Physician Compare Quality Measurement Technical Expert Panel (TEP) Summary Report

July 2019

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1 INTRODUCTION

As CMS's primary platform for patients and caregivers to review Medicare-accepting clinicians, the Physician Compare website publishes information about eligible groups and clinicians that is carefully selected by CMS on a systematic iterative basis. In 2019, performance data from the first year of the Quality Payment Program (QPP) (Performance Year (PY) 2017) will be published on Physician Compare, which will begin incorporating the requirements outlined in the Medicare Access and CHIP Reauthorization Act of 2015 (MACRA) and the CY 2017 QPP Final Rule (81 FR 77390 through 77399). The CY 2017 QPP Final Rule establishes that, in addition to other types of QPP indicators and affiliations, all performance measures submitted through the Merit-Based Incentive Payment System (MIPS) are available for public reporting on Physician Compare (81 FR 77390 through 81 FR 77399). MACRA builds upon the Physician Compare requirements outlined in section 10331 of the Affordable Care Act. To meet the rigorous public reporting standards established in section 10331 (a)(1) of the Affordable Care Act and through rulemaking, all QPP performance data published to Physician Compare must be accurate, valid, reliable, and comparable across available collection types, and any information included on Physician Compare profile pages must also resonate with patients and caregivers, as shown through user testing. CMS has contracted the Physician Compare support team to ensure the PY 2017 performance information published to Physician Compare aligns with statutory and regulatory obligations. As part of that process, the team convened two meetings in 2018 with the Physician Compare Technical Expert Panel (TEP) to obtain feedback on the intended approach to public reporting for PY 2017.

The remainder of this report summarizes the discussions and conclusions from both the May and August 2018 TEP meetings. Section 2 introduces the Physician Compare TEP. Section 3 reviews the issues CMS has considered to promote a seamless transition for public reporting on clinician performance as the QPP begins. Section 4 describes topics addressed during the May TEP meeting. Section 5 discusses topics covered during the August TEP meeting. Section 6 outlines the final PY 2017 public reporting plan.

2   ABOUT THE TEP

The Physician Compare support team consults with the Physician Compare Technical Expert Panel (TEP) for guidance on how to choose and display performance information on Physician Compare in a way that accurately and robustly reflects clinical performance and supports actionable comparisons. The TEP consists of clinicians, purchasers, and other experts with a broad range of experience in publicly reporting performance measures, improving health care quality, and developing and testing quality measures (Table 1). The Physician Compare support team convened the TEP on May 30, 2018 to discuss the data analysis plan and concept testing for publicly reporting PY 2017 MIPS measures. After receipt of the PY 2017 performance data, the Physician Compare support team and TEP members reconvened on August 22, 2018 to discuss preliminary results and review considerations surrounding policy decisions. Table 1 lists TEP participants and which TEP meeting they attended.

Table 1. TEP Members

TEP Member	Position(s),Organization	Location	TEP Attended
A.J. Yates, MD	Associate Professor, Department of Orthopedic Surgery/University of Pittsburgh School of Medicine	Pittsburgh, PA	May & August
Dale Shaller, MPA (TEP Chair)	Principal, Shaller Consulting Group	Stillwater, MN	May & August
Eric Holmboe, MD	Internist, Senior Vice President, Milestones Development and Evaluation of the Accreditation Council for Graduate Medical Education (ACGME)	Philadelphia, PA	-
Gregory Dehmer, MD	Professor of Medicine at the Texas A&M University College of Medicine and Director of the Cardiology Division at the Scott & White Clinic	Temple, TX	May & August
Jeffrey P. Jacobs, MD	Director of ECMO Program at All Children's Hospital, Professor of Cardiac Surgery (PAR) in the Division of Cardiac Surgery of the Department of Surgery at Johns Hopkins University, Surgical Director of the Heart Transplantation Program at All Children's Hospital, and Clinical Professor in the Division of Thoracic/Cardiovascular Surgery at University of South Florida College of Medicine.	St. Petersburg, FL	August
Michael Mihlbauer, MS	Practice Administrator, Anesthesiology Associates of Wisconsin	Milwaukee, WI	-
Robert Krughoff, JD	Founder and President, Center for the Study of Services/Consumers' Checkbook	Washington, DC	May
Sara Scholle, DrPH	Assistant Vice President, Research & Analysis/National Committee for Quality Assurance	Washington, DC	August

