least one 6-week workshop in the year in which they were trained.
- Must come from the same communities the CDSMP intends to serve.

In-person workshops:

- Are offered to group sizes of 10–16 participants.
- Are offered 2.5 hours a week over 6 weeks.
- Are conducted by two lay leaders who received training from a Master Trainer.

B. Other CDSMP Models

There are other self-management models for chronic disease, including such disease-specific Stanford-based models as the Arthritis Self-Management (Self-Help) Program (<http://patienteducation.stanford.edu/programs/asmp.html>), the Diabetes Self-Management Program (<http://patienteducation.stanford.edu/programs/diabeteseng.html>), and the Expert Patients Programme (EPP). The EPP is a central component of chronic disease management policy in the United Kingdom (Rogers et al., 2008), and is expected to target over 100,000 people in England and Wales by 2012 (Richardson et al., 2008). The EPP also is evidence-based and designed to help people with chronic disease self-manage their conditions, improve their health status, and reduce medical costs. Similar to Stanford's CDSMP, the EPP consists of six weekly workshops conducted in community settings, and is also available as an online tool.

In addition, a number of CDSMP programs based on the Stanford model have made modifications to the program (see the literature review in Appendix B). For example, a program concerned with health disparities targeted older African Americans in Philadelphia (Gitlin, Chernett, et al., 2008). The Stanford CDSMP was “translated” for delivery by a senior center, and the evaluation examined whether participants derived benefits similar to those of middle-class Caucasian clients. Nine program modifications were made: name change to Harvest Health; orientation session one week before start of sessions; use of culturally grounded language for key words; reference to “Black church” in instructor’s manual replaced with “spirituality”; use of culturally appropriate music during aerobic phase of session; serving of healthy snacks and emphasis on avoiding sweets and salt; introduction of a moment of silence at beginning of each session; additional unit on communicating with a health care provider of a different race; and certificate of completion of program. Other studies reported modifications in delivery mode (i.e., telephone or online) or type of instructor (health professional). It is still unclear whether changes to the traditional Stanford CDSMP influence participant outcomes, but approximately half of the studies reviewed included at least one modification.

C. AoA’s Role in Funding CDSMPs

AoA administers the CDSMP grant program and gives funding to Aging and/or Public Health Departments in the states. While Aging and/or Public Health Departments are the state grantees, there are also host and implementation sites in the state. A host site is the organization that oversees program operations (i.e. AAA) and may also manage recruitment and enrollment of participants. An implementation site (also known as a program delivery site)

is where the workshop is conducted (i.e. senior center, YMCA, church). In some locations the host and implementation site are the same place and in others they are not.

AoA has a contract with the National Council on Aging (NCOA) to provide technical assistance to CDSMP grantees (as well as to other EBDDP grantees). Technical assistance includes web-based training, on-site visits, targeted teleconferences, peer-to-peer mentoring, strategies and models for developing statewide CDSMP distribution systems, and strategies to sustain programs beyond the grant cycle (<http://www.healthyagingprograms.org>). NCOA maintains a database on programs, including location, number and characteristics of participants served, and workshops offered.

CHAPTER 3: INPUTS TO DESIGNING THE CDSMP EVALUATION

The design team carried out a number of activities in order to inform the evaluation design. These activities included: A literature review, a review of gray literature and data sources, and conference calls with CDSMP representatives, review of program data from the CDSMP technical assistance contractor, and a conference call with the Technical Expert Panel. Below, we describe these activities along with implications for the evaluation design.

A. Review of CDSMP Studies, Gray Literature, and Data Sources

A literature review of CDSMP evaluation studies was completed in Fall 2010 to inform this research design. The primary sources of literature were PubMed/MEDLINE and EBSCOHost, but also included the reference list of the Centers for Disease Control and Prevention (CDC) meta-analysis and other materials provided by AoA. Articles in three primary areas were reviewed: 1) program evaluation methodology; 2) outcome variables and the tools used to collect and measure them; and 3) program characteristics. Overall, 44 peer-reviewed articles were abstracted. Key findings from this review are summarized below.

A.1 Results of Past CDSMP Evaluations

Overall, the studies reviewed for this report provide evidence supporting the utility of the CDSMP and similar self-management programs in improving self-efficacy, health status, and health behaviors. In addition, while fewer studies investigated the effects on health care utilization, those that did found significant reductions in physician visits and hospital stay duration, suggesting that savings to health care financing programs such as Medicare and Medicaid may be possible. However, there are few subgroup analyses of the population AoA is mandated to serve, that is, people aged 60 or older, despite the group's participation in the randomized controlled trials of the intervention. The few studies that looked at the effects of CDSMP on older individuals either did not find positive effects or found only weak effects. Findings described below are for the general age group.

Though the 2001 study by Lorig and colleagues is one of few to report health and other outcomes for up to two years (most reported just 4-, 6-, or 12-month outcomes), this study did not have a formal control group after the first 6 months of the study. Because, after 6 months, the wait-listed control participants were offered the CDSMP intervention. Smeulders et al. (2009) found that 6-month findings on health behaviors and health care utilization were no longer present at 12 months, suggesting that, to understand the long-term effects of the program, follow-up data collection should be extended beyond 6 months. The ethical and practical considerations regarding how long the CDSMP intervention can be withheld is one that AoA and the national evaluation contractor will have to grapple with. It may be more realistic to follow the Lorig model by offering the CDSMP intervention to control participants after the first 6 months and tracking their outcomes over time.

Studies that tracked health care utilization found statistically significant results after 6 months (e.g., Ahmed & Villagra, 2006), 10 months (Ahmed & Villagra, 2006), 12 months (Ersek et al., 2008; Lorig, Ritter, & Jacquez, 2005; Lorig et al., 2001), and 2 years (Lorig et al., 2001). However, these studies did not use a specific instrument to measure health care utilization. Rather, participants self-reported health care utilization on questionnaires, which were then checked against participants' clinical records. One study relied on claims data to analyze CDSMP effects on utilization.

Some studies reported decreases in health services utilization (e.g., inpatient hospital use) and potentially reduced costs for CDSMP participants; however, these studies have several limitations, including the reliance on participant self-reports. Self-reporting of health care utilization should be viewed with caution because they vary in accuracy. According to Bhandari and Wagner (2006), several factors affect the accuracy of self-reported utilization information, including sample population and cognitive abilities, recall time frame, and type of utilization of interest.

Many studies reviewed and analyzed participant demographics such as age, sex, and race/ethnicity to determine program effects by demographic variable. Though most studies sampled Caucasian and female CDSMP participants, several studies specifically targeted Hispanic participants (two studies) and rural African American older adults (one study). Similarly, most studies reported age as a sample descriptor, rather than age in relation to outcomes.

Findings were mixed with regard to whether disease-specific or generic CDSMPs result in better outcomes for participants. Since AoA funds have mostly been used to implement the generic CDSMP, we recommend that the national evaluation focus on generic CDSMPs.

A.2 Strategies Used to Evaluate CDSMP Programs

Two main methodological techniques have been used to evaluate CDSMP programs: pre-post test design and randomized controlled trials (RCTs). We recommend an RCT design. In this subsection, we provide a description of both approaches, and our rationale for recommending the RCT design.

A.2.1 Randomized Controlled Trial (RCT) Design Approach

An RCT design is the most rigorous approach to evaluating outcomes of CDSMP participants. In this approach, participants are randomized into treatment or control groups. Participants in both groups are surveyed at baseline and follow-up periods. Control group participants are usually placed on wait lists and receive the intervention at a later date. Populations included in these trials have ranged from persons with chronic disease in general to those with specific chronic diseases such as diabetes, stroke, heart disease, or inflammatory bowel disease. Several RCTs have also been conducted on programs similar to or adapted from the CDSMP intervention model. Adaptations were made to accommodate the method of identifying

participant eligibility (e.g., self-report rather than physician diagnosis) or for use as a disease-specific program, such as hypertension.

The advantages of this approach for the evaluation contractor are that RCT will be the most rigorous approach possible and that all participants will eventually have the opportunity to receive program training after a 6-month waiting period. However, the evaluation contractor will need to justify, to a review board, the practice of denying a widely available intervention with significant benefits, and may determine the maximum wait time based on input from stakeholders and program experts. Furthermore, to find statistically significant results, a minimum number of CDSMP applicants should agree to be randomized and waitlisted if assigned to the control group. Willingness of CSMP applicants to accept randomization is unknown. There should also be a robust mechanism to keep track of people assigned to the control group and to prevent them from taking the workshop throughout the specified intervention period. Still, an RCT design approach is an extremely reliable form of scientific evidence in the hierarchy of evidence that influences health care policy and practice, because RCTs minimize spurious causality and bias.

A.2.2 Pre-Post Test Design Approach

Ten of the 25 studies examined for the literature review employed a pre-post design approach. This technique is a non-randomized intervention design, with no control group, that assesses outcome measures at baseline and again at various intervals after the implementation of the self-management program intervention. This type of longitudinal design compares changes in outcomes over time, but it is limited in the inference of causality since the sample is neither randomized nor contains a control group. Rather, participants in the study constitute the intervention group, all of whom are followed before and after the CDSMP. Lack of a robust comparison group is a weakness of this approach.

Each pre-post study tracked participants and administered follow-up questionnaires after the intervention of the CDSMP. While some studies followed up with participants only once, at 4–6 months post-intervention, other studies utilized multiple data collection intervals to evaluate the self-management program—in these cases, data collection occurred at baseline, 4–6 months post-intervention, and 12 months post-intervention. Two reviewed studies tracked the participants up to 2 years following the onset of intervention.

This approach may be considered if an RCT design is determined to be infeasible. This approach avoids the ethical implications of withholding treatment from participants, potential contamination in the control sample, and losing participants who are unwilling to be randomized. It would also allow the evaluation contractor to include sites with extremely small CDSMP programs that have trouble with recruitment. Finally, CDSMPs have program quotas that they must meet and report to AoA in order to maintain funding, and the use of a pre-post design would not pose any additional challenge or difficulty to sites to make their quotas for each reporting period. However, it is likely that AoA will relax quotas to accommodate study

participation. However, despite these advantages, the analytical strengths of the RCT approach still outweigh the flexibility of the pre-post design approach.

A.2.3 Randomization and Sampling Strategies

Randomization occurred after baseline data collection, often using a blinded randomizer strategy. Five of the CDSMP trials used a straightforward wait-list method for control group participants, such that all control group members had the option to enroll in a CDSMP after the study period, usually 6 months. Barlow and colleagues (2009) enhanced their treatment and wait-list control group design by adding a second, non-randomized control group consisting of individuals who explicitly reported disinterest in participating in the CDSMP, for whom baseline and follow-up data were collected and analyzed. This strategy creates a third group that may not traditionally participate in the program, but provides an interesting comparison group to those receiving the intervention. However, given the sample size requirements needed to generate statistically significant results, we do not recommend inclusion of this second control group in the evaluation study.

To ensure that a sufficient number of participants were selected into CDSMP treatment groups, many studies used specific randomization ratios, such as a 3:2 or 2:1 treatment-control ratio. Our recommendation for the evaluation contractor is for a 1:1 treatment-control ratio. We make this recommendation to minimize the sample size required for a given level of statistical precision.

One study investigating outcomes of CDSMP participants across different ethnic groups used a stratified sampling approach before randomizing participants into treatment and control groups. In this design, eligible prospective participants were stratified based on spoken language and geographic area, and then randomization occurred at a 2:1 treatment-control ratio.

