

Part C and D Complaints Resolution Performance Measure
CMS-10308

OMB Supporting Statement – Part B

September 3, 2010

Contents

Collection of Information Involving Statistical Methods	1
1. Respondent Universe and Sampling.....	1
2. Procedures for the Collection of Information.....	8
3. Methods to Maximize Response Rates and Data Reliability	13
4. Tests of Procedures or Methods	14
5. Individuals Consulted on Statistical Methods	17

Collection of Information Involving Statistical Methods

1. Respondent Universe and Sampling

CMS is interested in gathering information to determine the possibility of developing performance measures associated with beneficiaries' satisfaction with the complaints resolution process. This effort will emphasize that the measures are developed separately for each contract. The survey population is made up of beneficiaries with closed urgent or immediate need complaints that were filed against their respective contracts during the period covering the months of January and February 2011. This data collection period was chosen because CMS is interested in the months with the largest number of complaints in order to achieve the most statistically valid sample. All Medicare Advantage and Prescription Drug contracts will be surveyed regardless of their enrollment size, and the sampling will be carried out from the Complaints Tracking Module (CTM) database. However, members of 800 series contracts will be excluded from selection. 800 series contracts are MA Organizations, PDP sponsors, and Section 1876 Cost Plan Sponsors that offer, sponsor, or administer certain types of employer sponsored group contracts (employer/union-only group waiver contracts also referred to as EGWPs). Not only in this case, but also in many other situations, CMS excludes EGWPs as they are overseen differently than other contracts. Additionally, if the data collected from this effort is used for the development of a performance rating, CMS does not currently post performance ratings for EGWPs. This is primarily because these contracts are not open to the public but only to the relevant employer/union organization members.

This survey will collect data about beneficiaries' experience with the contract sponsor complaint resolution processes and the effectiveness of the resolution (a discussion of the indicators and preliminary measures from the survey instrument is included in Supporting Statement A, section B.16.a. Tabulations). The use of a short recall period will allow beneficiaries to have the best possible recollection of their experiences. The sampling of complaints will be carried out from the CTM database every week on a flow basis as they are closed. The data collection period will allow for a waiting period of 7 days for CMS and contract records to be updated and attempts to communicate with the beneficiary to be completed.

To ensure a good representation of the complaint population, a total sample of 6,500 complaints will be allocated across weeks, proportionally to the expected weekly count of closed complaints. This total sample size of 6,500 was determined based on precision requirements and budgetary constraints, as discussed below in subsection 2.a. The sampling strategy will exclude complaints that are outside of the scope of the contract, particularly some complaints related to enrollment issues (e.g., when a beneficiary enrolled after the deadline for enrollment). Further review of the complaints will take into consideration that there are certain actions that may have been within CMS guidelines but required further actions from agents other than the contract and these may have caused dissatisfaction on the part of the beneficiary (e.g., involvement by the Retro Processor Contract, which adds several days to a resolution.)

CMS will collect information on all contracts (except for 800 series) including those contracts with small enrollment and/or with a small number of complaints. CMS is interested in developing preliminary measures that can be calculated for all contracts. For this purpose, all contracts will be included in the data collection, and CMS will determine later what strategies will be used to address small samples and limited information.

The 2011 survey population is unknown at the present time and will remain unknown until the end of the survey, due to the rolling sampling approach adopted to minimize recall bias. For the purpose of designing the sampling, we used 2008 and 2009 CTM data for the period spanning January 1 through March 4. Although the 2011 complaint counts are expected to be different from those of 2008 and 2009, we expect the overall 2011 weekly distribution of closed complaints—all contracts combined—to have the same pattern as was observed in 2008 and 2009. That is, the highest volume of closed complaints is expected to be observed in the second week of collection, with a gradual decrease thereafter. The CMS staff supports this assumption, and available data on complaints for the first quarter of 2010 provide supporting evidence for this argument. Table B.1.a. shows the weekly distribution of complaints closed each week, based on 2008, 2009, 2010 CTM data for the period from January 1 through March 4. The weekly distribution of complaints includes only those complaints that were closed during the week.

