Pay for Reporting: Quality Assessments Only Methodology

Outcome and Assessment Information Set (OASIS) Quality Measure Development and Maintenance," Contract No. HHSM-500-2008-00019I/HHSM-500-T0001, Modification No. 000002

Submitted to:

Caroline Gallaher (CMS/OCSQ)
Centers for Medicare & Medicaid Services
7500 Security Blvd.
Baltimore, Maryland 21244

Submitted by

Eugene J. Nuccio, PhD

Division of Health Care Policy and Research
University of Colorado Denver
13199 East Montview Blvd
Aurora, CO 80045-7201

July 2, 2014
Date of Submission

Table of Contents

List of Tables and Figures

Pay for Reporting: Quality Assessments Only Methodology

The Home Health Quality Reporting Program (HH QRP) was implemented on January 1, 2007, with home health quality data being collected with the Outcome and Assessment Information Set ("OASIS") data collection instrument. Section 1895(b)(3)(B)(v)(I) of the Social Security Act ("the Act") states that ''for 2007 and each subsequent year, in the case of a home health agency that does not submit data to the Secretary in accordance with subclause (II) with respect to such a year, the home health market basket percentage increase applicable under such clause for such year shall be reduced by 2 percentage points."

The mandate to report quality measure data to the Centers for Medicare & Medicaid Services (CMS) with a resulting reduction in Medicare payments for non-performance is commonly referred to as a "pay-for-reporting program" or "pay-for-reporting requirement." To date, the quantity of OASIS assessments each HHA must submit to meet this requirement has never been proposed and finalized through rulemaking or through the sub-regulatory process. HHAs that have submitted at least one completed OASIS assessment per each 12 month reporting period (i.e. – 07/01 to 06/30) were considered to have met their reporting obligation as stated in section 1895(b)(3)(B)(v)(I) of the Act. The challenge of how to raise home health agency (HHA) performance standards for OASIS quality reporting has been an on-going effort during the past three years.

We believe that defining a more explicit performance requirement or standard for the submission of OASIS data by HHAs would better meet section 5201(c)(2) of the Deficit Reduction Act of 2005 (DRA), which requires that "each home health agency shall submit to the Secretary such data that the Secretary determines are appropriate for the measurement of health care quality. Such data shall be submitted in a form and manner, and at a time, specified by the Secretary for purposes of this clause."

Also, in February 2012, the Department of Health & Human Services Office of the Inspector General (OIG) performed a study to: (1) Determine the extent to which home health agencies (HHAs) meet Federal reporting requirements for the Outcome and Assessment Information Set (OASIS) data; (2) to determine the extent to which states meet federal reporting requirements for OASIS data; and (3) to determine the extent to which the Centers for Medicare & Medicaid Services (CMS) oversees the accuracy and completeness of OASIS data submitted by HHAs. In a report entitled, "Limited Oversight of Home Health Agency OASIS Data"[1], the OIG stated their finding that "CMS did not ensure the accuracy or completeness of OASIS data." The OIG recommended that we "identify all HHAs that failed to submit OASIS data and apply the 2-percent payment reduction to them". We believe that establishing a defined performance requirement for submission of OASIS quality data would be responsive to the recommendations of the OIG.

To create useful quality data, an HHA must submit at a minimum two matching OASIS assessments per patient. These matching assessments together create what is considered a "quality episode of care." A "quality episode of care" ideally consists of a Start of Care (SOC) or Resumption of Care (ROC)

[1] http://oig.hhs.gov/oei/reports/oei-01-10-00460.asp

assessment and a matching End of Care (EOC) assessment — (i.e., transfer to an inpatient facility, death, or discharge from the agency). As will be described in the next section of this report, there are seven types of OASIS assessments that can fit this definition of a "quality assessment."

The report is presented in two sections. The first section, titled "Defining Quality Assessments", presents a discussion of what constitutes quality data and provides a national summary of the quality submission performance for all HHAs. The second section, titled "Quality Assessment Submission Performance Results", presents an analysis of quality assessment submission rates for 12,347 HHAs during the 2012-2013 OASIS reporting period based on the proposed criterion of measuring the submission of Quality Assessments Only (QAO). Together, these two sections identify a strategy for CMS to replace the current performance criterion with one that ensures better compliance with the DRA 2005 requirement for HHAs to submit quality data.

Section One: Defining Quality Assessments

Why are Quality Data Important?

OASIS assessments are combined to form quality episodes and home health quality measures are calculated based on each HHA's quality episodes. When an HHA submits OASIS data forming quality episodes for nearly all of its patients, we can be confident that the resulting quality measure reflects the care that the HHA's patients receive. However, the current "single assessment" standard does not allow for computation of a quality measure.

There are several ways that HHAs could manipulate their submission of assessments that would skew their reported quality scores. First, the HHA could fail to submit initial (SOC/ROC) assessments – perhaps systematically for patients deemed unlikely to improve in dimensions captured by the OASIS quality measures. This would create an additional challenge for the HHA as the payment grouper value for the payment episode is based on the value generated by the submission of the SOC/ROC assessment. Second, the HHA could fail to submit end of care (EOC) assessments – again, perhaps systematically for patients who do not actually improve. By failing to submit the EOC assessment, a quality episode of care could not be computed for the patient and poor quality of care would potentially be undetected. Third, an HHA could submit the SOC/ROC and EOC assessment pairs needed to compute quality measures on only a small proportion of their patients. As with the previous way to manipulate quality performance reporting, an HHA could choose to report only those quality episodes that were successful in terms of patient improvement or health care utilization. Each of these scenarios, whether intentional or unintentional, would result in an inaccurate portrayal of HHA quality performance.

What are Quality Data for the Purpose of the HH QRP/ OASIS Pay-For Reporting Performance Standard?

To ensure that accurate quality data are available for each HHA, the definition of quality data must be broadened to include not only those assessments that can be formed into *Normal Quality Episodes* of care (i.e., an SOC/ROC followed by an EOC for the same patient) but also assessments that may result in quality episodes of care if the performance period were extended either forward or backward. Given that the HHA OASIS reporting period is circumscribed (July 1 of one year to June 30—with an additional

grace period of 30 days to allow for data submission--of the next year), submitted assessments that cannot be formed into *Normal Quality Episodes* of care would still be considered a quality assessment if they can be characterized as any of the following:

- beginning of an episode of care that is not yet complete--an SOC/ROC that occurs in the last 60 days of the performance period (identified as a *Late SOC/ROC*);
- end of an episode of care that began in the previous reporting period--an EOC that occurs in the first 60 days of the performance period (identified as an *Early EOC*);
- an extension of an as yet incomplete quality episode of care--a SOC/ROC assessment that is followed by one or more Follow-up assessments, the last of which occurs in the last 60 days of the performance period (identified as an *SOC/ROC Pseudo Episode*);
- a continuation of a previously begun quality episode of care--an EOC assessment is preceded by one or more Follow-up assessments, the last of which occurs in the first 60 days of the performance period (identified as an *EOC Pseudo Episode*); or
- SOC/ROC assessment that is part of a known one-visit episode (identified as a *One-visit episode*).

Each of these situations would indicate that the assessment submitted by an HHA has the potential to become a quality episode of care or is acceptable as a known one-visit episode.

Table 1 provides a summary of the labels used to describe and categorize each type of assessment as well as the number of assessments required to create each label. That is, to form either a *Normal Quality Episode* or *SOC/ROC Pseudo Episode* or *EOC Pseudo Episode* requires two assessments, whereas only single assessments are required for all of the other labels used to categorize assessments. The total number of assessments for each of these assessment labels is presented in the next section of the report. Table 1 also provides information about other Non-quality or Neutral types of assessments that an HHA might submit. For example, an HHA might submit a SOC/ROC assessment more than 60 days prior to the end of the reporting period. This SOC/ROC assessment should have a Follow-up assessment that occurs within the last 60 days of the reporting period. If the SOC/ROC assessment does not have any Follow-up assessments or if the SOC/ROC assessment does have Follow-up assessments but not one that occurs during the final 60 days of the reporting period, then this SOC/ROC assessment as well as any associated Follow-up assessments would be non-quality assessments because they do not represent quality data.

