Supporting Statement B

Collections of Information Employing Statistical Methods
Evaluation of the Medicare Patient Intravenous Immunoglobulin Demonstration

The Strengthening Medicare and Repaying Taxpayers Act of 2011 mandated that the Centers for Medicare & Medicaid Services (CMS) implement a 3-year Demonstration project that introduced a bundled payment for the items and services used for in-home administration of intravenous immunoglobulin (IVIG) for Medicare beneficiaries diagnosed with primary immunodeficiency disease (PIDD). Prior to the Demonstration, Medicare only covered the costs of administering IVIG treatments in medical settings (such as hospitals and infusion clinics). CMS was also mandated to evaluate the extent to which participating in the in-home Demonstration improved access to care for beneficiaries with PIDD and to assess the market dynamics for immunoglobulin.

Eligible Medicare beneficiaries—i.e., beneficiaries receiving IVIG or subcutaneous immunoglobulin (SCIG) for the treatment of PIDD—were able to enroll in the Demonstration by submitting an application, with signed consent from their treating physician. Applications for the IVIG Demonstration began in August of 2014 and continue to be accepted. As of September 2, 2016, there were 9,300 eligible beneficiaries, of whom 1,372 were enrolled in the Demonstration.

To evaluate whether the in-home IVIG Demonstration has improved access to care, CMS plans to conduct a survey of Medicare beneficiaries eligible for the Demonstration (enrolled and unenrolled). CMS also plans to undertake a series of semi-structured interviews with physicians, nurses, other caregivers of PIDD patients, and with representatives of immunoglobulin suppliers, manufacturers, and PIDD patient advocacy groups to assess immunoglobulin market dynamics.

Survey

The primary objective of the survey is to collect information on the immunoglobulin treatment experiences of Medicare beneficiaries enrolled in the Demonstration and those that are not enrolled. The survey contains questions on the following areas that comprise a beneficiary's immunoglobulin treatment experience:

- Difficulty finding a treatment provider.
- Difficulty obtaining the correct medication.
- Missed treatments and reasons for missed treatments.
- Travel time to receive treatment.
- Treatment duration.
- Waiting time for treatment.
- Side effects of treatment.
- Changes in dosage and reasons for such changes.
- Changes in mode of treatment (e.g., from intravenous to subcutaneous) and reasons for change.
- Self-assessment of overall health.
- Exposure to advice from non-physicians regarding whether to enroll in the Demonstration.

In addition to the immunoglobulin treatment experiences of beneficiaries, the survey will also provide information on:

- Reasons for enrolling in the Demonstration.
- Reasons for not enrolling in the Demonstration.
- Reasons for not using the in-home benefit.
- General availability of immunoglobulin (anecdotal shortages of immunoglobulin were reported in the past).
- Negative health effects arising from difficulties obtaining immunoglobulin.
- Reasons for apparent increases in the use SCIG.
- Awareness of the availability of the Demonstration among eligible beneficiaries.

The primary research hypothesis of the survey is that participation in the Demonstration improves the IVIG treatment experience (e.g., time spent traveling to and waiting for treatment, side effects, health issues, changes in treatment location, etc.) of enrollees compared to those not enrolled in the Demonstration. The study's null hypotheses are:

Null Hypothesis 1: The experiences of enrollees getting their IVIG treatments at home under the Demonstration are not statistically different than the experiences of those unenrolled.

Null Hypothesis 2: The experiences of enrollees getting their IVIG treatments at home under the Demonstration are not statistically different in comparison to their IVIG treatment experiences before the Demonstration.

Interviews

The overall purpose and use of information using interviews with providers, nurses, informal caregivers and patient advocates is to provide their perspectives on the Demonstration program, including the advantages and disadvantages of different treatment options, perspectives on access, quality, cost, convenience, ease of use, safety, and health outcomes for beneficiaries using the services provided under the Demonstration. The purpose and use of interviews with manufacturers, distributors, large GPOs, and home infusion therapy companies is to assess changes to date in the current IVIG market dynamics, including IVIG supply, distribution, demand, and access.

The study protocols provide guidelines for conducting interviews in order to increase consistency and reliability of the findings. The semi-structured interview guides include a list of open-ended questions that are asked of all interviewees. Additional follow-up interview questions and probes will be asked to clarify or to provide additional information based on the participants' response. This approach facilitates targeted interviews that can be easily analyzed and compared across while allowing for new or unexpected information to be provided by participants, thereby retaining the flexibility needed to delve deeper into the respondents' opinions. With an emergent and iterative design, our research is exploratory and we will continue to update our interview question probes to follow-up on any unexpected trends and insights. We will also validate findings and themes identified during early interviews in our later interviews through member checking, a method of soliciting feedback about particular findings from other members of the groups participating in interviews (Maxwell, 2005; Grossoehme, 2014).

1   Potential Respondent Universe and Sampling Method

1.1   Potential Respondent Universe

1.1.1   Survey

The universe for the survey is all Medicare beneficiaries with PIDD who are receiving immunoglobulin (based on paid historical Medicare claims) defined by the specific codes shown in Appendix C. The universe is composed of:

- **Group 1**: Those who have submitted an application to enroll in the Demonstration whose IG providers have submitted one or more claims[1] under the Demonstration (approximately 974 beneficiaries, as of September 2, 2016).
- **Group 2**: Those who have not enrolled in the Demonstration (approximately 7,928 beneficiaries, as of September 2, 2016).

Additionally, the universe includes 398 beneficiaries (as of September 2, 2016) who have enrolled in the Demonstration, but have not yet utilized the in-home benefit (Group 3).

1.1.2   Interviews

The potential respondent universe includes physicians, nurses, caregivers and advocates for patients with PIDD receiving IVIG; and IVIG manufacturers, distributors, Group Purchasing Organizations (GPOs), and infusion companies. To identify physicians, nurses, and caregivers for Medicare beneficiaries with PIDD receiving IVIG, we will conduct purposive and snowball convenience sampling to identify respondents in these groups, initially based on recommendations of physicians.