Table B.1.a. shows that while 12,392 complaints filed against 499 contracts were closed during the first 9 weeks of 2010, a total of 19,801 complaints filed against 541 contracts were closed during the same period in 2009. The weekly proportions of complaints in these 3 years remain very similar, with week 2 holding the highest number of complaints closed (21% in 2010, 16% in 2009, and 20% in 2008). Starting in week 3 these percentages decrease gradually through week 9. This distribution is expected to vary substantially from one contract to another, with some small contracts having no more than one closed complaint. Table B.1.b. demonstrates a nearly consistent distribution of complaints by complaint categories in the first nine weeks of the year. Table B.1.c. presents the distribution over 12 months, which, when compared to Table B.1.b., shows that the distribution of the project sample is close to the distribution of complaints over the year.

The data collection period was selected primarily for the expected high complaint volume during the first three months of the calendar year as beneficiaries and contracts work out benefits/services and operational issues. CMS expects that contracts are efficient in resolving enrollment and other immediate issues and, thus, the complaints topics are better examined during this time. Complaints issues encountered later on during the calendar year are also captured in the first quarter of the contract year, therefore there is only a moderate bias in the included complaint types and numbers. CMS will note in its results the period of data collection to limit the representativeness of the selected complaints and prevent confusion over generalizations to the entire contract year.

This time period was also selected to provide flexibility for analysis and production of performance measures (if CMS chooses to do so) per contract by mid-July. This timeline would allow time for CMS to produce the performance star rating and post online by mid-September for beneficiary reference in their selection of Parts C and D contracts.

Table B.1.a. Distribution of Complaints by Year and Week for the First 9 Weeks of the Year

Year	Number of Contracts	Number of Complaints									
		WK 1	WK 2	WK 3	WK 4	WK 5	WK 6	WK 7	WK 8	WK 9	Total
2010	499	845	2,572	1,544	1,477	1,395	1,367	1,107	1,065	1,021	12,392
2009	541	1,492	3,183	2,443	2,276	2,166	2,524	2,060	1,864	1,792	19,801
2008	470	2,340	6,402	5,816	4,798	3,932	2,741	2,148	1,702	1,701	31,580
Proportion											
2010		7%	21%	13%	12%	11%	11%	9%	9%	8%	100%
2009		8%	16%	12%	11%	11%	13%	10%	9%	9%	100%
2008		7%	20%	18%	15%	12%	9%	7%	5%	5%	100%

Table B.1.b. Distribution of Complaints by Week (the first 9 weeks) and Category (2008-2010)

