CASE STUDY: NATIONAL HEALTHCARE QUALITY AND DISPARITIES REPORT (QDR) SENSITIVITY ANALYSIS ON DEVELOPING AHRQ QUALITY INDICATOR ESTIMATES FOR 2015 USING ONLY ICD-9-CM DATA

Recommended Citation: Barrett ML, Heslin KC, Yoon F, Moore BJ. *Case Study: National Healthcare Quality and Disparities Report (QDR) Sensitivity Analysis on Developing AHRQ Quality Indicator Estimates For 2015 Using Only ICD-9-CM Data.* ONLINE. April 7, 2017. U.S. Agency for Healthcare Research and Quality. Available: https://www.hcup-us.ahrq.gov/datainnovations/icd10_resources.jsp.

TABLE OF CONTENTS

INDEX OF FIGURES

EXECUTIVE SUMMARY

The following case study describes the methodology and findings of an analysis investigating how best to report 2015 national and State-level rates for the AHRQ Quality Indicators (AHRQ QIs™) in the *National Healthcare Quality and Disparities Report* (QDR). The QDR provides a comprehensive overview of the quality of health care received by the general population and disparities in care experienced by different racial, ethnic, and socioeconomic groups. The QDR uses the QIs to examine trends in measures of quality associated with processes of care that occur in an outpatient or inpatient setting.

The most recent QDR examined national and State-level trends in the QIs from 2000–2014 using the Healthcare Cost and Utilization Project (HCUP) State Inpatient Databases. The data and QI programs were based on ICD-9-CM diagnoses and procedures. The 2015 HCUP databases include ICD-9-CM codes in the first three quarters of the year (Q1–Q3) and ICD-10-CM/PCS codes in the fourth quarter (Q4). The transition to ICD-10-CM/PCS introduces variation into the annual trends that complicates the ability to evaluate substantive differences over time. In addition, although there is a version of the AHRQ QI software for ICD-10-CM/PCS, it does not have the option to calculate risk-adjusted rates (as of the time that this case study was undertaken). The QDR reports risk-adjusted QI rates for comparative purposes.

Because of differences in the coding schemes and the need for risk-adjusted rates, the QDR team needed to determine the best approach to calculate QI estimates for 2015 based solely on ICD-9-CM data. To test alternative approaches, we compared QI estimates for calendar year 2014 ("true" 2014 estimates) with three alternative estimates:

- Fiscal year estimates (12 months of data from Q4 2013 to Q3 2014)
- Partial year estimates (9 months of data from Q1 2014 to Q3 2014)
- Weighted partial year estimates (9 months of data from Q1 2014 to Q3 2014, weighted to represent a full year based on quarterly variation in the prior year).

Because the QDR reports national QI rates by select patient characteristics (race and ethnicity, age, sex, income quartile, patient location, and expected primary payer) and hospital characteristics (region, ownership, teaching status, bed size, and location), we calculated national QI rates by all reporting categories for the calendar year 2014 data and the three alternative approaches. In addition, we calculated State-level QI rates for the calendar year 2014 data and the three alternatives. The case study limited the analysis to the 41 QIs reported in the QDR (out of 88 total QIs).

Within each reporting category, we calculated the absolute percentage difference between the true 2014 QI estimates and each of the three alternative estimates. The rationale for performing this calculation was that the optimal alternative approach yields QI rates that are closest to the true rate. Differences of more than 1 and 3 percent were counted. To summarize "closeness," we defined a discordance diagnostic that characterizes the closeness across the QIs within a reporting category as a percentage of comparisons above two separate thresholds: greater than 1 percent difference and greater than 3 percent difference.

Comparing the fiscal year national QI estimates to the true 2014 estimates:

- The difference was greater than 1 percent in 26 of the 41 QIs (63.4 percent).
- The difference was greater than 3 percent in 10 of the 41 QIs (24.4 percent).

Comparing the partial year national QI estimates to the true 2014 estimates:

- The difference was greater than 1 percent in 23 of the 41 QIs (56.1 percent).
- The difference was greater than 3 percent in 7 of the 41 QIs (17.1 percent).

The results for comparing the weighted partial year national QI estimates to the true 2014 estimates were similar to the partial year comparison:

- The difference was greater than 1 percent in 23 of the 41 QIs (56.1 percent).
- The difference was greater than 3 percent in 7 of the 41 QIs (17.1 percent).

The partial year estimates were less likely than the other two alternative approaches to be more than 3 percent different than the true 2014 estimates for the 41 QIs in the QDR when comparing national QI estimates by race and ethnicity, age, sex, income quartile, patient location, and hospital characteristics. In addition, the partial year estimates were less likely to be more than 3 percent different than the true 2014 estimates for the 41 QIs in the QDR when comparing State-level QI estimates.

Because the QDR focuses on examining trends in health care in the United States, it is essential to be able to test whether the 2015 estimates are significantly different from the calendar year 2014 estimates already used in the QDR trends. The current approach to testing significant differences between pairs of QI estimates in the QDR is a two-sided χ^2 test that assumes independent samples. This approach would continue to work for the partial year estimates, but not for the fiscal year estimates because of the overlapping quarter of data.

On the basis of this case study, the QDR team decided to calculate 2015 QI estimates using partial year data for the final year of reporting QI trends under ICD-9-CM. The QDR will document that the 2015 QI rates are not actual annual rates, but instead are based on 9 months of data from Q1 2015 to Q3 2015.

It is important to note that this case study was specific to the reporting of national and State-level QIs using ICD-9-CM data. Most users of the QI software derive hospital-level QI rates. It is unknown whether there would be similar results if comparing hospital-level QI rates.

INTRODUCTION

The following case study describes the methodology and findings of an analysis of how best to report 2015 national and State-level rates for the AHRQ Quality Indicators (AHRQ QIs™)[1] in the *National Healthcare Quality and Disparities Report* (QDR).[2] The QDR provides a comprehensive overview of the quality of health care received by the general population and disparities in care experienced by different racial, ethnic, and socioeconomic groups. The QDR uses the QIs to examine trends in measures of quality associated with processes of care that occur in an outpatient or inpatient setting. The most recent QDR examined national and State-level trends in the QIs from 2000 to 2014 using the Healthcare Cost and Utilization Project (HCUP) State Inpatient Databases (SID).[3]

The QI software relies on ICD diagnosis and procedure codes to create quality measures of hospital care. The software assigns hospital encounters to the numerators and denominators of each measure. The transition from ICD-9-CM[4] to ICD-10-CM/PCS[5] coding in October 2015 meant that a potential discontinuity would emerge in observed trend lines, owing solely to the change in coding and independent of true shifts in medical care. Therefore, the QDR team conducted the following analysis to investigate the best approach to calculating the QI measures commonly reported in the QDR for data year 2015 using only ICD-9-CM data.

METHODS

To test approaches to calculating annual QI estimates using partial year (PY) data, we compared the observed calendar year (CY) 2014 estimates to three alternative approaches. We considered key reporting categories included in the QDR reports: national, national by subpopulation, and State.