For patient advocates, the universe is all patient advocacy groups who are focused on patients with PIDD. We will conduct purposive and convenience sampling to identify respondents in these groups.

For interviews with manufacturers, distributors, Group Purchasing Organizations (GPOs), and infusion companies, the universe is all of the manufacturers and primary and secondary distributors of IVIG, the largest GPOs, and home infusion therapy companies. We will conduct purposive and convenience sampling to identify respondents in these groups.

1.2   Sampling Method

1.2.1   Sample Frame

1.2.1.1   *Survey*

The sample frame for the survey will be Medicare claims data for beneficiaries with PIDD who are receiving IVIG; including Demonstration participants who used the in-home benefit (Group 1) and non-participants (Group 2). The frame will also include those Medicare beneficiaries who have enrolled in the Demonstration but did not use the in-home benefit (Group 3), which constitute 29 percent of those enrolled in the Demonstration (398 out of 1,327). The data collected from this group will yield pertinent information on the reasons for non-use, such as inability to find an in-home IVIG service provider.

---

[1] A total of 12,932 claims had been submitted for IVIG provided to these beneficiaries under the Demonstration as of September 2, 2016.

Survey respondents will self-identify which group they are a member of using the survey instructions and screening questions.

1.2.1.2 *Interviews*

For interviews, we will conduct non-probability purposive sampling based on selection criteria. This sampling method provides advantages for gaining insights for this demonstration project due to 1) simplicity and ease of sampling, 2) capturing basic trends to support further research, 3) rapid data collection and analysis and, 4) low cost and ease of implementation.

We will use Medicare claims data to identify physicians who treat beneficiaries with PIDD who are receiving IVIG and either participating or not participating in the IVIG Demonstration. The criteria for selecting providers includes the volume of IV/SCIG administrations in a typical month, type of setting, geographic region, and urban/rural location to achieve maximum variation in the sample in order to provide a balanced and representative perspective of the program (Harris, et al., 2009).

To identify nurse respondents, we will ask physicians and participating organizations during the physicians' interviews to recommend nurses who are directly involved with the care of patients receiving IG. We will also contact provider organizations identified in the claims data that were not selected for physician interviews to identify additional nurses to participate in the interviews if needed. Snowball sampling will be used to identify informal caregivers based on recommendations from physicians and nurses.

Caregivers of beneficiaries will be selected via purposive/snowball sampling who fit the inclusion criteria based on physician and nurse recommendations.

Patient advocates will be selected based on publicly available data and inquiries.

Medicare claims data will be used to identify interviewees from home infusion companies. We will also select interviewees from a list of manufacturers and distributors who participated in the 2016 IVIG Access Demonstration Webinar and the list of participating manufacturers and distributers from the previous ASPE report.

1.2.2 Sample Allocation

1.2.2.1 *Survey*

All beneficiaries in Group 1 and Group 3 will be targeted for the survey. A stratified random sample of Group 2 beneficiaries will be targeted for the survey.

Table 1 shows the total number of beneficiaries, sample size, expected response rate, and number of completes for each survey estimation cell (i.e., Group 1, Group 2, and Group 3), and the derivation of these numbers is discussed below.

Table 1: Survey Universe

Respondent Group	Number of Beneficiaries	Sample Size	Number of Completes	Expected Response Rate [a]
Group 1: Enrolled beneficiaries with one or more Demonstration claims	974	974	390	40%
Group 2: Non-enrolled beneficiaries	7,928	1,398	559	40%

Respondent Group	Number of Beneficiaries	Sample Size	Number of Completes	Expected Response Rate [a]
Group 3: Enrolled beneficiaries with no Demonstration claims	398	398	159	40%
Total	**9,300**	**2,770**	**1,108**	**40%**

[a] Based on Klein, et al., 2011; Centers for Medicare and Medicaid Services, 2015; Amaya, et al., 2015; National Research Council, 2015; and Thorpe, et al., 2015.

Sample Allocation for Groups 1 and 3

All beneficiaries in Group 1 and Group 3 will receive the survey.

Sample Allocation for Group 2

A random sample stratified by geographic region will be used for Group 2. The Group 2 sample will be sufficiently large to yield statistically valid estimates of beneficiary experience with +/- 4 percent margin of error, e, at a 95 percent confidence level (i.e., α =5 percent). The desired sample size for Group 2, $n_{Group\,2}$ is calculated as (Stat Trek, 2015):

$$n_{Group\,2} = \frac{\left(z_{\alpha/2}^2\right)\left(p_{Group2}\right)\left(1-p_{Group2}\right)+e^2}{e^2 + \dfrac{z_{\alpha/2}^2\left(p_{Group2}\right)\left(1-p_{Group2}\right)}{N_{Group2}}}$$

where $n_{Group\,2}$ is the desired sample size; z is the critical value (or z score) associated with the desired confidence level α; e is the margin of error; $p_{Group\,3}$ is the response distribution; and $N_{Group\,2}$ is the population size.

Assuming that the proportion of those beneficiaries in Group 2 that report having experienced IVIG access problems is 50 percent (i.e., $p_{Group\,2} = 0.5$),[2] the desired sample size for Group 2 based on the above equation and our statistical precision target is:

$$n_{Group\,2} = \frac{(1.96)^2(0.50)(1-0.50)+(0.04)^2}{(0.04)^2 + \dfrac{(1.96)^2(0.50)(1-0.50)}{7,928}} = \frac{0.9620}{0.00172114} = 559$$

Given that our expected response rate to the survey is 40 percent for all three groups, 559 completes requires a target sample of 1,398 (= 559 ÷ 0.40) beneficiaries for the Group 2 survey estimation cell.