As noted previously, HHAs also submit OASIS assessments that do not contribute to the forming of quality episodes. For example, an HHA might submit multiple Follow-up assessments for a patient who is under care for an extended period of time (e.g., 180 days). These Follow-up assessments are neutral in that they document care is on-going, but do not directly contribute to creation of a quality episode of care.

There are two situations where an HHA's submission could be suspect (i.e., a non-quality assessment). First, an HHA could submit an SOC/ROC assessment more than 60 days prior to the end of the performance period, but not submit either a Follow-up assessment within 60 days of the end of the reporting period or an EOC assessment as is required if more than one assessment is anticipated for this patient. There are two types of unmatched SOC/ROC assessments. These assessments are labeled as

either *Unmatched SOC/ROC with Follow-up* or *Unmatched SOC/ROC but no Follow-up*. The *Unmatched SOC/ROC with Follow-up* are SOC/ROC assessments that are followed by one or more Follow-up assessment, but the final Follow-up assessment does not fall within 60 days of the end of the reporting period. The *Unmatched SOC/ROC but no Follow-up* assessments are SOC/ROC assessments for which no Follow-up assessment could be found. As noted previously, there are some instances where only one assessment is expected (i.e., a *One-visit episode*). Because the OASIS item M0100 Reason for Assessment, Option 2 (Start of Care—no further visits planned) was removed from the OASIS instrument, there is no way to document this situation using OASIS data. However, claims data associated with these SOC/ROC assessments do indicate whether only one assessment (*One-visit episode*) is expected and is used to identify this type of quality assessment.

TABLE 1: Quality, Non-Quality, and Neutral Assessment Labels.

Quality Labels	Assessments Required	Non-Quality Labels	Assessments Required	Neutral Label	Assessments Required
Normal Quality Episode	2	Unmatched SOC/ROC with Follow-up	1	Follow-up Only	1
Late SOC/ROC	1	Unmatched SOC/ROC but no Follow-up	1		
Early EOC	1	Unmatched EOC with Follow-up	1		
SOC/ROC Pseudo Episode	2	Unmatched EOC but no Follow-up	1		
EOC Pseudo Episode	2				
One-Visit Episode	1				

The second non-quality assessment situation is when an HHA submits an unmatched EOC assessment more than 60 days after the start of the performance period. As with the SOC/ROC assessments presented previously, there are two types of unmatched EOC assessments--*Unmatched EOC with Follow-up* or *Unmatched EOC but no Follow-up*. With the *Unmatched EOC with Follow-up* assessments, one or more Follow-up assessments that preceded the EOC assessment were found for the patient; however, the earliest of these Follow-up assessments did not occur in the first 60 days of the reporting period. For the *Unmatched EOC but no Follow-up* assessments, neither a SOC/ROC nor a Follow-up assessment associated with the EOC assessment could be found in the data base. All four of these non-quality assessments -- *Unmatched SOC/ROC with Follow-up*, *Unmatched SOC/ROC but no Follow-up, Unmatched EOC with Follow-up,* or *Unmatched EOC but no Follow-up* -- would represent assessments that should start or end a quality episode of care (i.e., produce quality data) but do not.

Parsing the 2012 – 2013 OASIS Assessments into Quality, Non-Quality, and Neutral Groups

The first step in determining the quality performance level of individual HHAs is to categorize all assessments from each HHA into one of the six quality, four non-quality, and one neutral assessment groups. This process can be described as parsing the set of all assessments that were downloaded as

part of the "low-bar" pay-for-reporting performance standard analysis using multiple algorithms to identify the quality group designation of the assessment. In addition to the use of OASIS data to identify the first five types of the quality assessment, claims data were matched against an initial set of unmatched SOC assessments to identify if any of these assessments would qualify as a "One Visit Episode"—the sixth quality assessment type.

Table 2: Parsing 2012-2013 OASIS Assessments Using Updated Quality, Non-Quality, and Neutral Definitions, provides a summary of the results of applying these algorithms to identify the number of assessments nationally for each of these quality, non-quality, and neutral designated assessments. The results presented reflect multi-part, analytic activities to generate the identified numbers of assessments from the 2012-2013 OASIS reporting period. There are six, represented in Table 2:

1. Identifying Normal Episodes
2. Identifying Late SOC/ROCs and Initial Unmatched SOC/ROCs
3. Identifying Early EOCs and Initial Unmatched EOCs
4. Creating SOC/ROC Pseudo Episodes and Identifying Unmatched SOC/ROC groups
5. Creating EOC Pseudo Episodes and Identifying Unmatched EOC groups
6. Identifying One Visit Episodes using both OASIS and Claims data

As expected, the total number of assessments submitted between July 1, 2012 and July 31, 2013 (n=17,522,648) was slightly larger than the number of assessments that were deemed legitimate based on the ability to associate an assessment with an active, certified HHA (n=17,393,155). This latter group of assessments is the starting population of assessments used to identify the five groups of quality, four groups of non-quality, and one group of neutral assessments. Once the population of assessments was identified, assessments where the RFA was a Follow-up assessment (04 or 05) were removed from the data set. This allowed the computation of quality episodes of care that were contained within the reporting period. Each of these normal episodes is comprised of two assessments: an SOC/ROC (01 or 03) and an EOC (06, 07, 08, or 09). For the purpose of these analyses, the frequency of each of the specific combination of RFA values is not important. There were 5,446,389 Normal Quality Episodes of care involving 10,892,778 assessments identified. As expected, the majority of all legitimate assessments submitted during the performance period could be combined into Normal Quality Episodes.

TABLE 2: Parsing 2012-2013 OASIS Assessments Using Updated Quality, Non-Quality, and Neutral Definitions.

	# of Assessments/ Episodes
Identifying Normal Episodes	
All downloaded assessments	17,522,648
Total assessments with legitimate CMS Certification Number (CCN)	17,393,155
Legitimate Assessments with no reason for assessment (RFA) 04,05	13,407,316
Assessments without unmatched RFA 01,03 to RFA 06-09	10,892,778
Normal quality episodes in 2012-2013 (Normal Quality Episode)	5,446,389
Identifying Late SOC/ROCs and Initial Unmatched SOC/ROCs	
# Legitimate Assessments with RFA 01,03	6,856,856
RFA 01,03 Assessments not in Regular Episodes	1,410,466
Unmatched RFA 01,03 that occurred in last 60 days (>= 050113) (Late SOC/ROC)	707,485
Unmatched RFA 01,03 that should have RFA 04,05 match	702,981
Identifying Early EOCs and Initial Unmatched EOCs	
# Legitimate Assessments with RFA 06-09	6,550,460
RFA 06-09 Assessments not in Regular Episodes	1,104,071
Unmatched RFA 06-09 that occurred in first 60 days (<=090112) (Early EOC)	596,767
Unmatched RFA 01,03 that should have RFA 04,05 match	507,304
Creating SOC/ROC Pseudo Episodes and Identifying Unmatched SOC/ROC groups	
# Legitimate Assessments with RFA 04,05	3,985,217
Add 702,981 RFA 01,03's that should have match	4,688,198
# of assessments after removing all RFA 04,05's that did not occur in the last 60 days of reporting period and all not sequenced and matched RFA 01,03's & RFA 04,05's	411,292
Pseudo episodes of either RFA 01,03 with either RFA 04,05 (SOC/ROC Pseudo Episode)	205,646
Remaining RFA 01,03 assessments that cannot be matched	497,335
Remaining unmatched RFA 01,03 assessments with at least one Follow-up (Unmatched SOC/ROC with Follow-up)	160,445
Remaining unmatched RFA 01,03 assessments with no Follow-up (Unmatched SOC/ROC but no Follow-up investigated as possible One-visit episode assessments based on claims data analyses.]	336,890
Unmatched RFA 01,03 assessments identified as "One Visit Episodes"	39,788
Final unmatched RFA 01,03 assessments with no Follow-up (Unmatched SOC/ROC but no Follow-up	297,102
Creating EOC Pseudo Episodes and Identifying Unmatched EOC groups	
# Legitimate Assessments with RFA 04,05 + 507,304 RFA 06-09 that should have match	4,492,521
# of assessments after removing all RFA 04,05's that did not occur in the first 60 days of reporting period and all not sequenced and matched RFA 04,05's & RFA 06-09's	674,608
Pseudo episodes of either RFA 04,05 followed by any RFA 06-09 (EOC Pseudo Episode)	337,304
Remaining RFA 06-09 assessments that cannot be matched	170,000
Remaining unmatched RFA 06-09 assessments with at least one Follow-up (Unmatched EOC with Follow-up)	36,771
Remaining RFA 06-09 assessments with no Follow-up that cannot be matched (Unmatched EOC with no Follow-up)	133,229