Measures

The AHRQ QIs are measures of health care quality that use readily available hospital inpatient administrative data. The QDR uses the QIs to highlight potential quality concerns and track changes over time. The AHRQ QIs used for the QDR include four sets of measures—Inpatient Quality Indicators (IQIs), Prevention Quality Indicators (PQIs), Patient Safety Indicators (PSIs), and Pediatric Quality Indicators (PDIs, formerly known as PedQIs). Only 41 of the 88 QIs are used in the QDR (see Appendix A for the list of QIs used in the QDR).

There are two types of QIs used in the QDR:

- Provider-level QIs in which both the numerator and denominator for the quality measure are based on discharge data. The numerator is the number of events and the

[1] Further information on the AHRQ QIs, including documentation and free software downloads, is available at http://www.qualityindicators.ahrq.gov/ .

[2] The *National Healthcare Quality and Disparities Report* (QDR) measures and tracks trends in quality and disparities in seven key areas of health care: patient safety, person-centered care, care coordination, effective treatment, healthy living, care affordability, and access to health care. AHRQ designs and produces the QDR, with support from the Department of Health and Human Services (HHS) and private sector partners. Please see https://www.ahrq.gov/research/findings/nhqrdr/index.html for further details.

[3] For more information on the HCUP inpatient databases, please see the HCUP User Support Web site.

[4] ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification.

[5] ICD-10-CM/PCS: International Classification of Diseases, Tenth Revision, Clinical Modification/Procedure Coding System.

denominator is the number of discharges at risk for the event. An example of a provider-based QI is death per 1,000 adult admissions with pneumonia.

- Area-level QIs have only the numerator number of events derived from discharge data. The denominator is based on population at risk. An example of an area-based QI is hospital admissions for short-term complications of diabetes per 100,000 population aged 18 years and older. For the national rate for this QI, the population aged 18 years and older would be specific to the U.S. population. For a State-level rate for this QI, the population aged 18 years and older would be specific to the State.

Of the 41 QIs reported in the QDR, 23 are provider-level QIs and 18 are area-level QIs. For the QDR, the QI software has been modified to calculate national and State-level rates, even for the provider-level QIs. Provider-level QI rates are not calculated or used by the QDR.

Data Source

QI estimates for the QDR are derived using the HCUP SID.[6] The HCUP SID were used for reporting State-level QI rates, as well as for creating an analysis file designed to provide national estimates for overall and by subpopulation such as by race and ethnicity, age, sex, and community income quartile. Additional information on how the QI software was applied to the HCUP data for the QDR is available in the *HCUP Methods Series Report #2016-01*.

Three Alternative Approaches

To test alternative approaches to calculating QI estimates using limited data in a year, we compared QI estimates for CY 2014 (i.e., the "true" 2014 estimates) to three alternative approaches:

- Fiscal year (FY) estimates (12 months of data from Q4 2013 to Q3 2014)
- PY estimates (9 months of data from Q1 2014 to Q3 2014)
- Weighted partial year (wPY) estimates (9 months of data from Q1 2014 to Q3 2014, weighted to represent a full year based on quarterly variation in the prior year).

Table 1 details how the three alternative approaches were calculated for the two different types of QIs.

[6] An overview of the SID is available on the HCUP User Support Web site at http://www.hcup-us.ahrq.gov/sidoverview.jsp.

Table 1. Three Alternative Approaches to Calculating AHRQ QI Estimates

Estimate Type	Provider-Level QI	Area-Level QI
Fiscal year	Numerators and denominators used 12 months of HCUP data from October 2013 to September 2014.	Numerators used 12 months of HCUP data from October 2013 to September 2014. Denominators used one-fourth of the national or State population from 2013 and three-fourths from 2014.
Partial year	Numerators and denominators used 9 months of HCUP data from January 2014 to September 2014.	Numerators used 9 months of HCUP data from January 2014 to September 2014. Denominators used three-fourths of the national or State population from 2014.
Weighted partial year	Numerators used 9 months of HCUP data from January 2014 to September 2014, weighted to account for the missing 3 months of numerator data. Denominators used HCUP data from January 2014 to September 2014, weighted to account for the missing 3 months of denominator data.	Numerators used 9 months of HCUP data from January 2014 to September 2014, weighted to account for the missing 3 months of data. Denominators used the national or State population from 2014.

Abbreviations: AHRQ QI, AHRQ Quality Indicators; HCUP, Healthcare Cost and Utilization Project.

The weights for the wPY used prior year data (from 2013) to estimate the missing quarter of data in the numerator and denominator counts. The weights were calculated separately for numerator and denominator counts. For example, the weights for the numerator of QI_i were calculated as the following:

$$\frac{\text{Sum of counts for } QI_i \text{ from Jan to Dec 2013 (12 months of data)}}{\text{Sum of counts } QI_i \text{ from Jan to September 2013 (9 months of data)}}$$

Table 2 shows an example of calculating wPY estimates using the following sample data for 2013 and 2014.

Table 2. Sample Data for Example of Calculating Weighted Partial Year Estimates

Year	Quarter	Denominator Counts	Numerator Counts
2013	1	500	5
	2	300	4
	3	200	3
	4	500	2
2014	1	600	4
	2	400	2
	3	350	2

The weight for the denominator uses 2013 denominator counts:

(500 + 300 + 200 + 500) / (500 + 300 + 200) = 1.50

The weighted denominator for 2014 is equivalent to the denominator counts for Q1–Q3 2014 times the denominator weight:

(600 + 400 + 350) * 1.50 = 2025

The weight for the numerator uses 2013 numerator counts:

(5 + 4 + 3 + 2) / (5 + 4 + 3) = 1.17

The weighted numerator for 2014 is equivalent to the numerator counts for Q1–Q3 2014 times the numerator weight:

(4 + 2 + 2) * 1.17 = 9.33

The wPY estimate for 2014 would be 9.33 / 2025.

Discordance Diagnostic

Our quantitative analysis of the alternative approaches is based on a comparison of QI rates calculated from FY, PY, and wPY estimates against the true 2014 estimate that is based on the CY data. We first calculated the absolute percentage differences in QI rates between the alternative approach and the CY estimate. The motivation is that the optimal alternative approach yields QI rates that are closest to the true rate. To summarize “closeness” we defined a discordance diagnostic that characterizes the closeness.

For the discordance diagnostic, we first determined the absolute percentage difference of each of the alternative estimates from the CY estimate. For example, the CY national estimate for PSI 18 *Obstetric trauma in instrument-assisted deliveries* was 121.42 per 1,000 deliveries and the FY estimate was 124.55 per 1,000 deliveries. The absolute percentage difference was 2.6 percent (absolute value of [124.55 – 121.42] / 121.45 * 100). We used the absolute percentage because we were focused on how close the estimates were, not whether one was lower or higher.[7]

Next, we flagged whether the CY estimate was different from the alternative estimate by two different thresholds: greater than 1 percent difference and greater than 3 percent difference. (The "flag" for each QI was a dichotomous variable that was equal to 1 if the absolute percentage difference exceeded the threshold and equal to 0 if it did not exceed the threshold.) In the above example for PSI 18, the FY estimate would be flagged as more than 1 percent different, but not more than 3 percent different. If the percentage difference was greater than 3 percent, it would be flagged as both 1 percent and 3 percent different.