Minimum Sample Size Needed for Group 2

While we plan to sample 1,398 members of Group 2 in order to achieve 559 completes, this section explores the *minimum* sample size necessary to achieve a desired power and effect size for hypothesis testing. For most surveys, 80 percent power and 20 percent effect size are typical assumptions used for these calculations. Note that the minimum sample size calculation presented below is not relevant for those enrolled in the Demonstration, i.e., those in Groups 1 and 3, because we plan to use a census approach rather than a sampling approach for those groups.

---

[2] The assumption yields the maximum sample size estimate.

For beneficiaries in Group 2, let p_i be the proportion of beneficiaries enrolled in the Demonstration who indicate having experienced IVIG access problems in the survey. Further assume that $p_i = p_{Group\,1} = p_{Group\,3}$ for simplicity. Then the sample size needed to compare a proportion of beneficiaries not enrolled in the Demonstration ($n_{Group\,2}$) to p_i will be given by

$$n_{Group\,2} = \left(\frac{Z_{1-\alpha/2} + Z_{1-\beta}}{ES} \right)^2$$

where

$$ES = \frac{\left| p_{Group\,2} - p_i \right|}{\sqrt{p_i(1-p_i)}}$$

and $p_{Group\,2}$ is the proportion of beneficiaries not enrolled in the Demonstration who indicate having experienced IVIG access problems in the survey. If we assume that the proportion of those beneficiaries in both Group 1 and Group 3 populations that report having experienced IVIG access problems is

$$p_{Group\,1} = p_{Group\,3} = 0.50$$

A 20 percent effect size, $ES = 0.20$ (i.e., ability to detect a 20 percent difference in the proportion of beneficiaries experiencing IVIG access problems between the Group 2 and Group 3 populations or between the Group 1 and Group 2 populations) implies:

$$0.20 = \frac{\left| p_{Group\,2} - p_i \right|}{\sqrt{p_i(1-p_i)}} = \frac{\left| p_{Group\,2} - 0.50 \right|}{\sqrt{0.50(1-0.50)}} = \frac{\left| p_{Group\,2} - 0.50 \right|}{\sqrt{0.25}}$$

$$p_{Group\,2} = 0.20 \times 0.50 + 0.50$$

$$p_{Group\,2} = 0.60$$

Given

$$\alpha = 0.05$$

$$\beta = 0.80$$

The required minimum sample size for the unenrolled beneficiary group (Group 2) will be[3]

$$n_{Group\,2} = \left(\frac{Z_{1-\alpha/2} + Z_{1-\beta}}{ES} \right)^2 = \left(\frac{1.96 + 0.84}{0.20} \right)^2 = 194$$

Note that different assumptions about the proportion of enrolled beneficiaries (i.e., groups experiencing IVIG access problems) will lead to different minimum sample size estimates for the unenrolled beneficiaries (Group 2) even if the power, significance, and effect size figures are unchanged. Below (Table 2) we present the power afforded by different sample sizes for the unenrolled group, Group 2, at different effect size, *ES*, levels. Eighty percent power and 20 percent effect size are typical standards

---

[3] Note that the actual calculations are based on unrounded numbers.

used in similar surveys, which imply a minimum sample size of 194 for this survey. However, our proposed sample size of 561 completes affords an effect size of 15 percent at 95 percent power, far exceeding these typical standards.

Table 2: Power Associated with Different Size Samples for Group 2 at Varying Effect Sizes (*ES*)

Sample Size [a]	Power					
	ES = 10%	*ES* = 15%	*ES* = 20%	*ES* = 25%	*ES* = 30%	*ES* = 35%
50	10%	18%	29%	42%	57%	71%
75	14%	25%	41%	58%	75%	87%
100	17%	32%	52%	71%	86%	95%
125	20%	39%	61%	81%	93%	98%
150	23%	45%	69%	87%	96%	99%
175	26%	51%	76%	92%	98%	100%
194 [b]	**28%**	**55%**	**80%**	**94%**	**99%**	**100%**
200	29%	56%	81%	95%	99%	100%
225	32%	62%	86%	97%	100%	100%
250	35%	66%	89%	98%	100%	100%
275	38%	70%	92%	99%	100%	100%
300	41%	74%	94%	99%	100%	100%
325	44%	77%	95%	100%	100%	100%
350	46%	80%	97%	100%	100%	100%
375	49%	83%	97%	100%	100%	100%
400	52%	85%	98%	100%	100%	100%
425	54%	87%	99%	100%	100%	100%
450	56%	89%	99%	100%	100%	100%
475	59%	91%	99%	100%	100%	100%
500	61%	92%	99%	100%	100%	100%
525	63%	93%	100%	100%	100%	100%
550	65%	94%	100%	100%	100%	100%
575	67%	95%	100%	100%	100%	100%
600	69%	96%	100%	100%	100%	100%
625	71%	96%	100%	100%	100%	100%
650	72%	97%	100%	100%	100%	100%

[a] This represents the sample size needed for the second group, i.e., beneficiaries not enrolled in the Demonstration.
[b] This is the desired sample that would yield 80 percent power at a 20 percent effect size level, as noted in the earlier discussion.

1.2.2.2 *Interviews*

A purposive sampling method will be used to identify participants for interviews. We propose potential sample sizes to provide a diverse cross-section of participants and to reach data saturation: when no new information, trends, codes, or themes are gained from additional interviews (Fusch & Ness, 2015; Guest, et al., 2006). We will oversample caregivers who are caring for beneficiaries who are participating in the Demonstration compared to caregivers who are caring for eligible beneficiaries who are not participating in the Demonstration at a rate of 2 to 1. Maximum sample size for each group of participants is provided below:

- 36 physicians: 24 physicians treating beneficiaries participating in the Access Demonstration and 12 physicians treating beneficiaries with PIDD receiving IVIG not participating in the Demonstration.