Only results from *Normal Quality Episodes* are reported in the public (e.g., Home Health Compare) and private (e.g., CASPER reports) performance reports provided to HHAs. However, as was described previously, this is insufficient for identifying all of the assessments submitted by an HHA that could lead to quality data (i.e., assessments, that if the reporting period was longer, we would have the ability to compute a *Normal Quality Episode* of care). The next step in identifying the other quality assessments was to identify all SOC/ROC assessments and eliminate those assessments that already had been used to create the normal episodes of care. The result of these analyses identified 1,410,466 assessments that were unused previously. Of these, 707,485 were identified as being *Late SOC/ROC* assessments; that is, these assessments occurred in the final 60 days of the performance period. These assessments can be viewed as quality assessments, but incomplete with regard to forming an episode of care for two reasons. First, because the payment period for these SOC/ROCs was still in effect, there was no need to submit a Follow-up assessment. Second, no associated EOC assessment was submitted as yet for these assessments. The remaining SOC/ROC assessments (n=702,981) that were submitted earlier in the reporting period should have an associated Follow-up assessment. These assessments were used in an analysis that occurred later in the analytic process.

A similar set of analyses were computed to identify *Early EOC* assessments and the set of EOC assessments that occurred later in the reporting period and should have an associated Follow-up assessment that preceded the EOC assessment. The results of these analyses showed that there were 1,104,071 not previously used EOC assessments, of which 596,767 were *Early EOC's* and 507,304 were EOC assessments that should have had an associated Follow-up assessment that preceded the EOC assessment.

Once each of the SOC/ROC and EOC assessments that should have had associated Follow-up assessments were identified, pseudo episodes of care for each of these groups were computed. A pseudo episode of care is an incomplete episode of care based on the definition of a normal quality episode of care. That is, an *SOC/ROC Pseudo Episode* of care begins with an SOC/ROC assessment and ends with a Follow-up assessment that occurred in the last 60 days of the reporting period. Similarly, an *EOC Pseudo Episode* of care begins with a Follow-up assessment that occurred in the first 60 days of the reporting period and ends with an EOC assessment. As with the *Normal Quality Episode*, there are two assessments that comprise each of the two types of pseudo episodes. If the pay-for-reporting performance period was longer, then there is evidence based on the existence of an associated Follow-up assessment for each of these groups of pseudo episodes that an associated SOC/ROC assessment would have been submitted prior to the performance period or an EOC assessment would be submitted after the end of the current performance period.

To identify the *SOC/ROC Pseudo Episodes*, the 3,985,217 Follow-up assessments were reviewed to identify those Follow-up assessments that occurred in the final 60 days of the reporting period. These Follow-up assessments were combined with the 702,981 unmatched SOC/ROC assessments that should have an associated Follow-up. Using a modified version of the algorithms used to create the *Normal Quality Episodes*, a total of 205,646 *SOC/ROC Pseudo Episodes* of care were identified. There were a

total of 411,292 assessments used to create these *SOC/ROC Pseudo Episodes*. The assessments used to construct the *SOC/ROC Pseudo Episodes* are considered quality assessments.

The remaining 497,335 SOC/ROC assessments were designated as unmatched SOC/ROC assessments (potentially non-quality assessments) and two additional analyses were computed. These 497,335 SOC/ROC assessments were matched against all of the 3,985,217 Follow-up assessments that did not occur in the last 60 days of the reporting period using a modified version of the algorithms used to create the *Normal Quality Episodes*. A total of 160,445 SOC/ROC assessments were associated with at least one Follow-up assessment. These SOC/ROC assessments were labeled as *Unmatched SOC/ROC with Follow-up*. The remaining 336,890 assessments were designated as *Unmatched SOC/ROC with no Follow-up*. Subsequent to this parsing analysis of OASIS assessments, a second analysis involving matching these unmatched SOC/ROC assessments against HHA claims data from this same period was conducted. The result of these analyses identified 39,788 SOC/ROC assessments (approximately 0.23% of all assessments) as quality assessments based on the *One-Visit Episode* criteria. Conversely, there were 297,102 assessments that remained unmatched from this group and these could not be categorized as quality assessments. The results of these analyses are included in the table, even though they were not conducted at the same time as all of the other analyses.

Therefore, there were 457,547 SOC/ROC assessments that could be defined as non-quality assessments made up of two groups—those with some Follow-up assessments (160,445) and those without Follow-up assessments (297,102). While one could argue that the 160,445 *Unmatched SOC/ROC with Follow-up* assessments provide some indication that the HHA was trying to document on-going service to the patient, but failed to follow through with all required Follow-up assessments. For the purpose of some analyses that follow, the 497,335 unmatched SOC/ROC assessments that includes the subsequent identification of One Visit Episodes will be identified as non-quality assessments, recognizing that this is an over-estimation (by 0.23%) of the actual number of non-quality assessments from the unmatched SOC/ROC group.

The processes of identifying *EOC Pseudo Episodes* and unmatched EOC assessments paralleled the processes described in the previous two paragraphs. Follow-up assessments were combined with the 507,304 previously unmatched EOC assessments that occurred in the first 60 days of the reporting period. A different modified set of algorithms that paralleled those used to create *Normal Quality Episodes* was applied to these data. The result of these analyses identified 337,304 *EOC Pseudo Episodes* comprised of 674,608 assessments. There were 36,771 EOC assessments for which one or more Follow-up assessments not occurring in the first 60 days of the reporting period could be found (*Unmatched EOC with Follow-up* assessments). This left 133,229 *Unmatched EOC with no Follow-up* assessments. A total of 170,000 EOC assessments are unmatched and will be designated as non-quality assessments for this analysis. While these assessments fit the definition of non-quality assessments, there is some evidence that for a portion of these assessments HHAs had documented evidence that they were trying to comply with the requirement for on-going follow-up.

Prevalence of Quality, Non-Quality, and Neutral Assessments

Assessing the prevalence of quality, non-quality, and neutral assessments based on the parsing of assessments displayed in Table 2 is challenging. However, Table 3: Raising the Bar Analyses for 2012 - 2013: Summary Characterization as Quality, Non-Quality, and Neutral Assessments, provides a summary and further analysis of how well HHAs performed nationally in submitting quality data. All percentages are based on number of assessments.

TABLE 3: Summary Characterization as Quality, Non-Quality, and Neutral Assessments.