The flagged differences then were summed across QIs within the following QDR reporting categories: national rates overall and by race and ethnicity, age, sex, income quartile, patient location, payer (for provider-level QIs only), and five hospital characteristics (region, ownership, teaching status, bed size, location) as one group. We also summed the flagged differences across the State-level rates.

To evaluate the three alternative approaches, we calculated the discordance diagnostic as the percentage of comparisons with differences above the two different thresholds. The denominator of the percentage is the total number of possible comparisons; the numerator is the number of comparisons above the threshold. If the CY estimate was unavailable because it had been suppressed because of small sample size or a relative standard error greater than 30 percent, then there was no possible comparison and the QI was not counted in the numerator or denominator.

Ability to Test for Significant Differences in QI Trends

Because the QDR focuses on examining trends in health care in the United States, it is essential to be able to test whether the 2015 estimates are significantly different from the CY 2014 estimates. The current approach to testing significant differences between pairs of QI estimates in the QDR is a two-sided χ^2 test that assumes independent samples. In addition to looking at the discordance diagnostics, we examined the ability to test significant differences between the 2015 QI rates derived using the alternative approaches and the existing CY 2014 QI estimates used for the QDR trends.

RESULTS

Comparisons of the Calendar Year Estimates With the Three Alternative Approaches

Results are reported across the 41 QIs reported in the QDR and then separately for area-level and provider-level QIs. Figures 1–9 show the CY estimates compared with the three alternative approaches within reporting categories.

---

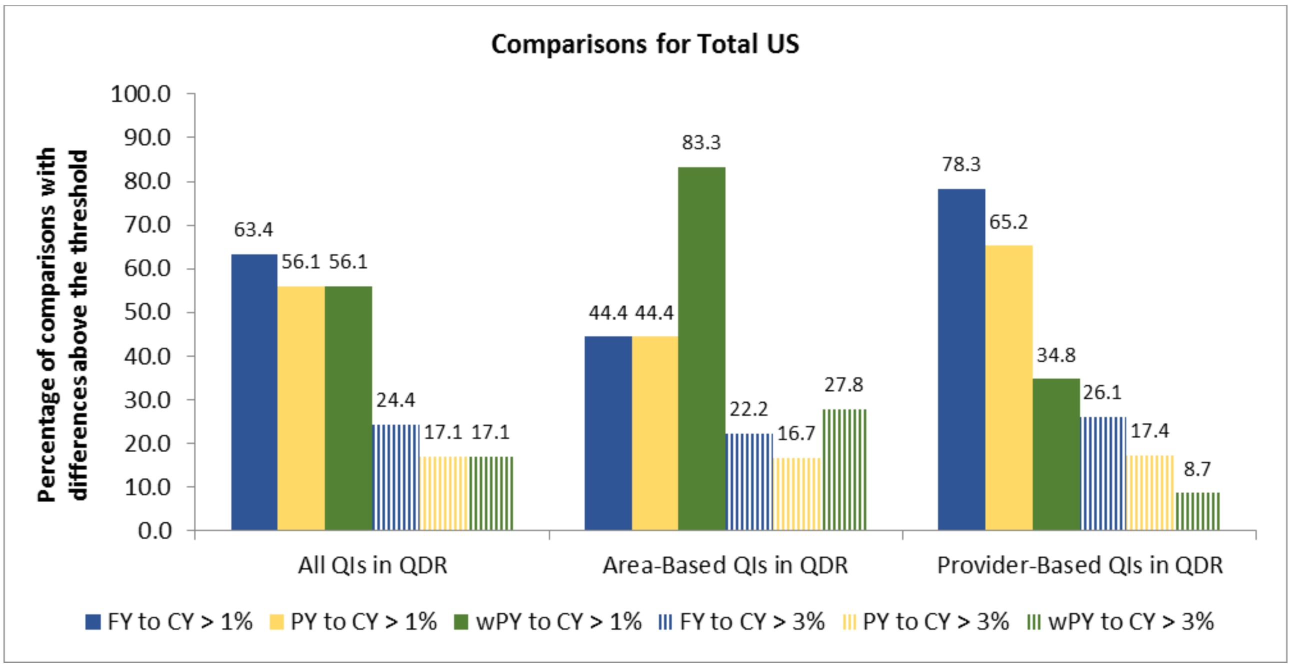
[7] It should be noted that the use of the percentage difference is more sensitive to variation in QIs with small rates. For example, the true 2014 estimate for PSI 14 *Postoperative abdominal wound dehiscence* is 1.61 per 1,000 adult abdominopelvic surgery discharges and the FY estimate is 1.65 with an absolute percentage difference of 2.5 percent. The percentage difference for PSI 14 (2.5 percent) is very similar to the percentage difference for PSI 18 (2.6 percent), but the true 2014 and FY estimates for PSI 14 are numerically much closer (from 1.61 to 1.65) than PSI 18 (from 121.42 to 124.55). Given that the QDR focuses on testing trends, we think that using the criteria of a percentage point difference is appropriate.

Figure 1 shows the comparisons of the CY 2014 national estimates with the three alternative approaches. The percentage is the number of QIs with estimates that were greater than the two thresholds (1 percent and 3 percent). For example, in 63.4 percent of the 41 QIs, the absolute difference between the FY and CY estimates was greater than 1 percent.

The results showed the following:

- For the area-based QIs, the PY estimates were most similar to the CY estimates (i.e., had the lowest percentages of QIs failing the 1 and 3 percentage point threshold).
- For the provider-based QIs, the weighted PY estimates were most similar to the CY estimates.

Figure 1. Comparisons of the CY 2014 National Estimates With the Three Alternative Approaches



Abbreviations: CY, calendar year; FY, fiscal year; PY, partial year, QDR, Quality and Disparities Report; QIs, quality indicators; wPY, weighted partial year.

For the FY estimates that were greater than the CY estimates by at least 3 percent, the average percentage difference was 4.9 percent.[8] The similar average was 3.7 percent for the PY to CY comparison and 4.5 percent for the wPY to CY comparison.

