

PRA Clearance Package

Prospective Evaluation of Evidence-Based Community Wellness and Prevention Programs

Part B. Collections of Information Employing Statistical Methods

Version: September 5, 2014

Centers for Medicare & Medicaid Services

Table of Contents

Part B. Collections of Information Employing Statistical Methods.....3

- 1. Respondent Universe and Sampling Methods.....3
- 2. Information Collection Procedures.....8
- 3. Methods to Maximize Response Rates.....9
- 4. Tests of Procedures.....10
- 5. Statistical Consultants.....11

List of Attachments.....11

Part B. Collections of Information Employing Statistical Methods

1. Respondent Universe and Sampling Methods

At this time, eight wellness programs will be included in the evaluation. The selected wellness programs include: the Chronic Disease Self-Management Program (CDSMP), the Diabetes Self-management Program (DSMP), EnhanceFitness, EnhanceWellness, Fit & Strong!, A Matter of Balance, Stepping On, and Walk with Ease. Sites that implement the programs will enroll new participants and identify Medicare beneficiaries; these individuals will be asked to complete a baseline survey as well as a 6 month and 12 month followup. A national survey of Medicare beneficiaries aged 65 or older will also be conducted to provide estimates of readiness and potential demand for wellness programs. This national sample will also be used to identify control samples, one for each of the 8 “treatment” samples selected from the identified wellness programs. The control samples will be determined through the development of models predicting propensity to participate in a wellness program, using data from the national survey as well as claims data.

The national survey and the program participant surveys will include a baseline, a 6-month follow-up, and a 12-month follow-up. For the participants, the 6-month sample will include both wellness program completers and non-completers; whereas, the 12-month sample will include only the program completers who also submitted their 6-month follow-up data. The six-month follow-up survey of wellness program participants also includes questions to identify those participants who indicate that they stopped participation. This information will be used to identify samples for the average treatment effect on the treated (ATT) analysis (where only program completers will be compared with the matched control group) and the complier average causal effect analysis (where characteristics of the non-completers will be matched with those from the control group. Data on reasons why non-completers dropped out of a wellness program will also be collected. A final component of the data collection effort includes unstructured, in-depth interviews with program staff.

Wellness Program Participant Sample

As part of the overall evaluation study, CMS will select and formally contract with wellness evaluation partners. At this time, the included wellness programs are CDSMP, DSMP, EnhanceFitness, EnhanceWellness, Fit & Strong!, A Matter of Balance, Stepping On, and Walk with Ease. The selection criteria for the wellness evaluation partners included:

- Offering an evidence-based wellness program
- Being in operation for at least one year
- Having an established sustainability plan or funding stream
- Having a demonstrated capability to enroll a sufficient number of Medicare beneficiaries as new participants within a one year period
- Being willing and able to support data collection and reporting

Wellness evaluation partners will (1) provide access to Medicare beneficiaries who may participate in the Program Participant Survey and (2) participate in the in-depth interviews about program operations and costs. Refer to Exhibit 4 for sample size assumptions.

Of course, programs will vary in the number of enrollees, program completion rates, and survey response rates. For discussion purposes we have assumed that the largest program will have complete data for 1,500 enrollees, and the smallest program will have complete data for 300 enrollees. Exhibit 3 shows the 95 percent confidence interval half-widths for estimates of proportions for program sample sizes ranging between 300 and 1,500.

Exhibit 3. 95 percent confidence interval half-widths for estimates of proportions given expected wellness program achieved sample sizes

Estimate	Achieved Sample Size	95 percent confidence interval half-width
0.5	300	0.06
	600	0.04
	900	0.03
	1,200	0.03
	1,500	0.03
0.7	300	0.05
	600	0.04
	900	0.03
	1,200	0.03
	1,500	0.02
0.9	300	0.03
	600	0.02
	900	0.02
	1,200	0.02
	1,500	0.02

National Survey of Medicare Beneficiaries Sample

The national sample will be drawn from Medicare enrollment and claims files. Beneficiaries meeting survey eligibility criteria (non-institutionalized, living in the United States) may be assigned to sample strata based on characteristics available through claims and enrollment data, such as dual-eligible beneficiaries or an oldest age category, depending on the characteristics of program partners. After sorting on variables such as census region, state, Zip code, or other useful stratification variables available on the sample frame, we will select equal probability systematic samples of beneficiaries within strata with sampling rates varying by strata as needed. The sample will be selected in waves with 12 replicate subsamples for monthly releases, matching the 12-month program enrollment period.

