SUPPORTING STATEMENT

Part B

Evaluation of Patient Satisfaction and Experience of Care for Medicare Beneficiaries with End-Stage Renal Disease (ESRD): Impact of the ESRD Prospective Payment System (PPS) and ESRD Quality Incentive Program (QIP)

Version: July 22, 2013

Centers for Medicare & Medicaid Services

Contents

B. Collections of Information Employing Statistical Methods

1. Respondent Universe and Sampling Methods

There are two data collection efforts included in this submission: one is qualitative in nature (with stakeholders in the renal health community) and one is quantitative in nature (with Medicare beneficiaries receiving certain ESRD services). We will discuss each respondent group in turn below.

Stakeholders. The respondent universe includes key stakeholders in the renal health care community. There is no sampling frame available and purposive sampling methods will be used to identify and recruit stakeholders in the renal community. Findings from the stakeholder interviews will not be statistically generalized to the respondent universe. However, findings will be relevant to inform CMS about potential impacts of the ESRD PPS/QIP legislation on health outcomes. Stakeholders will also assist in identifying topics and domains that may be missing from the existing ICH-CAHPS instrument. The information regarding any possible missing domains or topics in the current ICH-CAHPS could be used as an initial step to guide CMS in any process to update the ICH-CAHPS survey in the future.

The 40 stakeholder respondents will be distributed across a mix of provider and stakeholder types. CMS, through its contractors Acumen and Westat, will conduct interviews with front-line dialysis (non-physician) healthcare providers, such as pharmacists, nurses, care managers, social workers, or dieticians. CMS will also conduct interviews with practicing nephrologists who provide direct care to patients. The recruitment list will be developed to capture the categories of participants described above and include as broad a range of expertise as possible. CMS will collaborate with its research team to develop a list of potential participants. During the respondent identification phase, CMS will tabulate the categories of participants so we can track how the cells for each category are allocated to ensure representation among various stakeholder and care provider groups.

Beneficiaries. The universe will include all ESRD Medicare beneficiaries, with the following exclusions:
- Transplant patients;
- Those receiving inpatient hemodialysis;
- Those receiving home hemodialysis or peritoneal dialysis under "Method 2" (getting equipment and supplies from a Durable Medical Equipment (DME) provider);
- Pediatric patients; and
- Those who enrolled after January 1, 2011, when the PPS/QIP was already implemented.

These exclusions are in place because the survey is intended to measure beneficiary satisfaction and experience under the PPS/QIP implementation, and these populations did not experience a change in care from the legislation. Another group that will be excluded

are those beneficiaries receiving hospice care; the rationale is that we believe it would be insensitive to trouble these end-of-life beneficiaries with a survey about ESRD services.

The sample frame will be constructed using Medicare claims and enrollment data and data from databases, including the Medicare Beneficiary Database (MBD), and a commercial database for residential contact information matching. We will use a stratified sample design to facilitate oversampling of populations of interest to ensure adequate sample sizes for some important subpopulations as shown in Table 1 and subsequent text.

The CMS databases available for research purposes do not provide telephone contact information. This information must be obtained through matching the population with other commercially available sources. Commercial databases are expected to yield contact information for 60 percent of Medicare beneficiaries with ESRD. Since we expect 40 percent non-coverage, we will attempt to correct the non-coverage bias using the rich frame data using techniques such as calibration (Särndal and Lundström, 2005; Kott, 2009) or propensity score method (Lee, 2006) so that the survey results can be extrapolated to the whole population.

Because it is important to consider the difference in applicability of the PPS and QIP measures to the different treatment modalities and population types in the adult ESRD population, we propose to construct a sampling plan that will permit analysis of beneficiary experiences representing the following types of beneficiaries:
   •   Beneficiaries receiving in-center dialysis
   •   Beneficiaries receiving home or peritoneal dialysis

We will also sort the files to ensure proportional representation through systematic sampling (i.e., implicit stratification, which does not set up "hard" boundaries, as opposed to explicit stratification, which does).