		Assessments		
Quality Group Title	**Description of Quality Group**	**N =**	**% of Total**	
	Total assessments with legitimate CCNs	17,393,155	100.00%	**Quality**
Normal Quality Episode	Normal quality episodes in 2012-2013	10,892,778	62.63%	76.60%
Late SOC/ROC	Unmatched RFA 01,03 that occurred in last 60 days (>= 050113)	707,485	4.07%	
Early EOC	Unmatched RFA 06-09 that occurred in first 60 days (<=090112)	596,767	3.43%	
SOC/ROC Pseudo Episode	Pseudo episodes of either RFA 01,03 with either RFA 04,05	411,292	2.36%	
EOC Pseudo Episode	Pseudo episodes of either RFA 04,05 followed by any RFA 06-09	674,608	3.88%	
One Visit Episode	Single SOC OASIS assessments identified as not requiring an additional visit based on HHA claims data	39,788	0.23%	
				Non-Quality
Unmatched SOC/ROC	Remaining RFA 01,03 assessments that cannot be matched	457,547	2.63%	3.61%
	Unmatched SOC/ROC assessments with Follow-up	160,445	0.92%	
	Unmatched SOC/ROC assessments with No Follow-up	297,102	1.81%	
Unmatched EOC	Remaining RFA 06-09 assessments that cannot be matched	170,000	0.98%	
	Unmatched EOC assessments with Follow-up	36,771	0.21%	
	Unmatched EOC assessments with No Follow-up	133,229	0.77%	
				Neutral
Neutral	Assessments (04 or 05) that do not contribute to episodes	3,442,890	19.79%	19.79%

Approximately 62.6% of all legitimate assessments submitted were used to compute Normal Quality Episodes. By definition, all of these assessments are quality assessments. Late SOC/ROCs and Early EOCs represented 4.0% and 3.4% of assessments, respectively. About 6.2% of assessments were classified as pseudo episodes. As stated previously, *One-Visit Episodes* (a single quality assessment) accounted for 0.23% of all assessments. Based on OASIS assessment data analysis only, a total of 76.6% of all legitimate assessments submitted during the pay-for-reporting performance period are categorized as Quality Assessments.

The remaining unmatched SOC/ROC and unmatched EOC assessments are categorized as Non-Quality Assessments, and represent only 3.6% of all legitimate assessments. Finally, 19.8% of assessments (i.e., all unused Follow-up assessments) were designated as Neutral Assessments.

Defining Quality Performance

There are two different requirements for HHAs regarding submitting assessments. These are:

1. Submit all OASIS assessments (CMS CoP requirement), or
2. Submit all OASIS assessments needed to create quality outcomes (DRA 2005 requirement).

The current HHA pay-for-reporting performance criterion of "submit one assessment during a 12-month reporting period" is far below either of these requirements. Because this report focuses on the requirements for the DRA 2005, the CMS CoP requirement will not be addressed.

Using the expanded definition of which assessments would be categorized as quality, non-quality, or neutral assessments presented in the previous section, the DRA 2005 requirement could be addressed in two ways—one that included both neutral and quality assessments in the performance equation, and one that would only include quality assessments in the HHA performance criteria. Based on multiple analyses of both approaches and multiple presentations to CMS management personnel, the pay-for-reporting performance criterion based on Quality Assessments Only (QAO) was chosen as the pay-for-reporting performance standard. One of the major objections to including neutral (Follow-up) assessments in the pay-for-reporting performance criterion was that an HHA could score 100% on the pay-for-reporting performance criterion by submitting only neutral assessments—and no quality assessments. The annual QAO standard will be raised incrementally to achieve the goal/requirement for an HHA to submit all OASIS assessments.

The Quality Assessments Only (QAO) metric is based on the proportion of Quality and Non-Quality assessments submitted by the HHA and ignoring the number of Neutral assessments submitted by the HHA. Hence, the QAO formula based on this definition would be as follows:

$$\text{QAO} = \frac{(\#\text{ of Quality Assessments})}{(\#\text{ of Quality Assessments } + \#\text{ of NonQuality Assessments})} * 100$$

The QAO definition has advantages and disadvantages. An advantage is that the QAO metric uses the same types of assessments that are used to compute a quality episode of care. Currently, Follow-up (neutral) assessments are ignored in all computations of episodes of care. Indeed, in the cases of a *SOC/ROC Pseudo Episodes* or *EOC Pseudo Episodes,* a Follow-up assessment is used as the proxy for the EOC assessment in the former case or the SOC/ROC assessment in the latter case and counted as Quality rather than Neutral assessment. *Late SOC/ROC* assessments are deemed quality assessments because we would assume that an EOC assessment would be forthcoming at some point after the reporting period. Similarly, *Early EOC* assessments are deemed quality assessments because we would assume that an SOC/ROC assessment occurred at some point prior to the reporting period. A single SOC that satisfies the criterion as a One-Visit Episode also is categorized as a quality assessment.

Additionally, using the QAO metric to assess an HHA's pay-for-reporting performance is likely to have the added value of reducing the number of multiple Follow-up assessments performed on the same patient as these assessments provide no value to the HHA in the QAO computation. A disadvantage of the QAO approach would be that this approach does not encourage submission of all OASIS assessments—at this point in time. However, CMS has other efforts underway that will require that all claims requests are accompanied by the associated SOC/ROC or Follow-up assessment. This matching activity should reduce any pattern by HHAs of not submitting all OASIS assessments.

If an HHA submitted no non-quality assessments, then 100% of the HHA's assessments would be deemed "quality assessments." As an HHA's number of non-quality assessments increases relative to the number of quality assessments submitted, the overall performance of the HHA is lower relative to the DRA 2005 requirement to submit quality assessments. One special case when computing an HHA's pay-for-reporting quality submission performance is when an HHA has no submitted assessments during the reporting period. The current criterion, "submit one assessment during a 12-month reporting period," provides a simple dichotomy (i.e., those HHAs with zero submissions vs. those HHAs with at least one submission). For the proposed new definition of pay-for-reporting performance based on the QAO metric, HHAs with zero submissions would simply be given a quality submission value of 0% as division by zero is mathematically impossible.

The QAO model of measuring HHA quality submissions has several valuable features. First, the goal performance levels (i.e., 100%) based on this model would clearly address the DRA 2005 and address the CMS CoP requirement better than the current "single assessment" standard now in place. Second, the operational definitions used to categorize assessments as quality, non-quality, and neutral are explicit and consistent with current uses for the OASIS assessment information. Third, the model adds no burden on HHAs. Fourth, HHAs can effectively monitor their progress toward achieving the goal of 100% quality assessments using the formula used to compute this value. Fifth, by eliminating the neutral assessments (i.e., unused Follow-up assessments), HHAs cannot simply keep patients on-care indefinitely and demonstrate that they are submitting quality data. Sixth, CMS can incrementally increase the HHA pay-for-reporting performance criterion (i.e., proportion of quality assessments submitted) annually to encourage more HHAs to achieve the 100% quality assessment goal.

Using only the OASIS data that does not include the very small percentage of *One-Visit Episodes* in the computation, the results for HHA pay-for-reporting quality submission performance nationally are encouraging. Based on the QAO definition, an impressive 95.22% of all submitted assessments were designated as quality assessment. This extremely high pay-for-reporting performance using the QAO metric suggests that, in aggregate, HHAs are submitting quality data at very high rates. Certainly there is variation in quality assessment submission performance by CMS Region, state, and for individual HHAs. The determination of this variation by CMS Region, by state, and across individual HHAs will be the focus of Section Two of this report.