TEP Member	Position(s),Organization	Location	TEP Attended
Sherrie Kaplan, PhD, MSPH, MPH	Professor of Medicine and Assistant Vice Chancellor, Healthcare Evaluation and Measurement Executive Co-Director, Health Policy Research Institute School of Medicine/ University of California, Irvine	Irvine, CA	-
Ted von Glahn, MS	Consultant	San Francisco, CA	May & August
Thomas Smith, MD, MS	Medical Director, Division of Managed Care, NYS Office of Mental Health/New York State Psychiatric Institute	New York, NY	-

3 QUALITY PAYMENT PROGRAM BACKGROUND

As part of CMS's continued phased approach to public reporting on Physician Compare, CMS is publicly reporting a subset of the 2017 Quality Payment Program (Year 1) information submitted under the Merit-based Incentive Payment System (MIPS) and Alternative Payment Models (APMs). Information on the Physician Compare website is publicly reported on clinician and group profile pages and in the Physician Compare Downloadable Database on data.medicare.gov. Similar to how data were reported under the legacy programs, all data must meet the established public reporting standards to be publicly reported on Physician Compare (§414.1395(b)).

The remainder of this section discusses considerations pertaining to the publication of performance information on Physician Compare as CMS transitions to the QPP. Section 3.1 summarizes the PY 2016 quality measures released on Physician Compare in late 2017, including the launch of Physician Compare star ratings. Section 3.2 describes the Quality Payment Program and the mandates establishing which QPP performance information must be published on Physician Compare. Section 3.3 outlines the PY 2017 public reporting plan implemented for Physician Compare.

3.1 PY 2016 Performance Data Release

In late 2017, CMS initiated a phased-in approach to star ratings on Physician Compare by publishing star rated performance scores for 13 PY 2016 PQRS measures on group profile pages. Per the 2016 PFS Final Rule (80 FR 71116 through 71135), the Physician Compare support team calculated the ABC^{TM} benchmark for the 5-star cut-off and determined the 1- to 4-star cut-offs using the Equal Ranges method.[1] PY 2016 Qualified Clinical Data Registry (QCDR) measures were published to profile pages as percentages, and Consumer Assessment of Healthcare Providers and Systems (CAHPS) for Physician Quality Reporting System (PQRS) summary survey measures were published to profile pages as top-box scores. Performance scores for PQRS measures that did not receive a star rating but did meet the public reporting standards were included in the Physician Compare downloadable database. Table 2 outlines the number of measures and display type published on Physician Compare in late 2017, by measure and reporting entity.

Table 2. Number of Measures Reported on Physician Compare for PY 2016

Entity Type	Measure Type	Publication Location	Number of Measures	Display
Group	PQRS	Live-Site Profile Pages	13	Star Rating
Group	PQRS	Downloadable	56	Performance Rate Percentage and Star Rating
Group	QCDR	Live-Site Profile Pages	1	Performance Rate Percentage
Group	QCDR	Downloadable	1	Performance Rate Percentage

---

[1] [Physician Compare Benchmark and Star Ratings Fact Sheet](#)

Entity Type	Measure Type	Publication Location	Number of Measures	Display
Group	CAHPS for PQRS	Live-Site Profile Pages	8	Top-Box Percentage Score
Group	CAHPS for PQRS	Downloadable	8	Top-Box Percentage Score
Individual	PQRS	Live-Site Profile Pages	0	N/A
Individual	PQRS	Downloadable	55	Performance Rate Percentage
Individual	QCDR	Live-Site Profile Pages	34	Performance Rate Percentage
Individual	QCDR	Downloadable	39	Performance Rate Percentage

Appendix A details the measures published in late 2017 on public-facing profile pages, including information about the star rating cut-offs and percentage of groups assigned to each star rating category.

3.2 The Quality Payment Program

The Medicare Access and CHIP Reauthorization Act of 2015 (MACRA), an amendment to Title XVII of the Social Security Act, established the Quality Payment Program (QPP) (81 FR 77009). Aiming to reward high value, high quality Medicare clinicians with payment increases, the Merit-based Incentive Payment System (MIPS) consolidates three legacy programs: the Physician Quality Reporting System (PQRS), the EHR Incentive program, and Value-Based Modifier, as specified under Section 101(b) in the MACRA. First, the PQRS was the legacy pay-for-reporting quality reporting program that aimed to incentivize provision of quality care to Medicare fee-for-service beneficiaries. The PQRS program has been replaced by the MIPS Quality performance category. Second, the pay-for-reporting EHR Incentive program aimed to incentivize clinicians to utilize EHR technology when providing care. In order to receive incentive payments, clinicians had to demonstrate that they were meaningfully using her technology. This legacy program was replaced by the MIPS Advancing Care Information[2] category. Finally, the Value-Based Payment Modifier (VM) Program, which has been replaced by the MIPS Cost category, adjusted Medicare payments based on the quality of care provided and the cost for items and services under the Medicare Physician Fee Schedule (PFS).