Through the data analysis, we want to be able to provide descriptions of the experience of the following sub-populations. [1]
   • Dialysis treatment
      o In-center treatment hemodialysis (93%)
      o Home/peritoneal treatment (7%)
   • Race/ethnicity
      o White, non-Hispanic (56.0%)
      o Hispanic, any race (15.3%)
      o African American, non-Hispanic (36.9%)
      o Other (includes Asian, Native American) (6.3%)
   • Gender
      o Male (56.6%)
      o Female (44.4%)
   • Age

_____

[1] The percentages in parentheses are ESRD patient distribution obtained from the 2011 USRDS Annual Report, based on 2009 data. Some categories are not available in the report.

- o 20-44 (14%)
- o 45-64 (41%)
- o 65-74 (22%)
- o 74+ (21%)
- Beneficiary residence location
  - o Urban
  - o Non-urban
- Facility Ownership
  - o Profit (approximately 95%)
  - o Facility Non-profit (approximately 5%)
- Size/Facility Type (patient distribution)
  - o Large Dialysis Organization (LDO) Chain (63.4%)
  - o Small Dialysis Organization (SDO) Chain (11.6%)
  - o Hospital-based (10%)
  - o Independent (15%)

We consider 300 as the minimum respondent sample size needed for reasonable precision for a given sub-population. Using an equal probability sampling method, no special treatment is needed if a sub-population is large enough to ensure the minimum sample size for the sub-population. Since the overall number of completes is fixed at 2,500 respondents, a rough guide whether a sub-population would reach the threshold of 300 respondents is whether it accounts for over 12% of the total population. Based on the available distributional data on the ESRD population, the sub-populations that need special treatment such as oversampling would include:

1. Beneficiaries receiving home/peritoneal treatment;
2. Beneficiaries receiving care in hospital-based facilities;
3. Beneficiaries receiving care in non-profit facilities; and
4. Beneficiaries representing the "other" racial group (Asians and American Natives).

Sample Design for Beneficiaries Receiving Home/Peritoneal Treatment. Considering that treatment modality is one of the key variables in analysis, CMS plans to explicitly stratify the population by this variable. The sample has been planned to yield an estimated 2,000 completed surveys with beneficiaries receiving in-center ESRD treatment and 500 surveys with beneficiaries receiving home/peritoneal treatment.

Sample Design based on Types of ESRD Facilities. We anticipate that beneficiary experiences will vary by facility ownership and type. Therefore, we propose to select facility-related samples from the in-center sub-population. From the in-center sub-population group (n=2,000), we propose to allocate 1,000, 300, 300, and 400 to the LDO, SDO, hospital-based, and independent facility groups, respectively. This allocation is efficient because the selection probabilities vary less, resulting in less variable weights. These four facility groups are then explicit strata within the in-center modality stratum. If we use the same strategy to boost the sample size for the non-profit facilities, we run into a problem of fragmented strata. As a result, we propose to select beneficiaries receiving treatment in non-profit facilities with a probability that is three times larger

than the probability for beneficiaries treated by for-profit facilities within each stratum. This will greatly enhance the chance of reaching the minimum sample size for the sub-population.

Sample Design Summary: The explicit stratification discussed above is summarized in Table 1. We expect a 65 percent response rate for the target population – ESRD beneficiaries who meet the criteria outlined previously. The initial field sample size in Table 1 is obtained by inflating the target sample size by a factor of 1.54 (= 1/0.65).

Table 1. Explicit strata and sample allocation

Sampling Stratum		Population Size	Target Sample Size	Initial Field Sample Size
In-center	LDO	64,891	1,000	1,538
	SDO	11,873	300	462
	Hospital	10,235	300	462
	Independent	15,353	400	615
Home-Peritoneal		7,704	500	769
Total		110,056	2,500	3,846

Within each explicit stratum, the list will be sorted by the following variables to achieve implicit stratification:

•   Race/ethnicity;
•   Age;
•   Gender;
•   Urban/rural indicator;
•   Facility profit status;
•   Size of facility.