The sample must be of adequate size to support the construction of comparison groups for the outcome analyses. For the propensity score matching process described in Part A, a large sample is required to ensure a sufficient number of completed responses for each of the 8 control samples to be selected for comparison purposes with the 8 program participant samples. The sample size will be large enough to allow for the possibility that some beneficiaries may be considered ineligible to serve as a control for a given program. For example, some participants may report enrolling in a program similar to one being evaluated, and thus may be excluded from the corresponding comparison group. Additionally, when frequency matching between program participants and the eligible national sample subsampling within propensity strata (e.g., five groupings based on the quintiles of the propensity scores of program participants) with large numbers of national survey participants may be undertaken to help limit costs.

The power to identify differences between program participants and the corresponding comparison sample hinges on a number of factors including the number of new participants expected for a given program over the period of data collection, the size of the difference to be detected, and the resources available for the study.

Allowing for subsampling, nonresponse, and ineligibility, Westat plans to select a sample of roughly 15,500 Medicare beneficiaries for the national sample. This is expected to provide more than sufficient precision for the national estimates of readiness. Exhibit 4 outlines the total sample size to be selected for the national sample (from which the comparison samples will be drawn). The targeted number of initial participants per program is roughly 2,000, but this is expected to vary. With 2,000 initial participants, it is anticipated that about 600 will complete a program and all surveys although with lower numbers of initial participants the number might go as low as 300. However, in addition to analyses of individual programs, analyses are planned after pooling data gathered from programs of a similar nature. For the purposes of power calculations we have considered power in terms of 450 full-term respondents for a single program as well as 900 and 1,200 full-term respondents, representing what might be obtained from a single large program or the pooling of two or more programs. In terms of controls we expect to obtain roughly 3,000 respondents from the initial national sample who provide data through the 12-month follow-up. Comparisons may be done with subsamples of these respondents, so we also consider 2,000 and 1,000 sample alternatives for comparison purposes. It should be noted that both unweighted and weighted analyses are planned. Unweighted analyses provide greater power while the use of sample weights in analyses provides protection against the potential for bias arising from model misspecification. The power calculations undertaken here apply to unweighted analyses.

Exhibit 5 shows the expected level of power required to detect differences of two specified sizes (.05, and .06) assuming that 450 program participants for a given wellness program completing the program and all surveys, compared to a corresponding number of controls of 3,000, 2,000, and 1,000. For differences of size .05, the power ranged from .40 to .52 while for differences of .06 the range was from .54 to .68. Exhibit 6 considers the power that might be gained from pooling multiple programs for a given set of analyses for the two sets of program participants (900 and 1,200) and three sets of controls (1,000; 2,000; and 3,000). For differences of .05 and 900 program participants, power ranged from .56 to .78 depending on the number of controls. For 1,200 program participants the range was from .63 to .86. For differences of .06 the range for 900 program participants was from .72 to .91 while for 1,200 program participants the range was from .78 to .95.

Exhibit 4. Expected sample performance, national surveys and participant surveys

Program Participants	
Expected number of enrollees annually (across all 8 programs)	16,000
Expected percentage new enrollees completing a baseline survey	0.70
Expected number of baseline survey responses	11,200
Six Month Followup Survey (all those completing a baseline)	
Percentage of new enrollees reporting program completion	0.65
Expected 6-month completion rate among program completers	0.8
Expected 6-month completion rate among program non-completers	0.65
Expected number of program completers with a 6-month survey	5,824
Expected number of program non-completers with a 6-month survey	2,548
Total number of 6-month surveys expected	8,372
Twelve Month Followup Survey* (completers with a 6-month survey response)	
Expected 12-month completion rate among program completers	0.9
Expected number of program completers with a 12-month survey	5,242
Size of Comparison Sample Needed	
Expected 12 month participants at largest site	1,500
Targeted ratio of comparison sample to program participants	1
Estimated design effect due to differential weights	2
Estimated size of 12 month comparison sample needed	3,000
Total Number to be Selected Initially for the National Sample	
Estimated size of 12-month comparison sample needed	3,000
Expected 12-month response rate	0.9
Expected total of 12-month comparison sample to be fielded	3,333
Expected percentage of six month participants sampled for the 12 month evaluation (sampling rates may vary by strata)	0.5
Expected number of completed 6-month respondents	6,667
Expected 6-month response rate	0.8
Expected number of completed baseline respondents	8,333
Expected response rate to national survey/baseline instrument	0.63
Expected number of sampled Medicare beneficiaries eligible for the study	13,227
Expected eligibility rate among sampled Medicare beneficiaries selected for the national sample	0.85
Total number to be selected for the national sample	15,562