Complaint Category	Year	Total	% Total Volume	% Wk 1	% Wk 2	% Wk 3	% Wk 4	% Wk 5	% Wk 6	% Wk 7	% Wk 8	% Wk 9
Enrollment/Disenrollment	2010	6920	55.8%	50.7%	52.2%	58.7%	55.7%	58.8%	57.7%	56.8%	56.8%	56.5%
Pricing/Co-Insurance	2010	1753	14.1%	10.5%	12.1%	14.6%	16.6%	15.4%	15.2%	14.4%	14.8%	13.8%
Benefits/Access	2010	1287	10.4%	13.0%	9.5%	11.3%	10.6%	10.5%	9.8%	10.1%	10.0%	10.2%
Plan Administration	2010	812	6.6%	16.4%	16.4%	3.6%	3.4%	2.5%	2.5%	2.9%	2.1%	2.2%
Formulary	2010	650	5.2%	4.4%	4.0%	4.7%	5.0%	5.6%	5.9%	6.9%	5.7%	6.7%
Exceptions/Appeals	2010	269	2.2%	1.1%	1.4%	1.4%	2.0%	1.6%	2.7%	2.8%	3.8%	3.9%
Customer Service	2010	254	2.0%	2.1%	1.9%	1.7%	2.3%	2.2%	1.8%	2.1%	2.0%	2.5%
Marketing	2010	224	1.8%	0.5%	0.8%	2.0%	2.4%	2.1%	1.8%	2.3%	2.7%	2.4%
Other	2010	223	1.7%	1.3%	1.6%	2.0%	2.0%	1.2%	2.5%	1.8%	2.1%	1.8%
Total	2010	12392	12392	845	2572	1544	1476	1395	1367	1107	1065	1021
Enrollment/Disenrollment	2009	11989	60.5%	62.5%	62.1%	59.7%	60.3%	59.4%	63.4%	60.0%	58.2%	58.0%
Benefits/Access	2009	2488	12.6%	15.8%	14.6%	14.3%	12.7%	11.7%	10.6%	11.1%	10.2%	11.7%
Pricing/Co-Insurance	2009	2454	12.4%	10.3%	11.7%	13.1%	12.7%	13.7%	11.6%	12.9%	12.6%	12.7%
Formulary	2009	899	4.5%	2.9%	2.7%	4.2%	4.5%	5.0%	4.6%	5.4%	6.8%	5.8%
Plan Administration	2009	781	3.9%	4.2%	4.4%	3.2%	4.7%	3.6%	4.0%	3.7%	3.6%	4.0%
Customer Service	2009	395	2.0%	1.9%	1.8%	2.1%	2.2%	2.2%	1.6%	1.7%	2.6%	2.1%
Exceptions/Appeals	2009	287	1.4%	0.9%	0.7%	1.4%	0.9%	1.3%	1.3%	2.0%	2.3%	2.8%
Marketing	2009	233	1.2%	0.7%	0.6%	0.8%	1.0%	1.6%	1.4%	1.5%	1.9%	1.4%
Other	2009	275	1.3%	0.8%	1.3%	1.2%	1.0%	1.4%	1.4%	1.7%	1.9%	1.6%
Total	2009	19801	19801	1492	3183	2443	2276	2166	2524	2061	1864	1792
Enrollment/Disenrollment	2008	20774	65.8%	63.2%	68.9%	64.7%	64.8%	65.0%	66.5%	66.6%	65.2%	64.3%
Pricing/Co-Insurance	2008	4346	13.8%	16.8%	12.8%	15.4%	14.1%	13.2%	13.4%	12.3%	12.1%	11.9%
Benefits/Access	2008	2503	7.9%	9.3%	9.0%	7.5%	7.3%	8.1%	7.2%	7.4%	7.4%	7.1%
Formulary	2008	1388	4.4%	3.0%	3.2%	3.5%	5.2%	4.8%	4.9%	5.7%	6.0%	6.6%
Customer Service	2008	1203	3.8%	4.6%	2.8%	3.8%	4.2%	4.2%	3.8%	3.4%	4.4%	4.7%
Plan Administration	2008	504	1.6%	1.4%	1.5%	2.0%	1.8%	1.5%	1.2%	1.5%	1.2%	1.8%
Grievances	2008	326	1.0%	0.4%	0.7%	1.4%	0.7%	1.5%	1.2%	1.2%	1.0%	1.1%
Exceptions/Appeals	2008	272	0.9%	0.5%	0.3%	1.0%	1.0%	0.9%	1.0%	0.9%	1.2%	1.5%
Other	2008	264	0.8%	0.8%	0.7%	0.4%	1.0%	1.0%	0.9%	1.0%	1.6%	0.9%
Total	2008	31580	31580	2340	6402	5816	4798	3932	2741	2148	1702	1701

Note: the first 8 major categories are listed and the rest are represented by “other.” Percentages are based on column totals.

Table B.1.c. Distribution of Complaints by Month and Category (2009)

Complaint Category	Total	% Total Volume	% Month 1	% Month 2	% Month 3	% Month 4	% Month 5	% Month 6	% Month 7	% Month 8	% Month 9	% Month 10	% Month 11	% Month 12
Enrollment/Disenrollment	36999	59.7%	60.8%	60.3%	60.1%	64.8%	63.2%	61.0%	57.4%	54.9%	53.3%	53.2%	54.7%	52.7%
Benefits/Access	7708	12.4%	13.9%	11.0%	9.8%	10.0%	10.1%	12.6%	15.1%	15.8%	15.0%	15.1%	14.7%	15.9%
Pricing/Co-Insurance	7696	12.4%	12.4%	12.3%	12.4%	10.5%	11.5%	11.3%	12.0%	14.2%	14.4%	14.6%	13.4%	16.1%
Formulary	3142	5.1%	3.8%	5.4%	6.4%	4.7%	4.9%	5.2%	4.8%	5.1%	6.3%	5.6%	5.8%	4.2%
Plan Administration	1963	3.2%	4.1%	3.9%	3.8%	2.4%	2.6%	2.4%	2.6%	1.9%	2.7%	2.9%	3.1%	3.3%
Customer Service	1250	2.0%	2.0%	2.0%	1.9%	1.8%	2.0%	1.8%	2.0%	2.3%	1.8%	2.1%	2.4%	2.7%
Exceptions/Appeals	1241	2.0%	1.0%	1.9%	2.3%	2.0%	2.2%	2.4%	2.1%	2.1%	2.6%	2.9%	2.2%	2.6%
Marketing	792	1.3%	0.8%	1.6%	1.6%	1.8%	1.3%	1.2%	1.4%	1.0%	1.4%	0.6%	0.9%	0.6%
other	10050	1.9%	1.1%	1.5%	1.5%	2.0%	2.2%	2.0%	2.4%	2.7%	2.4%	2.9%	2.5%	2.0%
Total	61982		10387	8480	7749	7608	5790	4863	4133	3189	2841	2586	2008	2348

Note: the first 8 major categories are listed and the rest are represented by “other.” Percentages are based on column totals.