There are three sampling strategies for the evaluation contractor to consider. *Strategy 1* poses a minimal disruption to the host site. In this option, the site would obtain consent when the participant shows up for the first workshop. This is potentially infeasible because participants who have made the decision to take the workshop may be unwilling to be randomized once they arrive at the site. Even if participants agree to be randomized and end up in the control group, they may change their mind or take another workshop. This places a burden on the host site to convince participants to join the evaluation, to track those interested in the program, and to monitor whether the control group is contaminated.

Strategies 2 and 3 require the evaluation to make a significant contribution to the infrastructure development of the CDSM program. In *Strategy 2*, AoA would provide additional funds to participating host sites for an enrollment system that collects baseline information and obtains consent before the participant arrives at the workshop. In this approach, there would be a standard tracking and evaluation system across sites. *Strategy 3* builds on Strategy 2 by

increasing AoA funding also for recruitment activity and by incorporating an advanced marketing strategy, which may address the issue of sample size for the evaluation and insufficient number of participants for the host site. The reviewed literature includes several strategies for recruitment of participants, such as advertisements in doctors' offices, senior centers, and religious institutions, as well as public and radio service announcements.

The evaluation contractor should also take into account statewide enrollment systems. A statewide enrollment system occurs when there is a central location, such as a website or phone number, where states can enroll or recruit participants for the program. If a site is selected for participation in the evaluation from a state with a statewide enrollment system, the evaluation contractor should consider revising the recruitment strategy for the site and consider where baseline information is collected.

A.3 Outcomes of Interest

We recommend that the evaluation contractor study health status, health behavior, self-efficacy, quality of life (QOL) outcomes, including social/role activity limitations and psychological well-being, cognitive symptom management, and health care utilization and costs.

The literature reviewed by the design team demonstrates that investigators agree overwhelmingly that self-efficacy is a critical concept to measure when evaluating a CDSMP program; scales developed by Lorig and colleagues (1999) have been widely used to measure this outcome. Self-efficacy theory, as developed by Albert Bandura, states that high-level self-efficacy is a prerequisite to realizing self-management goals, as well as critical in determining whether individuals will maintain or improve their health status (Bandura, 1989 as cited in Du & Yuan, 2010). For this reason, a participant's belief in his or her ability to manage the condition can act as a predictor of health outcomes.

Measures of health status and health behaviors are also common, with studies looking specifically at self-rated health, degree of pain and discomfort, role limitations, time spent engaging in exercise and other indicators through the use of validated scales. One study included the validation of a measurement developed specifically for self-management program outcome evaluation—the Health Education Impact Questionnaire (HeiQ), which was shown to have strong psychometric properties including validity (Osborne, Elsworth, & Whitfield, 2007).

To allow AoA to compare the findings of the national evaluation to previous CDSMP studies and to provide cost estimates for AoA's use, we recommend that the health, QOL, and self-efficacy measures be derived using instruments and methods used by Stanford University (see the literature review in Appendix B as well as <http://patienteducation.stanford.edu/research>).

The Stanford Patient Education Research Center has developed a Chronic Disease Sample Questionnaire (included in Appendix C), which is intended to be used as a mail survey to CDSMP participants. The survey questions above cover most of the course topics. The questionnaire includes the following scales, most of which have good psychometric properties:

- Health status indicators: general self-rated health status; health distress; and measures of fatigue, shortness of breath, and pain
- Health behaviors: exercise scale
- QOL: social/role activity limitations
- Self-efficacy: 6-item chronic disease self-efficacy scale

The questionnaire does not include measures of cognitive symptom management or disability status, which we would recommend be included in the national CDSMP survey. Item scales for each of these may be found on the Stanford Patient Education Research Center website.

Health care utilization has also been examined as an outcome, primarily through the use of self-reported visits to physicians and emergency rooms, and the number and duration of hospital stays. One study relied on analysis of claims data from participants enrolled in a managed care organization (Ahmed & Villagra, 2006). In addition, the cost-effectiveness of self-management programs has been assessed to demonstrate potential reductions in health care utilization (Kennedy et al., 2007; Richardson et al., 2008). We recommend the evaluation contractor study health care utilization outcomes including physician visits, emergency department visits, and hospitalizations; and health care expenditures including costs to Medicare, Medicaid (if feasible, see below), and out-of-pocket costs. Since AoA's target population is older adults (60 years or older), we recommend that Medicare (and Medicaid, if feasible) data obtained for individuals 65 years of age or older. Medicaid enrollees of this age group will be dual eligibles.

Although the Stanford Chronic Disease Sample Questionnaire includes a four-item health care utilization scale, the design team recommends that utilization be measured via health care claims. Medicare and perhaps also Medicaid claims data will allow both utilization and costs for fee-for-service beneficiaries¹ to be measured together, with perhaps more accuracy than the self-reported measures used in prior CDSMP studies.

To construct a longitudinal data set of treatment and control participants that links self-reported health status and behavior indicators to utilization and expenditures, it will be necessary to obtain, from each study participant, identifying information beyond name and date of birth (which are items included in the Stanford Chronic Disease Sample Questionnaire). We would suggest that the CDSMP enrollment application and pre-consent process obtain name, social security number (SSN), Medicare health insurance claim number (HICN), Medicaid enrollee identifiers if available (SSN or state-assigned MSIS-ID), gender, and date of birth to ensure that adequate identifiers are available for linking individuals across time and data type (e.g., claim, survey). Medicare and other health records should be collected at least for the period 12-months prior to enrollment, as well as 6 and 12-months post enrollment. Details on data collection and data analysis methods are described below in Chapter 5.

¹ Medicare and Medicaid managed care enrollees will need to be excluded from the utilization and cost analyses due to data limitations.

A.3.1 Data Limitations

As described earlier, CDSMPs are open to older adults with chronic conditions. According to NCOA, 27 percent of current CDSMP participants are below the age of 60 (NCOA, demographic report prepared in March 2011). Though self-reported health status data can be obtained and evaluated for this cohort of participants, the utilization and cost analyses will likely be limited to the Medicare, and perhaps also Medicaid, beneficiaries. With sufficient personal identifiers, the evaluation contractor can develop a longitudinal data set to track participation, health status, and physician and hospital utilization and related costs. This is particularly feasible for Medicare beneficiaries, but less so for dual-eligible Medicaid enrollees, for the reasons described below.

Several key parameters must be met in order to evaluate health-related utilization and costs for CDSMP treatment and control participants: (1) individual participants must be identifiable and linkable across time and datasets; (2) claims data must be available for the same time periods as the randomized assignment study; (3) privacy-protected data housed by CMS must be made available; and (4) sufficient time and resources must be applied to create and analyze the longitudinal data file. Although we are confident that a file can be constructed linking the treatment and control group participants to Medicare claims for the study period (presuming Medicare enrollment data are available for accurate linking), we are less confident that Medicaid claims can be linked in with sufficient accuracy and timeliness. Following are the currently available Medicaid data:

- Medicaid Statistical Information System (MSIS) data are submitted by each state's Medicaid program and contain Medicaid enrollment and claims paid information. They can be obtained from CMS by fiscal quarter. However, each state has a different manner of identifying eligible beneficiaries and services reported, and how they are labeled in the MSIS files may vary by state. The state-specific anomalies (or nuances) limit the cost-effectiveness of their use and thus limit the utility of these data for the national CDSMP evaluation. Furthermore, there appears to be at least two years time lag between the Medicaid covered service and data availability.
- Medicaid Analytic eXtract (MAX) data are built from the MSIS and reflect the services used by Medicaid enrollees during a calendar year. The most recently available data are for calendar year 2008. The MAX data are desirable for this CDSMP evaluation, as they are analytic extracts that enable analyses of enrollment, utilization, and expenditures at the person level (Wenzlow, Schmitz, & Shepperson, 2008). However, the files require time for MSIS data to be validated and for the files to be built, cleaned, and made available; thus, the time lag may be too great for MAX data to be included in these analyses.

The long time lag makes it infeasible to utilize either MSIS or MAX in a study that is less than 6 years long. Another way to obtain Medicaid data – without long delays – is to identify a small number of states that are willing to directly provide Medicaid data. Due to significant inconsistencies across states, we recommend that only two states are identified. Note that these states will have to have robust CDSMP outreach to Medicaid enrollees and will have to

have significant number of Medicaid enrollees taking CDSMP workshops. Through environmental scan and focus groups, the evaluator is recommended to assess states for robustness of their CDSMP's Medicaid focus, if any, and also get buy in from states to make their Medicaid data available. The design team recommends that the evaluation contractor weighs the strengths and potential limitations of various Medicaid data acquisition strategies to ultimately decide whether and how the evaluation should include analyses of Medicaid utilization and costs as a CDSMP outcome. We also recommend that some items regarding out-of-pocket costs to individual CDSMP participants be added to the self-reported survey questionnaire.

A.4 Dual Eligibles and CDSMP

Even though Medicaid data may not be feasible to utilize in the evaluation, there may still be ways, albeit limited, to assess the impact of CDSMP on some aspects of Medicaid through availability of Medicare data on dual eligibles. Dual eligibles include 9 million low-income elderly and disabled Medicare beneficiaries who qualify for coverage based on their low income. Dual eligibles account for 18% of Medicaid enrollees but 46% of Medicaid spending. The dual eligible population is also growing. Medicaid coverage rates for the community among the over 65 population increased from 7.6 percent in 1987 to 14.1 percent in 1996. The management of chronic conditions in this group is likely to result in substantial savings. Furthermore, health care reform (Affordable Care Act, 2010) stipulated a number of initiatives related to provision of CDSMP to dual eligibles. Medicaid covers important services and co-pays that Medicare limits or does not cover, such as long-term care. Copays that would be paid by Medicaid for doctor visits and hospitalizations are recorded in Medicare data and that may facilitate a limited evaluation of the impact of CDSMP on Medicaid. The evaluation contractor should research and assess such opportunities.

After determining the potential of the CDSMP program to reduce Medicaid costs among dual eligibles, we considered whether and to what extent states involve this population in the CDSMP. Our research including conference calls with a number of CDSMP grantees and analyzing CDSMP technical assistance data suggested that dual eligibles may not be robustly targeted yet. While it is unclear what proportion of CDSMP participants are Medicaid beneficiaries, there may be more data on this population in the near future. American Recovery and Reinvestment Act (ARRA) funding for CDSMPs requires that the State Medicaid Agency is involved in the development and implementation of the program (AoA, 2010). The new CDSMP programs are required to give special attention to serving low-income, minority and limited English speaking older adults, including Medicaid eligible individuals.

An online search of CDSMP programs and mention of the Medicaid population, found that states such as New York are beginning, while others (Maryland) intend, to include the population in later years of the program (NACDD, 2010). With their state data indicating that 5% of Medicaid chronic care population accounts for 50% of the Medicaid health care expenses (Goehring, 2010), Washington State now offers reimbursement for diabetes SMP and aims to provide CDSMP reimbursement for Medicaid.