Ultimately, the way that clinicians and groups interact with these updated programs depends on what track of the QPP they participate in. Per MACRA, clinicians can report their QPP data through either the (1) Merit-based Incentive Payment System (MIPS) or (2) an Advanced Alternative Payment Model (APM) (81 FR 77009). The remainder of this section discusses how performance data are submitted and used through these two options.

3.2.1 The Merit-Based Incentive Payment System

Clinicians participating in the MIPS must qualify as a MIPS eligible clinician (81 FR 77040 through 77041). MIPS performance is evaluated through four performance categories: Quality, Improvement Activities (IA), Advancing Care Information (ACI),[2] and Cost.

---

[2] Referred to as Promoting Interoperability in Performance Year 2018 onward.

Under the Quality Payment Program, certain performance categories will be analogous to legacy programs. The MIPS Quality category replaces the legacy PQRS quality measures and will be represented on Physician Compare as MIPS quality measures, CAHPS for MIPS measures, and QCDR quality measures. MIPS quality measures will continue to be displayed as star ratings and percentages on live-site profile pages and in the downloadable database. The MIPS ACI (Promoting Interoperability) performance category replaces the EHR Incentive Program and attestations, and some measure performance rates in this category will be published in the downloadable database. Successful participation in the MIPS ACI performance category will be displayed through a checkmark on the live-site profile pages. The MIPS Quality, ACI, and Improvement Activities performance category scores will be published in the downloadable database but not the live-site profile pages. Physician Compare will not publicly report 2017 cost data because they do not meet the Physician Compare established public reporting standards.

The MIPS final score is the sum of the products of each performance category score and each performance category's assigned weight multiplied by 100 (81 FR 77319). For PY 2017, the following weights apply: Quality (60%), Cost (0%), IA (15%), and ACI (25%) (81 FR 77322). These performance categories may be redistributed depending on the measures available to a particular group or individual clinician or extenuating circumstances that occur during the reporting period (81 FR 77325 through 77329).

3.2.2 Alternative Payment Models

Eligible clinicians can participate in the QPP through Alternative Payment Models (APMs), which can be classified as APMs, Advanced APMs, or MIPS APMs. MIPS APMs do not meet the statutory definition of an Advanced APM are not eligible for the APM Incentive Payment, and will earn a minimum score of 50% in the MIPS IA performance category (81 FR 77246l; 81 FR 77253 through 77255). MIPS APMs are also subject to the APM Scoring Standard under MIPS (81 FR 77246 through 77249). Clinicians participating in an Advanced APM bear additional financial risk for their patient population, specific clinical condition, or care episode in order to receive an APM incentive payment. Advanced APM participants are excluded from MIPS participation and payment adjustments if they meet the threshold for payments or patients to become a Qualifying APM Participant (QP) or Partial Qualifying APM Participant (Partial QP) (81 FR 77399 through 77400). Partial QPs will not be eligible for APM incentives but remain exempt from MIPS, unless they voluntarily report to the MIPS program to receive payment adjustments (81 FR 77433). In order to qualify as an Advanced APM, participants must (1) use certified electronic health record technology (CEHRT), (2) provide payment for covered professional services based on quality measures comparable to MIPS quality measures, and (3) either be a Medical Home Model expanded under CMS Innovation Center authority or bear a significant financial risk (81 FR 77406). Notably, the quality performance data collected under non-ACO Advanced APMs is exempt from public reporting on Physician Compare under PY 2017.

MIPS APM participants include MIPS eligible clinicians who are eligible for the APM scoring standard under MIPS. MIPS APMs must participate in the APM under an agreement

with CMS, include one or more MIPS eligible clinicians on a Participation List, and base payments on performance, cost/utilization, and quality measures (81 FR 77249). MIPS APM participants scored under the APM scoring standard will receive a MIPS final score and associated MIPS payment adjustment based on the APM entity's combined performance. Per Section 1848(q)(9)(A)(i)(I) that all MIPS final and performance category scores are publicly reported,[3] MIPS scores achieved by MIPS APM participants will be made available on Physician Compare.