Exhibit 5. Power calculations for comparisons between program participants and national survey comparison samples for 450 program participants and varying numbers of controls

Difference to be Detected	Program Participants		Comparison Sample		Power
	estimated proportion	n	estimated proportion	n	
0.05	0.7	50 ⁴	0.65	3,000	0.52
			0.55		0.49
			0.45		0.49
	0.6		0.65	2,000	0.50
			0.55		0.46
			0.45		0.46
	0.5		0.65	1,000	0.43
			0.55		0.40
			0.45		0.40
0.06	0.7		0.64	3,000	0.68
			0.54		0.64
			0.44		0.64
	0.6		0.64	2,000	0.65
			0.54		0.61
			0.44		0.61
	0.5		0.64	1,000	0.58
			0.54		0.54
			0.44		0.54

Exhibit 6. Power calculations for comparisons between program participants and national survey comparison samples for 900 and 1,200 program participants and varying numbers of controls

Difference to be Detected	Program Participants		Comparison Sample		Power
	estimated proportion	n	estimated proportion	n	
0.05	0.7	900	0.65	3,000	0.78
	0.6		0.55		0.74
	0.5		0.45		0.73
	0.7	1,200	0.65	3,000	0.86
	0.6		0.55		0.83
	0.5		0.45		0.82
	0.7	900	0.65	2,000	0.74
	0.6		0.55		0.69
	0.5		0.45		0.69
	0.7	1,200	0.65	2,000	0.82
	0.6		0.55		0.77
	0.5		0.45		0.77
	0.7	900	0.64	1,000	0.62
	0.6		0.54		0.57
	0.5		0.44		0.56
	0.7	1,200	0.64	1,000	0.68
	0.6		0.54		0.64
	0.5		0.44		0.63
0.06	0.7	900	0.64	3,000	0.91
	0.6		0.54		0.88
	0.5		0.44		0.87
	0.7	1,200	0.64	3,000	0.95
	0.6		0.54		0.94
	0.5		0.44		0.93
	0.7	900	0.7	2,000	0.87

	0.6		0.6		0.84
	0.5		0.5		0.84
	0.7	1,2 00	0.7	2,000	0.93
	0.6		0.6		0.90
	0.5		0.5		0.90
	0.7	9 00	0.7	1,000	0.77
	0.6		0.6		0.73
	0.5		0.5		0.72
	0.7	1,2 00	0.7	1,000	0.83
	0.6		0.6		0.79
	0.5		0.5		0.78

In-Depth Interviews with Wellness Program Staff

The respondent universe for the in-depth interviews with program staff is based on the programs participating in the study, which includes CDSMP, DSMP, EnhanceFitness, EnhanceWellness, Fit & Strong!, A Matter of Balance, Stepping On, and Walk with Ease. We anticipate approximately 3-6 staff will be interviewed per program. We will work with each program to identify the best individuals in each program capable of providing information about the key domains in the in-depth interview discussion guide. We anticipate that the Program Director or Coordinator will be able to provide most of the information about program implementation challenges and best practices, and operational costs, supplemented with in-depth interviews with trainers or instructors who provide the wellness sessions. Findings from the program staff in-depth interviews will not be statistically generalizable to the respondent universe. However, findings will be relevant to inform CMS about wellness program implementation and estimated operational costs.