Section Two: Quality Assessment Submission Performance Results

The DRA 2005 requirement and its associated 2% penalty in reduced future payments focuses on the performance of individual HHAs, not the performance of all HHAs in aggregate. However, where to set the standard for applying the 2% penalty must be placed in context. As reported in the previous section, the national quality submission performance rate using the QAO definition was very impressive (95.22%). Based on this statistics alone, there is support for the belief that the vast majority of OASIS assessments meet the definition of quality assessments.[2] This national value based on all eligible assessments (note: neutral assessments are not eligible for consideration) provides that most global characterization of the pay-for-reporting performance standard for individual HHAs regarding submission of quality assessments (i.e., the DRA 2005 requirement). Two more refined estimates based on CMS Region and individual state performance would provide additional information regarding how this high level of quality assessment submissions varies based on national geography. These CMS Region and state assessment data and the results are presented in this section of the report.

CMS Region Quality Assessment Submission Performance

Table 4:CMS Region Quality, Non-Quality, and Neutral Assessment Distributions, provides the distribution using percentages for each group of assessments. The percentages[3] are based on the national totals for each group of assessments (e.g., *Normal Quality Episode, Late SOC/ROC*) using the definitions presented in Section One of this report. As expected, CMS Regions 4, 5, and 6 account for 57.6% of all assessments submitted during the 2012-2013 reporting period. In comparison, CMS Regions 7, 8, and 10 accounted for only 7.6% of all submitted assessments during the reporting period.

One of the interesting findings is to compare the percentage of all assessments submitted by a CMS Region with the percentage of Quality or Non-Quality assessment submitted by that same region. For example, CMS Region 1 submitted approximately 6.9% of all assessments and the same percentage of Quality assessments, but only 5.6% of the Non-Quality assessments. CMS Region 1 can be viewed as performing as expected with the submission of Quality assessments, but better than expected for the Non-Quality assessments because they submitted a smaller percentage of these types of assessments (i.e., Non-Quality) than their overall submission percentage. Conversely, CMS Region 2 submitted 9.1% of all assessments, 9.2% of Quality assessments, and 11.2% of Non-Quality assessments. CMS Region 2 can be viewed as performing as expected with the submission of Quality assessments, but worse than expected for the Non-Quality assessments because they submitted a larger percentage of these types of assessments (i.e., Non-Quality) than their overall submission percentage. CMS Region 9 had approximately the same percentage of all assessments submitted (9.3%) as CMS Region 2, but performed better than expected on Quality assessments (10.2%) and worse than expected on Non-Quality assessment (11.0%).

[2] The term quality assessment refers to whether the assessment can contribute to creating a quality episode of care, not the clinical accuracy of the information contained in the assessment.

[3] The One-Visit Episodes (0.23% of assessments) are not included in these totals.

TABLE 4: CMS Region Quality, Non-Quality, and Neutral Assessment Distributions.

	Overall Percentages			% Quality Assessments				
CMS Region	All	Quality	Non-Quality	Normal Episodes	Late 0103	Early 0609	Pseudo 0103	Pseudo 0609
1	6.9	6.9	5.6	7.0	6.5	6.6	6.0	5.8
2	9.2	9.1	11.2	9.4	8.2	8.6	6.5	7.9
3	9.4	10.5	10.1	11.2	9.7	9.3	4.8	4.6
4	23.5	24.2	22.6	24.4	24.8	24.4	21.8	22.4
5	17.2	17.5	17.9	17.6	18.1	17.9	15.9	15.5
6	16.9	13.5	15.1	11.3	15.4	16.2	33.3	33.0
7	3.5	3.7	2.8	3.9	3.2	3.3	2.7	2.5
8	2.0	2.0	1.9	2.1	1.9	1.9	1.6	1.4
9	9.3	10.2	11.0	10.7	10.1	9.5	6.1	5.7
10	2.1	2.3	1.9	2.5	2.1	2.3	1.3	1.3

	Overall Percentages			% Non-Quality Assessments		% Neutral
CMS Region	All	Quality	Non-Quality	No Pseudo 0103	No Pseudo 0609	Leftover 0405
1	6.9	6.9	5.6	5.6	5.6	7.2
2	9.2	9.1	11.2	10.2	14.0	9.1
3	9.4	10.5	10.1	10.2	9.9	5.2
4	23.5	24.2	22.6	22.8	22.1	20.9
5	17.2	17.5	17.9	18.2	17.0	15.8
6	16.9	13.5	15.1	15.5	13.7	30.3
7	3.5	3.7	2.8	2.7	3.3	3.0
8	2.0	2.0	1.9	1.8	2.2	1.7
9	9.3	10.2	11.0	11.3	10.0	5.6
10	2.1	2.3	1.9	1.8	2.3	1.3

The percentage of Quality assessments submitted is not consistent across the subcategories of submissions that define a Quality assessment. CMS Region 6 provides the most dramatic differences among these subcategories. This Region submitted approximately 13.5% of all Quality assessments nationally. However, CMS Region 6 accounts for only 11.3% nationally of all Normal Episodes, but 33.3% and 33.0% nationally of the Pseudo0103 and Pseudo0609 episodes, respectively. The interpretation is that CMS Region 6 has fewer completed episodes of care (i.e., quality episodes) than would be expected, and more than twice what would be expected of incomplete/on-going episodes of care (i.e., Pseudo0103 (*SOC/ROC Pseudo Episodes)* and Pseudo0609 (*EOC Pseudo Episodes)*). Other unpublished analyses conducted by the author for patients from CMS Region 6 have indicated that their length of stay is substantially longer than for other CMS regions.

The other two broad categories of OASIS assessments, Non-Quality and Neutral, display different results when compared to each other. In general, there are few differences between the overall percentages of Non-Quality assessments and the two subcategories that comprise this general category across all 10 CMS regions. Conversely, there are large fluctuations among CMS regions when their overall percentage of assessments is compared with the percentage of neutral assessments. CMS Region 2 has

only about half as many neutral assessments (5.2%) than would be expected (9.4%). CMS Region 9 displays a similar pattern. Conversely and not unexpectedly, CMS Region 6 has nearly twice as many neutral assessments (30.3%) based on their overall percentage of assessments submitted (16.9%).

Based on these results, there is variation among the CMS regions with regard to the number (percentage) of Quality, Non-Quality, and Neutral OASIS assessments submitted. Even simple comparisons (i.e., overall vs. major quality category and major quality category vs. subcategory) can provide insight regarding quality submission performance across CMS regions.

Table 5: CMS Region Quality Submission (Pay-for-Reporting) Rates Based on the QAO Metric Definition, provides results of applying the Quality Assessment Only (QAO) metric definition of quality assessment submissions presented in Section One of this report for each CMS region. Additionally, the table contains the pay-for-reporting performance ranks for each CMS region based on the QAO metric definition of quality submissions. Finally, the table displays the minimum, maximum, and ranges for QAO percentages across these 10 regions.

TABLE 5: CMS Region Quality Submission (Pay-for-Reporting) Rates Based on the QAO Metric Definition.

CMS Region	QAO Definition %	Rank QAO
1	96.1	3
2	94.2	10
3	95.4	6
4	95.5	5
5	95.1	7
6	94.7	9
7	96.3	1
8	95.5	4
9	94.9	8
10	96.1	2
Min	94.2	
Max	96.3	
Range	2.1	

There is little difference among the worst (lowest percentage) and best (highest percentage) CMS regions. Even the worst performing CMS region (CMS Region 2) submitted quality assessments at a rate of 94.2% based on the QAO metric. Conversely, the best performing CMS region (CMS Region 7) quality submissions rate was only slightly higher at 96.3% QAO, although what is needed to move a region's rate by this 2.1% difference is unknown. Lower rates (compared with the current pay-for-reporting performance requirement methodology) and variability for the QAO metric suggest that using this metric as the definition of pay-for-reporting performance would provide both more room for

improvement in quality submissions across the 10 CMS regions. While CMS Region 6's overall ranking of 9th using the QAO definition is not surprising given their excessive use of Neutral assessments that do not contribute in the QAO computations, the ranking of Region 2 as having the lowest pay-for-reporting ranking based on the QAO metric is evidence that more investigation of state level differences would be appropriate.