- 60 nurses: 40 nurses treating beneficiaries participating in the Access Demonstration and 20 nurses caring for beneficiaries with PIDD receiving IVIG not-participating in the Demonstration
- 36 informal caregivers: 24 caregivers providing care for beneficiaries who are participating in the Access Demonstration and 12 caregivers providing care for beneficiaries who are not-participating in the Demonstration
- 6 patient advocate representatives
- 33 manufacturers, distributors, and/or providers: 9 IVIG manufacturers, 9 primary or secondary distributors of IVIG, 9 of the largest GPOs, and 6 home infusion therapy companies

1.2.3   Expected Response Rate

1.2.3.1   *Survey*

To estimate respondents' responsiveness to the survey, we examined the rates of response to other surveys similar in length, mode(s) of administration, and population sampled. Although Medicare has not previously performed a survey of beneficiaries with PIDD, Medicare does administer the Medicare Consumer Assessment of Healthcare Providers and Systems (MCAHPS) Survey. This survey is directed at beneficiaries using Medicare fee-for-service, Medicare Advantage, and the Medicare Prescription Drug Plan. The MCAHPS is similar in length (13 pages), format, purpose, and target population to our current survey. The 2007 MCAHPS reports an overall response rate of 49 percent (Klein, et al., 2011). In contrast, the 2014 Hospital CAHPS survey, which includes, but is not limited to, Medicare beneficiaries, reports a response rate of 30 percent (Centers for Medicare and Medicaid Services, 2015). Based on: 1) these two reported response rates, 2) the consideration of the general downward trend in survey responsiveness over the last 10-15 years (Amaya, et al., 2015; National Research Council, 2015; Thorpe, et al., 2015), and 3) the possibility that beneficiaries with a chronic disease might be, on average, less likely to respond to a voluntary survey than the overall Medicare population (as sampled by MCAHPS), we estimated the response rate to this survey at 40 percent (i.e., the average of the two reported response rate estimates, 49 percent and 30 percent).

1.2.3.2   *Interviews*

Based on prior research and experience conducting qualitative interviews with providers and nurses, response rates are estimated to be about 60 percent (Pit, et al., 2014). Interviews with caregivers may have a lower response rate, ranging from 19-60 percent (Fowler Jr., et al., 2002; Seow, et al., 2016). We expect a high response rate from patient advocacy groups based on their investment in this topic and knowledge of the Demonstration. Based on previous interviews with manufacturers and distributors for the ASPE report, we expect a response rate of 100 percent.

2   Procedures for the Collection of Information

2.1   Statistical Methodology for Stratification

2.1.1   Survey

The survey will target all beneficiaries in Groups 1 and 3. Hence, sample stratification is not relevant for Groups 1 and 3. A statistical method for stratification will be used for Group 2 based on geographical regions.

We will conduct proportional stratified random sampling among the ten Health and Human Services (HHS) regions where beneficiaries with PIDD reside.

- Region I: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
- Region II: New Jersey, New York, Puerto Rico, Virgin Islands
- Region III: Delaware, District of Columbia, Maryland, Pennsylvania, Virginia, West Virginia
- Region IV: Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee
- Region V: Illinois, Indiana, Michigan, Minnesota, Ohio, Wisconsin
- Region VI: Arkansas, Louisiana, New Mexico, Oklahoma, Texas
- Region VII: Iowa, Kansas, Missouri, Nebraska
- Region VIII: Colorado, Montana, North Dakota, South Dakota, Utah, Wyoming
- Region IX: Arizona, California, Hawaii, Nevada (American Samoa, Guam, Northern Mariana Islands, Trust Territory of the Pacific Islands)
- Region X: Alaska, Idaho, Oregon, Washington

The design reflects simple proportionate sampling such that the sample size of each Group 2 geographic region stratum is proportional to the size of the universe for that stratum. In other words, if a given geographic region stratum contains 20 percent of all unenrolled beneficiaries (Group 2) in the study universe, the sample size for that stratum will account for 20 percent of our overall sample size for Group 2. We will combine those geographic regions in which there is an insufficient number of Group 2 beneficiaries.

Assuming a response rate of 40 percent and a statistical precision target of +/- 4 percent margin of error at 95 percent confidence level, we will use stratified random sampling by geographic region to select 1,403 beneficiaries for Group 2.

2.1.2 Interviews

No statistical methodology for stratification nor sample selection will be used.

2.2 Statistical Methodology for Sample Selection

2.2.1 Survey

The survey will target all beneficiaries in Groups 1 and 3. Hence, sample selection considerations are not relevant for these populations.

The statistical method for selecting Group 2 beneficiaries within each geographic region stratum will involve assigning each beneficiary a random index number, using a random number generator. The beneficiaries in each stratum will then be arranged in ascending order according to their random index number. If S_j is the size of the solicited sample in the jth stratum, then those S_j beneficiaries with the smallest index numbers will be selected and included in the sample.

2.2.2 Interviews

No statistical methodology will be used for the sample selection for the interviews.

2.3   Estimation Procedure

2.3.1   Analytic Methods

2.3.1.1   *Survey*

Survey data will be collected and maintained using an online survey system (Qualtrics). Final survey data will be downloaded in comma-delimited format for data cleaning and analysis. We will perform data cleaning and descriptive analysis in SAS v.9, and text analysis (for those questions that require verbatim responses) in MS Excel.[4]

Using the survey algorithms in SAS v.9 (e.g., PROC SURVEYFREQ, PROC SURVEYMEANS, etc.), the data analysis to be conducted will involve:

- A non-response bias analysis using variables such as age, gender, HHS region, and dual eligibility status to assess any non-response bias (i.e., whether and how the non-respondents are different than the respondents).
- For each respondent, computation of:
- Simple weights which are the inverse of the selection probability multiplied by the probability of response in the absence of non-response bias, or
- Adjusted weights that account for non-response bias using the variables age, gender, HHS region, and dual eligibility status, if determined to influence response based on the findings of the non-response bias analysis.
- Tabulating weighted proportions and corresponding standard errors for each survey question in groups 1, 2, and 3 (e.g., weighted proportion of respondents who responded "Yes," "No," or "Don't Know" for a given survey item).
- Testing to see if there are statistically significant differences in responses to each survey item among groups 1, 2 and 3.