2. Information Collection Procedures

National Survey of Medicare Beneficiaries

The national survey sample will be divided into 12 monthly replicates; each replicate will be handled the same way. We will first mail the questionnaire and cover letter to all sampled beneficiaries. The materials will include brief instructions in Spanish providing a phone number the beneficiary may call to request Spanish-language materials; in response to such a request, the data collection contractor will send Spanish versions of the questionnaire and cover letter, and send Spanish versions of all subsequent materials. After one week, we will either send a thank you/reminder postcard or place an automated thank you/reminder telephone call to every sampled beneficiary. Three weeks after the initial mailing, we will send a second copy of the questionnaire and a different cover letter to those sample members who have not yet returned a survey. After another two weeks, we will attempt to survey non-responders by

telephone if a number is available. We anticipate obtaining telephone numbers for about 60 percent of the sample from reverse directory services.

Data collection for the 6- and 12-month follow-up surveys will be conducted in the same way.

Wellness Program Participant Survey

The Wellness Program Participant Baseline Survey will be self-administered at the start of the wellness program. Survey materials will be distributed to eligible program participants by program staff. Spanish versions of the survey will be available upon participant request. After completing the survey, respondents will be asked to seal the survey in an envelope and return the sealed envelope to the program staff. Program staff will then return batches of completed surveys to Westat. No follow-up procedures are planned.

The Wellness Program Participant 6- and 12-month followup surveys will be delivered by mail. Data collection for the 6- and 12-month follow-up surveys will be conducted in the same way as the National Survey of Medicare Beneficiaries.

In-depth Interviews with Wellness Program Staff

In-depth interviews with staff from each wellness program will be conducted on-site or by telephone, depending on the organizational structure of each program and will take place towards the end of the first year of program participation in the study. The evaluation contractor will work with each selected program to determine the appropriate approach, particularly for those programs where the program is national in scope and implemented in several sites. We anticipate that the following types of program staff will be interviewed: Executive Director, Outreach/Recruitment Coordinator, Program Coordinator, Instructor(s), and the Financial Officer. The estimated time for each in-depth interview is between 45 and 60 minutes. In order to prepare for the in-depth interview, the evaluation contractor will request prior to the interview readily available secondary material that describes the program, including information on policies and procedures, outreach and enrollment processes, training materials or program implementation toolkits, and other descriptive information. Secondary documents that address the topics in the in-depth discussion guide will be confirmed during the in-depth interview. We will also inform the wellness program in advance of the in-depth interview that we will be discussing and seeking information about the program's operational costs in order to estimate a per-beneficiary cost for the program. The program component costs will be discussed and reviewed as part of the in-depth interview in order to ensure consistency in level of detail and definition of each cost component.

3. Methods to Maximize Response Rates

Surveys

Baseline program participant surveys

During recent discussions with potential wellness programs, CMS was advised by the wellness program experts that surveys be self-administered on-site, at the start of the wellness program. This would result in a higher response rate than surveys delivered by mail, which was originally planned. It was determined that the earlier response rate estimate for mailed surveys was lower than what will be achieved with delivery during a class session. Thus, the response rate was increased, and therefore the total number of completed baseline surveys was also increased.

Mailed Surveys - Reminders and follow up calls

Several methods will be used to maximize response rate for mailed surveys. One week following the mailing of the survey and cover letter, we will either send a thank you/reminder postcard or place an automated thank you reminder telephone call to every sampled beneficiary. Three weeks after the initial mailing, Westat will send a second copy of the survey and a different cover letter to those sample members who have not yet returned a survey. After another two weeks, Westat will attempt to survey non-responders by telephone if a number is available.

Other processes

In addition, Westat employs systematic processes for conducting mail and telephone surveys that will maximize response rate and these processes are described below:

- *Mail Surveys:* Westat will print and assemble materials for the initial mailing including the introductory letter, the questionnaire, and a postage-paid return envelope. Westat employs well-established practices that include well-organized assembly operations as well as automated processes and checks to ensure that the cover letter and mailing labels match, that the expected number of packets is assembled for mailing, and that all materials expected in the package are included. Correspondence and surveys will be generated in the same production runs and stuffed into window envelopes so that all materials go into the appropriate envelope. Westat will verify a sample of packages to ensure that materials are appropriately assembled. Our tracking system will monitor the status of each mail-out, postal non-returns, and telephone call status.