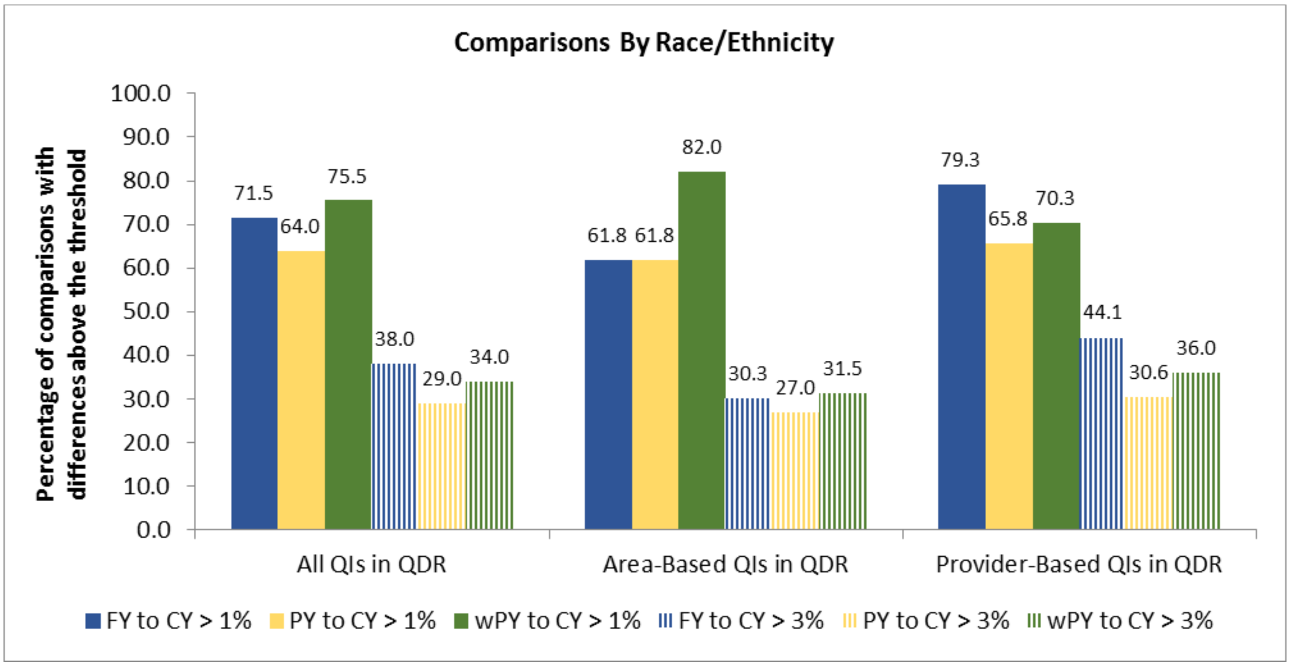
[8] The added PQI for immunization-preventable influenza was excluded from these averages because of expected annual variation in the incidence of influenza.

Figure 2 shows the CY 2014 national estimates by race and ethnicity compared with the three alternative approaches. With 41 QIs and 5 race categories (White, Black, Hispanic, Asian-Pacific Islander, and other), there were a maximum of 205 comparisons possible. Because of suppressed CY estimates (due to small sample size or a relative standard error greater than 30 percent), there were 200 comparisons (89 for the area-based QIs and 111 for the provider-based QIs).

The results showed the following:

- For the area-based QIs, the PY estimates were most similar to the CY estimates.
- For the provider-based QIs, the PY estimates were most similar to the CY estimates.

Figure 2. Comparisons of the CY 2014 National Estimates With the Three Alternative Approaches, by Race and Ethnicity



Abbreviations: CY, calendar year; FY, fiscal year; PY, partial year, QDR, Quality and Disparities Report; QIs, quality indicators; wPY, weighted partial year.

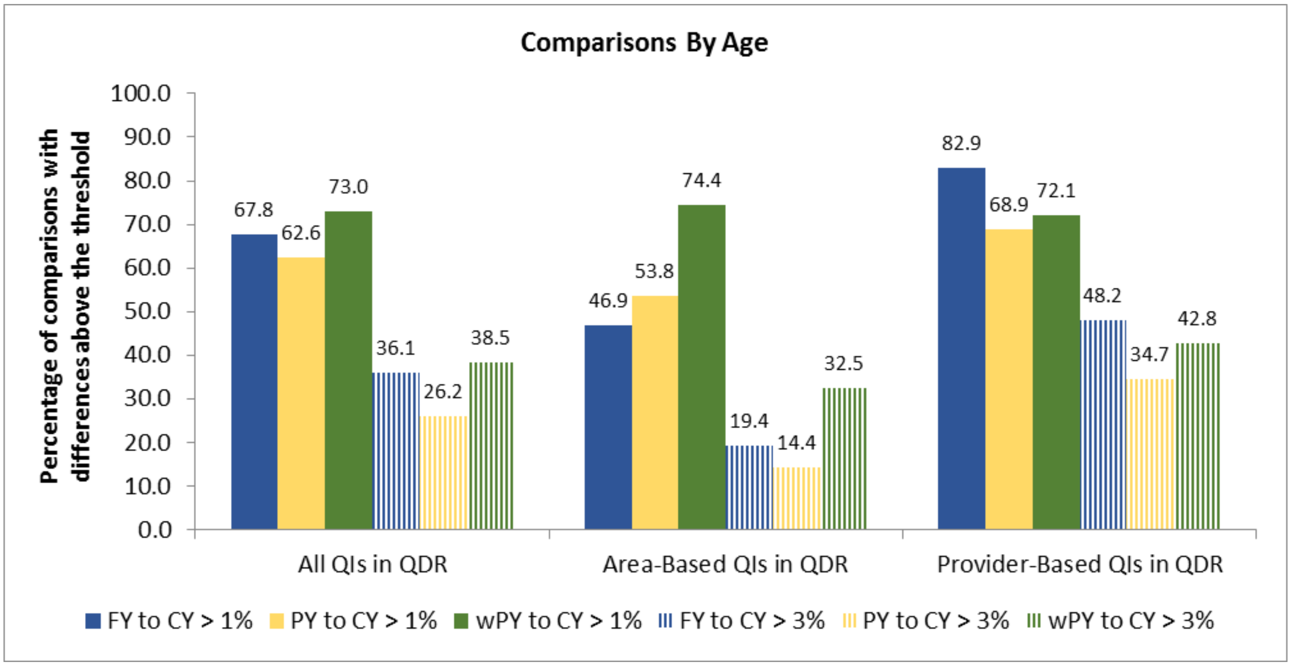
For the FY estimates that were greater than the CY estimates by at least 3 percent, the average percentage difference was 6.8 percent. The similar average was 6.0 percent for the PY to CY comparison and 7.1 percent for the wPY to CY comparison.

Figure 3 shows the CY 2014 national estimates by age compared with the three alternative approaches. With 41 QIs and various age categorizations,[9] there were 382 comparisons (160 for the area-based QIs and 222 for the provider-based QIs).

The results showed the following:

- For the area-based QIs, the FY estimates were most similar to the CY estimates at the 1 percent threshold, but the PY estimates were most similar to the CY estimates at the 3 percent threshold.
- For the provider-based QIs, the PY estimates were most similar to the CY estimates.

Figure 3. Comparisons of the CY 2014 National Estimates With the Three Alternative Approaches, by Age



Abbreviations: CY, calendar year; FY, fiscal year; PY, partial year, QDR, Quality and Disparities Report; QIs, quality indicators; wPY, weighted partial year.

For the FY estimates that were greater than the CY estimates by at least 3 percent, the average percentage difference was 6.9 percent. The similar average was 6.7 percent for the PY to CY comparison and 7.4 percent for the wPY to CY comparison.