Sample Design for Beneficiaries Based on Race. Any beneficiary belonging to the "other race" group (which includes anyone reporting a race or ethnicity other than White, Black, and Hispanic) will be selected with a probability three times larger than the probability for other beneficiaries. The exception will be those "other race" beneficiaries who are treated by a non-profit facility – in that case their probability of selection will be based on membership in a non-profit, and will not be based on their race. This means that beneficiaries in non-profit facilities or belong to other race group will be selected with a probability three times larger than the probability for other beneficiaries to ensure the minimum sample size (300) for these subgroups. This will be achieved by using a Probability Proportional to Size (PPS) sampling method after giving a size of three to these rarely occurring units and a size of one to all others. Since the sampling list is sorted by implicit stratification variables, the sampling method becomes systematic PPS sampling to be performed within each explicit stratum independently. The method first determines the sampling interval by the ratio of the sum of the size measures of the units

in the stratum to the stratum initial field sample size. From the first interval in the sorted list, one unit is selected randomly, and then other units are selected by adding the sampling interval to the random start successively until the list is exhausted.

Beneficiaries will be interviewed in English and in Spanish; those speaking other languages will be excluded from the data collection.

<u>Power Analysis for Subgroup Comparisons:</u> The survey will collect many categorical variables, for which various population proportions will be estimated. These estimates will be compared for different subgroups (e.g., Male vs. Female). For these comparisons, the minimum detectable differences are computed for a one-sided normal test comparing two proportions with a significance level (alpha) of 5 percent and a power of 80 percent for various subgroup sample sizes. We expect considerable variation in the sampling weights due to unequal probability sample design and nonresponse weighting adjustment. Variable weights increase the design effect, which is projected to be 1.3 based on expected variability of the final weights.[2] Assuming that an equal sample size for the comparing subgroups and the two subgroup samples are uncorrelated, Table 2 shows the minimum detectable differences for two population proportions, 50 percent and 30 percent. The effective sample size is the design effect adjusted sample size, that is, the sample size divided by the design effect of 1.3.

We would require the minimum sample size for subgroup comparisons is 300. If one of the subgroup sample size is increased by x percent, then the minimum detectable difference will be decreased by about $0.13x$ percent. This approximation will work reasonably well up to $x=100$. For example, if subgroup 1 has a sample size 300 and subgroup 2 has a sample size of 600, then minimum detectable difference for $P_1=50\%$ would decrease to 10.07 percentage points from 11.57 percentage points. Therefore, if $P_2=60.07\%$, such difference can be detected with 80 percent of chance by the one-sided test with alpha = 5%.

The weighting factor to the design effect is given by $1+\mathrm{CV}_w^2$, where CV_w^2 is the squared coefficient of [2] variation (CV) of the final weights, which is projected to be 0.3.

Table 2. Minimum detectable differences

Subgroup 1		Subgroup 2		Minimum Detectable Effect Size	
Sample Size	Effective Sample Size	Sample Size	Effective Sample Size	$P_1 = 30\%$	$P_1 = 50\%$
300	231	300	231	10.61%	11.57%
400	308	400	308	9.19%	10.02%
500	385	500	385	8.22%	8.97%
600	462	600	462	7.50%	8.18%
700	538	700	538	6.94%	7.58%
800	615	800	615	6.50%	7.09%
900	692	900	692	6.12%	6.68%
1000	769	1000	769	5.81%	6.34%
1100	846	1100	846	5.54%	6.04%

We did not prepare power analysis for continuous variables because we do not have a clear picture of their population variances unlike for binomial variables. Power analysis is sufficient to meet CMS's informational needs given time and budget considerations.

2. Information Collection Procedures

Data collection procedures will vary by population; we discuss each in turn below.