For beneficiaries for whom the project team has a telephone number, Westat will use a telephone prompt using Interactive Voice Response (IVR) Notification. Message content may vary based on how the call is answered (e.g., one message and set of options for a live answer and a different or no message for answering machines). Westat will develop and record scripts for both answering machine and live contacts, and our group will program the IVR system to deliver these appropriately. Beneficiaries without telephone numbers will be sent a thank you/reminder post card.

- *Telephone Surveys:* Westat anticipates completing about 140 telephone interviews in each of the 12 waves. The baseline survey will average about 30 minutes, and Westat will make a minimum of six call attempts per sample member.

Non-Response Analysis

For the national survey, we will have information on both responders and non-responders from Medicare enrollment and claims files. We will prepare sample weights initially calculated as the inverse of the probability of selection and then adjusted for nonresponse using available data items such as age, gender, and level of service utilization that are correlated with readiness measures. We will identify non-response adjustment cells using CHAID (a categorical search algorithm), establishing groupings (adjustment cells) showing distinct response propensities. Each grouping is a marginal or a cross-classification of variables where data are available for both respondents and non-respondents. We will calculate standard errors for estimates of readiness that include consideration of the design effect introduced by the variation in sample weights introduced either through the sample selection process or from nonresponse adjustments.

For the participant survey, we anticipate having some data on participants available from program partners, either at the individual or aggregate level. If the data are at the individual level, we will conduct nonresponse analysis and incorporate the results into the sample weighting as described for the national sample. If data are available only at the aggregate level, we will use the program figures as control totals for a post-stratification adjustment of the sample weights. Each program partner may also have multiple sites in the evaluation; each site will be treated as a sampling stratum for weighting purposes, with nonresponse adjustments and/or post-stratification performed at the stratum level.

In-depth Interview with Wellness Program Staff

The subcontracts that the evaluation contractor negotiates with each wellness program will incorporate provisions for each program to cooperate and participate in the in-depth interviews, provide readily available secondary materials on policies and procedures and aggregated financial and operational cost information, as well as provide evaluation data on program participants (e.g., participant attendance).

4. Tests of Procedures

Surveys

The data collection procedures proposed are well-tested; the data collection contractor has used similar procedures for many surveys of Medicare beneficiaries and other populations. The mail instruments will be carefully reviewed by multiple staff, and the CATI instruments will be tested to ensure that they match the mail survey insofar as feasible and that the CATI database correctly captures the reported values.

As described in Part B, Attachment 1 most of the survey items except the readiness questions have been validated and used in multiple surveys of similar populations. The newly developed survey items were reviewed by Westat's survey design team with years of experience designing Medicare beneficiary surveys new survey items were also cognitively tested with nine beneficiaries to assess the administration of the mail instrument and understanding of the readiness items. Results of the cognitive testing showed no significant issues with basic understanding of the questions, the overall ease of answering the questions, and the interpretation of "community". Results of the cognitive testing suggested two minor modifications to the newly developed readiness questions: 1) A definition of wellness program should be included to facilitate consistency in the interpretation of the term "program" across respondents. 2) A response option of "I don't need to make a change like this" should be included to capture the possibility that the change in question is not applicable for the respondent. The survey was also administered to five older adults (ages 60-75 years) to assess the time required for completion. This test indicated that the time to complete ranged from 12 to 21 minutes, with the older respondents requiring more time to complete the survey.

In-depth Interviews with Wellness Program Staff

The domains and topics in the program staff interview guide were discussed with leaders of wellness programs, and experienced wellness program researchers (described in Part A, Section 8). These experts confirmed that program would be able to speak freely during the interview and address the areas of wellness program implementation and provide information about the program operational costs. Interview guides will be refined based on feedback from early respondents.

5. Statistical Consultants

Two senior statisticians were consulted this data collection effort. Dr. Paul Zador is a Westat senior statistician and study design methodologist with expertise in propensity modeling and more than 30 years of experience. Mr. Ralph DiGaetano, a Westat senior sampling statistician, has more than 30 years of experience in survey research and data analysis, including development of sample designs, case-control studies, sampling procedures, weights, and variance estimation plans.

List of Attachments

- Part B, Attachment 1 – Sources of Survey Items