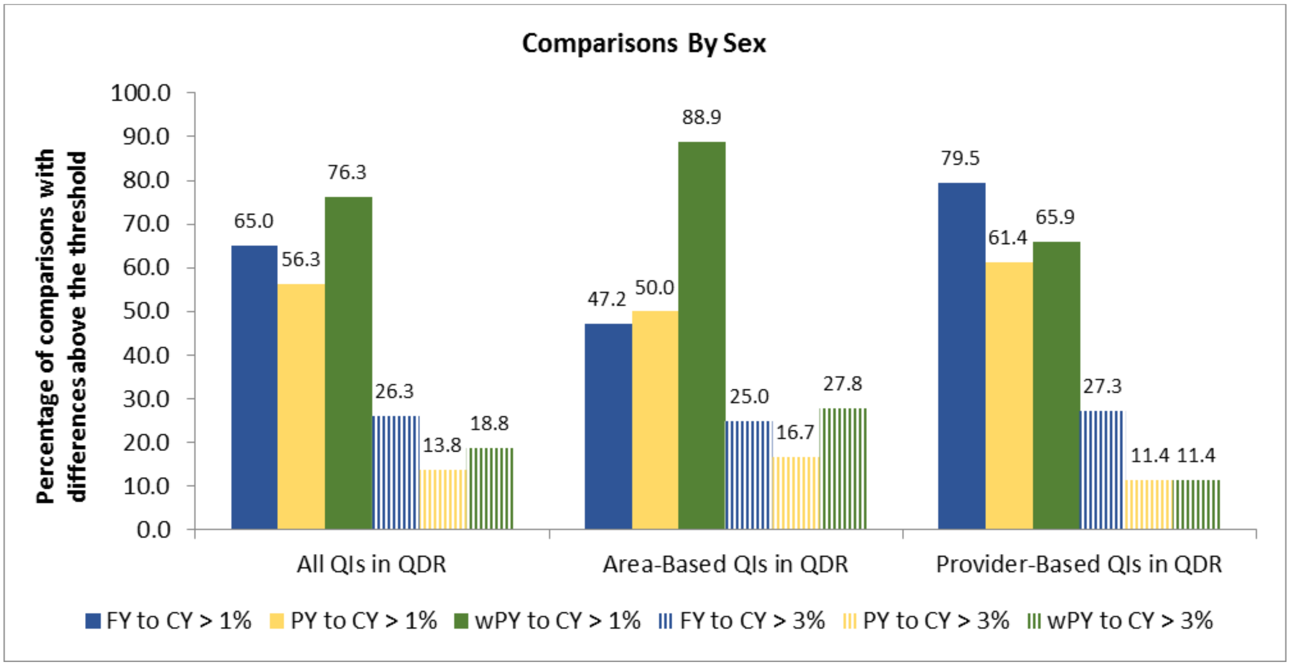
[9] There are four different age categorizations used for the QDR reporting: (1) across all ages, (2) age groups for pediatric measures, (3) age groups for obstetric measures, and (4) age groups for patients aged 65 years and older.

Figure 4 shows the CY 2014 national estimates by sex compared with the three alternative approaches. With 41 QIs and 2 sex categories (male and female), there were a maximum of 82 comparisons possible. Because there are two obstetric QIs specific to females, there were 80 comparisons (36 for the area-based QIs and 44 for the provider-based QIs).

The results showed the following:

- For the area-based QIs, the FY estimates were most similar to the CY estimates at the 1 percent threshold, but the PY estimates were most similar to the CY estimates at the 3 percent threshold.
- For the provider-based QIs, the PY estimates were most similar to the CY estimates.

Figure 4. Comparisons of the CY 2014 National Estimates With the Three Alternative Approaches, by Sex



Abbreviations: CY, calendar year; FY, fiscal year; PY, partial year, QDR, Quality and Disparities Report; QIs, quality indicators; wPY, weighted partial year.

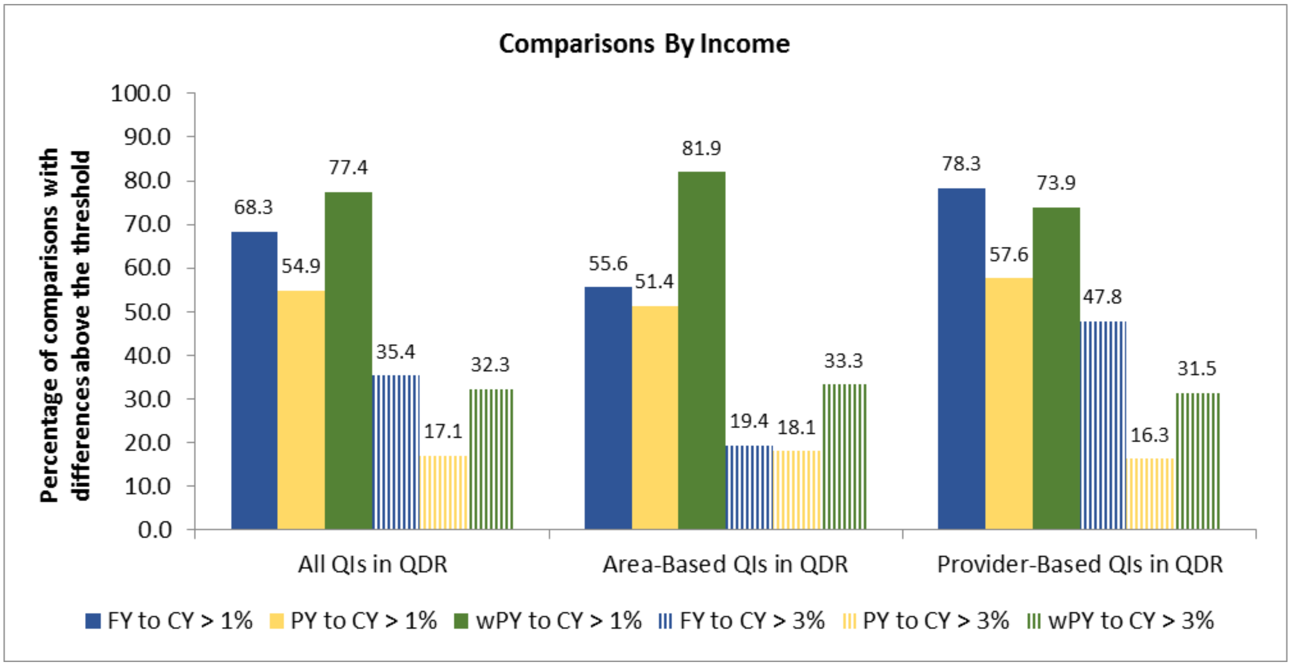
For the FY estimates that were greater than the CY estimates by at least 3 percent, the average percentage difference was 5.0 percent. The similar average was 4.9 percent for the PY to CY comparison and 4.9 percent for the wPY to CY comparison.

Figure 5 shows the CY 2014 national estimates by income quartile compared with the three alternative approaches. With 41 QIs and four income quartiles, there were 164 comparisons (72 for the area-based QIs and 92 for the provider-based QIs).

The results showed the following:

- For the area-based QIs, the PY estimates were most similar to the CY estimates.
- For the provider-based QIs, the PY estimates were most similar to the CY estimates.

Figure 5. Comparisons of the CY 2014 National Estimates With the Three Alternative Approaches, by Income Quartile



Abbreviations: CY, calendar year; FY, fiscal year; PY, partial year, QDR, Quality and Disparities Report; QIs, quality indicators; wPY, weighted partial year.