According to a recent CDC brief, a few states are moving toward Medicaid reimbursement for CDSMP (Gordon & Galloway, 2010; www.healthyagingprograms.com). While this has been occurring on a relatively small scale to date, the brief reports that one state has Medicaid clinics specializing in asthma and diabetes, and these patients receive referrals to CDSMP programs. Another strategy has been to train Medicaid managers to run CDSMP programs within their clinics. A Partners in Care Foundation conference in 2010 argued from a Social Enterprise Reimbursement Model that once Medicaid accepts CDSMPs as a reimbursable benefit, they can cover the benefit under the Medicaid Waiver program. The state of Washington amended their Aged/Disabled Waiver to include provision of CDSMPs and California is pursuing a similar strategy. Through Oregon's "Living Well" program, CDSMP coaches refer Medicaid Fee-for-Service clients to programs in their area.

We conclude that the evaluation contractor should look at Medicaid outcomes if feasible. In the near future – when the evaluation contract is awarded – it is expected that states will be more heavily engaging the dual eligibles. Even though there does not yet seem to be wide-spread (i.e., nationwide) engagement of dual eligibles with CDSMP yet, there potentially are a few states which are far along. The evaluation contractor should identify such states and explain to them benefits of the evaluation and try to get their cooperation to provide Medicaid data directly to the evaluator. At the very least, the evaluator should identify viable ways of indirectly looking at Medicaid outcomes through Medicare data for dual eligibles.

B. Conference Calls with CDSMP Representatives

The research team facilitated conference calls in early March 2011 with "on-the-ground" CDSMP program representatives to gain a better understanding of the program's organizational and financial structure, participant tracking and data management systems, and marketing efforts. AoA program staff identified three target states (Illinois, Michigan, and North Carolina), and contacted potential respondents in those states to brief them about the purpose of the calls and to request their participation. In total, 11 respondents participated in six conference calls that included representatives of state grantee agencies and their partner agencies, and representatives from host sites such as coordination directors, master trainers, and workshop leaders. A fuller summary of the interviews is presented in Appendix D.

The grantee in all three participating states is either the State Department on Aging or the Department of Public Health; however, these two agencies work closely to administer the CDSMP. The primary CDSMP funding in all three states currently comes from ARRA grants supplemented by Title III-D awards, which are made directly to the AAAs. In addition, state-specific sources of support include local funds, in-kind contributions (e.g., staff time), funding from AoA's Sustaining Evidence-Based Health Promotion Programs, and support from organizations such as the CDC and diabetes and arthritis associations. The program is not supported in any way by clients' contributions: there are no client cost-sharing plans or fees to participate in the program. Concern was frequently expressed regarding future funding and consequently the sustainability of the program. All states expressed the need for additional

funding to expand recruitment, including a focus on gaining provider buy-in to increase referrals to the program. In general, the respondents believe there is a need to reach a critical mass of programs and regions in order to develop a stronger cross-referral system.

The AAAs serve as the primary CDSMP host sites and receive funding directly from the State CDSMP grantees. The amount of funding that an AAA receives is based on the projected number of workshops and participants expected to be served. Some AAAs also use Title III-D funding to support the CDSMP, although they are not required to do so by the grant.

The host site is responsible for all aspects of program coordination including leader training, recruitment, materials, and coordinating workshops. In some states, the AAA subcontracts to other organizations to serve as additional coordination sites. For example, in Illinois, Rush University, White Crane Wellness Center, University of Illinois, and the Affordable Assisted Living Coalition are subcontractors to the AAA and are responsible for conducting trainings and/or hosting workshops in their areas. A variety of entities serve as implementation sites including senior centers, home health agencies, libraries, hospitals, clinics, and some AAAs. Given that the host site is the lowest common entity managing the program, we propose host site as the evaluation site. AoA and the evaluator will recruit host sites for study participation possibly through respective states. Then, the evaluation contractor will work with the recruited AAAs to set up and run the evaluation at local implementation sites.

Reaching workshop capacity can be relatively easy or extremely difficult, depending on the location of the workshop, and the season of the year in which it is held. For example, rural AAA catchment areas tend to cover wider geographic areas than urban catchment areas; it is therefore more difficult to recruit participants, as well as workshop leaders, because they have to travel longer distances to reach an implementation site. Moreover, compounding the difficulty of older adults and persons with disabilities to travel to a workshop site is the fact that services, such as transportation, often are limited in rural areas. In addition, it is also difficult to recruit and sustain a full class during the winter months, when many older adults and persons with disabilities find it difficult to travel. Furthermore, the sites described their older adults as “snow birds,” who travel south for the winter.

These findings have direct implications for the evaluation design and should be taken into consideration by the evaluation contractor when planning the study’s sampling strategy. To maximize the sample size, oversampling from urban areas or omitting extremely rural catchment areas (e.g., frontier catchment areas) from the sampling frame should be considered. In addition, planning around the winter months should be taken into consideration, although this will be a challenge since the proposed evaluation plan suggests a 6-month treatment/control group study, thus encompassing all 12 months of the year. If the time frame allows, one suggestion would be to include only participants who do not anticipate traveling south during the winter, and to start the study in the early spring so that the control group would be starting in the fall, before the harshest winter months prohibit safe travel. Another method for increasing sample size is to expand program recruitment both directly through face-to-face presentations, the distribution of flyers and brochures, and television and

radio advertisements; and, indirectly by obtaining the buy-in of providers as a way to promote referral to the program. As previously discussed, all states included in the conference calls reported the need to expand their recruitment efforts, but noted that additional funding would be necessary to do so. To ensure an adequate sample size, AoA should consider funding additional recruitment efforts at selected study sites.

Enrollees have to register for the workshops, but there is no application process or eligibility screening for enrolling in the CDSMP; all potential participants are eligible to enroll. At some host sites, when participants call to register, they are given detailed information about the program, including the expectation of their attendance at six workshop sessions. This process allows potential enrollees to self-screen in or out of the program before registering or attending any of the sessions. Interviewees described a higher participant retention rate in catchment areas where this process has been implemented. However, the process of informing potential enrollees about the program has not been implemented consistently, even within host sites.

This finding also has implications for the evaluation design. In order to assign participants to treatment or control groups, there has to be a point, prior to attending the first session, at which all potential participants contact the host (or implementation) site. It is feasible that at study sites, all potential enrollees would be processed through the host site (regardless of their first point of contact). To facilitate this, it may be necessary for AoA to provide additional funding so that study sites can conduct a systematic eligibility screening for the purpose of recruiting participants and assigning them to treatment or control groups. We propose developing a “Participant Tracking System” (PTS) for the host sites to facilitate, among other things, uniform enrollment across sites, randomization, and tracking participation of enrollees and non-enrollees. PTS is described in detail in Chapter 5.

The states that participated in the design team interviews follow the AoA mandate that defines a CDSMP “completer” as an enrollee who attends four of the six workshop sessions. Overall, the average program completion rate of 75 percent (range = 72–77 percent) across the three participating states was relatively high, considering that the program serves a chronic disease population. Several reasons were given by participants for dropping out of the program including health-related issues, inclement weather, transportation barriers, and lack of initial understanding that the program is a full six weeks, not a drop-in series of classes. As noted above, systematically administering an eligibility screening prior to enrollment may help to decrease the dropout rate.

On occasion, potential CDSMP enrollees have had to be put on wait lists for the next available workshop. This generally occurs in large urban areas, but for the reasons discussed above, it can also occur in rural areas. There are two primary reasons for generating a wait list. First, on rare occasions, it is difficult to engage a workshop leader. In rural areas, the distance that they have to travel to the implementation site at least once a week for six weeks can be a difficult challenge to overcome. In urban areas, they may be stretched too thin; that is, there are not enough leaders for the number of workshops necessary to serve the number of enrollees. Second, the number of clients who enroll at one time may be too many for one workshop, but

not enough for a second workshop to begin concurrently. As noted previously, filling classes can be very easy or extremely difficult, and getting participants to stay committed is also a challenge. At least one state reported that the wait for program participation can be 6 months or longer.

This finding is very useful for the evaluation design since we have proposed that the design include a 6-month wait-list control group. Most of the host sites in the states that participated in the design team interviews have found it necessary to utilize a wait list on occasion, and most respondents agreed that 6 months would not be an unreasonable amount of time for potential enrollees to wait for the next available program. However, one respondent was concerned that given the chronic health condition of CDSMP enrollees, it is likely that a number of participants would be lost over the 6-month waiting period.

In all states, participants are tracked for attendance, and leaders are monitored for program fidelity. Host sites send enrollment figures and attendance records to the state so that the number of participant completions and the number of workshops held can be tracked. During the first workshop session attended, all enrollees are required to fill out the standardized NCOA enrollment form, which requests demographic information including gender, ethnic background, and geographic region, as well as information on enrollees' chronic condition(s). Enrollees also fill out a brief exit form that asks which, if any, of the tools that they obtained during the workshop they expect to use in the future. Enrollment and exit data are collected via paper and pencil, and forwarded (depending on the state) to the state grantee agency, the host site, or the program evaluator, where they are electronically entered into the NCOA database. In all states, hard copy data are retained for a minimum of 1 year, and in one state the data are entered into an Access database and retained by the state. All data are reflected in a national database maintained by a Seattle-based subcontractor of NCOA. The NCOA database is de-identified.

Since workshop leaders already administer data collection forms to all program enrollees, it is possible to request that they also administer the national evaluation survey prior to beginning the first workshop. However, since many workshop leaders are volunteers or serve as a leader as part of their agency position, asking them to collect these additional data may be too burdensome. Another option is to ask workshop leaders to distribute packets that contain a fact sheet about the study, an informed consent form, and the survey. Time might be allotted at the start of the first session for enrollees to complete the pre-implementation survey. The third option is mail the packets to participants who were screened for eligibility during registration, as discussed above.

None of the states interviewed have an online CDSMP, but they indicated an interest in piloting one if additional funds were made available. The evaluation contractor might consider a small sub-study that tests the validity of the online CDSMP. Provided that enrollees have access to a computer, several of the issues that result in their dropping out of the program would be alleviated through the online program, such as some health-related issues, inclement weather,

transportation barriers, and the unwillingness to attend at least four of the six sessions in person.

Lastly, interviews with state and local CDSMP representatives confirmed our gray literature findings on the extent of states' current outreach to Medicaid enrollees. None of the interviewees tracked or knew the proportion of Medicaid beneficiaries participating in CDSMP. However, some of them mentioned current and planned initiatives such as involving Medicaid partners, recruiting Medicaid enrollees, and trying to integrate CDSMP into Medicaid.

C. Data from the CDSMP Technical Assistance Contractor

Recommendations in this evaluation design are informed from conference calls and data from the Technical Assistance (TA) contractor, the National Council on Aging (NCOA). NCOA has a cooperative agreement with AoA to provide TA to CDSMP grantees and other evidence based programs. NCOA administers a grantee data collection platform via www.salesforce.com and uses www.healthyagingprograms.org as a document sharing tool. The IMPAQ team and AoA representatives spoke with NCOA to understand the intake process and data collection procedure for CDSMP participants. After an initial call, the IMPAQ team also participated in a webinar for the CDSMP data collection tool. Following the webinar, a list of requests for NCOA data and other supporting documents was prepared, and they provided valuable information for the evaluation design. This section provides a brief overview of key NCOA data documents and demographic information on states and participants.

During communication with NCOA Representatives, we received and reviewed the following documents:

- *Workshop Forms.* These include the attendance log, participant information survey, and the workshop information cover sheet.
- *CDSMP Demographic Report.* This spreadsheet includes national data on the size of the program and breakdown of participants by key demographics.
- *CDSMP Grantee Reach.* This document details by state the number of workshops hosted, number of participants enrolled, and number of participants who completed the program.
- *Demographic Reports for the 24 original AoA grantees.* Similar to the CDSMP Demographic Report, these reports include state level data on the CDSMP program and demographic characteristics of participants.