Stakeholder Interviews. Westat will conduct in-depth stakeholder interviews with ESRD experts, including front-line dialysis staff. Interviews with stakeholder respondents are planned to begin two months after OMB approval. The interviews will be conducted to provide information about stakeholder perception of the impacts of the ESRD PPS/QIP legislation on health outcomes for ESRD beneficiaries and to provide information to CMS regarding any possible domains or topics that may be missing from the existing ICH-CAHPS instrument. The information regarding any possible missing domains or topics in the current ICH-CAHPS could be used as an initial step to guide CMS in any process to update the ICH-CAHPS survey in the future.

The interview will be conducted using an interview methodology that blends ethnographic interview techniques with a focused protocol that is more commonly seen in research data collection activities such as focus groups. The respondent will be asked to conceptualize domains and topics by discussing the most common issues related to the ESRD PPS/QIP. The goal is that the respondent speaks freely without the biasing or contaminating effects of directive questions. On the other hand, the protocol will indicate broad, general areas that the respondent will be encouraged to address.

The estimated time of the interview is between 50 and 60 minutes. The stakeholder interviews will be conducted using an interview protocol. A copy of the interview protocol is provided in Attachment C. General discussion topics for the stakeholder interviews include:
- Current or recent past experience with PPS/QIP
    - Context of respondent's prior work
    - Current work processes or care delivery processes
- Patient care-related issues
    - Pharmaceutical (e.g. transition to oral medications)
    - Availability, cost, and patient receptivity
- Patient educational issues, especially related to the mode of dialysis
    - Options (e.g. home, transplant)
    - Patient outcomes, including:
    - Clinical (infection rates, complications, efficacy of dialysis, hospitalizations, readmissions, emergency department visits);
    - Satisfaction and quality of life;
    - Financial burden for patient (direct medical and indirect social); and
    - Functional (physical, mental, societal, other—e.g. pain).
- Issues that were unexpected or unanticipated, to the extent that they affect the care delivered to beneficiaries
- Overall impressions of benefits of PPS/QIP and opportunities for improvement
- Review of the ESRD Beneficiary Survey instrument
    - Knowledge/experience with the instrument
    - Critique of domains and items
    - Suggestions for improvement
        Deletions or additions
        Length
        Administration issues

Beneficiary Survey. The Beneficiary Survey will use primarily closed-ended questions. In addition, 30 of the 2,500 respondents will be asked to respond in their own words to four open-ended questions covering each of the Beneficiary Survey domains (as described in Supporting Statement Part A). The administration time for the Beneficiary Survey will average 15 minutes. Survey administration will rely on a Computer Assisted Telephone Interview (CATI) system using professionally-trained telephone interviewers. The Beneficiary Survey will encompass the following domains: access to services, quality of care, outcomes, and cost. A more detailed description of the survey domains and topics are presented in the table below and are cross-walked to survey items and research questions in Attachment D.

Table.1: Beneficiary Survey Domains and Topics

Domain	Topic
Access to care	Current treatment modality and facility
	Changes to modality and/or facility
	Information/education regarding treatment modalities
	Frequency of treatment (times per week and length

	of treatment)
	Frequency of interactions with doctors, other medical staff
	Access to prescription medicines
	Access to testing/labs
	Access to medical supplies
	Shared decision making
Quality of care	Experience with doctors
	Experience with other providers
	Management of physical comfort
	Treatment oversight/management of problems
	Self-management support
	Coordination of care
	Knowledge of ESRD PPS/QIP
Outcomes	Quality of life measures
	Physical health
	Mental health
	ER visits
	Hospital stays
	Infections
Costs	Cost of treatment, out of pocket expenses
	Cost of prescription meds
	Part D coverage, other insurance coverage (Medigap)
	Reimbursement process
	Discussion of costs with providers
Demographics	Income
	Race/ethnicity
	Education
	Severity of illness
	Living arrangements
	Support structure
	Comorbidities (cardiovascular disease, diabetes, anemia management)

Once finalized, the Beneficiary Survey will be translated into Spanish. A Spanish language research team that is conversant in questionnaire design and testing methodologies and survey translation will translate the Beneficiary Survey that will undergo evaluation by a different translation expert. Discrepancies in translation will be adjudicated and corrected by a team of bilingual survey methodologists. In addition, the Spanish translation will be reviewed by translators who are native speakers of different Spanish dialects, since general vocabulary and specific jargon can differ significantly among Cuban, Mexican, Puerto Rican, and various Central and South American dialects.