Based on the results from Table 5, pay-for-reporting performance measure using the QAO metric indicates that the rate of quality assessment submissions is quite high. While there is only a small amount of variation between the best and worst performing CMS regions, the effect of including or not including Neutral assessments was strongly evidenced in the rankings of these CMS regions—notably CMS Region 6.

State-by-State Quality[4] Assessment Submission Performance

Table 6: State-by-State Quality, Non-Quality, and Neutral Assessment Distributions, and Table 7: State-by-State Quality Submission Rates Using the QAO Metric Definition, are the state-by-state counterpart tables to the CMS region Tables 4 and 5 presented previously. To facilitate comparison between the information presented in all four tables, the states in the state-by-state tables are arranged by CMS Region. As with Tables 4 and 5, Tables 6 and 7 report the percent of national assessments within the quality categories that were submitted by each state. For example, Connecticut (CT) submitted 1.95% of all assessments, 1.66% of all Quality assessments, 1.69% of all Non-Quality assessments, and 3.12% of all Neutral assessments submitted nationally. This means that CT submitted about the percentage of Quality and Non-Quality assessments that would be expected based on their overall national submission rate, but somewhat more Neutral assessments than would be expected based on this national rate.

TABLE 6: State-by-State Quality, Non-Quality, and Neutral Assessment Distributions.

		Overall Percentages			% Quality Assessments					% Non-Quality Assessments		% Neutral
CMS Region	St ID	Overall	Quality	Non-Quality	Normal	Late 0103	Early 0609	Pseudo 0103	Pseudo 0609	No Pseudo 0103	No Pseudo 0609	Leftover 0405
1	CT	1.95	1.66	1.69	1.6	1.73	1.68	2.29	2.2	1.71	1.61	3.12
1	MA	3.42	3.55	2.63	3.66	3.31	3.39	2.81	2.7	2.58	2.77	3.08
1	ME	0.45	0.50	0.35	0.54	0.41	0.45	0.2	0.22	0.36	0.33	0.25
1	NH	0.42	0.46	0.20	0.49	0.4	0.42	0.23	0.25	0.23	0.12	0.29
1	RI	0.42	0.46	0.51	0.49	0.45	0.45	0.21	0.2	0.5	0.56	0.24
1	VT	0.23	0.24	0.19	0.24	0.22	0.22	0.22	0.2	0.2	0.17	0.22
2	NJ	2.10	2.45	1.72	2.68	2.01	2.01	0.75	0.68	1.57	2.18	0.82
2	NY	6.57	6.12	8.83	6.15	5.56	5.98	5.38	6.83	7.99	11.29	7.88
2	PR	0.51	0.53	0.60	0.54	0.62	0.59	0.39	0.33	0.63	0.52	0.4
2	VI	0.00	0.00	0.01	0	0	0.01	0	0	0.01	0.02	0
3	DC	0.10	0.11	0.25	0.11	0.14	0.1	0.05	0.06	0.27	0.21	0.05
3	DE	0.21	0.23	0.19	0.25	0.21	0.18	0.13	0.11	0.19	0.2	0.14
3	MD	1.20	1.37	1.66	1.48	1.25	1.19	0.46	0.41	1.76	1.36	0.47

[4] The One-Visit Episodes (0.23% of assessments) are not included in these totals.

TABLE 6: State-by-State Quality, Non-Quality, and Neutral Assessment Distributions. (cont'd)

		Overall Percentages			% Quality Assessments					% Non-Quality Assessments		% Neutral
CMS Region	St ID	Overall	Quality	Non-Quality	Normal	Late 0103	Early 0609	Pseudo 0103	Pseudo 0609	No Pseudo 0103	No Pseudo 0609	Leftover 0405
3	PA	5.08	5.79	4.95	6.27	5.01	4.88	2.13	1.95	4.78	5.44	2.38
3	VA	2.17	2.30	2.56	2.37	2.38	2.23	1.52	1.56	2.7	2.15	1.63
3	WV	0.64	0.69	0.48	0.7	0.71	0.7	0.49	0.5	0.46	0.54	0.48
4	AL	2.12	2.07	1.35	1.97	2.31	2.33	2.74	2.78	1.33	1.39	2.45
4	FL	8.66	9.42	10.09	9.82	9.58	8.82	5.76	5.56	10.39	9.22	5.47
4	GA	2.42	2.51	2.86	2.54	2.56	2.66	1.96	2.19	2.83	2.95	2.01
4	KY	1.72	1.73	1.40	1.72	1.79	1.82	1.8	1.83	1.4	1.39	1.75
4	MS	1.65	1.44	0.86	1.24	1.65	1.66	3.27	3.25	0.89	0.76	2.61
4	NC	2.96	3.01	3.07	3.1	2.88	2.79	2.37	2.29	3.13	2.9	2.78
4	SC	1.31	1.44	1.08	1.5	1.37	1.46	0.84	0.9	0.98	1.37	0.85
4	TN	2.63	2.57	1.88	2.47	2.66	2.88	3.07	3.6	1.8	2.11	2.99
5	IL	4.78	4.77	5.06	4.56	5.32	5.47	5.82	6.21	5.34	4.24	4.78
5	IN	1.61	1.69	1.42	1.72	1.73	1.69	1.35	1.35	1.44	1.34	1.33
5	MI	4.00	4.51	4.72	4.77	4.63	4.55	1.8	1.82	4.64	4.98	1.9
5	MN	1.26	1.11	0.91	1.1	0.94	0.97	1.47	1.3	0.93	0.86	1.92
5	OH	4.63	4.48	4.89	4.47	4.69	4.41	4.84	4.25	5	4.6	5.19
5	WI	0.91	0.97	0.92	1.02	0.81	0.85	0.62	0.55	0.9	1	0.69
6	AR	1.03	0.99	0.73	0.96	1	1.07	1.26	1.28	0.7	0.81	1.22
6	LA	2.55	2.06	2.25	1.71	2.31	2.44	4.87	5.39	2.33	2	4.52
6	NM	0.46	0.45	0.56	0.44	0.49	0.49	0.52	0.49	0.52	0.68	0.49
6	OK	2.07	1.61	1.50	1.33	1.74	1.87	4.34	4.06	1.56	1.32	3.96
6	TX	10.79	8.41	10.02	6.87	9.81	10.35	22.32	21.73	10.4	8.9	20.12
7	IA	0.85	0.74	0.47	0.73	0.66	0.62	1.17	0.94	0.46	0.51	1.33
7	KS	0.61	0.66	0.49	0.69	0.56	0.62	0.43	0.41	0.45	0.61	0.45
7	MO	1.71	1.91	1.56	2.04	1.61	1.7	0.89	0.94	1.46	1.82	0.94
7	NE	0.37	0.39	0.30	0.42	0.34	0.36	0.24	0.22	0.3	0.32	0.27
8	CO	1.05	1.08	0.96	1.12	1.03	0.96	0.93	0.73	0.86	1.27	0.94
8	MT	0.13	0.15	0.06	0.16	0.12	0.13	0.05	0.06	0.06	0.07	0.06
8	ND	0.08	0.09	0.03	0.1	0.06	0.07	0.04	0.04	0.03	0.04	0.06
8	SD	0.08	0.09	0.04	0.09	0.06	0.07	0.03	0.04	0.04	0.05	0.05
8	UT	0.55	0.55	0.76	0.56	0.57	0.57	0.44	0.48	0.78	0.69	0.47
8	WY	0.07	0.08	0.05	0.08	0.07	0.07	0.07	0.06	0.04	0.08	0.07
9	AZ	1.12	1.24	1.41	1.33	1.08	1.03	0.59	0.52	1.39	1.48	0.61
9	CA	7.42	8.14	8.60	8.52	8.1	7.65	4.8	4.56	8.96	7.56	4.41
9	HI	0.11	0.13	0.15	0.14	0.12	0.11	0.04	0.06	0.12	0.25	0.06
9	NV	0.67	0.70	0.82	0.7	0.79	0.7	0.68	0.58	0.88	0.66	0.54
10	AK	0.04	0.05	0.04	0.05	0.04	0.04	0.03	0.03	0.03	0.05	0.03
10	ID	0.33	0.36	0.23	0.37	0.33	0.34	0.23	0.23	0.18	0.35	0.23
10	OR	0.71	0.77	0.75	0.8	0.77	0.75	0.5	0.46	0.71	0.86	0.46
10	WA	1.03	1.16	0.87	1.23	0.99	1.13	0.54	0.57	0.82	1.02	0.56