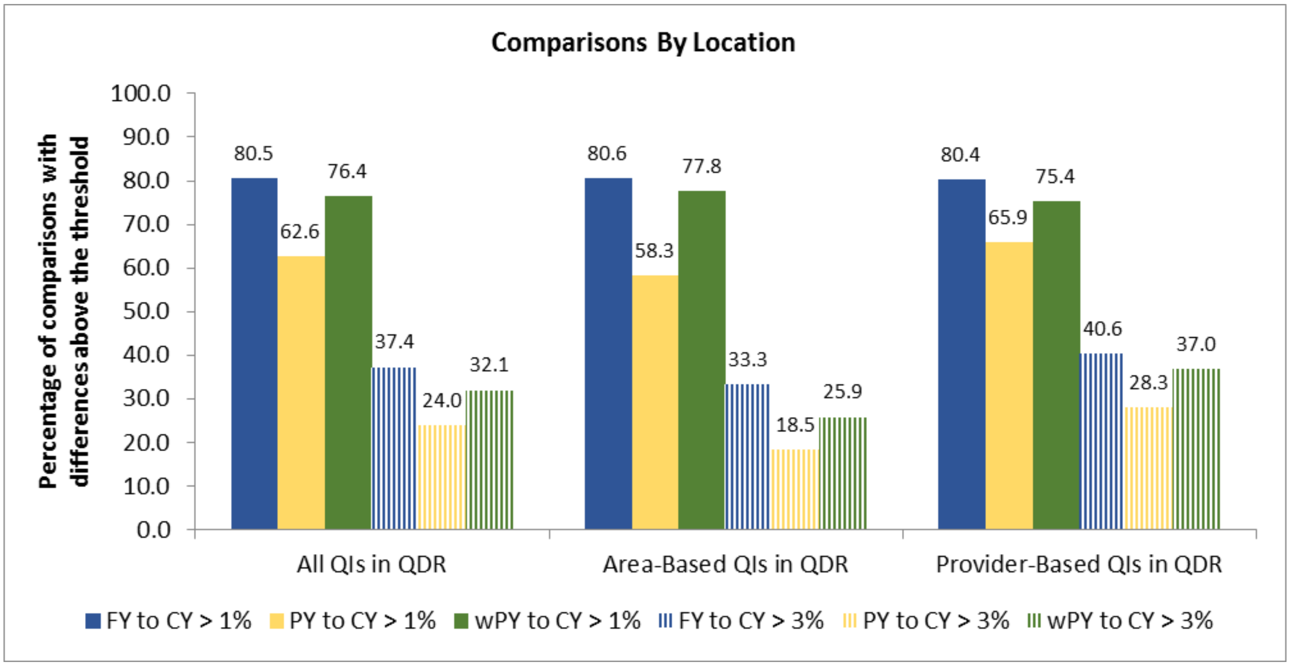
For the FY estimates that were greater than the CY estimates by at least 3 percent, the average percentage difference was 5.6 percent. The similar average was 5.0 percent for the PY to CY comparison and 5.5 percent for the wPY to CY comparison.

Figure 6 shows the CY 2014 national estimates by patient location compared with the three alternative approaches. With 41 QIs and 6 National Center for Health Statistics (NCHS) location categories (e.g., large central metropolitan), there were 246 comparisons (108 for the area-based QIs and 138 for the provider-based QIs).

The results showed the following:

- For the area-based QIs, the PY estimates were most similar to the CY estimates.
- For the provider-based QIs, the PY estimates were most similar to the CY estimates.

Figure 6. Comparisons of the CY 2014 National Estimates With the Three Alternative Approaches, by Patient Location



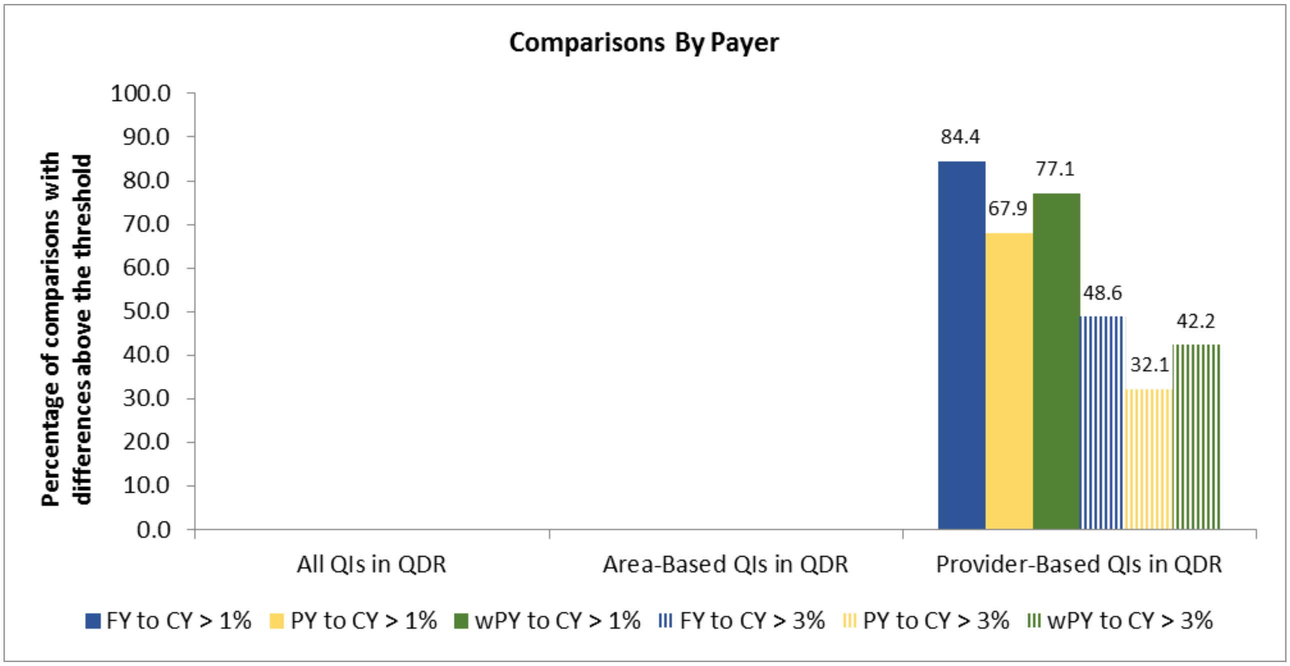
Abbreviations: CY, calendar year; FY, fiscal year; PY, partial year, QDR, Quality and Disparities Report; QIs, quality indicators; wPY, weighted partial year.

For the FY estimates that were greater than the CY estimates by at least 3 percent, the average percentage difference was 5.9 percent. The similar average was 5.6 percent for the PY to CY comparison and 7.0 percent for the wPY to CY comparison.

Figure 7 shows the CY 2014 national estimates by payer compared with the three alternative approaches. With 23 provider-based QIs and 5 payer categories (Medicare, Medicaid, private, uninsured, and other), there were a maximum of 115 comparisons possible. Because of suppressed CY estimates, there were 109 comparisons.

For the provider-based QIs, the results showed that PY estimates were most similar to the CY estimates.

Figure 7. Comparisons of the CY 2014 National Estimates With the Three Alternative Approaches, by Payer



Abbreviations: CY, calendar year; FY, fiscal year; PY, partial year, QDR, Quality and Disparities Report; QIs, quality indicators; wPY, weighted partial year.