3. Methods to Maximize Response Rates

Stakeholder Interview. Participants are not being selected via probability-based sampling methods. A "response rate" has no clear meaning in the context of this effort. However, to encourage participation, multiple calls will be made to complete the interview (if contact is not made during the initial call attempt). While there will be no pre-notification letter sent to potential respondents, stakeholder recruitment will occur through phone and email. Interviews will be scheduled in advance such that the participants may make adequate time available.

Beneficiary Survey. Advance letters will be sent to respondents to explain the survey and alert respondents to a telephone call invitation to do the survey. A toll-free number will be included in letter for those who wish to establish the legitimacy of the survey or have some questions answered. CMS' research vendor will also leave periodic voice messages for respondents not found at home at the time of the call.

The survey team will use established calling procedures designed to maximize the number of interviews completed among a Medicare beneficiary population. Calls are spread out over different types of days (weekday versus weekend) and over different times of the day. The automated CATI scheduler will distribute the calls to the most productive calling periods for that particular type of case. In addition, CMS' research vendor will re-contact most refusal cases once, in an effort to complete the interview. Interviewers who are judged to be particularly skillful receive special training, and only they have access to refusal conversion cases. The interviewers judge the nature of refusals as mild, firm, or hostile; hostile refusals are not re-contacted. The refusal conversion interviewers also coach each other by sharing successful refusal conversion techniques. Calls with no response or no answer will be called a maximum of twenty-one times if needed to achieve the set number of completed surveys.

A national telephone survey of the ESRD population has not recently been undertaken by the U.S. government. Estimates based on cognitive testing conducted in support of this survey as well as experience in conducting surveys of Medicare beneficiaries suggest that we may expect a response rate of about 65 percent. At OMB's recommendation we can conduct a non-response bias analysis to compare known characteristics of respondents to those of the non-respondents. We will use the propensity score method to analyze the nonresponse bias and make nonresponse adjustment. We will further apply the calibration weighting to the nonresponse-adjusted weight to reduce the non-coverage bias.

4. Tests of Procedures

CMS's research vendor has staff with expertise in data collection and evaluation methods and senior researchers have reviewed the materials to identify any unclear, ambiguous, and uncomfortable questions.

Westat conducted cognitive testing of the ESRD Beneficiary Survey with nine ESRD Medicare beneficiaries. A supporting crosswalk of all the survey items tested is included as Attachment D. In addition, Appendix D also lists the survey items that came directly from the ICH-CAHPS, items that were modified, and items that were developed for the

current survey. A copy of the cognitive testing (pilot test) report is included as Attachment F.

The stakeholder interviewer protocol performed as expected based on four preliminary in-depth interviews. Respondents spoke freely during the interview and respondents were encouraged to address broad areas of ESRD care.

5. Statistical Consultants

Dr. Hyunshik James Lee of Westat was consulted for this data collection effort. Dr. Lee is a senior statistician at Westat with almost 30 years of experience in survey methods.

6. References

Kott, P. (2009). Calibration weighting: combining probability samples and linear prediction models, in D. Pfeffermann and C.R. Rao (Eds.), *Handbook of Statistics, Sample Surveys: Design, Methods and Application*, 29B, Amsterdam: Elsevier.

Lee, S. (2006). Propensity Score Adjustment as a Weighting Scheme for Volunteer Panel Web Survey. *Journal of Official Statistics*, 22, 329–349.

Särndal, C.E., and Lundström, S. (2005). *Estimation in Surveys with Nonresponse*. New York: Wiley.