To understand the national universe of CDSMP participants, there are several useful demographic statistics. For example, 27% of participants are under the age of 60, and the two leading chronic conditions are arthritis and hypertension. The program is disproportionately female (78%) and Caucasian (69%), and 17% report Latino ethnicity. Only 2% of all participants have taken the course previously.

C.1 Evaluation Universe

According to recent data from the CDSMP technical assistance provider, in the last 12 months, 47 states provided CDSMP workshops to a little over 30,000 participants, with a workshop completion rate of 74 percent. Of these workshops, 81 percent were for the English version of the generic CDSMP, 8 percent were for the Spanish version of the generic CDSMP, and the remaining workshops were for various disease-specific CDSMPs. 27 percent of participants were under 60 years of age. AoA is not the sole funder of this program; the Centers for Disease Control and Prevention (CDC) is also a funder of CDSMPs. Our team proposes the following inclusion and exclusion criteria:

States: AoA is considering for evaluation only the States that are the original 24 AoA grantees (i.e., the states that received AoA funding in 2006 or 2007). These original grantees conducted about 75 percent of CDSMP workshops nationwide in the past 12 months.

Type of workshop: We recommend limiting the evaluation to the generic CDSMP (both English and Spanish versions), which constitutes about 88 percent of all workshops.

Participants: We recommend limiting the population to persons aged 60 and older. AoA's target population is adults aged 60 or older. Approximately 70 percent of all participants fall into this category. Dual eligible are also included in this category.

Source of funding: We recommend that the evaluation be limited to host sites that received primarily AoA funding. Some of the host sites are primarily funded by sources other than AoA (e.g., CDC). We were unable to obtain data on the proportion of participants whose workshops were funded by non-AoA sources. The technical assistance contractor has indicated that sites with large proportions of younger participants are more likely to be funded by CDC.

Population subject to the evaluation: After taking into account the inclusion and exclusion restrictions listed above, we estimate that a maximum of 14,500 individuals constitute the evaluation universe.

C.2 Challenges of Recruiting Adequate Sample Size

TA contractor data indicate that distribution of workshops over host sites is sparse and heavily skewed. Our team obtained host-level enrollment data for some of the largest eight grantees (California, Florida, Illinois, Michigan, North Carolina, Oregon, New Jersey, and New York) out of 24 states subject to the evaluation. The data show that half of the participants belong to only 27 host sites out of 172 total host sites. Most of the sites are very small: 102 sites are expected to provide less than 50 study participants in a random assignment evaluation.² In order to be

² Key assumptions are: 1) Intake period is 12 months, 2) 30% of workshop participants will agree to participate in the study and thus agree to be potentially denied the workshop with 50% chance, and 3) 10% of workshop

able to recruit an evaluation sample of 3,000, we estimate that the largest 85 sites will have to be included in the evaluation, which is logistically quite challenging. Furthermore, Recovery Act funding will expire in March 2012 causing more challenges to recruit adequate number of subjects from a reasonable number of host sites. These observations form the basis of our recommendations that 1) At least the largest 20 sites should be picked as evaluation sites to ensure adequate sample size 2) AoA provides additional funds to evaluation sites to maintain *and* double enrollment in CDSMP. Even though including only largest sites will weaken the representativeness of the evaluation, this seems to be necessary choice to have a feasible study. The evaluation contractor should explore other options as they may become more feasible due changes in policy, funding, and implementation.

As part of an environmental scan task, the evaluation contractor is strongly recommended to reconsider the key assumptions and reassess the feasibility of obtaining adequate sample size. It may also be necessary to interview some of the sites to get a sense of their projected enrollments during the intake.

In light of these potential challenges in recruiting evaluation subjects, we are also proposing an alternative design (Propensity Score Matching) which does not require as many evaluation subjects and does not involve random assignment. This alternative design is described in Chapter 5.

D. Key Findings from Technical Expert Panel (TEP)

On March 30, 2011, the IMPAQ team, in collaboration with AHRQ and AoA, held a Technical Expert Panel (TEP) meeting with three CDSMP experts. Generally, the experts thought that the evaluation design was feasible, and offered suggestions to improve the report, namely to provide greater detail in the approach to key tasks and to offer options for the evaluation contractor. This section will provide an overview of the TEP meeting discussions. The minutes of the TEP meeting are presented in Appendix E.

The TEP made the following suggested revisions to the Design Report:

- Despite the limitations and difficulty of using Medicaid data, TEP members thought that this should be included as an option in the evaluation. They stated that state governments are particularly interested in the impact of CDSMPs on Medicaid costs (for dual eligibles) and would be willing to comply, when possible. There are a number of new Medicaid initiatives funded by Affordable Care Act that potentially incorporate CDSMPs: Health Homes (Section 2703), Medicaid Incentives for Prevention of Chronic Diseases Program (Section 4108), and State Demonstrations to Integrate Care for Dual Eligible Individuals (Section 2602). It was recommended that, at the very least, the evaluator should identify and track dual eligibles. Medicare administrative data on dual eligibles may allow estimation of some components of Medicaid expenditures on dual

participants will not be eligible to participate in the study (repeat customers, have dementia, etc.). 4) All of the host sites will be willing to participate in the evaluation.

eligibles. The evaluator should assess any such relationships in order to try to estimate the impact of CDSMP on some Medicaid aspects.

- The evaluation should expand site selection to reflect the diversity of the CDSM program. All TEP members felt strongly that site selection should include sites that represent the range of experience in settings (e.g., sites from rural areas or frontier states) and sites featuring the Spanish-language version of the program, “Tomando Control de Su Salud.” TEP members thought that inclusion of both types of sites was necessary to capture the diversity of the CDSMP.
- In addition to the recommended control variables, TEP members encouraged the collection of other control variables, such as participant insurance status, receipt of Older Americans Act (OAA) services, and language of instruction for the course.
- The TEP also recommended that more details be given on exclusion factors for participation in the study, especially if the participant previously took the course (course repeater); was a proxy for a family member or friend with a chronic condition (proxy); or scored low on a cognitive screen for the study.
- A final TEP recommendation was that the evaluation contractor should consider the inclusion of “class zero,” or an introductory session, where participants can learn about the program and the evaluation contractor can collect baseline data.

The TEP also raised some high-level evaluation issues. One TEP member suggested expanding the study beyond the original 24 grantees to increase sample size potential. In addition, AoA and the TEP members thought it was necessary to detail an alternative plan for when ARRA funding ends. The design contract has been revised to include an alternative plan in the event that sites are unable to provide the number of participants necessary to derive statistically significant results.

All of above suggestions were considered and integrated into revisions of the design report. In particular, to address TEP comments, we expanded the section on the analysis of participant characteristics in Chapter 5. Within each chapter, revisions have been made to address key points raised by the TEP.

E. Further Exploration Needed

Although much information was obtained in the interviews with CDSMP representatives and the review of the published, peer-reviewed literature, as described above, questions remain for the evaluation design that may be best answered through an environmental scan. For example, in the interval between the current design study and the national CDSMP evaluation, policy, programmatic, or funding changes may influence how CDSMPs are administered or implemented. Another important area is to reassess the feasibility of the design and make any changes if necessary. This step is critical ensuring a viable design in light of uncertainties in future CDSMP funding. Furthermore, there are potentially sweeping changes initiated through the healthcare reform some of which have evidence-based interventions such as CDSMPs at

their center. Discussions with the policy makers and stakeholders may lead to revised or added research questions. The goal is to fully utilize this great opportunity of rigorously evaluating a key program and to ensure participation of various stakeholders who are utilizing CDSMPs as a way to change the way health care is delivered. Particular attention must be given to related concurrent policies that are being implemented. The evaluation contractor needs to understand and incorporate any such changes into the final evaluation design.

Other areas for which additional information could be obtained from an environmental scan include:

- Reassessing the feasibility of design based on fresh CDSMP program data and new policy environment
- Identifying Medicare outcomes that would be proxies to Medicaid outcomes
- Assessing the extent of States' outreach to Medicaid enrollees; identifying a number of states which has robust CDSMP outreach to Medicaid beneficiaries; among these states identifying/recruiting a few states interested in sharing Medicaid data,
- Conducting focus groups to facilitate decision making on design-related issues such as Medicaid-related issues and feasibility of implementing the RCT. It will also provide valuable input for the evaluation design.
- Identifying other evidence-based educational initiatives that could introduce bias or contamination to the site selection process (e.g., CMS-funded Quality Improvement Organization pilot study to test the feasibility of adding a CDSMP benefit to Medicare, and the Texas A&M/Stanford study to replicate the original Lorig studies).
- Identifying appropriate instruments with good psychometric properties appropriate for use in the elderly to capture self-reported health status, health behavior, and self-efficacy.
- Identifying sites with statewide enrollment systems to understand these CDSMP programs and their data or enrollment systems.
- Identify the number of sites utilizing a "class zero," or an information class about the program.
- Assessing the feasibility of sites performing a cognitive screen (i.e. Mini Mental State Examination) at time of enrollment into the course. While sites do not currently perform this screen, the evaluation contractor should consider excluding participants with cognitive impairment.

Another area of activity is helping CDSMP grantees which are participating in the evaluation develop a uniform application system. Activities include:

- Work with AoA, the sites, and possibly with the TA contractor to design a uniform registration and baseline data collection system at each evaluation site. This system is supported by a Participant Tracking System (see in Chapter 5).

- Reviewing and understanding current practices at each evaluation site.
- Developing a system into which each site with potentially different current practices will be integrated.

E.1 Measuring Program Differences

Given the diversity of the CDSM program, we recommend the evaluation contractor consider the below factors deemed to influence program success. This section identifies four domains that contribute to the success of the program and potentially impact participant outcomes. After a review of key organizational and resource factors, we suggest additional staffing and participation factors which arose during site interviews and discussions with the TA contractor.

The evaluation contractor should capture the following four domains at the host site level through a web-based survey:

- Organization and Resource Factors
 1. Organizational structure: Who is the individual or agency responsible for running the program? Who holds the license?
 2. Financing structure and strategy: What funding and additional resources are available? Is there funding instability?
- Staffing and Participation Factors
 3. Experience of facilitator or leader
 4. Number of course completers & repeaters

Currently, research is underway to explore how program factors influence CDSMP participant outcomes (Lorig, personal communication, Sept. 7, 2010). Key domains of interest include organizational structures and financial structures and strategies. The organizational or reporting structure is an important consideration, with variability in the agency or individual running the CDSMP. For example, the organization that supervises and administers the program ranges from small private organizations to large organizations with a longer history of with CDSMP programs. The organizational structure is also influenced by who holds the license and is legally responsible for the course. It is hypothesized that program success may be influenced by the type of individual or agency administering the program. Secondly, resources, such as financial and other resources available, are likely to create differences across programs. A single site may receive funding from four or five organizations, each with a potentially different interest in the program. It is expected that sites with greater funding will produce higher quality courses and attract more experienced facilitators/leaders. An indicator of funding instability should also be included in the web survey to capture programs that will lose their funding in the near future. We expect that these organizational and resource factors create disparities across sites and influence participant compliance and program success.