As expected, there is a large amount of variation in performance by states even within the same CMS Region. For example in CMS Region 2, New Jersey (NJ) contributed 2.10%, 2.45%, and 1.72% of overall, Quality, and Non-Quality assessments nationally. The NJ Quality assessment rate of 2.45% was consistent with its subcategories of Normal Assessment, Late0103, and Early0609 assessments, but much above (meaning that NJ did better than expected) for Pseudo0103 and Pseudo0609 assessments (0.75% and 0.68%, respectively). While NJ's Non-Quality overall and subcategory assessment rates were very consistent, NJ did much better than expected with its Neutral assessments (2.10% Overall vs. 0.82% Neutral). Conversely, New York's (NY) performance on quality submissions was much worse than NJ's. NY contributed 6.57%, 6.12%, and 8.83% of Overall, Quality, and Non-Quality assessments nationally, the latter percentage indicating a problem when compared with the other two percentages. Their performance on the two types of Non-Quality assessment also was somewhat uneven at 7.99% for No Pseudo0103 assessments and 11.29% of the No Pseudo0609 assessments nationally. Finally, NY accounted for 7.88% of all Neutral assessment, as compared with just 6.57% of all submitted assessments. Recall that Table 5 showed CMS Region 2 was ranked low in its pay-for-reporting performance—and NY was the largest contributor to this Region's poor performance given these state results.

A similar story of differences among states within a CMS region is found with CMS Region 6. The overall assessment submission rates for Arkansas (AR) and New Mexico (NM) were relatively similar to their submissions of Quality, Non-Quality, and Neutral assessments. However, Louisiana (LA), Oklahoma (OK), and Texas (TX) had a much different pattern. In general, each of these states submitted fewer than expected Normal Episodes and more than expected Pseudo0103 and Pseudo0609 assessments within the Quality category. Each of these states submitted twice as many Non-Quality and Neutral assessments when compared with their expected number. For example, TX submitted 10.79% of all assessments nationally, but 22.32%, 21.73% and 20.12% of No Pseudo0103, No Pseudo0609, and Neutral assessments, respectively. As with CMS Region 2, the poor pay-for-reporting performance of CMS Region 6 can be traced to these three states: LA, OK, and TX.

As with the CMS Regional results presented previously, this delineation of assessments into Quality, Non-Quality, and Neutral assessments provides a useful comparison tool using their national assessment submission rates as their performance benchmark. While there may be small differences between a state's national submission values and their Non-Quality or Neutral assessment submission rates (e.g., about 2% for LA and OK), proportionally submitting twice as many Non-Quality or Neutral as would be expected (i.e., 2% expected vs. 4% submitted) based on their national submission rates is certainly an area where improvement can be made by these states.

The pattern observed at the CMS region aggregation level regarding pay-for-reporting quality assessment submission performance based on the QAO metric is repeated at the state level in Table 7: State-by-State Quality Submission Rates Using the QAO Metric Definition. While quality assessment submission performance is still quite robust regardless of the measure, the three non-state entities--District of Columbia (DC), Puerto Rico (PR), and Virgin Islands (VI) did have the lowest rates (i.e., minimums) of quality assessment submission. When these three entities were excluded from the

analyses, the minimum state pay-for-reporting performance rate was a respectable 93.2% (NY) while the maximum was a robust 98.4% (ND) based on the QAO metric.

While differences in the QAO metric value among states within a single CMS region can be informative, the rankings of states within regions can also help illuminate the trends previously observed across CMS regions. For example, NJ's pay-for-reporting rank is 11th while NY's rank is 50th. The small difference in quality performance rates (i.e., 96.6 (NJ) vs. 93.2 (NY)) made a very large difference in their national rankings. Similarly, the national rankings for CMS Region 6 showed that three of the five states (LA, TX, and NM) had QAO-based rankings ranging from 38th (LA) to 47th (TX) to 49th (NM). As expected, the choice to exclude Neutral assessments submitted from the computation does influence the number of Quality and Non-Quality assessments included in the QAO computation.

TABLE 7: State-by-State Quality Submission Rates Using the QAO Metric Definition.

CMS Region	ST ID	QAO Definition %	Rank QAO
1	CT	95.1	32
1	MA	96.4	15
1	ME	96.6	12
1	NH	97.9	3
1	RI	94.7	40
1	VT	96.2	21
2	NJ	96.6	11
2	NY	93.2	51
2	PR	94.6	42
2	VI	88.4	53
3	DC	89.3	52
3	DE	96.0	25
3	MD	94.3	48
3	PA	95.9	27
3	VA	94.7	41
3	WV	96.6	10
4	AL	96.8	8
4	FL	94.9	37
4	GA	94.6	44
4	KY	96.1	22
4	MS	97.1	5
4	NC	95.1	33
4	SC	96.4	17
4	TN	96.5	13
5	IL	94.9	36
5	IN	96.0	26
5	MI	95.0	34
5	MN	96.0	24
5	OH	94.8	39
5	WI	95.4	30
6	AR	96.5	14
6	LA	94.8	38
6	NM	94.2	49
6	OK	95.5	29
6	TX	94.4	47
7	IA	96.9	6
7	KS	96.4	16
7	MO	96.1	23
7	NE	96.3	19
8	CO	95.7	28
8	MT	98.0	2
8	ND	98.4	1
8	SD	97.7	4
8	UT	93.6	50
8	WY	96.7	9
9	AZ	94.6	43
9	CA	95.0	35
9	HI	94.5	45
9	NV	94.4	46
10	AK	96.3	20
10	ID	96.9	7
10	OR	95.3	31
10	WA	96.4	18

Min	88.4
Max	98.4
Range	10.0

Min*	93.2
Max	98.4
Range	5.2

* Excludes PR, VI, and DC

What can we learn from the information presented in Tables 6 and 7?

1. Using either the CMS region or state as the unit of aggregation without regard for individual HHA performance, the large majority of assessments submitted can be categorized as quality assessments using the QAO metric for pay-for-reporting performance criterion. In general, the state-by-state pay-for-reporting performances are in the low- to upper-90% levels based on the QAO metric.
2. Relatively small differences in QAO values can result in large differences in national ranking as shown in CMS Region 2 when the QAO values and ranks for NJ and NY are compared.
3. The exclusion of Neutral assessments (i.e., extra or unmatched 04 or 05 assessments) from the computational equation can make a very large difference in the national ranking in both CMS region and state aggregations, especially in CMS Region 6.
4. Overall rates of assessments submitted, quality assessments submitted, and non-quality assessments submitted by both the CMS region and state aggregation levels can be used as benchmark values to see how submission rates for the individual subcategories that comprise these three summary values vary from the benchmark values. That is, if a state wanted to improve its overall quality submission rate, the Quality Improvement Organization (QIO) group in that state could identify which of the assessment subcategories (i.e., normal episodes, pseudo0103, pseudo0609, early0609, or late0103) were higher than the overall quality submission rate and work to reduce these percentages.