The recommended sampling approach is described as follows:

- The total number of sample complaints to be selected for all contracts under investigation is 6,500. This overall 2011 sample size will be allocated across the 9 weeks of the survey implementation as shown in Table B.2. This allocation is based on the weekly 2008–2009 average proportion of closed complaints. In week 1 of the year 2011, for example, 520 complaints will be sampled, while week 2 will provide 1,170 of the total 6,500 sample complaints. The last sampling phase will occur in week 9 with the selection of a total of 455 complaints from all contracts. An analysis of 2009 CTM data has demonstrated that an overall sample size of 6,500 is sufficient to achieve for each contract an error margin of 10% for a minimum confidence level of 85%. Moreover, the proposed allocation ensures a weekly sampling fraction (i.e., the ratio of the sample size to the 2009 population size) that varies from 24% to 40%. 2010 data confirm these sampling estimates.
- In 2011, the final number of complaints to be selected from each contract in any given week will be determined at the time of sampling on the basis of the actual observed counts. This will be achieved by allocating the predetermined overall weekly sample of Table B.2 across contracts, proportionally to the square root of the observed counts of complaints. The complaint sample weekly allocation to contracts is carried out proportionally to the square root of the observed complaint counts, as opposed to the plain counts, to avoid an underrepresentation of contracts with a small number of complaints. Each contract must be well represented in the total sample since performance measures will be calculated individually for each contract.

In week 1 for example, the number of sample complaints to be selected from a particular contract is calculated as follows:

$$\frac{\sqrt{C_{1j}}}{\sum \sqrt{C_{1j}}}$$

where C_{1j} is the predetermined number of complaints to sample in week 1 for all contracts, C_{1j} is the observed number of complaints filed against contract j in week 1, and $\sum \sqrt{C_{1j}}$ is the number contracts with at least one complaint in week 1.

- The square root rule will provide an initial allocation of the weekly complaint sample across contracts in 2011. This allocation will eventually be adjusted, primarily to increase the sample size in small contracts or decrease that of large contracts so as to meet the precision objectives for all contracts. Since the total number of complaints filed against a contract will not be known until after week 9, the achieved error margin will be monitored each week from week 3 and will be used to eventually adjust the weekly sample size as needed.

- Postponing the sample size determination at the contract level to 2011 is due to the difficulty of predicting the actual 2011 counts of complaints with any reliability. In the next section, we will discuss the precision level that we anticipate with the current sampling strategy.

Table B.2: Allocation of the 2011 Complaint Sample Across Weeks

Week	Wk 1	Wk 2	Wk 3	Wk 4	Wk 5	Wk 6	Wk 7	Wk 8	Wk 9	Total
Allocation of the 2011 Complaint Sample	520	1,170	975	845	780	715	585	455	455	6,500
Weekly Proportions of 2011 Sample Complaints (2008-2009 Avg. Proportions)	8%	18%	15%	13%	12%	11%	9%	7%	7%	100%

Since the actual number of complaints is unknown until the end of the survey (March 4, 2011). The sample size of 6,500 is estimated based on the observed number of closed complaints between January 1, 2009 and March 4, 2009. Specifically, it was estimated through the following steps:

- Obtain the population size (total number of complaints closed during Jan 1-March 4 excluding uninterested complaints such as complaints from provider, non urgent and immediate etc) for each contractor
- Estimate the required sample size based on population size, required precision level (error of margin as 0.10 and confidence level as 85%), and estimated response rate (80%)
- Sum the sample size over all contracts that have at least one complaint.

A proactive sampling design has been developed to minimize low response rates or oversample contracts with a small number of complaints. Therefore, we may achieve confidence intervals of 95% for some contracts and an 85% confidence interval for all contracts. After the completion of the first full-scale data collection, CMS may choose to revisit and increase the confidence interval for future data collection efforts.