Note: Area-based QIs are not reported by expected payer in the QDR (and therefore not included in this comparison) because the denominators are specific to population counts and population insurance coverage information does not correspond directly to HCUP expected payer categories in some States.

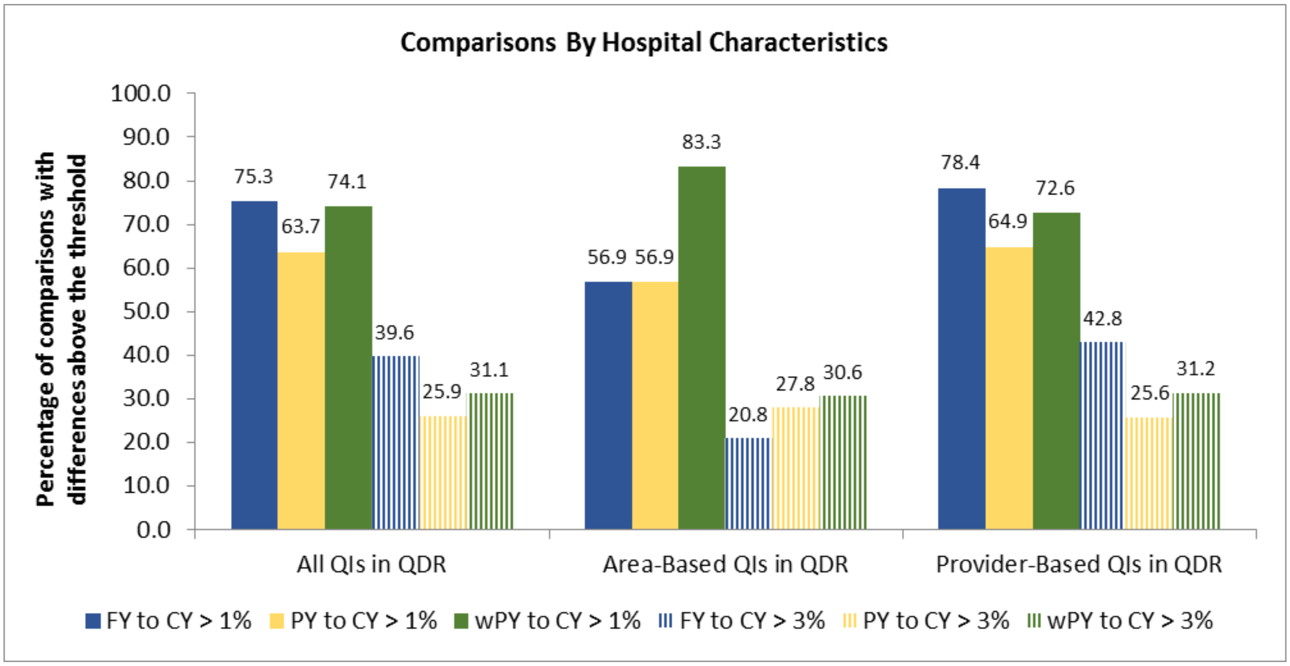
For the FY estimates that were greater than the CY estimates by at least 3 percent, the average percentage difference was 8.2 percent. The similar average was 7.9 percent for the PY to CY comparison and 7.7 percent for the wPY to CY comparison.

Figure 8 shows the CY 2014 national estimates by hospital characteristics compared with the three alternative approaches. With 41 QIs and 5 different hospital categorizations (region, ownership, teaching status, bed size, location), there were 502 comparisons (72 for the area-based QIs and 430 for the provider-based QIs).

The results showed the following:

- For the area-based QIs, the FY estimates were most similar to the CY estimates.
- For the provider-based QIs, the PY estimates were most similar to the CY estimates.

Figure 8. Comparisons of the CY 2014 National Estimates With the Three Alternative Approaches, by Hospital Characteristics



Abbreviations: CY, calendar year; FY, fiscal year; PY, partial year, QDR, Quality and Disparities Report; QIs, quality indicators; wPY, weighted partial year.

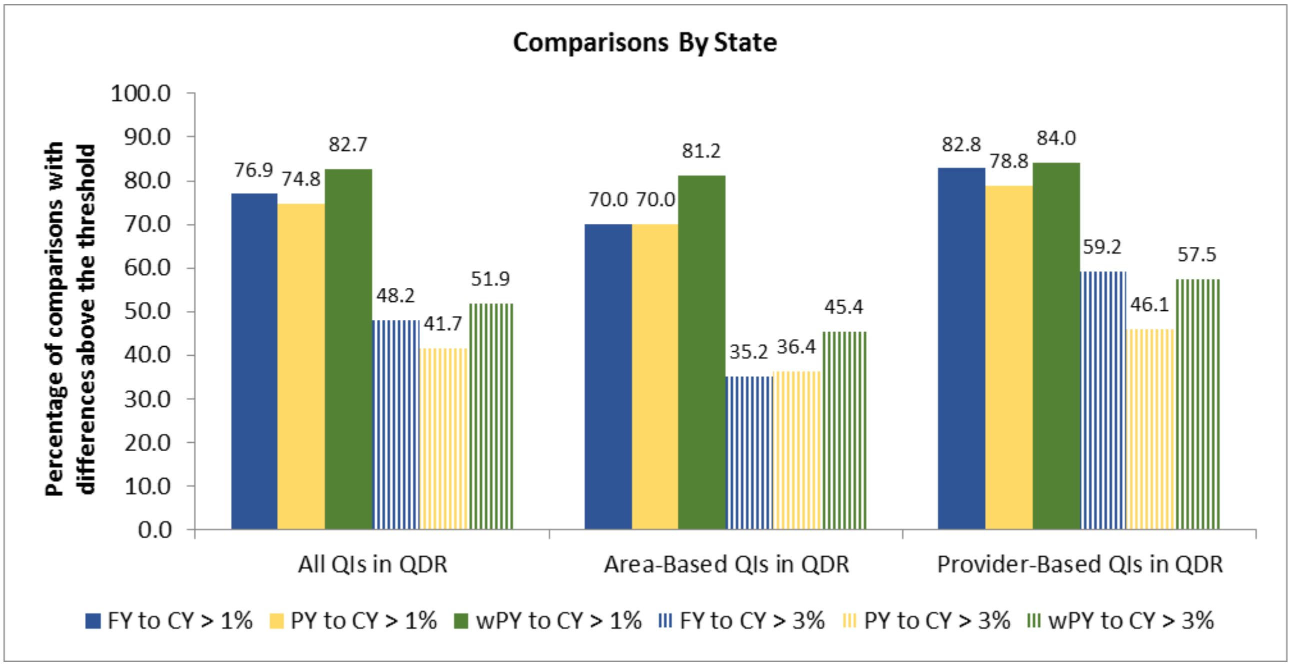
For the FY estimates that were greater than the CY estimates by at least 3 percent, the average percentage difference was 7.2 percent. The similar average was 6.2 percent for the PY to CY comparison and 7.1 percent for the wPY to CY comparison.