Staffing and participation are a second group of domains to include in the web-based survey. We expect that facilitators or leaders with stronger training, longer experience, and similar background to participants will produce programs that have a greater and longer influence on individual outcomes. Variables of interest include years of experience, number of previous CDSM programs taught, educational background, and fluency in other languages. We also recommend the web-survey collect information about course completers and repeaters. The number of course completers serve as a proxy for site strength and the number of course repeaters measures whether participants at each site have previously taken a CDSMP course. We also suggest the evaluation contractor analyze the concentration of course completers and repeaters across sites to better understand high performing sites. We expect that individual-level experience will be influenced by the domains described in this section. Given the changing policy environment and dynamic nature of the program, we suggest that the evaluation contractor consider additional program factors at the time of the evaluation.

CHAPTER 4: EVALUATION DESIGN – OUTCOME EVALUATION

The key objective of this project is to design the most rigorous yet feasible evaluation of the impact of the CDSMP on the health status, health behavior, self-efficacy, quality of life, cognitive symptom management, and health care utilization and expenditures of individuals with chronic conditions. To address key research questions about the effectiveness of the CDSMP, the IMPAQ/Abt team proposes the implementation of a random assignment evaluation (i.e., experimental, Randomized Controlled Trial - RCT) design. This design, which is widely regarded as the gold standard for evaluating program impacts, will provide a comprehensive and rigorous assessment of the efficacy of CDSMP in improving the outcomes of program participants. Even though the experimental design is the gold standard, we cannot say with certainty that it will be feasible to implement at the time of evaluation. Due to uncertainties in future availability of funds for CDSMP³ and in applicants' willingness to accept being put on a waiting list for CDSMP workshops, we are also proposing a quasi-experimental evaluation design as an alternative. The alternative design utilizes propensity score matching (PSM) methodology which is one of the most rigorous methodologies in the absence of a feasible experimental methodology. Furthermore, sample size requirements are much more easily met in comparison to an experimental design since PSM does not involve randomization of applicants (i.e., no denial of service) and it also involves a matched administrative control group that provides more additional sample. An optimal strategy would be to start implementing the experimental design and to switch to the PSM design if it becomes apparent halfway through the intake process that the target sample size will not be reached.

In this chapter, we provide a detailed overview of our proposed evaluation design for conducting the evaluation of the effectiveness of the CDSMP in AoA-funded settings. Section A lists the key research questions to be addressed by the evaluation. Section B describes the two approaches to evaluation: a RCT and an alternative quasi-experimental PSM design. Section C provides a power analysis to guide the determination of minimum sample size required to detect statistically significant and qualitatively meaningful results. Site selection and recruitment considerations are presented in Section D. Section E provides details on baseline data collection and the participant tracking system that is necessary to ensure uniform data collection and critical monitoring of evaluation site data. Random assignment procedures are described in Section F. Section G provides details on follow-up data collection and Section H concludes the section with the details of the data analysis plan. Where appropriate, we note how the implementation of the quasi-experimental design differs from the experimental design.

³ ARRA funding for CDSMP is set to expire in March 2012.

A. Research Questions

The objective of the impact evaluation is to produce rigorous analyses of the CDSMP impacts on participant outcomes. Specifically, the evaluation will address a number of key research questions, including:

- What is the impact of CDSMP on the physical and mental health status of program participants?
- Is CDSMP effective in assisting program participants to better manage their conditions and achieve self-efficacy?
- What is the impact of CDSMP on health care services utilization?
- Does CDSMP lead to a reduction in health care expenditures?
- What is the impact of CDSMP on Medicare costs for participants who are Medicare beneficiaries?
- What is the impact of CDSMP on Medicaid costs for dual eligibles?
- How do the impacts of the program differ across key groups (age group, Medicare status, chronic disease, etc.)?
- How do the implementation factors affect impacts? Do factors such as the organizational structure, financial structure, experience of facilitator or leader, and site-level course completion rates affect the success of CDSMP?
- What are the characteristic of CDSMP participants? How do they differ from AoA's usual customers? How is their use of long term care services and support? What percent are Medicaid beneficiaries (dual eligibles)?

B. Evaluation Methodologies

The impact of the CDSMP on participant outcomes is essentially the difference between participant outcomes after program participation (treated outcomes) and participant outcomes had they not participated in the program (counterfactual outcomes). The typical evaluation problem is that, while the treated outcomes are observed, the counterfactual outcomes are not observed. The solution to this problem is straightforward when random assignment (RA) is used to determine program participation. Under RA, eligible program applicants are randomly assigned to the treatment group (receive program services) or to the control group (do not receive program services). RA ensures that individuals in the treatment and the control groups are equivalent in their observed and unobserved characteristics at the time they apply to participate in the program. Thus, any subsequent differences between the treatment outcomes (treated outcomes) and the control outcomes (counterfactual outcomes) are attributed to the program. This explains why experimental evaluation designs are preferable for rigorously

assessing program efficacy relative to non-experimental designs that rely on statistical methods to identify appropriate comparison groups.

As explained before, we are also presenting an alternative rigorous design as a backup plan in case the evaluation sites will not be able to reach recruitment targets necessary for a meaningful evaluation. The alternative design is a quasi-experimental design where the control group members are selected from the extensive pool of Medicare administrative records. Separately in each host site service area, Medicare beneficiaries with similar characteristics will be identified and matched to the beneficiaries who are taking the CDSMP workshops. In this design, there is no randomization and hence everybody taking the workshops can be in the evaluation sample. We propose Propensity Score Matching (PSM) to ensure controls and experimental cases are matched as best as they can by an extensive set of variables on individual demographic characteristics, Medicare utilization and expenditures over the 12 months preceding the study. Propensity Score Matching offers an ideal solution to the problem of trying to simultaneously match on a multitude of variables. In this method, this extensive set of variables is used to predict the probability that an individual participates in the CDSMP. Individuals are then matched using these estimated probabilities.

PSM is a relatively costly technique – if implemented rigorously - due to intense data and analytic requirements. However, it is proven to be very effective and regarded often as the second best option after RCTs when a large pool of individuals exists to obtain matches (Medicare administrative records). There will be separate matching exercises for each host site area ensuring that the comparison group matches reside in the same area as the treatment group individuals (“hard matches” by geographic area). Furthermore, within these geographic areas we further hard-match certain key subgroups such as dual eligibles. The nature of the technique requires multiple iterations of matching by varying certain parameters and then carrying out extensive “balance tests” for each such iteration. Balance tests are used to assess how well a particular matching works by assessing how close the statistical distributions of the matched characteristics of the comparison group are to those of the program group. The drawback of the PSM is that comparison cases may not be matched with program cases on unobservable characteristics such as motivation to improve health. The validity of the comparison group depends on how important these unobservable factors are in explaining outcomes of interest and how well they can be predicted by the extensive set of Medicare administrative records. Further advantages and disadvantages of the PSM vis-à-vis RCT are described in the following sections.

C. Required Sample Sizes

An important consideration in implementing a design is to include a sufficiently high number of participants to ensure that the evaluation detects statistically significant estimates of the CDSMP’s impacts. Depending on the number of program applicants from the selected evaluation sites and a set of assumptions about the means and standard deviations of the

outcomes of interest, one can calculate the minimum detectable effect (MDE). The MDE is the smallest program impact that the evaluation will likely be able to detect for a given outcome of interest. If the actual impact of the CDSMP on a given outcome is higher than the MDE, one can be reasonably confident that the impact will be detected by the evaluation. Since the actual program impact is not known a priori, smaller MDEs are desirable; the smaller the MDE, the greater the chance that the actual impact will be detected.

We are considering the 6-month impact of the CDSMP on a number of key variables that are derived from survey data or Medicare administrative records. Key variables derived from the surveys: (1) likelihood of self-efficacy and (2) likelihood of good self-rated health; key variables derived from Medicare administrative records: (3) likelihood of hospitalization (inpatient stay), (4) number of hospital stays, (5) days hospitalized, (6) hospital expenditures, and (7) total Medicare expenditures. Exhibit 1 presents the MDEs for these outcomes using different baseline sample sizes (i.e., the number of treatment and control group participants).^{4,5,6,7,8}

As shown in Exhibit 1, if the evaluation includes 1,500 participants (750 treatment and 750 control), the MDE for the likelihood of self-efficacy will be 7.5 percentage points. This means that if the impact of the CDSMP program is at least 7.5 percentage points, an evaluation sample of 1,500 program participants will enable the detection of the program's impact. By the same token, an evaluation sample of 1,500 participants would enable the detection of a 7.1 percentage-point positive impact on the likelihood of good health, a 5.1 percentage-point reduction in the likelihood of hospitalization, a 0.12 days reduction in the number of hospital stays, and a 0.8 reduction in the number of days hospitalized. Furthermore, a sample of 1,500

⁴ The typical standard used in program evaluation is that the estimated impact is statistically significant at the 5 percent level. The typical standard for statistical power (i.e., the probability that a statistically significant impact is detected) is 80 percent. Based on our review of several reports and data sources, we make the following assumptions about the outcomes of interest: (1) 50 percent of program applicants achieve self-efficacy, (2) 65 percent of applicants report that they are in good health, (3) 12 percent of Medicare beneficiaries has an inpatient stay, (4) the standard deviation of the number of hospital stays is 0.783, (5) standard deviations of days hospitalized is 5.4, (6) the standard deviation of hospital expenditures is \$5,073, and (7) the standard deviation of total Medicare expenditures is \$11,530.

⁵ Our team did not have access to any Medicare data specific to this design project. We utilized a 5% Medicare annual sample and only included beneficiaries who were 65 or older, had 12 months of Part A and Part B coverage, and did not have Medicare Advantage enrollment in any month. The data did not allow for calculation of standard deviation of total Medicare expenditures for a 6-month period. We approximated the 6-month standard deviation as 75% of the 12-month standard deviation. Further details can be provided separately if requested. The true MDEs can be larger or smaller.

⁶ We assume that 70 percent of baseline study participants will respond to the 6-month follow up survey. As described in Section 3.C, 60% of baseline participants are assumed to have Medicare claims records as some of the participants will be younger than 65, be Medicare Advantage enrollees, or will not be able to be matched to their Medicare records for various reasons. Since there is no non-response in administrative records, we also assume that everyone for whom Medicare data is available at baseline will have Medicare data available at month 6 as well.

⁷ The estimates are based on simple random drawing. We did not consider between-site variation and the implied "design effect". Thus, it is possible that the presented MDEs are biased downwards.

⁸ We assumed that 5% of the variation in the survey-based outcomes of interest will be explained by covariates. For Medicare claims-based outcomes of interest, this was set at 10%.

study participants would enable the detection of reductions of \$756 and \$1,719 in Medicare hospital expenditures and total Medicare Part A and Part B expenditures, respectively.

If the evaluation sample includes a larger number of participants, the MDE would be lower. For example, a sample of 2,000 participants (1,000 treatment and 1,000 control) would enable the detection of at least a 6.5 percentage-point impact on the likelihood of self-efficacy. If the evaluation sample consists of 3,000 applicants, the MDE would be even lower (5.1 percentage points). As shown in Exhibit 1, the MDEs decline with larger sample sizes, indicating that large evaluation samples would enable the detection of small program impacts.