Table B.3. below summarizes the distribution of contractor and complaints by complaint range.

Table B.3: Distribution of Contracts and Complaints Count Range (2009)

	Range of Complaint Count									
	0 – 19		20 – 51		52 – 84		85+		TOTAL	
	Contracts	Complaints	Contracts	Complaints	Contracts	Complaints	Contracts	Complaints	Contracts	Complaints
Population Size	436	2156	51	1557	19	1191	35	14897	541	19801
Sample Size	436	2156	79	2509	26	1711	0	0	541	6376

The sample size of 6,500 is rounded up from the estimated sample size of 6,376 to ensure a sufficient sample size.

2. Procedures for the Collection of Information

a) *Statistical Methodology, Estimation, and Degree of Accuracy*

The primary objective of this survey is to collect data to determine the possibility of developing performance indicators that measure the beneficiary’s satisfaction with the complaint’s final outcome and complaint process. The current study design is optimized for performance measures that are expressed in the form of percentages. The sample size () for each contract will depend on the complaint population size (), the desired confidence level (CL), and the error margin (E) associated with the performance measure. The three quantities , CL, and E are interrelated in such a way that two of them must be known to determine the third. Therefore, our desire to determine the sample size requires the knowledge of CL and E, which must be hypothesized.

As indicated in section B.1, the disproportional distribution of complaints by week requires a weekly selection of complaints with different selection probabilities. These differential selection probabilities must be accounted for when quantifying the precision of performance measures. The use of different selection probabilities will result in an increase in the variance associated with survey statistics by a factor known as the Design Effect (DEFF). For a given value of DEFF, the sample size for a particular contract is calculated as follows:

$$n = \frac{(z_{\alpha/2}/2E)^2 \times DEFF}{1 + \frac{(z_{\alpha/2}/2E)^2 \times DEFF - 1}{N}} \tag{B.2}$$

where $z_{\alpha/2}$ is the critical value representing the influence of the confidence level on the error margin. The subscript α associated with the critical value represents the lack of confidence in the magnitude of the error margin (i.e., $\alpha = 1 - \text{Confidence Level}$) and is assigned a small value during the study design.

Table B.4 shows the minimum sample size required by population size, and for various values of the confidence level and the error margin. These estimated sample sizes are based on a

hypothesized design effect of 1.2, which represents an increase of 20% of the variance due to the complexity of the sample selection protocol. Design effects of 1.2 or less are common in many statistical surveys that are based on complex samples. The 85%/10% column contains the minimum sample size requirements that will be implemented.

Table B.5 shows the size of the initial sample required to obtain the minimum number of respondents of Table B.4. The numbers in Table B.5 are based upon the assumption of a response rate situated around 80%, and give an indication of the number of complaints required per contract in the sample to meet the specified precision requirements for different values of the complaint population size. Using 2009 CTM data, we were able to determine that a total initial sample size of 6,500 complaints allows us to guarantee, for each contract, a maximum error margin of 10% and a minimum confidence level of 85%. To ensure a minimum confidence level greater than 85% for the same error margin would require a sample size that is greater than 6,500.

The sample size of 6,500 could achieve a maximum error margin of 10% and a minimum confidence level of 85% for each contract. The sample size is based on the total complaint population of 19,801 and its observed distribution pattern among contracts. However, a ceiling sample size of 6,500 will not be able to ensure the achievement of the same precision if there is dramatic difference between 2011 and 2009 complaint data in terms of total complaint population size or distribution patterns among contracts. For example, if the 2011 complaint population size is significantly larger than 19,801, given the same distribution pattern, but the sample size ceiling is set to 6,500, the achieved precision level will be lower than 85%/0.10. Another possible scenario is that the complaint population size of 2011 could be similar to 19,801, but the distribution pattern could change (i.e., the number of contracts with low volume complaints increases while the number of contracts with large volume complaints decreases). This could also decrease the level of precision if the total sample size is set to 6,500. Last, the level of precision would increase under a total sample size of 6,500 if the changes in the population size and the distribution pattern are in the opposite direction.