3.3   Public Reporting Under the Quality Payment Program

Physician Compare will publicly report the PY 2017 Quality Payment Program data in 2019. Section 3.2.1 outlines new provisions for public reporting, as stipulated in the QPP Year 1 Final Rule. Section 3.2.2 outlines the data analysis plan for assessing that measures meet public reporting standards.

3.3.1 New Provisions Under the Quality Payment Program

Sections 1848(q)(9)(A) and (D) of the Social Security Act require Physician Compare to publicly report (i) MIPS eligible clinicians' final scores, (ii) MIPS eligible clinicians' performance scores for each MIPS category (i.e., quality, cost, improvement activities, and advancing care information), (iii) APM affiliations, and, to the extent feasible (iv) the names and performance of APMs.

To guide the public reporting of QPP data on Physician Compare, the CY 2017 QPP Final Rule (81 FR 77390 through 77399) outlines the approach for public reporting MIPS and APM data on Physician Compare, as well as any other information required by MACRA. This includes all measures and activities reported under MIPS via all available collection types, and applies to both MIPS eligible clinicians and groups; it does not, however, apply to voluntary reporters for PY 2017. CMS has the flexibility to stipulate, via regulation, what data collected under the QPP are published to Physician Compare, where information is published (i.e., live-site profile pages or downloadable database), and how the information is displayed to the public.

The CY 2017 QPP Final Rule stipulates that new objectives, measures, and activities will not be publicly reported that have been in use for less than 1 year (81 FR 77395; 81 FR 77396; 81 FR 81 FR 77397, respectively). Given this, the MIPS Improvement Activities performance category, which is a new performance category for PY 2017, will not be published on the Physician Compare website for PY 2017. Physician Compare will not publicly report 2017 cost data because they do not meet the Physician Compare established public reporting standards. According to the QPP Year 1 Final Rule, clinician- and group-level performance reported under the MIPS Quality and MIPS ACI performance categories remain available for publication on Physician Compare.

---

[3] The MIPS Cost performance category was not used for scoring under the first year of the QPP and therefore, was not recommended for public reporting on Physician Compare.

As specified in section 101(c)(q)(9)(A)(ii) of MACRA and finalized in the CY 2017 QPP Final Rule, CMS is required to publicly report an indicator on clinician or group profile pages if participating in an APM, as well as a link to that clinician's or group's dedicated APM website, as technically feasible (81 FR 77398). Clinician- and group-level affiliations to APMs (including Advanced APMs) will be made available on Physician Compare, as technically feasible, when the PY 2017 data is released in July 2019 (see Appendix). Further, quality performance data submitted by Shared Savings Program Accountable Care Organizations (ACOs) and NextGen ACOs will be available for publication on Physician Compare ACO profile pages.[4] The Physician Compare team helps to inform public reporting decisions by conducting in-depth analyses on measure data to identify measures that meet public reporting standards. These analyses are provided to CMS in order to facilitate selection of measure data published on the Physician Compare website. Section 3.3.2 outlines the measure analysis approach that helps to inform CMS decision making.

3.3.2 Data Analysis Plan

The Physician Compare support team conducts a standard set of analyses to identify performance measures that comply with mandated public reporting standards and are available for CMS to publish to Physician Compare. These standards specify that published data must be reliable, valid, comparable, and accurate across available data collection types. Additionally, testing is conducted to ensure that patients and caregivers can use the information published for their decision-making.

ABC™ Benchmarking and Star Rating

A subset of group MIPS quality measures will be eligible for display as star ratings on live-site profile pages. The 2016 PFS Final Rule (80 FR 71116 through 71135) established that Physician Compare use the ABC™ benchmark methodology to calculate a 5-star rating benchmark for individual measures. Through discussions at the 2017 TEP meeting,[5] it was determined that the equal ranges method would be used to produce meaningful and reliable 1- to 4-star rating assignments. A more detailed overview of CMS's star rating and benchmarking methodology can be found on the Physician Compare Initiative Page.[6]

Reliability Tests

Measure reliability refers to the extent to which differences in performance rates for each quality measure are due to actual differences in performance versus variation that arises from measurement error. In order to determine that measure reliability standards are met, reliability is assessed using two methods: (1) the beta binomial test and (2) a split half reliability test. If a measure[7] passes both of these reliability tests, the performance scores for quality measures are

---

[4] Data for the Shared Savings Program and NextGen ACOs have historically been made publicly available on Physician Compare under the PQRS legacy program.
[5] Physician Compare TEP Summary Report, December 2017
[6] Physician Compare Benchmark and Star Ratings Fact Sheet
[7] Evaluated at the measure-, measure stratum-, reporting entity-, collection type-level.

considered meaningfully different across reporting entities, rather than due to measurement error. Further, in order to be published as a star rating, the ABC^{TM} benchmark, star rating cut-offs, and star rating assignments must pass further reliability testing.