2.3.1.2   *Interviews*

Interviews will be conducted by telephone and digitally audio-recorded with the permission of the interviewee. Interviews will be conducted by an experienced facilitator and a note-taker to capture responses should the interviewee decline to be audio-taped. Audio-recordings will be transcribed and de-identified. We have developed a preliminary coding structure based on previous literature, patient surveys, and pilot tests of the interviews including conceptual, participant, relationship, and setting codes (Bradley, et al., 2007). Following the transcription of interviews, we will update our coding structure to reflect information collected during the interviews. Content and thematic analyses will be used to identify key findings and to make comparisons across interviewees' responses. Our team will use NVivo, a computer software package, for qualitative data management and analysis. NVivo qualitative software allows our researchers to examine all of the text that is presented in the transcripts, to identify excerpts that contain content meaningful to the research questions, and – finally – to apply any number of appropriate code(s) to the excerpt. The themes uncovered via excerpting and coding become the framework that we use to understand how all themes and concepts are related to each other and to the overarching evaluation objectives.

---

[4] Text analysis will involve a review and analysis of the verbatim responses to those questions that include an "Other – Please Specify" response category.

2.3.2 Simple Weights

2.3.2.1 *Survey*

Each respondent to the survey will be assigned a weight based on the inverse of the selection probability of the respondent's corresponding stratum multiplied by the probability of response. Below we discuss the method we will use in computing simple weights for respondents in each survey estimation cell (i.e., Group 1, Group 2, and Group 3) that account for probability of selection and response, but do not incorporate the possibility of non-response bias. Thus, the derivation of these simple weights is based on the assumption that there are no significant differences with respect to such factors as dual eligibility status, age, race, original reason for enrolling in Medicare, between respondents and non-respondents to the survey in any of the survey estimation cells. Weights that deal with the possibility of non-response bias are discussed in Section 3.2 below.

For survey estimation cell, i (where i = Group 1, Group 2, and Group 3), the probability of selection, $P^i_{S,j}$ for the jth geographic region is given by:

$$P^i_{S,j} = \frac{S^i_j}{U^i_j}$$

where U^i_j is the number of beneficiaries in Group i and geographic region j; S^i_j is the size of the solicited sample in Group i and geographic region j. Because we will select all beneficiaries in Groups 1 and 3 for the survey, the size of the solicited sample in geographic region j will be equal to the number of beneficiaries in the same geographic region, i.e.:

$$S^i_j = U^i_j$$

for i = Group 1 and Group 3. Therefore,

$$P^i_{S,j} = \begin{cases} 1 & i = \text{Group 1} \\ \dfrac{S^i_j}{U^i_j} & i = \text{Group 2} \\ 1 & i = \text{Group 3} \end{cases}$$

Additionally, for survey estimation cell, i, the probability of response, $P^i_{R,j}$, for the jth geographic region can be calculated by dividing the solicited sample size in each stratum by the actual number of responses from the corresponding stratum, i.e.:

$$P^i_{R,j} = \frac{R^i_j}{S^i_j}$$

where S^i_j is the size of the solicited sample in Group i and geographic region j; R^i_j is the actual (responded) sample in Group i and geographic region j. Then the simple sample weights, W^i_j for Group i and geographic region j are computed as:

$$W^i_j = \frac{1}{P^i_{S,j} \times P^i_{R,j}}$$

where the terms are as defined above. Since for Group 1 and Group 3,

$$P_{S,j}^{Group\,1} = P_{S,j}^{Group\,3} = 1$$

the simple weights for each survey estimation cell will be:

$$W_j^i = \begin{cases} \dfrac{1}{P_{R,j}^i} & i = \text{Group 1} \\ \dfrac{1}{P_{S,j}^i \times P_{R,j}^i} & i = \text{Group 2} \\ \dfrac{1}{P_{R,j}^i} & i = \text{Group 3} \end{cases}$$

2.3.2.2 *Interviews*

This does not apply to interviews.

2.3.3 Degree of Accuracy Needed for the Purpose Described in the Justification

2.3.3.1 *Survey*

The accuracy required of the respondents poses no special demands on them. All data being requested can be readily supplied by respondents. The sample size was calculated to enable us to generate weighted sample estimates of proportions of interest in each group in the +/- 4 percent range of the true proportion with 95 percent confidence (i.e., $\alpha = 5$ percent).

2.3.3.2 *Interviews*

The accuracy required of the respondents poses no special demands on their behalf. All data being requested can be readily supplied by respondents.

2.4 Unusual Problems Requiring Specialized Sampling Procedures

2.4.1 Survey

There are no unusual problems anticipated.

2.4.2 Interviews

There are no unusual problems anticipated.

2.5 Use of Periodic (Less Frequent than Annual) Data Collection Cycles to Reduce Burden

2.5.1 Survey

This is a one-time data collection, which will minimize the burden on survey respondents.

2.5.2   Interviews

This is a one-time data collection, which will minimize the burden on interviewees.

3   Methods to Maximize Response Rates and Deal with Issues of Non-Response

3.1   Methods to Maximize Response Rates

3.1.1   Survey

The survey will be implemented both by mail and on line. Survey respondents will receive a pre-notification letter from CMS describing the survey and providing each respondent with the URL of the online survey and their unique username and password. A week later, respondents who have not submitted an online survey will receive the full survey package (see Table 3). This package will include a cover letter from CMS, a stamped return envelope, and two survey booklets—one survey for enrollees and one for non-enrollees. Online participation will again be encouraged.