Challenges in Achieving the Pay-for-Reporting Quality Performance Goal

As stated previously, the pay-for-reporting performance standard using the QAO metric should encourage continuous improvement activities by HHAs to achieve the goal of 90% of HHAs submitting at least 90% quality assessments (90/90). Thus far, all of the analyses have used aggregated data at the CMS Region or state. However, the pay-for-reporting performance criterion is an HHA-level metric. Therefore, while these pay-for-reporting analyses of CMS regional and state aggregations are informative, they cannot be prescriptive regarding the national performance goal for individual HHAs. The specific pay-for-reporting performance standard based on the QAO metric needs to take into consideration performance at the HHA-level, and the challenge in assessing quality data submissions for HHAs with very small numbers of assessments submitted. While the problem is side-stepped with the current low-bar requirement of "one assessment," the problem cannot be ignored with the QAO metric.

This section will present agency-level pay-for-reporting performance based on the QAO metric and display the results for national, CMS region, and state-by-state aggregations. Additionally, the pay-for-reporting performance values will be presented for each level of aggregation (i.e., nation, CMS regions, and state-by-state) for all HHAs, HHAs with at least 20 assessments, and HHAs with fewer than 20 assessments as one attempt to identify the challenge of applying pay-for-reporting to smaller HHAs.

Table 8: National Distribution of HHA Quality Scores, displays the QAO values for a range of percentile values from the 0.5th percentile through the 90th percentile, with greater delineation at the lower

performance levels for all 11,718 HHAs with at least one submitted OASIS assessment in pay-for-reporting reporting year 2012-2013. These results show that fewer than 5% of HHAs have a QAO value of <70%; or stated positively, more than 95% of all HHAs have a QAO value of greater than 70%. Similarly, more than 90% of HHAs have a QAO value of greater than 80%, and slightly less than 80% of all HHAs have a QAO value of 90%. All these results further support the results using aggregated data from the CMS Region and state levels that were presented previously that when these data are analyzed at the HHA-level there is strong evidence that HHAs are submitting quality data at a high rate nationally. Based on this information, a CMS goal of getting 90% of HHAs to submit 90% quality assessments ("90/90") does not appear to be out of reach if an incremental approach is used.

TABLE 8: National Distribution of HHA Quality Scores.

	All HHAs	
	(N=11718)	
Percentile	**# in Group**	**QAO**
.5th	59	0
1st	117	22.2
2nd	234	49.0
5th	586	70.2
10th	1172	81.1
20th	2344	89.1
30th	3515	92.4
40th	4687	94.3
50th	5859	95.7
60th	7031	96.7
70th	8203	97.6
80th	9374	98.3
90th	10546	99.1

The issue of HHA size is considered when displaying quality performance data on Home Health Compare (HHC). If an HHA has fewer than 20 quality episodes of care for a particular quality outcome measure, their score is suppressed from being displayed on HHC. Given the challenges of small sample sizes, computing stable QAO values for smaller HHAs (defined as those HHAs submitting fewer than 20 assessments annually) also will be challenging.

One approach to challenging HHAs to improve their pay-for-reporting performance based on their QAO metric would be to increase the required QAO score incrementally each year. For example, CMS could begin with a requirement that all HHAs attain a QAO score of at least 70%, and then increment this by 10% each year to a goal QAO value of 90%. This would appear to be reasonable based on national data (n=11,718 HHAs), but an analysis of how this would differentially impact larger HHAs (defined as those HHAs submitting fewer than 20 assessments annually; n=11,372 HHAs) and smaller HHAs (n=346 HHAs).

Table 9: Number and Percentage of HHAs Below Three Pay-for-Reporting Performance Standards Based on 2012-2013 Assessment Reporting, provides an analysis of how these two groups of HHAs are affected by the 70%, 80%, and 90% standards based on current performance data.

Table 9: Number and Percentage of HHAs Below the Pay-For-Reporting Performance Standards, Based on 2012-2013 Assessment Reporting

			Large HHAs	Small HHAs
		All HHAs	(>= 20 Assessments)	(< 20 Assessments)
	Total	11,718	11,372	346
QAO Scores <70%	**# HHAs**	574	435	139
	% HHAs	4.9%	3.8%	40.2%
QAO Scores <80%	**# HHAs**	1066	912	154
	% HHAs	9.1%	8.0%	44.5%
QAO Scores <90%	**# HHAs**	2589	2399	190
	% HHAs	22.1%	21.1%	54.9%

Using the initial pay-for-reporting performance criterion of at least a70% QAO value, 574 HHAs (4.9%) nationally would receive the 2% DRA market basket reduction because these HHAs failed to meet the pay-for-reporting standard base on their 2012-2013 assessment submissions. However, these HHAs would be split differently among the larger and smaller HHAs. More (in absolute numbers) larger HHAs would receive the market basket reduction (n=435); however, this would represent a smaller percentage (3.8%) than the national percentage. The reverse would be true for the small HHAs. Only 139 smaller HHAs would receive the market basket reduction; however, this represents 40.2% of all smaller HHAs. This pattern holds as the pay-for-reporting performance criterion based on the QAO metric continues to increase to achieve the 90/90 performance goal. As both Tables 8 and 9 show, currently HHAs are performing at a 90/80 rate (i.e., 90% of HHAs are submitting at least 80% quality assessments based on the QAO metric computation).

CMS Region and State Differences in Attaining the Initial Pay-for-Reporting Performance Criteria of QAO at least 70%

Table 10: Summary Distribution of HHAs with a QAO Score of <70% by Region and State by HHA Size, displays the frequency distribution of HHAs that would fail to meet the initial proposed minimum pay-for-reporting performance standard by CMS region, by state, and by HHA size. All CMS Regions have at least a few HHAs that do meet the proposed initial pay-for-reporting standard of QAO greater than or equal to 70%. Only a few states (PA, FL, MI, OH, and TX) have double digit numbers of small HHAs failing to meet the proposed initial pay-for-reporting criteria.

Table 10: Summary Distribution of HHAs with a QAO Score of <70% by Region and State by HHA Size.

		QAO <70%		
CMS Region	State	All HHAs	Large HHAs	Small HHAs
1	CT	1	1	0
1	MA	12	9	3
1	RI	1	1	0
2	NY	3	3	0
2	PR	1	1	0
3	DC	5	3	2
3	MD	3	2	1
3	PA	27	12	15
3	VA	14	8	6
3	WV	2	2	0
4	FL	74	62	12
4	GA	1	0	1
4	TN	3	2	1
5	IL	33	30	3
5	IN	5	5	0
5	MI	48	37	11
5	MN	11	6	5
5	OH	73	60	13
5	WI	3	2	1
6	LA	8	8	0
6	OK	9	8	1
6	TX	141	95	46
7	IA	6	3	3
7	MO	3	2	1
7	NE	1	1	0
8	CO	5	3	2
8	UT	10	8	2
9	AZ	10	6	4
9	CA	53	47	6
9	NV	5	5	0
10	OR	1	1	0
10	WA	2	2	0
Total		574	435	139

Summary

This document is comprised of two major sections. In the first section HHA submission of OASIS assessments is categorized into three groups: Quality, Non-quality, and Neutral. The methodology used to parse all assessments submitted during the reporting period as well as the results from this effort were presented. In the second section, the Quality Assessments Only (QAO) was defined and tested using national, CMS region, and state level aggregations. Patterns of performance were identified for each of these levels of aggregation. Additionally, the application of results to the HHA was explored, specifically as they related to larger and smaller HHAs. A hypothetical CMS policy of an initial pay-for-reporting performance standard based on HHAs achieving a QAO value of at least 70% was tested using OASIS assessment submissions from the 2012-2013reporting period to assess the potential impact of this policy decision.