Reliability of Star Rating Cut-offs

To assess the reliability of the benchmark and star rating cut-offs, the Physician Compare support team tested how stably each method calculates star rating cut-offs when presented with changes in the performance rate distribution that could be expected due to chance, given the sample size (e.g., the number of reporters) and amount of variation in performance across reporting entities. To investigate each method's capacity to produce reliable cut-offs, a bootstrapping analysis was conducted where reporters' performance scores were randomly sampled with replacement until the sample size was equal to the number of reporters for that measure. This process was repeated 500 times for each measure/mechanism combination. For each simulated data set, the star rating cut-offs were recalculated and the simulated cut-offs were used to reassign each reporter to a simulated star rating category, using their original performance score.

Star Rating Assignment Reliability

In addition to producing robust cut-offs, an ideal star rating method should categorize reporters meaningfully, given the precision of the observed performance rates. To ensure that star assignments were not influenced heavily by random error, each reporter's actual performance rate and patient population sizes were used to simulate counterfactual performance rates using a binomial distribution. The ABC^{TM} benchmark and the star rating cut-offs were recalculated for each simulation, and reporters were assigned a simulated star rating category based on their simulated performance rate. After running 500 simulations for a measure/mechanism combination, the frequency of reporters receiving the same and different star ratings across simulations was determined.

Reliability Thresholds

The following thresholds are used to assess reliability: 80% accuracy of assignments across simulations; multi-star shift of less than five percent; and a Fleiss' Kappa of less than or equal to 0.6. Positive results on these tests imply that reporters assigned to different star categories are meaningfully different and can be compared to each other.

Validity

Validity refers to the degree to which a metric measures what it purports to measure. Validity for each quality measure is assessed by evaluating the extent to which observed performance rates on measures are impacted by factors unrelated to true performance, such as characteristics of the reporter's patient population (i.e., case-mix) or selective reporting of the patient population. The validity of measure data is addressed in three ways: (1) evaluation of outcome measure risk adjustment strategies, (2) investigation into the possibility of selective reporting, and (3) review of specifications to ensure they align with clinical best practice

guidelines. If outcome measures are not risk adjusted, measure data could be influenced by confounding factors, such as patient population characteristics. This would impact the observed performance rates, which would therefore not accurately reflect the true quality of clinical care provided. Additionally, the Physician Compare team analyzes reporting rates and performances rates for each measure using a Pearson correlation to assess if there is evidence suggesting that reporters could be selectively reporting data in order to inflate performance scores. Finally, as clinical guidelines and best practices change over time, our team ensures that only measures that align with current best practice guidelines are selected for public reporting.

Comparability

Comparing the performance score distributions from data collected via different collection types (e.g., CMS Web Interface vs. Qualified Registry) for the same measure shows that raw performance data submitted through distinct pathways are not always comparable. Thus, measure analyses and the establishment of benchmarks and star ratings are performed for each measure/collection type[8] combination, rather than aggregating data from different mechanisms. Further, the use of benchmarks helps to ensure performance, relative to what is achievable for a given collection type, can be validly compared across collection types. While measure data could be reported for a variety of collection types if benchmarks and star ratings are used, CMS opted to select one collection type for a given measure to maintain continuity through the roll out of the QPP. When data collected via multiple collection types for a single measure meet all of the Physician Compare public reporting standards, the collection type that represents the highest number of distinct reporters is selected for public reporting. To ease the transition to the QPP, CMS opted to prioritize data collected via CMS Web Interface when technically feasible, as this prioritization method was implemented in in the final public reporting year of the PQRS legacy program.

Accuracy

Measure accuracy refers to the degree to which a measure correctly assesses what it purports to measure. To allow for ample time for measure testing and validity assessments of newly introduced measures, we have consistently implemented a policy that we do not publicly report measures during the first year of their inclusion in a CMS quality reporting program. After a measure's first year in use, we evaluate the measure again to see if and when the measure is suitable for public reporting. As such, in the CY2017 QPP Final Rule, we finalized to not publicly report first year objectives, measures, and activities will not be publicly reported that have been in use for less than 1 year (81 FR 77395; 81 FR 77396; 81 FR 81 FR 77397, respectively). Consequently, only measures that have been carried over to the QPP from the legacy PQRS or Meaningful Use programs are eligible for public reporting in PY 2017. Additionally, measures with specifications that are found to misalign with best clinical practice

---

[8] Evaluated at the measure-, measure stratum-, reporting entity-, collection type-level.

are excluded from public reporting, as data submissions for these measures are less likely to accurately reflect true clinical performance.