Table 3: Overview of Data Collection Steps to Maximize Response Rates

Contact	Contact Type (Date)
Initial Contact	Mail / Survey Cover Letter (i.e., pre-notification letter) / Survey Link / Incentive
First Reminder	Mail / Survey Reminder Cover Letter / Survey Hardcopy / Survey Link (1-2 weeks after initial mailing)
Second Reminder	Postcard / Survey Link (1-2 weeks after first reminder)
Third Reminder	Mail / Survey Reminder Cover Letter / Survey Hardcopy / Survey Link (1-2 weeks after second reminder)
Fourth Reminder	Reminder phone call with an option to complete the survey via phone (1-2 weeks after third reminder)

Table 3 illustrates the multiple strategies we will employ to maximize response rates, including multiple contacts (i.e., an initial contact and several reminders), pre-notification letters, the availability of a Spanish-language survey, multiple modes of administration, and incentives. Text of the notifications and reminders are provided in the Appendix B.

Simplification of survey into two forms: To minimize non-response rates we will include two separate survey booklets in each mailed survey package—one booklet for enrollees and one for non-enrollees.

Multiple contacts. In this data collection, we plan to follow the Dillman Total Design survey method (Dillman, et al., 2014), which emphasizes multiple contacts with members of the sample as being one of the most successful techniques to increase response rates. This technique is now considered standard methodology for any survey. In this survey, we will use a pre-notification message, which also includes the survey link, followed by a copy of the survey with a cover message (a reminder sent to all respondents shortly after they receive the copy of the survey), followed, finally, by one or more contacts with non-respondents using a combination of messages (post card) and complete survey packages. Phone calls to non-responsive beneficiaries will only be made as a fourth reminder, four to eight weeks after the initial contact.

Pre-notification letters/emails that provide more information on the study increase respondent confidence in the validity and the importance of the study resulting in higher response rates. As such, we will send out pre-notification letters as part of this data collection effort.

Translation availability. The survey package and accompanying mailings will include text in Spanish that provides the phone number to request a Spanish version of the letter and survey.

Multiple mode administration (phone and mail, mail and Web, etc.) of a survey has been shown to increase response rates (Dillman, et al., 2014). Additionally, the use of multiple modes can also reduce non-response error and data collection costs. In this survey, respondents will be offered the option of completing the survey on-line.

Incentives. All survey respondents will receive a $2 cash incentive enclosed in the survey package as a means to increase their likelihood of response. Although some earlier studies reported ambiguous results from enclosing small, noncontingent (i.e., the incentive is not contingent on participation) cash incentives with mailed surveys, more recent studies and meta-analyses have refined this view. Mercer et al. (2015) analyzed results from over 40 studies and reported that surveys with a $2 noncontingent incentive resulted in a 10 percent higher response rate than those with no incentive. Parsons and Manierre (2014), in their study of incentives, response rates, and nonresponse bias, referenced Millar and Dillman (2011), who found that a mailed survey notification with a $2 incentive elevated response rates to a web survey among college students from 21.2 percent (notification without incentive) to 38.2 percent. Parsons and Manierre further stated: "Overall, the research in this area suggests that unconditional incentives reduce nonresponse bias." Gneezy and Rey-Biel (2014) assessed the impact of incentives in increments from zero to $30 and found that a $1 incentive doubled the response rate obtained without incentive (i.e., a 100 percent increase), and that a $2 incentive increased the no-incentive rate by over 140 percent.

Since widely accepted data collection techniques are being used and substantial resources are being devoted to minimize non-response, we expect the response rate to this survey to be comparable or better than that achieved for surveys of similar size and scope.

3.1.2   Interviews

To maximize response rates of interviews, we will conduct pre-notification phone calls and schedule interviews at times that are convenient to participants, including early morning or evening phone interviews. Additionally, we will use multiple contacts to recruit and follow-up with participants (Pit, et al., 2014; VanGeest & Johnson, 2011).

3.2   Methods to Deal with Issues of Non-Response

3.2.1   Survey

Potential reasons for non-response include refusal, health status, language barrier, and other circumstances, as well as the inability to contact the respondent. After the survey has been conducted, we will perform an analysis of non-response bias in the survey estimates. If non-response bias is detected, we will create adjusted weights based on age, gender, race, HHS region, dual eligibility, original reason for Medicare, and Medicare Advantage enrollment.

Using standard procedures, we will first construct a logistic model of the propensity for survey completion based on the following exogenous variables available for each target respondent (Lohr, 1999; Abraham, et al., 2006):

- Age,
- Gender,
- Race,

- HHS region
- Dual eligibility,
- Original reason for Medicare (aged, disabled, ESRD, combination), and
- Medicare Advantage Enrollment (on a monthly basis).

The general form of the logistic function (omitting the group superscripts for simplicity) is expressed as,

$$P_R = \frac{e^{x\beta}}{\left(1+e^{x\beta}\right)} = \frac{1}{\left(1+e^{-x\beta}\right)}$$

where P_R in this context is the probability of a respondent completing the survey and x and β are the vectors of explanatory variables (e.g., age, gender, race, HHS region and location, dual eligibility status, etc.) and their respective coefficients. Given the above equation, the probability of survey nonresponse can be written as,

$$1 - P_R = 1 - \frac{1}{\left(1+e^{-x\beta}\right)} = \frac{1}{\left(1+e^{x\beta}\right)}$$

The odds of a positive survey response are, therefore,

$$\left(\frac{P_R}{1-P_R}\right) = e^{x\beta}$$

Taking the natural log of both sides, the above equation becomes,

$$\ln\left(\frac{P_R}{1-P_R}\right) = x\beta$$

For the purposes of nonresponse analysis for this survey, the logit model to be estimated can be specified as,

$$\ln\left(\frac{P_{R,k}}{1-P_{R,k}}\right) = a + b_1 x_{1k} + b_2 x_{2k} + \ldots + b_n x_{nk} + \epsilon_k$$

where $P_{R,k}$ is probability of response for target respondent k; a is the intercept term; b_i are the associated coefficient vectors for explanatory variables (e.g., age, gender, race, HHS region and location, dual eligibility status, etc.); ϵ_k is the error term; and x_i are vectors of dichotomous dummy variables from the sampling frame corresponding to target respondent k.