Figure 9 shows the CY 2014 State-level estimates compared with the three alternative approaches. With 41 QIs and 39 States used in 2014, there were a maximum of 1,599 comparisons possible. Because of suppressed CY estimates, there were 1,527 comparisons (701 for the area-based QIs and 826 for the provider-based QIs).

The results showed the following:

- For the area-based QIs, the FY estimates were most similar to the CY estimates.
- For the provider-based QIs, the PY estimates were most similar to the CY estimates.

Figure 9. Comparisons of the CY 2014 State-Level Estimates With the Three Alternative Approaches



Abbreviations: CY, calendar year; FY, fiscal year; PY, partial year, QDR, Quality and Disparities Report; QIs, quality indicators; wPY, weighted partial year.

For the FY estimates that were greater than the CY estimates by at least 3 percent, the average percentage difference was 8.0 percent. The similar average was 7.4 percent for the PY to CY comparison and 9.1 percent for the wPY to CY comparison.

Ability to Test for Significant Differences in Trends

The PY estimates for 2015 would be based on data that will not include any of the CY 2014 data. The current approach to testing significant differences between pairs of estimates from independent samples (two-sided χ^2 tests) would continue to be appropriate. Likewise, we can assume that the wPY estimates will be based on an independent sample; specifically, we assume that the numerator and denominator weights are known, fixed constants.

In contrast, estimates for FY 2015 would include the last quarter of data from 2014 (i.e., the FY 2015 data overlap with the CY 2014 data). Our QDR approach to testing differences would not be appropriate. There are statistical tests for dependent samples, but we would need to research the best approach.

RECOMMENDATION FOR THE 2015 QI ESTIMATES FOR THE QDR USING ONLY ICD-9-CM DATA

On the basis of the discordance diagnostics and the ability to test significant differences in trends, we concluded that the best approach to calculating QI estimates for the QDR based solely on ICD-9-CM data for 2015 is to calculate PY estimates using the first three quarters of data from 2015. The QDR will document that the 2015 QI rates are not actual CY rates, but instead they are based on 9 months of data from January 2015 to September 2015.

The PY estimates were less likely than the other two alternative approaches to be more than 3 percent greater than the CY estimates for the 41 QIs in the QDR when comparing national QI estimates by race and ethnicity, age, sex, income quartile, patient location, and hospital characteristics. In addition, the PY estimates were less likely to be more than 3 percent greater than the true 2014 estimates for the 41 QIs in the QDR when comparing State-level QI estimates. Although the PY estimates often were closer to the actual CY estimates (than the FY and wPY estimates), it is clear that the PY estimates still were different. The PY estimates for the national rates were more than 1 percent different than the CY estimates for 56.1 percent of the 41 QIs in the QDR and more than 3 percent different than the CY estimates for 17.1 percent of the 41 QIs. When the PY estimates differed by more than 3 percent, the average difference was 3.7 percent.

It is important to note that this case study was specific to the reporting of national and State-level QIs using ICD-9-CM data. Most users of the QI software derive hospital-level QI rates. It is unknown whether there would be similar results if comparing hospital-level QI rates.

APPENDIX A. QUALITY INDICATORS USED IN THE QDR

Quality Indicator	Description
IQI software	
IQI 11	Deaths per 1,000 admissions with abdominal aortic aneurysm repair
IQI 12	Deaths per 1,000 adult admissions with coronary artery bypass surgery, age 40 years and over
IQI 15	AMI mortality rate (number of deaths per 1,000 discharges for AMI)
IQI 16	Deaths per 1,000 adult admissions with congestive heart failure
IQI 20	Deaths per 1,000 adult admissions with pneumonia
IQI 30	Deaths per 1,000 adult admissions with percutaneous transluminal coronary angioplasty, age 40 years and over
PDI software	
PDI 1	Accidental puncture or laceration during procedure per 1,000 discharges, under age 18 years
PDI 14	Hospital admissions for pediatric asthma per 100,000 population, under age 18 years
PDI 15	Hospital admissions for short-term complications of diabetes per 100,000 population, ages 6–17 years
PDI 90	Overall PDI composite per 100,000 population, ages 6–17 years
PDI 91	Chronic PDI composite (asthma, diabetes) per 100,000 population, ages 6–17 years
PDI 92	Acute PDI composite (gastroenteritis, urinary tract infections) per 100,000 population, ages 6–17 years
PQI software	
PQI 01	Hospital admissions for short-term complications of diabetes per 100,000 population, age 18 years and over
PQI 02	Perforated appendices per 1,000 admissions with appendicitis
PQI 03	Hospital admissions for long-term complications of diabetes per 100,000 population
PQI 05	Avoidable admissions for chronic obstructive pulmonary disease per 100,000 population, age 18 years and over
PQI 07	Avoidable admissions for hypertension per 100,000 population, age 18 years and over
PQI 08	Hospital admissions for congestive heart failure per 100,000 population
PQI 11	Avoidable admissions for bacterial pneumonia per 100,000 population
PQI 13	Avoidable admissions for angina per 100,000 population, age 18 years and over
PQI 14	Hospital admissions for uncontrolled diabetes per 100,000 population
PQI 15	Hospital admissions for asthma per 100,000 population, age 18 years and over
--	Hospital admissions for immunization-preventable influenza per 100,000 population, age 65 years and over (measure added specifically for the QDR)
PQI 90	Overall PQI composite per 100,000 population, age 18 years and over
PQI 91	Acute PQI composite per 100,000 population, age 18 years and over
PQI 92	Chronic PQI composite per 100,000 population, age 18 years and over
PSI software	
PSI 02	Deaths per 1,000 admissions in low-mortality diagnosis-related groups
PSI 04	Deaths per 1,000 discharges with complications potentially resulting from care (failure to rescue)
PSI 06	Iatrogenic pneumothorax per 1,000 discharges

Quality Indicator	Description
PSI 07	Central venous catheter bloodstream infections per 1,000 discharges
PSI 08	Postoperative hip fractures per 1,000 surgical discharges, age 18 years and over
PSI 09	Postoperative hemorrhage or hematoma with surgical drainage or evacuation per 1,000 surgical discharges
PSI 10	Postoperative physiologic and metabolic derangements per 1,000 elective surgeries
PSI 11	Postoperative respiratory failure per 1,000 elective surgical discharges
PSI 12	Postoperative pulmonary embolus or deep vein thrombosis per 1,000 surgical discharges
PSI 13	Postoperative septicemia per 1,000 elective surgical discharges of 4 or more days
PSI 14	Postoperative abdominal wound dehiscence per 1,000 discharges
PSI 15	Accidental laceration or puncture during procedure per 1,000 discharges
PSI 17	Birth trauma injury to neonate per 1,000 selected live births
PSI 18	Obstetric trauma per 1,000 instrument-assisted deliveries
PSI 19	Obstetric trauma per 1,000 vaginal deliveries without instrument assistance

Abbreviations: AMI, acute myocardial infarction; IQI, Inpatient Quality Indicator; PDI, Pediatric Quality Indicator; PQI, Prevention Quality Indicator; PSI, Patient Safety Indicator; QDR, Quality and Disparities Report.

Note: Dashes indicate there is no specific QI number for this measure. The measure was added specifically for the QDR.