Table B. 4: Minimum Number of Respondents by Desired Confidence Levels and Error Margins

Desired Level	Confidence	95%	90%	85%	80%	95%	90%	85%	80%
		Target Error Margin		5%	5%	5%	5%	10%	10%
Population Size	10	10	10	10	10	10	10	9	9
	20	20	19	19	19	18	17	16	15
	50	46	44	42	41	36	32	28	26
	100	83	77	72	67	54	46	39	34
	200	140	124	112	100	74	58	48	40
	500	241	198	167	142	94	70	56	45
	1,000	316	246	200	165	104	76	59	48
	2,000	375	280	222	180	110	79	61	49
	5,000	423	305	237	190	113	80	62	49
	10,000	441	315	243	194	114	81	62	50
	20,000	451	320	246	196	115	81	62	50
50,000	457	323	248	197	115	82	63	50	

Table B. 5: Estimated Initial Sample Sizes Based on an 80% Response Rate

Desired Level	Confidence	95%	90%	85%	80%	95%	90%	85%	80%
		Target Error Margin		5%	5%	5%	5%	10%	10%
Population Size	10	10	10	10	10	10	10	10	10
	20	20	20	20	20	20	20	20	19
	50	50	50	50	50	45	40	35	33
	100	100	97	90	84	68	58	49	43
	200	175	155	140	125	93	73	60	50
	500	302	248	209	178	118	88	70	57
	1,000	395	308	250	207	130	95	74	60
	2,000	469	350	278	225	138	99	77	62
	5,000	529	382	297	238	142	100	78	62
	10,000	552	394	304	243	143	102	78	63
	20,000	564	400	308	245	144	102	78	63
50,000	572	404	310	247	144	103	79	63	

The population size in tables B.4 and B.5 refer to the total number of complaints received during the research time period per contractor. The sample size of a contractor then is determined based on the population size, the desired precision level (both error margin and confidence level), design effects as well as response rate, as displayed in table B3 and table B4.

CMS has chosen to use a non-standard confidence interval of 85% due to budgetary constraints since a 90% or 95% confidence interval will require a larger sample given the same error margin and desired response rate (See Table B.6). This is in alignment with other CMS reported monitoring and performance measures which are also calculated using an 85% confidence interval. CMS may adjust the confidence interval target at a later time.

Table B.6: Required Sample Sizes per Desired Confidence Interval

Confidence Interval	Required Sample Size
80%	5831
85%	6376
90%	7092
95%	8097

Note: error margin=0.10 DEFF=1.2 Response rate=0.8

Another implication of the use of differential selection probabilities is the need to weight the performance measures using weights obtained as the inverse of the complainant's selection probability. The beneficiary is the unit of analysis that should be weighted, and the complaint is the sampling unit that receives the initial sampling weight. Therefore, the beneficiary weight will be sum of the sampling weights of all complaints associated with the same beneficiary. If n is the number of beneficiaries in the sample, and n_i the number of beneficiaries with a specific characteristic of interest, the proportion of beneficiaries with the characteristic of interest is given by:

$$p = \frac{\sum_{b=1}^m w_b}{\sum_{b=1}^M w_b} \quad (\text{B.3})$$

where w_b is the weight associated with beneficiary b .

b) Unusual Problems Requiring Specialized Sampling Procedures

This survey will collect data about immediate-need complaints, which must be closed within 48 hours, and urgent complaints, which must be closed within 7 to 10 days. To account for the delays needed by health contracts to close the complaints filed during a week, the weekly sampling will select complaints filed during the 7-day period that ended 10 days prior to the beginning of the sample selection. The last sample also would be selected 10 days after the last week of February 2011. This delay in data collection would allow for allow time for beneficiaries to receive notification of their complaint resolution or for data to be updated in the electronic systems.

c) Periodic Cycles to Reduce Burden

We will implement the survey over a period of 2 months in order to collect data regarding beneficiaries' recent experience with their health contract's complaint resolution process. The

need for each interview to target one specific complaint makes a cyclical collection of data unfeasible.

3. Methods to Maximize Response Rates and Data Reliability

a) Response Rates

We estimate an initial sample of 6,500 beneficiaries to result in 5,200 completed surveys (80% response rate). To achieve this target, we will utilize a mixed-mode approach that utilizes telephone as the primary mode of data collection, with mail follow-up. We believe that an 80% response rate is achievable for three reasons: (1) this is a government-sponsored survey related to Medicare; (2) we will be surveying a motivated population of people who have taken a stance and filed a complaint by calling 1-800-Medicare; and (3) we are using a mixed-mode approach that gives beneficiaries two options for participating in the survey. In addition to offering two modes of completion, several other strategies will be used to achieve this high response rate.