Using maximum likelihood methods, we will estimate the above logistic relationship for target respondents and determine which, if any, estimated coefficients, are statistically significant. If none of the coefficient estimates are statistically significant, no adjustments to weights would be necessary as this would indicate lack of non-response bias. On the other hand, if some or all coefficient estimates are found to be significantly related to the probability of responding to the survey, then it will be necessary to adjust the weights for non-response.

Given the above regression model, the predicted probability of a positive survey response for a given potential respondent k will be calculated as:

$$\widehat{P_{R,k}} = \frac{1}{\left(1 + e^{-\left(\hat{a} + \hat{b}_1 x_{1k} + \hat{b}_2 x_{2k} + \ldots + \hat{b}_n x_{nk}\right)}\right)}$$

We will then use these predicted probability estimates to recalculate the nonresponse bias adjusted weight to be applied to each respondent's responses. The nonresponse adjusted weights can be expressed as follows:

$$\widehat{\widehat{W}}_{jk} = \alpha_j \left(\frac{1}{P_{S,j} \times \widehat{P_{R,j}}}\right)$$

where $\widehat{\widehat{W}}_{jk}$ is the adjusted weight for respondent k in geographic region j; $\widehat{P_{R,j}}$ is the estimated response probability for a respondent derived from the logistic regression, and α_j is the normalization factor. These factors are calculated to normalize the estimated response probabilities so that the set of nonresponse adjusted weights have the following property for each stratum within a given estimation group:

$$U_j = \sum_k \widehat{\widehat{W}}_{jk} R_{jk} = \alpha_j \sum_k \left(\frac{1}{P_{S,j} \times \widehat{P_{R,j}}}\right) R_{jk}$$

where k is summed over all respondents in stratum j within a given survey estimation cell (i.e., Group 1, Group 2, and Group3).

Depending on the results of the above analysis, we will also consider using the multivariate regression-based imputation approach, to impute estimated values for non-respondents to address non response bias.

3.2.2   Interviews

We will keep the phone interviews brief to minimize non-response. We will monitor response rates of all of the interview groups and assess the diversity of the sample according to our pre-determined selection criteria. We will continue to follow-up with groups of participants as needed to improve response rates. Additional methods to deal with non-response are not applicable.

3.3   Generalizing to the Universe Studied

3.3.1   Survey

Because we will target all of the beneficiaries in Groups 1 and 3 and a stratified random sample of beneficiaries in Group 2, we expect that the information collected will yield reliable data that can be generalized to the universe studied.

3.3.2   Interviews

Purposive convenience sampling limits generalizing to the universe, so purposely selecting participants to provide maximum variation in key characteristics will help improve the generalizability of the findings. Although the goal of qualitative interviews is not statistical generalizability, we expect the results will provide valuable insights to supplement the beneficiary survey and secondary data analyses of

Medicare claims data. There is a point of diminishing return to a qualitative sample—as the study goes on, more data does not necessarily lead to more information. This is because one occurrence of a piece of data, or a code, is all that is necessary to ensure that it becomes part of the analysis framework. Frequencies are rarely important in qualitative research, as one occurrence of the data is potentially as useful as many in understanding the process behind a topic (Polit & Beck, 2010). We will continue interviews until we reach a point of saturation, which will indicate that when the collection of new data does not shed any further light on the issues under investigation (Green & Thorogood, 2009).

4 Test of Procedures or Methods

4.1 Survey

As part of developing the mail and online survey instruments, the project team has conducted cognitive testing to get initial feedback on respondents' understanding of questions, consistency in interpreting questions and response options, ability to recall necessary information, how well the items reflect the measurement domains, and the flow of the survey tools and interviews.

We first beta-tested the survey instruments with two contractor employees. Employees were provided with fictional "bios" providing made-up medical information consistent with the respondent universe, and were asked to refer to these bios when answering the survey items. This procedure more closely approximated the time that an actual respondent might require to complete the survey. On average, they required less than 30 minutes to complete the survey. For burden estimates, we assume that the survey will require 30 minutes to complete, whether in paper or online form.

We additionally conducted cognitive testing of the beneficiary surveys with up to nine members of the universe studied (Medicare beneficiaries with PIDD who receive IG, including both those enrolled and not enrolled in the Demonstration). In these interviews, respondents provided valuable feedback on how to improve question wording, simplify skip patterns, and otherwise make the elements of the survey package more interpretable. Based on respondent feedback during cognitive testing, we revised the survey to improve the questions – make them easier to comprehend and reduce the complexity of skip patterns.

4.2 Interviews

We modified previous used data collection instruments for the ASPE 2007 study based on an updated literature review. The content validity of the interview guides was pilot tested using two physicians and two nurses knowledgeable of IVIG and SCIG treatment. Furthermore, we conducted pilot tests of the interviews with a physician, nurse, and informal caregiver (a total of three participants) for interview timing, flow, structure, and clarity of questions and responses. We made revisions based on these pilot-tests to minimize burden and improve utility of the interview guides.

5 Individuals Consulted on Statistical Aspects of Design

Table 4 below provides the names, affiliation, and contact information for those consulted on the statistical aspects of the design and who will collect or analyze the information.

Table 4: Individuals Consulted on Statistical Aspects and Performing Data Collection & Analysis

Name	Affiliation	Contact Information
Allen Dobson, Ph.D.	Dobson DaVanzo & Associates, LLC	703-260-1762
Joan DaVanzo, Ph.D., M.S.W.	Dobson DaVanzo & Associates, LLC	703-260-1761
Audrey El-Gamil	Dobson DaVanzo & Associates, LLC	703-260-1764
Alexandra Collins	Dobson DaVanzo & Associates, LLC	703-722-8948

Name	Affiliation	Contact Information
Nikolay Manolov	Dobson DaVanzo & Associates, LLC	703-260-1768
Aylin Sertkaya, Ph.D.	Eastern Research Group, Inc.	781-674-7227
Carlie Knope	Eastern Research Group, Inc.	781-674-7343
Andreas Lord	Eastern Research Group, Inc.	781-674-7381
Elyse Levine, Ph.D.	Booz Allen Hamilton	240-453-5387
Anna Ettinger, Ph.D., M.S.W., M.P.H.	Booz Allen Hamilton	814-404-9315
Dimitrios Koutsonanos, M.D.	Booz Allen Hamilton	404-589-5181
Ada-Helen Volentine, Ph.D.	Booz Allen Hamilton	434-995-2045

Table 5 shows the name of CMS staff who advised on design.