Exhibit 1: 6-Month Minimum Detectable Effects by Target Sample Size

	Number of Participants in the Evaluation Sample				
	N = 1,500	N = 2,000	N = 3,000	N = 4,000	N = 5,000
1) Likelihood of Self-Efficacy	+7.5	+6.5	+5.3	+4.6	+4.1
2) Likelihood of Good Health	+7.1	+6.2	+5.1	+4.4	+3.9
3) Likelihood of Hospitalization	-5.1	-4.4	-3.6	-3.1	-2.8
4) Number of Hospital Stays	-0.12	-0.10	-0.08	-0.07	-0.06
5) Days Hospitalized	-0.80	-0.70	-0.57	-0.49	-0.44
6) Hospital Expenditures	-\$756	-\$655	-535	-463	-414
7) Total Medicare Expenditures	-\$1,719	-\$1,488	-\$1,215	-\$1,052	-\$941

There are two important considerations in determining the number of participants that will be included in the CDSMP evaluation. First, a sufficiently high number of participants must be included to ensure that the evaluation will detect the CDSMP impacts. Exhibit 1 shows the MDEs for various sample sizes. Second, the number of participants in the evaluation cannot exceed the projected number of CDSMP applicants in the selected evaluation sites who will be willing to participate in the evaluation during the intake period. As explained in Section 3.C, it does not seem realistic to expect more than 3,000 study participants even with supplementary grants from AoA to host sites that are aimed at doubling the number of CDSMP participants.

Based on the power calculations presented in Exhibit 1, and the feasibility considerations, an evaluation sample of 3,000 individuals (1,500 treatment and 1,500 comparison) would enable the detection of sufficiently low MDEs for the key outcomes of interest except Medicare expenditures. It does not seem feasible to detect reasonable differences in Medicare expenditures due to very high variability in them. A meaningful effect size to detect would be \$300 per participant as this has recently been estimated as the per participant cost of holding

the workshop.⁹ \$300 also corresponds to 7.5% of average Medicare expenditures in a period of six months. Ability to detect an effect size of \$300 requires a sample size of 50,000. The evaluation contractor should propose methods to increase precision in the estimate of Medicare expenditure which may include modeling logarithm of Medicare expenditures rather than actual expenditures.

It is our estimation that available sample size for the PSM option will be triple the available sample size for the RCT option. MDEs for the PSM option are expected to be at most half of those for the RCT. Thus, the PSM option will have significantly higher statistical power than the RCT option.

D. Site Selection and Recruitment

An important aspect of this evaluation is to identify CDSMP program sites that are strong candidates for inclusion in the evaluation. The evaluation contractor must work closely with AoA to identify CDSMP evaluation sites that satisfy certain criteria, including the following:

- *Size* – As discussed earlier, large evaluation sample sizes will ensure that the evaluation will be able to detect small program impacts. To achieve high statistical power in assessing CDSMP impacts, the evaluation contractor and AoA should select large CDSMP evaluation sites to be included in the evaluation. The selection of large sites—sites that receive a high number of applications for program participation—will ensure that large sample sizes will be achieved and that the evaluation will be able to detect program impacts with high accuracy.
- *Representativeness* – If feasible, selected evaluation sites will represent a wide range of program sites, based on a number of community-level characteristics, including but not limited to (1) rural/urban location, (2) average household income, (3) socioeconomic characteristics, and (4) average health care expenditures. If evaluation sites are representative of a wide range of program sites, AoA will be able to use the results of the evaluation to draw inferences on the effectiveness of the CDSMP program on a large scale.
- *Wait Lists* – If feasible, the evaluation will include sites that keep wait lists due to a large demand for CDSMP services. Such sites would presumably be more willing to participate in the evaluation, since randomly assigning a portion of their applicants to the control group will not affect their usual operations.
- *Tomando de Su Salud* – The evaluation will include sites implementing the Spanish-speaking version of the CDSMP.

⁹ One of the CDSMP grantee states we interviewed estimated the per participant cost of providing the workshop as \$300. Some of the Lorig articles provide estimates in the range of \$70-\$200; however, these are likely to be outdated.

The evaluation contractor should use criteria, such as those outlined above, to identify a preliminary list of sites for inclusion in the evaluation. This list may include 70–80 CDSMP sites. Once this list is approved by AoA, the contractor must then make all possible efforts to ensure that a sufficient number of selected sites agree to participate in the evaluation. We propose that the evaluation includes 50 sites; these numbers will ensure a sufficiently high number of participants even though there will be challenges for coordinating and managing a 50-site RCT evaluation. If the PSM option is selected, the evaluator should consider a smaller number of sites. The evaluation contractor must make all possible efforts to secure the participation of the target number of sites. Specifically, we recommend that the evaluation contractor implement and support the following activities to ensure site participation:

- *Step 1 – Inform sites of the study and generate support for participation.* The evaluation contractor should begin the site recruitment process by contacting the sites on the preliminary list. The contractor should draft and transmit letters of introduction to the sites; these letters should be sent to a high-level site administrator who is likely to have sufficient authority for committing the site to participate in the study. We recommend that the evaluation contractor enlist AoA's and Medicare's support in publicizing the study and encouraging site participation. Thus, AoA should provide a letter supporting the study and asking for the site's cooperation. The letter of introduction is intended to provide a method for ensuring that contacts understand the following: the objective of the study, the random assignment design of the study, and the need for participating sites to cooperate with the evaluation contractor to facilitate random assignment of program applicants and the collection of baseline data. The letter will also ask sites to consider participating in the study to assist AoA in gaining a better understanding of the effectiveness of the CSMP. Finally, the letter should provide contact information for the evaluation contractor and the AoA project officer.
- *Step 2 – Conduct teleconferences with interested sites.* The contractor should schedule teleconferences with sites that expressed interest in participating in the study. During these conference calls, the contractor will provide detailed information to the sites about the project's objectives and the requirements for the evaluation sites. Specifically, the following topics should be discussed during the conference calls: (1) overview of the project, (2) key research questions, (3) study design, (4) criteria for selecting evaluation sites, and (5) requirements for evaluation sites (random assignment of applicants, collecting contact information on applicants, etc.). It is recommended that AoA participate in these calls, when possible, to indicate its support of the request and to provide further encouragement for the sites' participation and explain any plans for offering grant funding to help implement the study.
- *Step 3 – Secure site participation and select evaluation sites.* Once sites commit to participating in the program, the evaluation contractor and AoA will select the final list of sites to be included in the evaluation. The contractor will then contact the selected evaluation sites to inform them of their inclusion in the evaluation and to request letters of commitment.

E. Baseline Data Collection and Participant Tracking System

The evaluation sites will collect information on all individuals aged 60 or older who apply for CDSMP participation during the evaluation period. To facilitate the evaluation, sites will be asked to uniformly collect the following baseline information:

- Applicant contact information (name, address, phone number)
- Socioeconomic characteristics (gender, race, ethnicity, age, education, income, etc.)
- Health characteristics (health status, chronic conditions, illnesses, disabilities, etc.)
- Living arrangements (living alone, living with family, managed care facility, etc.)
- Health care expenditures prior to CDSMP application
- Medicare/Medicaid status (Medicare, Medicaid, and Dual eligibility)
- Health insurance status and Participation in a Medicare Advantage Plan
- Prior participation in the CDSMP or similar programs
- Type of Program Participant (individual or proxy/caregiver)
- Use of long-term care services or supports
- Language spoken

Note that while this survey will be administered to both treatment and control group members in the RCT design, it will only be administered to the treatment group members in the PSM design. The evaluator will not have contact information for the comparison group members in the PSM design. Furthermore, even with the contact information, response rates would be very low due to the “cold call” nature of the inquiry.

The above information will be used to produce detailed descriptive analyses of the characteristics of CDSMP applicants. In addition, this information will be used to produce control variables for the impact analyses (see Section H, below).

The evaluation contractor will be responsible for facilitating the process of collecting baseline information on program applicants by the evaluation sites. Specifically, the contractor will develop a baseline data collection instrument that will be used by the evaluation sites to collect information on applicant characteristics. The contractor will train all evaluation sites on how to use this instrument before the start of the implementation period.

Furthermore, the evaluation contractor will create a Participant Tracking System (PTS), which will be used to track all interactions of program applicants with the CDSMP, from program application through service delivery. The PTS will be designed to include the following features:

- Store baseline characteristics of all applicants – The PTS will be designed to allow evaluation sites to store the baseline applicant characteristics once they are collected. The PTS will also allow the evaluation sites and the evaluation contractor to access information on program applicants and produce custom reports using up-to-date data. Furthermore, the PTS will be designed to allow the evaluation contractor to obtain easy access to baseline characteristics for conducting the impact analyses.
- Facilitate random assignment (for the RCT study)– The PTS will be designed to allow the evaluation contractor to access up-to-date information on new program applicants for conducting the random assignment of applicants to the treatment group or the control group and notify sites of the assignment.
- Track participants’ program use – The PTS will be designed to allow sites to document program use by participants. Specifically, sites will be able to collect information on whether treatment group members showed up to receive program services, the types of services received, and whether they completed the program. In addition, sites can use the PTS to report whether a proxy participant attended the program instead of the actual applicant. This information can be used to assess the use of the program by treatment group members.
- Track contamination: For RCT, the PTS will track whether the control group members take CDSMP workshop within the first 6 months. For PSM, the PTS will track everyone in the matched Medicare comparison group whether they take the CDSMP workshop. This will take quite a bit of effort as it requires identifying and matching the Medicare beneficiaries with the CDSMP participants based on limited information such name, gender, age, and geographical area.

The PTS will provide the sites and the contractor with information necessary to effectively monitor program activities. For example, the PTS will provide real-time data on program enrollment and program attrition, enabling the monitoring of how far in the sequence of CDSMP services the participant progresses. Furthermore, the evaluation contractor will be able to identify early if sites are experiencing a shortfall in the expected number of applications.

F. Random Assignment

Random assignment is applicable to the RCT design only. A random assignment (RA) process will be used to assign program applicants to the treatment group or the control group. Those assigned to the treatment group will receive program services, whereas those in the control group will have to wait 6 months before receiving services. To achieve the required sample sizes for the evaluation, the intake period will be 12 months; if the required sample sizes are not reached, the evaluation may switch to the PSM option.¹⁰ Below, we discuss the

¹⁰ The exact implementation period will be estimated by the evaluation contractor once the recruitment of sites is complete.

recommended steps for successfully implementing RA of program applicants during the implementation period.

- *Step 1 – Informed consent form.* During the implementation period, sites will inform all program applicants aged 60 years or older that the site is participating in an RA study of the CDSMP program. Sites will explain to these applicants that each applicant will be randomly assigned to the treatment group or the control group and that assignment to the control group means that the individual will need to wait 6 months before receiving program services. Sites will also explain that applicants will be asked to respond to a follow-up survey, approximately 6 months after program application. Individuals who are still interested in program participation must sign an informed consent form which states that the applicant is aware of the evaluation, the random assignment process, and the follow-up data collection.
- *Step 2 – Random assignment of program applicants.* During the implementation period, evaluation sites will feed information about new program applicants (aged 60 or older) into the PTS. On a weekly basis, the evaluation contractor will access the PTS to identify new applicants and will use statistical software with random number generator capabilities to assign each new applicant a random number. Half of the numbers will then be randomly drawn and assigned to the treatment group, and the remaining half will be assigned to the control group. The evaluation contractor will then update the PTS with information about whether the individual was assigned to the treatment group or the control group.
- *Step 3 – Notify applicants and provide services.* Once the evaluation sites receive the treatment and control group information, they will notify applicants of their assignment. The sites will then provide services to all program applicants as follows: treatment group members will receive services once workshops become available; control group members will be provided services after the 6-month wait period and after the follow-up data collection is completed.