User Testing

In order to meet public reporting standards, data published on Physician Compare must resonate with users as determined through testing (81 FR 77395; 81 FR 77396; 81 FR 81 FR 77397, respectively). The Physician Compare team tests performance information available for public reporting using plain language to facilitate user understanding. Quality measures including MIPS, QCDR, and CAHPS for MIPS measures are tested using plain language titles and descriptions. Additionally, other PY 2017 QPP concepts for potential inclusion on Physician Compare group and/or clinician profile pages such as star rating and benchmark language, Medicare Assignment status, APM affiliations, ACO affiliations, and the ACI successful reporter indicator are tested along with website design layouts for their inclusion.

3.3.3 Feedback on Measure Recommendations

Results from the above analyses were used to curate a list of measures recommended for public reporting for PY 2017. This list, along with other considerations in regards to how and where to publish measure data, are presented to the TEP in order to garner expert feedback that CMS may utilize when rendering final measure publication decisions. The following sections of this report detail the discussions and TEP recommendations made in regards to measure publication on Physician Compare.

4 MAY TEP MEETING

TEP members and the Physician Compare team convened on May 30, 2018 to discuss the current state of PY 2017 implementation and pending sub-regulatory policy decisions. Section 4.1 discusses current concept testing findings for the Physician Compare live-site. Section 4.2 provides an overview of the clinician engagement approach for PY 2017. Section 4.3 outlines outstanding policy decisions, relevant TEP member commentary, and the Physician Compare team's ultimate public reporting recommendations.

4.1 Concept Testing Findings

As specified in the CY 2017 QPP Final Rule, CMS has indicated that, in addition to meeting all public reporting standards, measures selected for publication on Physician Compare profile pages must also resonate with Medicare beneficiaries and their caregivers (81 FR 77395; 81 FR 77396; 81 FR 81 FR 77397, respectively). Information that is not determined to be meaningful to beneficiaries and their caregivers through concept testing are published exclusively to the Physician Compare downloadable database.

In August, 2017 the Physician Compare team conducted concept testing with users on potential QPP language, the four performance categories, Alternative Payment Models (APMs), and live-site profile page layout designs. Section 4.1.1 provides an overview of user concept testing. Section 4.1.2 discusses potential implications of concept testing findings. Section 4.1.3 provides an overview of next steps.

4.1.1 User Concept Testing

Respondents were presented with plain language category attestations, category titles, and category introductory text for Quality, IA, and ACI categories. Users found the information interesting, but expressed a preference for short, concise language over lengthier descriptions. The phrase "Advancing Care Information" was not meaningful to users. Since the time of testing, CMS has changed "Advancing Care Information" to "Promoting Interoperability" beginning with PY 2018. The phrase "Electronic Health Record Technology" resonated with users significantly more than "Advancing Care Information". The phrase "Promoting Interoperability" was tested and determined to not be meaningful to respondents.

Users were able to clearly understand the IA and ACI checkmark plain language attestations. However, users felt that some attestations were more important than others and that several attestations sounded similar. Including the word "secure" in ACI attestations made the information resonate more with users. ACI star ratings were understood by users, however for a few measures, users attributed the star ratings to patient behavior. In general, users were more interested in whether their clinician used an EHR system. The overall checkmark for the ACI category resonated with users, but the distinction between a successful reporter versus a high performer was unclear.

The phrase "Alternative Payment Model" was not clear to users. Many users thought that APMs involved how patients paid clinicians (e.g., credit card, installment plans, etc.).

Subsequent testing emphasized that the "group is reimbursed by Medicare" through an APM. This, in conjunction with hover definitions, aided user understanding of the concept. Linked associations between APM affiliation and model participants were unclear to users. May testing added the phrase "participates through" the model participant, which helped understanding.

Variations in website layout on the Physician Compare live-site profile pages were tested. First, the Physician Compare team tested adding one QPP section to profile pages for all MIPS performance category information. Users felt that they were presented with too much information at once and that it was difficult to locate information of interest to them. Additionally, the "Quality Payment Program" section title was not meaningful to users. Users then evaluated two different layouts. The first layout, Design A, incorporated IA and ACI attestations into the General Information section of the profile and ACI measure performance information into the performance section. The second layout, Design B, added a Practice Improvement section with information on IA and ACI categories. Users preferred the first layout, citing the ease of having checkmark attestations readily accessible in the General Information section and the grouping of checkmarks and performance scores together, respectively.