First, before telephone interviewing begins, an advance letter describing the purpose and sponsorship of the survey will be mailed to potential respondents (the letter is presented in Appendix D). This advance letter will assure potential respondents that the caller is conducting a research interview and not soliciting donations or selling anything. Letters will be sent approximately one week before the sample is released to the phone survey scheduler. The letter will provide a toll-free call-in number.

Second, experienced interviewers will be assigned to the study and extensively trained. These interviewers will be thoroughly trained on data collection procedures, including methods for promoting cooperation among sample members. Interviewers are skilled at encouraging cooperation and will minimize the impact on responses resulting from the persuasion of reluctant respondents.

Third, call scheduling in CATI will allow respondents to select the time most convenient for them to be interviewed. We will make up to 10 attempts per complaint/beneficiary over a 3-week period.

Fourth, beneficiaries who do not respond to the telephone survey by the 3-week mark will be sent a paper copy of the questionnaire with a postage-paid return envelope.

Finally, a reminder postcard with a toll-free number will be sent to all nonrespondents approximately one week after the hard copy mailing.

Although both approaches will be employed, the primary mode of data collection is intended to be telephone administration. The 10 attempts to reach each beneficiary by phone and the 1-800 number provided to beneficiaries who receive mailed surveys both encourage phone

participation in this data collection. Surveys will be mailed to telephone nonrespondents after a period of calling through the CATI system.

b) Reliability of Data Collection

The beneficiary questionnaire was built on questionnaires developed for other studies, including the CAHPS Hospital Survey and the CAHPS Health Plan Survey (Adult Medicaid Questionnaire), both of which were reviewed and approved by OMB. Although the two CAHPS surveys served as the original framework for the questionnaire, PDP Customer Service measures were reflected in several questions. The J.D. Power and Associates “2009 National Health Insurance Plan Study” question topics regarding customer satisfaction were also incorporated. The questions were designed to ensure that they would be easily understood by respondents. Revisions were made to the draft questionnaire based on the results of the pretest, feedback from CMS stakeholders, and public comments received from the publication of the 60-day Federal Register Notice.

The use of computer-assisted telephone interviewing (CATI) to conduct the majority of interviews will help to ensure the consistency of the data. CATI controls question branching (reducing item nonresponse due to interviewer error), modifies wording (providing memory aids and probes and personalizing questions), and constructs complex sequences that are not possible to produce or are less accurate in hard-copy surveys. The probes, verifications, and consistency checks are built into the system and standardize the procedures. These procedures ensure the reliability of the data collection methods and the data collected through those methods. Issues regarding the uniformity of surveys completed through the two modes of data collection are detailed in Supporting Statement A (Section B.3. Use of Information Technology).

Last, IMPAQ International will monitor each interviewer’s work using silent call-monitoring equipment and video monitors that display the interviewer’s screen.

4. Tests of Procedures or Methods

We propose to conduct two tests of procedures/methods for this survey:

Pre-Test: While OMB review was underway in March 2009, we tested the survey instrument with a *convenience* sample of nine Medicare Part C and Part D beneficiaries. The pre-test design was based on a cognitive interviewing model. The goal of the cognitive interviews was to test the questionnaire content, ensure that the survey instructions and question wording are clear and understandable, and that response options are adequate. The cognitive interviews allowed us to determine the validity of the questions: Are respondents interpreting them as intended? Are the questions measuring the constructs of interest? Questions that are misunderstood by respondents or that are difficult to answer can be improved prior to fielding the main survey, thereby increasing the overall quality of survey data. Additionally, once survey data has been collected, cognitive testing results can provide useful information for users by documenting potential

sources of response error as well as providing a richer understanding of the type of data that has been collected.

Cognitive testing of the beneficiary questionnaire included conducting semi-structured interviews using verbal probing techniques. Each interview consisted of two components: (1) the interviewer administered the questionnaire and recorded the respondent’s answers, and (2) after each question, the interviewer engaged the respondent in a conversation exploring the meaning of the item and the respondent's answer.

The results of the cognitive test were not expected to be statistically significant due to the size of the sample and the results will not be added to the full-scale data collection. Some of the main findings of the pre-test are presented below:

- About half of the participants did not perceive their complaint to be resolved.
- Respondents may not differentiate between how long it took to resolve the complaint and how the complaint was handled. It appears they are mainly concerned with the way the complaint was handled.
- Believing your complaint was resolved seems to be a driving factor in how participants responded. The issue of resolution appears to set the tone in how satisfied the respondent is with the overall process regardless of how long the resolution took or who they perceived to resolve their complaint either Medicare or the contract.