This process will ensure that the random assignment of program applicants to the treatment or the control group is successfully implemented. Random assignment will enable the evaluation contractor to compare the outcomes of treatment and control group members 6 months after program application (using the follow-up survey and Medicare/Medicaid data) to rigorously estimate CDSMP impacts. Furthermore, this design ensures that all program applicants receive CDSMP services, with the caveat that control group members will need to wait 6 months before receiving services. This design alleviates ethical concerns since none of the evaluation site program applicants will be denied services.

Evaluation sites will be asked to screen applicants based on their age; only applicants aged 60 or older will be randomly assigned to the treatment or the control group, while those under age 60 will follow the normal process of receiving CDSMP services.¹¹

¹¹ As of March 2011, 27% of CDSMP program applicants in the last 12 months were younger than 60 years of age.

G. Follow-Up Data Collection

Given the relative scarcity of CDSMP participants (see Section 3.C) and the need to have as many study participants as possible to increase the likelihood of detecting statistically significant effects (see Section 4.C), it is critical to retain as many baseline participants for the follow up surveys as possible. To that end, we recommend administering the follow up surveys as Computer Aided Telephonic Interviews (CATI) while providing the paper-and-pencil option to participants to maximize response rates. Follow up survey response rates are assumed to be at least 70%. The primary goal of the planned outcome evaluation is to assess health and other outcomes for CDSMP participants compared to non-participants. Data collection procedures and instruments designed to assess outcomes will be identical for the treatment and control groups. Follow up survey data for the RCT study will be collected from both treatment and control groups at months 6 and 12 after random assignment. For the PSM study, survey data will be only be collected from the treatment group at baseline, 6 months, and 12 months.

G.1 Survey Procedures

Survey type and method: development and testing. As stated earlier, there are several well-developed and tested instruments for evaluating health outcomes, quality of life outcomes, and self-efficacy. We recommend that the evaluation contractor select an existing, psychometrically sound tool to measure self-reported health, quality of life (QOL), and self-efficacy outcomes based upon the environmental scan described previously. We recommend that the survey mode be a mail survey, as has been employed in past CDSMP evaluations.

Use of a validated instrument will eliminate the need to cognitively test the survey tools; however, pilot-testing will still be required to ensure smooth survey implementation across sites and valid survey results. As stated previously, for purposes of generating survey cost estimates, we have assumed that the instrument selected will be the Stanford Chronic Disease Questionnaire. The rationale for recommending this tool is that all domains of interest are included in one instrument, the scales included in the tool have documented psychometric properties,¹² and the degree of respondent burden is minimized with one tool containing approximately 48 items.

Sampling Frame. The sampling frame for the RCT design will be obtained from CDSMP sites participating in the national evaluation. Participant contact information for the treatment and control groups will be obtained by the site through the application process and forwarded to the survey contractor for input into the contact database. Copies of informed consent for survey participation will have been obtained by each participating site and copies forwarded to the evaluation contractor for record-keeping. The treatment arm of the PSM study will have the same sampling frame as the RCT. However, the comparison arm will have a different sampling frame. For a given host site, the sampling frame for the comparison group will be the Medicare

¹² See <http://patienteducation.stanford.edu/research/cdCodeBook.pdf>.

beneficiaries who are 65 years of age or older residing in the area served by the host site. The list of these beneficiaries will be obtained from Medicare administrative records through the Centers for Medicare and Medicaid Services. Then, PSM will be used to select among these beneficiaries the ones that match well to the treatment group by an extensive set of individual characteristics as described in Section 5.B.

G.2 Survey Instrument Development

Pilot testing. Pilot testing informs the evaluation contractor of difficulties in administering the survey (e.g., wording is difficult to read as written) as well as difficulties that respondents may have in interpreting and/or understanding specific questions. In addition, pilot testing can help the researcher test the length of the survey and adjust it as necessary to maintain the preferred survey length. Most importantly, pilot testing will ensure that the evaluation contractor is obtaining the desired data from the survey questions, for both the treatment and the control groups.

Given the design team's recommendations that the already-tested Stanford Chronic Disease Questionnaire be utilized, we do not believe that it will be necessary for the evaluation contractor to conduct a lengthy pilot test; however, we would recommend that the survey instrument and survey procedures be tested for control and treatment group participants from two CDSMP sites to assure AoA and the evaluation contractor that the instructions are clear and the process will work as intended prior to full-scale baseline data collection.

Survey administration. The study will be conducted through self-administered paper and pencil surveys for the baseline data collection. Two waves of follow-up measurements will be administered through CATI telephonic survey system. An item to capture whether control group participants have received any education about their disease since the baseline survey should be added to these follow-up surveys to enable assessment of control group contamination. Spanish-language surveys should be sent to those treatment and control group participants who are identified by the CDSMP site as reading only Spanish.

G.3 Administrative Data Review

As previously stated, the design team recommends that the effects of CDSMP programs on health care utilization and expenditures be assessed using Medicare claims data. These data may be obtained from CMS, using the following procedures:

- *Step 1* – Obtain understanding from AoA about any interagency agreements regarding data use/data acquisition.

- *Step 2* – Complete ResDAC data request package, including study protocol, request letter, and Data Use Agreement (DUA).¹³
- *Step 3* – Request access and specify media (e.g., DVD, external hard drive) in which to receive the following Medicare data: EDB, Inpatient SAF, Outpatient SAF. It would be desirable for the national evaluation contractor to be able to access the data via the CMS mainframe due to the potentially large size of the files; however, mainframe access is unlikely given other CMS data processing priorities.

It is likely that, during the development of the study protocol and data request package, the evaluation contractor will need to consider development of a “finder file” for ResDAC’s use. This file will provide ResDAC (or other CMS data contractor) with the appropriate identifiers for the CDSMP treatment and control samples. The finder file will allow ResDAC to pull claims for only those Medicare beneficiaries for whom survey data are available and thus streamline the resulting size and content of the administrative files. The evaluation contractor should anticipate that awaiting receipt of data from ResDAC will likely take the better part of Year 1 of the contract, and thus the data request package should be submitted as soon as possible after contract award.

Claims data for the year just prior to the evaluation, for the 12 months of recruiting and for the 12-month follow-up period, should be obtained for development of the longitudinal file. Our estimate at this point of the data required would be for calendar years 2012 through 2015.

G.4 IRB and OMB Clearance

To ensure the protection of human subjects used in the proposed research, the survey and evaluation plan must be approved by a certified Institutional Review Board (IRB) prior to commencement. In addition, it is anticipated that the data collection effort will require preparation and submittal of a Paperwork Reduction Act (PRA) package for Office of Management and Budget (OMB) approval. Based upon previous experience, a minimum of 9 months should be allowed in the proposed project schedule for OMB clearance; thus, the design and data collection protocols must be finalized within the first several months of the national evaluation in order to obtain clearance to proceed with the study by Year 2. Mechanisms for achieving the highest possible response rate should be articulated in the PRA package, as should the evaluation contractor’s plans for assessing non-response bias. We suggest that data on the study population be collected from the sites to help assess non-response bias and make any needed adjustments to the survey data. We also suggest that responders to follow-up surveys are compensated for their time with \$15 checks. We anticipate the need for OMB clearance both for the CDSMP participant survey and the CDSMP host site web survey.

¹³ See http://www.resdac.org/Medicare/requesting_data_NewUse.asp for details on the data request procedures and package.

H. Data Analyses

The objective of this evaluation is to estimate the CDSMP impact on participant outcomes, 6 months and 12 months following program participation. For the RCT design, since program applicants will be randomly assigned into the treatment group (receive services) or the control group (required to wait 6 months before receiving services), there will not be any observed or unobserved differences in characteristics between the treatment and the control groups. Thus, treatment-control differences in outcomes 6 months after RA can be attributed to CDSMP participation and represent accurate estimates of program impacts.¹⁴ At the 12th month, some control group participants may have taken the workshop after being put in the waitlist for 6 months. So, the 12 month impact estimates will provide a lower bound for the effectiveness of CDSMP. Even though the PSM design ensures that there are no observed pre-existing differences between the treatment group and the matched comparison group, it does not guarantee this for unobserved differences. Below, we provide an overview of the recommended data analyses for estimating the impact of CDSMP on participant outcomes. If there is a difference between the RCT and PSM designs in data analyses (after randomization for RCT and after matching for PSM), we indicate these below.

H.1 Descriptive Analyses of Evaluation Sample

Prior to conducting any impact analyses, descriptive analyses of all CDSMP applicants in the sample should be produced based on information provided at the time of application. These analyses will provide an overview of the baseline characteristics of treatment and control group individuals, including:

- Socioeconomic characteristics (gender, race, ethnicity, age, education, income, etc.)
- Health characteristics (health status, chronic conditions, illnesses, disabilities, etc.)
- Living arrangements (living alone, living with family, managed care facility, etc.)
- Health care expenditures prior to CDSMP application
- Medicare/Medicaid status (Medicare, Medicaid, and Dual eligibility)
- Health insurance status and Participation in a Medicare Advantage Plan
- Prior participation in the CDSMP or similar programs
- Use of long-term care services and supports
- Type of program participant (individual or proxy/caregiver)
- Language spoken

¹⁴ The focus is to estimate the program's intent-to-treat effect on participant outcomes, that is, the effect of being offered CDSMP services. From this point on, *impact* or *effect* refers to the intent-to-treat effect.

These descriptive analyses are necessary to provide an overall characterization of CDSMP applicants in each evaluation site. This information will be used to produce characteristics comparisons between CDSMP applicants and all individuals aged 60 or older in the state, and between CDSMP applicants and all Medicare beneficiaries in the state. Analysis will include completers and non-completers, as well as an analysis of how many and which classes they attended. Additionally, these analyses will provide some preliminary evidence on whether random assignment of CDSMP applicants to the treatment or the control group was performed effectively. The PSM, by design, makes sure that there are no existing differences by conducting balance tests. So, the following tests to assess randomization only apply to RCT.

- *Treatment-control comparisons in observable characteristics* – The evaluation contractor should produce a standard treatment-control group means comparison of each available characteristic. Using t-tests, the contractor should examine whether there are any statistically significant differences in characteristics between treatment and control group members in each evaluation site. If randomization was done effectively, there should not be any statistically significant differences in characteristics between the treatment and the control group in each evaluation site.
- *Estimate likelihood of treatment group assignment* – The evaluation contractor should estimate a linear regression model for each evaluation site, where the dependent variable is the likelihood of being assigned to the treatment group, and controls include all available characteristics at the time of application. If randomization was done effectively, the estimated parameters of each control variable will be statistically insignificant.

H.2 Impact Analyses Using Means Differences

Once it is verified that randomization of program applicants was successful, means comparisons in outcomes 6 months after RA between treatment and control group applicants will produce unbiased estimates of CDSMP program impacts. The first step in estimating program impacts is to produce comparisons between the mean outcomes of treatment group members and the mean outcomes of control group members for all outcomes of interest. T-tests should be used to assess if the treatment-control differences in post-RA outcomes are statistically significant, that is, if those in the treatment group had different outcomes than those in the control group.