4.1.2 Implication of User Testing

Based on the findings of May user concept testing, it was concluded that language will be needed to aid user understanding of the terms "Alternative Payment Model" and affiliations of APMs. Additionally, given that the term "Electronic Health Record Technology" resonated well with users, the Physician Compare team suggested using this to label the ACI (Promoting Interoperability) information on Physician Compare profile pages. The Physician Compare team proceeded with profile page development using Design A.

4.1.3 Next Steps

Testing will continue on QPP concepts available for public reporting on Physician Compare. These concepts include the following:

- Quality Information, IA attestation, ACI attestations and rates;
- Revised plain language for section headers and introductory text;
- Language around APMs and affiliations; and
- QCDR measure language.

4.2 Policy Discussion Topics

The Physician Compare team presented pressing policy considerations to TEP members for their input. This section provides an overview of the specific policy considerations, relevant background information, TEP member commentary, and the Physician Compare team's ultimate public reporting recommendations.

4.2.1 Limiting Public Reporting to Six Measures

A subset of stakeholders requested that CMS consider aligning public reporting with the MIPS payment program by limiting to the six quality measures that count towards the MIPS quality component score. Currently, clinicians are scored using the six measures that result in the highest number of achievement points under MIPS scoring criteria, and one measure must be an outcome or high-priority measure. Clinicians and groups can select from a list of hundreds of MIPS quality measures and, as a result of the flexibility in measure options for scoring purposes, it is possible that reporters select and are scored based on measures that ultimately are not selected for public reporting on Physician Compare (e.g., due to low overall reporting volume). Accordingly, subsetting exclusively to measures that meet all public reporting requirements and were used for scoring purposes could restrict the data made available to Medicare beneficiaries and their caregivers looking to make informed healthcare decisions.

TEP members acknowledged that it would be difficult to harmonize the MIPS and Physician Compare scoring standards. For example, clinicians may select measures that are commonly high performing, which are not ideal for public reporting as these metrics do not typically aid in beneficiaries' ability to distinguish between healthcare providers. Additionally, reporters may not consistently report on the same measures, which would prevent Physician Compare users observing changes in performance over time. One TEP member pointed out that given that this is the first year of the program, it is unclear whether the divergence in scoring between the two would result in a problem.

4.2.2 Inclusion of Malpractice Information on Physician Compare

CMS is investigating the possibility of incorporating malpractice information on Physician Compare. Patients and caregivers have consistently shown a high level of interest in adding malpractice information to Physician Compare. However, "malpractice" is a complex concept and properly communicating the nuance would be challenging. Additionally, malpractice data are not currently available to the Physician Compare team. Furthermore, some stakeholder groups would not be supportive of the inclusion of malpractice information on Physician Compare.

TEP members agreed that it would be difficult to accurately convey malpractice information to users. Malpractice payments can sometime be made for reasons that are unrelated to clinician care. Additionally, clinicians who are employed by vertically-tiered hospitals are typically not served with legal issues. Instead, legal teams will sue the medical center and the issue will be settled at the medical center level. This creates an unfair disadvantage for clinicians who work in small medical practices. TEP members advised against adding malpractice information to Physician Compare.

5 AUGUST TEP MEETING

The Physician Compare TEP convened in August 2018 to discuss several options for implementing QPP Year 1 policies and to review preliminary PY 2017 performance data analysis results. Policy implementation discussion topics ranged from the method for handling quality measures that underwent specification changes during the first year of the QPP to the options for public display of data available for the first time on Physician Compare (e.g., MIPS final and performance category scores). Section 5.1 expands upon the TEP members' opinions on these policy discussion topics and Section 5.2 summarizes the preliminary analytic results presented to the TEP, including the TEP members' commentary on these findings.

5.1 Policy Discussion

In August 2018, the Physician Compare team presented implementation options for five specific public reporting concepts that are not otherwise defined in federal law. The remainder of this section details the topics and implementation options, relevant TEP member commentary, and ultimate public reporting recommendations. Under the Social Security Act, CMS is required to report (i) MIPS final scores, (ii) MIPS performance category scores, and (iii) clinician and group affiliations with Alternative Payment Models. Otherwise, CMS has flexibility to stipulate what data collected under the QPP are published to Physician Compare, where the information is published (i.e., live-site profile pages or downloadable database), and how the information is displayed to the public.