Following the pre-test and receipt of public comments on the 60-Day Federal Register Notice, CMS made changes to the questionnaire, which are summarized in OMB Supporting Statement A (Section B.8.a. Federal Register Notice and Comments, *Survey Instrument*). However, the survey instrument changes listed below are meant to highlight the changes that were principally prompted by information from the pre-test and the comments of the nine participants:

General	In Q1, “ resolved ” was replaced with “ settled. ” In other questions, “ resolution ” was replaced with “ final out come or decision ” to prevent beneficiary bias and to guide the beneficiary towards the actions taken by the contract as opposed to the beneficiary's opinion of the decision.
Q5	Question 5 has been removed from the survey. (CMS has decided to drop questions about repeat complaints or multiple attempts to contact the contract)
New question	Add question to assess beneficiary satisfaction with aspects of the complaint handling process. Beneficiaries will rate their satisfaction with components of the handling process such as length of the complaint process and courtesy of the contract representative. In a simplified form,

	<p>this satisfaction question addresses issues from the original Q6 and Q8.</p> <p>This question is now Q2 in the new survey instrument.</p>
New question	<p>Add question to identify survey respondent. This is a demographic question to differentiate between respondent and proxy.</p> <p>This question is now Q9 in the new survey instrument.</p>

The changes that were made to the survey instrument were meant to clarify question wording while improving beneficiary understanding and response quality.

The pre-test also revealed a significant incidence of beneficiaries who file complaints through a representative or proxy. This issue was led to further analysis of complaints and CTM data to quantify the expected proportion of affected complaints and to adequately prepare for this population in the pilot test and full-scale data collections.

As the pre-test did not expose any insurmountable concerns, the survey instrument and data collection methods were deemed acceptable for the study.

Pilot Test: After receipt of OMB approval, we will conduct a pilot test with approximately 100 beneficiaries in Q4 2010. The sample will be selected randomly following the proposed sampling plan for the actual survey. The purpose of the pilot is to test the instrument on a broader population, refine the data collection process, and produce preliminary measure statistics – essentially, it is a dry run of all activities for the full-scale data collection. The pilot will also allow testing for strategies to achieve reliable data and remove complaints that are not within the contract’s domain. On issues of the data collection process, some of the testing will include:

- Sending a pre-notification letter to sampled beneficiaries;
- Loading sample information into the Computer-Assisted Telephone Interviewing (CATI) system;
- Administering the programmed instrument and ensuring that skip patterns are functioning correctly;
- Implementing the mail follow-up option for telephone nonrespondents;
- Reviewing the data collected to make sure the questions are performing as intended under real field conditions; and
- Testing the preliminary performance measures and conducting exploratory and risk-adjusted analyses.

Findings from the pilot test will be used to refine the data collection process to ensure seamless implementation of the main survey. Both quantitative and qualitative analyses will be conducted with pilot test data. These analyses will focus on two primary objectives: (1) identifying

questions in the survey instrument that require refinements and (2) noting any necessary changes to logistics and operations. Through qualitative analysis of open-ended questions, we will determine the utility of these questions and also whether any closed-ended questions ought to be added or modified to incorporate a common response or theme from beneficiary responses. Main logistical issues that will be tracked include any difficulties with receiving data from CMS in real time, timing of telephone and mailing communications with beneficiaries, and survey center operations such as the issues raised by the interviewers or adjustments to the FAQs. In reviewing these issue areas listed above, the pilot will be a test of all the aspects of the full-scale data collection ensuring that the study will run smoothly in 2011.

The answers from the pilot will not be added to the survey results from the actual data collection. At the end of the pilot test, we will submit a sample report reflecting the information collected from the pilot test. This sample report will assist CMS in refining the reporting requirements.

5. Individuals Consulted on Statistical Methods

The following persons outside of CMS contributed to, reviewed, and/or approved the design, instrumentation and sampling plan:

Name	Affiliation	Telephone Number
Philippe Gwet	IMPAQ International	301-326-9001
Gongmei Yu	IMPAQ International	443-539-9769
Oswaldo Urdapilleta	IMPAQ International	202-289-0004 x503