Table 5: CMS Staff who advised on Design

Name	Affiliation	Contact Information
Pauline Karikari-Martin, PhD, MPH, MSN	Center for Medicare and Medicaid Innovation	410-786-1040

6   References

Abraham, K. G., Maitland, A. & Bianchi, S. M., 2006. Nonresponse in the American Time Use Survey: Who Is Missing from the Data and How Much Does It Matter?. *Public Opinion Quarterly,* 70(5), pp. 676-703.

Amaya, A. F., Leclere, K. C. & Liao, Y., 2015. Where to start: An evaluation of primary data-collection modes in an address-based sampling design. *Public Opinion Quarterly,* 79(2), pp. 420-442.

Bradley, E., Curry, L. & Devers, K., 2007. Qualitative data analysis for health services research: Developing taxonomy, themes, and theory. *Health Services Research,* 42(4), pp. 1758-1772.

Centers for Medicare and Medicaid Services, 2015. *Summary of HCAHPS survey results: January 2014 to December 2014 discharges.* [Online]
Available at:
http://www.hcahpsonline.org/Files/October_2015_Summary_Analyses_Survey_Results.pdf

Cochran, W., 1977. *Sampling techniques.* New York: Wiley.

Creswell, J., 2007. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches.* Thousand Oaks, CA: Sage.

Dillman, D., Smith, J. & and Christian, L., 2014. *Internet, phone, mail, and mixed-mode surveys: The tailored design method.* New York: Wiley.

Fowler Jr., F. et al., 2002. Using telephone interviews to reduce nonresponse bias to mail surveys of health plan members. *Medical Care,* 40(3), pp. 190-200.

Fusch, P. & Ness, L., 2015. Are we there yet? Data saturation in qualitative research. *The Qualitative Report,* 20(9), pp. 1408-1416.

Gneezy, U. & Rey-Biel, P., 2014. On the relative efficiency of performance pay and noncontingent incentives. *Journal of the European Economic Association,* 12(1), pp. 62-72.

Green, J. & Thorogood, N., 2009. *Qualitative Methods for Health Research.* 2nd ed. Thousand Oaks, CA: Sage.

Grossoehme, D., 2014. Overview of qualitative research. *Journal of Health Care Chaplaincy,* 20(3), pp. 109-122.

Guest, G., Bunce, A. & Johnson, L., 2006. HOw many interviews are enough? An experiment with data saturation and variability. *Field Methods,* 18(1), pp. 59-82.

Harris, J. et al., 2009. An introduction to qualitative research for food and nutrition professionals. *Journal of the American Dietetic Association,* 109(1), pp. 80-90.

Klein, D. et al., 2011. Understanding nonresponse to the 2007 Medicare CAHPS survey. *The Gerontologist,* Issue First published online June 23, 2011.

Lohr, S., 1999. *Sampling design and analysis.* Pacific Grove: Duxburgy Press.

Maxwell, J., 2005. *Qualitative Research Design: An INteractive Approach.* 2nd ed. Thousand Oaks, CA: Sage.

Mercer, A., Caparoso, A., Cantor, D. & Townsend, R., 2015. How much gets you how much? Monetary incentives and response rates in household surveys. *Public Opinion Quarterly*, 79(1), pp. 105-129.

Millar, M. M. & Dillman, D. A., 2011. Improving Response to Web and Mixed-mode Surveys. *Public Opinion Quarterly,* Volume 75, pp. 249-269.

National Research Council, 2015. *Nonresponse in social science surveys: A research agenda. Committee on National Statistics, Division on Behavioral and Social Sciences and Education, Panel on a Research Agenda for the Future of Social Science Data Collection.* Washington, D.C.: National Academies Press.

Parsons, N. & Manierre, M., 2014. Investigating the relationship among prepaid token incentives, response rates, and nonresponse bias in a web survey. *Field Methods,* 26(2), pp. 191-204.

Pit, S., Vo, T. & Pyakurel, S., 2014. The effectiveness of recruitment strategies on general practitioner's survey response rates - a systematic review. *BMC Medical Research Methodology,* 14(1).

Polit, D. & Beck, C., 2010. Generalization in quantitative and qualitative research: Myths and strategies. *International Journal of Nursing Studies,* 47(11), pp. 1451-1458.

Seow, H. et al., 2016. The Caregiver Voice survey: A pilot study surveying bereaved caregivers to measure the caregiver and patient experience at end of life. *Journal of Palliative Medicine.*

Stat Trek, 2015. *Sample Size: Simple Random Samples.* [Online]
Available at: http://stattrek.com/sample-size/simple-random-sample.aspx

The National Alliance for Caregiving (NAC) & the AARP Public Policy Institute, 2015. *Caregiving in the U.S. 2015: Appendix A Detailed Methodology.* [Online]
Available at: http://www.caregiving.org/data/04methodology.pdf
[Accessed 18 September 2016].

Thorpe, L. et al., 2015. Rationale, Design and Respondent Characteristics of the 2013–2014 New York City Health and Nutrition Examination Survey (NYC HANES 2013–2014). *Preventive Medicine Reports,* 2015(2), p. 580–585.

VanGeest, J. & Johnson, T., 2011. Surveying nurses: Identifying strategies to improve participation. *Evaluation & the Health Professions,* Issue 0163278711399572.