These analyses should be produced for all available outcomes of interest, including health status, self-efficacy, total health care expenditures, Medicare/Medicaid expenditures, etc. Based on the results of these analyses, the evaluation contractor will be able to draw preliminary conclusions about the efficacy of CDSMP in improving participant outcomes.

H.3 Impact Analyses Using Regression Models

To estimate the impact of CDSMP on the post-RA outcomes of program participants with increased statistical efficiency, multivariate regression models should be used. Regression models are used to refine the impact estimates by eliminating differences in outcomes between

the treatment and the control groups that may have occurred by chance as a result of differences in observed characteristics. These models will also permit subgroup analyses that can identify impact differences across key participant groups (e.g., by age, chronic condition, Medicare status) and by program characteristics (e.g., program size, location, services).

The impact analysis regression model can be expressed by the following equation:

$$Y = \alpha \cdot T + X \cdot \beta + u \quad (1)$$

The dependent variable in this model (Y) is the participant post-RA outcome of interest (self-efficacy, health care expenditures, etc.). Control variables include:

T , which equals 1 if the individual was in the treatment group and 0 otherwise.

X , which includes all available individual characteristics at RA (e.g., socioeconomic characteristics, living arrangements, chronic condition, participation in prior workshop, and prior health care expenditures).

u , which is a zero mean disturbance term.

The parameter of interest in this model is α , the regression-adjusted treatment effect of the CDSMP program on the outcome. This parameter represents the impact of being assigned to the CDSMP treatment group (intent-to-treat effect) and not the impact of actually receiving services. The model will be estimated separately for each available post-RA outcome. Once the evaluation contractor estimates these models, t-tests should be used to assess whether estimated impacts are statistically significant.

The results of these analyses will address many of this evaluation's key research questions. For instance, based on these results, AoA will be able to assess if CDSMP led to significant improvements in the health status and self-efficacy of participants. AoA will also be able to assess whether CDSMP participation led to significant savings in Medicare and/or Medicaid expenditures.

H.4 Sensitivity Analyses

In addition to the above analyses, we recommend that the evaluation contractor performs sensitivity analyses that will provide better insight on program effectiveness and will ensure that program impacts were estimated accurately. Specifically, we suggest the following types of analyses.

Subgroup analyses. The above regression model can be modified to assess whether the CDSMP program had differential impacts across key applicant groups. For example, it would be interesting to assess if the program has differential impacts by chronic condition, living arrangements, and health status at the time of application. To conduct these analyses, the evaluation contractor can modify the regression model above to include interactions between the treatment indicator (T) and the characteristic of interest (age, chronic condition, etc.). T-

tests should then be used to examine if the interaction effects are statistically significant, that is, if there are important differences in program impacts across key groups. These analyses will provide important insight to AoA about program effectiveness.

Program differences analyses. There may be specific implementation factors that lead to higher CDSMP impacts. For example, it would be interesting to assess if the program has differential impacts across program characteristics, such as duration, class size, or language of instruction. To conduct these analyses, the evaluation contractor can modify the regression model above to include interactions between the treatment indicator (T) and the implementation factors of interest. Using t-tests, the contractor will be able to assess if there are important differences in program impacts across key implementation factors.

Weighting to account for sample attrition. We also recommend that the evaluation contractor produce weights that account for sample attrition. These weights would be used in the analyses to capture treatment-control differences in survey response and ensure that the impact estimates are not tainted by sample attrition. Using attrition weights, therefore, is critical to ensure that the evaluation produces consistent estimates of program impacts. The evaluation contractor could also use other methods to account for sample attrition, such as the PROC MIXED/PMM approach.

Sample contamination. In some states, CDSMP program services are available to individuals through an online tool. Thus, it is possible that in such states, control group members would receive the services online before the expiration of the 6-month waiting period. This would “contaminate” the control group and might jeopardize the evaluation. One strategy for avoiding this issue is to ensure that the evaluation excludes sites in states where an online CDSMP tool is available. Another strategy is to use responses to the follow-up surveys to control for whether individuals in the treatment or the control group received the online CDSMP services or some other type of services.

CHAPTER 5: TIMELINE

For planning purposes, we anticipate that an evaluation with either of the RCT or PSM design will take approximately 4 years to complete, as follows:

- In Year 1, the environmental scan and evaluation design would be completed, CDSMP sites would be selected, the survey instrument would be finalized, the IRB and PRA packages would be developed and submitted, CDSMP sites would be trained on data collection (for random assignment and administration of baseline survey), and survey administrators would be trained.
- In Year 2, upon OMB clearance, participants would be consented and randomly assigned to treatment and control status during a 12 month intake period for the RCT design. PSM design will not include randomization. Baseline survey data collection would take place for most of the participants, followed by any necessary survey revisions. Web-based host site surveys will be completed. CDSMP sites would be monitored for data submittal compliance, and technical assistance would be offered as needed. First follow up phone survey would be conducted (the 6-month survey) for most of the participants.
- In Year 3, the baseline and the 6-month surveys would be completed and the 12-month phone survey will be conducted with most of the participants. First round of Medicare administrative data with six months of claims run-out will be acquired. It is assumed that AoA will enter into an agreement with CMS that allows the evaluation contractor receive Medicare data as soon as the claims run-out period is over. For the RCT option, this data will be used for participation analysis. For the PSM option, this data will be used to find matching comparison cases. CDSMP sites would be monitored for data submittal compliance, and technical assistance would be offered as needed.
- In Year 4, the 12-month survey will be completed. Second round of Medicare data with six months of claims run-out will be acquired and analyzed. It is assumed that AoA will enter into an agreement with CMS that allows the evaluation contractor receive Medicare data as soon as the claims run-out period is over. Final report will be produced.

REFERENCES

- Agency on Aging. Evidence-Based Disability and Disease Prevention programs. http://www.aoa.gov/AoARoot/AoA_Programs/HPW/Evidence_Based/index.aspx
- Ahmed, T. & Villagra, V. (2006). Disease management programs: program intervention, behavior modification, and dosage effect. *Journal of Consumer Policy*, 29(3), 263-278.
- Bhandari, A. & Wagner, T. (2006). Self-reported utilization of health care services: improving measurement and accuracy. *Medical Care Research Review*, 63(2), 217-235.
- Center for Healthy Aging. National Council on Aging. <http://www.healthyagingprograms.org>
- Center for Healthy Aging. Evidence-Based Programs. National Council on Aging. <http://www.healthyagingprograms.org/content.asp?sectionid=32>
- Ersek, M., Turner, J. A., Cain, K. C., & Kemp, C. A. (2008). Results of a randomized controlled trial to examine the efficacy of a chronic pain self-management group for older adults. *Pain*, 138(1), 29-40.
- Gitlin, L. N., Chernett, N. L., Harris, L. F., Palmer, D., Hopkins, P., & Dennis, M. P. (2008). Harvest Health: translation of the Chronic Disease Self-Management Program for older African Americans in a senior setting. *Gerontologist*, 48(5), 698-705
- Goeppinger, J., Armstrong, B., Schwartz, T., Ensley, D., & Brady, T. J. (2007). Self-management education for persons with arthritis: managing comorbidity and eliminating health disparities. *Arthritis and Rheumatism*, 57(6), 1081-1088.
- Kennedy, A., Reeves, D., Bower, P., Lee, V., Middleton, E., Richardson, G., et al. (2007). The effectiveness and cost effectiveness of a national lay-led self care support programme for patients with long-term conditions: a pragmatic randomised controlled trial. *Journal of Epidemiology and Community Health*, 61(3), 254-261.
- Lorig, K. Personal communication, Sept. 7, 2010.
- Lorig, K., Ritter, P. L., & Jacquez, A. (2005). Outcomes of border health Spanish/English chronic disease self-management programs. *Diabetes Educator*, 31(3), 401-409.
- Lorig, K., Ritter, P. L., & Plant, K. (2005). A disease-specific self-help program compared with a generalized chronic disease self-help program for arthritis patients. *Arthritis and Rheumatism*, 53(6), 950-957.

- Lorig, K., Ritter, P., Stewart, A. L., Sobel, D. S., Brown, B. W., Bandura, A., et al. (2001). Chronic disease self-management program: 2-year health status and health care utilization outcomes. *Medical Care*, 39(11), 1217-1223.
- Lorig, K., Sobel, D. S., Stewart, A. L., Brown, B. W., Bandura, A., Ritter, P., et al. (1999). Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: a randomized trial. *Medical Care*, 37(1), 5-14.
- Lorig, K., Stewart, A., Ritter, P., Gonzalez, V., Laurent, D., & Lynch, J. (1996). *Outcome measures for health education and other health care interventions*. Thousand Oaks, CA: Sage Publications, Inc.
- Osborne, R. H., Elsworth, G. R., & Whitfield, K. (2007). The Health Education Impact Questionnaire (heiQ): an outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. *Patient Education & Counseling*, 66(2), 192-201.
- Personal communication with National Council on Aging staff, January 27, 2011.
- Powers, B. J., Olsen, M. K., Oddone, E. Z., & Bosworth, H. B. (2009). The effect of a hypertension self-management intervention on diabetes and cholesterol control. *American Journal of Medicine*, 122(7), 639-646.
- Research Data Assistance Center. Complete ResDAC data request package.
http://www.resdac.org/Medicare/requesting_data_NewUse.asp
- Richardson, G., Kennedy, A., Reeves, D., Bower, P., Lee, V., Middleton, E., et al. (2008). Cost effectiveness of the Expert Patients Programme (EPP) for patients with chronic conditions. *Journal of Epidemiology & Community Health*, 62(4), 361-367.
- Rogers, A., Kennedy, A., Bower, P., Gardner, C., Gately, C., Lee, V., et al. (2008). The United Kingdom Expert Patients Programme: results and implications from a national evaluation. *Medical Journal of Australia*, 189(10 Suppl), S21-S24.
- Smeulders, E. S. T. F., van Haastregt, J. C. M., Ambergen, T., Janssen-Boyne, J. J. J., van Eijk, J. T. M., & Kempen, G. I. J. M. (2009). The impact of a self-management group programme on health behaviour and healthcare utilization among congestive heart failure patients. *European Journal of Heart Failure*, 11(6), 609-616.
- Stanford School of Medicine. Arthritis Self-Management (Self-Help) Program. Patient Education Research Center. <http://patienteducation.stanford.edu/programs/asmp.html>
- Stanford School of Medicine. Chronic Disease Questionnaire. Patient Education Research Center. <http://patienteducation.stanford.edu/research/cdCodeBook.pdf>

Stanford School of Medicine. Diabetes Self-Management Program. Patient Education Research Center. <http://patienteducation.stanford.edu/programs/diabeteseng.html>

Stanford School of Medicine. Research instruments developed, adapted, or used by the Stanford Patient Education Research Center. Patient Education Research Center. <http://patienteducation.stanford.edu/research>

Stanford School of Medicine. CDSMP Fidelity Toolkit. Patient Education Research Center. [http://patienteducation.stanford.edu/licensing/Fidelity Toolkit2010.pdf](http://patienteducation.stanford.edu/licensing/Fidelity_ToolKit2010.pdf), 2010.

Wenzlow, A., Schmitz, R. & K. Shepperson (2008). A profile of Medicaid institutional and community-based long-term care service use and expenditures among the aged and disabled using MAX 2002. Final Report. Washington, DC: Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services, January 18.

Zaslavsky, A. (2001). Statistical issues in reporting quality data: small samples and casemix variation. *International Journal for Quality Health Care*, 13(6), 481-488.