5.1.1 APM Clinicians Who Report MIPS Measures

The Physician Compare team consulted with the TEP members to develop a process for handling data submitted for the MIPS program by clinicians participating in an APM under a one TIN and participating in MIPS under a separate affiliated TIN. Under the QPP, eligible clinicians can opt to participate in Advanced Alternative Payment Models (Advanced APMs) or in the Merit-Based Incentive Payment System (MIPS). The CY2017 QPP Final Rule indicates that all performance data collected under MIPS is eligible for public reporting (81 FR 77395; 81 FR 77396; 81 FR 81 FR 77397, respectively), but the Centers for Medicare and Medicaid Innovation (CMMI) advised that only MSSP and Next Generation ACO performance data were available for 2017. Because MIPS APM participation occurs at the individual eligible clinician TIN-NPI level, it is technically possible that a single eligible clinician could participate in an APM under one affiliated TIN and participate in MIPS under another affiliated TIN.

One TEP member suggested that it would be beneficial to publicly report individual clinician data when possible, even if collected under an Advanced APM, because website users find individual clinician quality performance data valuable. Another member voiced the opinion that clinicians participating in APMs might not be appropriately tracking to the requirements for participating under MIPS, and to include MIPS data submitted by clinicians also participating in APMs might provide an inaccurate account of the clinicians' priorities.

5.1.2 Measure Specification Changes

Under the final year of the legacy PQRS program, the Physician Compare team began the phased approach for publicly reporting quality performance as star ratings on live-site profile pages with the intention to publicly report the same subset of group-level quality measures as star ratings in the subsequent year, the first year of the QPP, with the goal of maintaining consistency through the transition to the Quality Payment Program. CMS subsequently released plans for future measure specification changes and retirement in the future years of the QPP. While none of the star-rated measures underwent specification changes between the final year of PQRS and the initiation of the QPP, TEP members' were asked if measures that might be retired or undergo specification changes in the future should be publicly reported as star ratings. Generally, the TEP members did not express strong opinions about handling measures that might be retired or undergo specification changes in future years of the QPP.

5.1.3 MIPS Final and Performance Category Scores: Publication Format

All data, including MIPS final and performance category scores, collected under the QPP is eligible for public reporting on Physician Compare, but it is up to CMS to decide the most appropriate data to report and where to report the data, based on the Physician Compare support team's recommendations. Data that resonate well with Medicare patients and their caregivers are typically publicly reported on public-facing group and individual profile pages, whereas data that are not easily understood by or useful to Medicare patients and their caregivers are published in a downloadable database, utilized most widely by a more technical audience, including researchers. As specified in MACRA under section 101(c)(q)(9)(A), CMS must make available on Physician Compare, either on the downloadable database or on profile pages, final scores and performance category scores. Accordingly, the Physician Compare team surveyed the TEP to garner their opinion on the most appropriate publication location for these data.

The TEP members strongly favored publishing MIPS performance category and final scores in the downloadable database, given this data does not resonate well with Medicare patients and their caregivers.

5.1.4 CAHPS and QCDR Data

Historically, CAHPS for PQRS data and data collected via qualified clinical data registries (QCDRs) has been publicly available on both public facing profile pages and in the Physician Compare downloadable database. Under the QPP, the Physician Compare team anticipates that the data collected via the CAHPS for MIPS survey and QCDRs can be treated in the same manner as they have historically, in that all data that meet the public reporting standards are published to Physician Compare. CAHPS and QCDR data that resonate with website users will be reported on public facing profile pages as performance rate percentages. The TEP members were in favor of maintaining the status quo for publishing CAHPS and QCDR measures.

5.1.5 MIPS Categories Eligible for Public Reporting

The MIPS arm of the QPP requires that clinicians and groups report data for four performance categories: Quality, Cost, Improvement Activities (IA), and Advancing Care Information (ACI). While some of these performance categories replace previous CMS initiatives, including the meaningful use and PQRS legacy programs, other performance categories such as IA are new. In the CY 2017 QPP Final Rule, established that they will not publicly report any measures or activities in their first year of use (81 FR 77396). The Physician Compare team surveyed the TEP on publicly reporting measures that were previously reported under QPP legacy programs and not publicly reporting IA measures, given that they were in their first year of use.

The TEP members agreed that data collected for the first time under the QPP, but had been available under a legacy CMS program, should be eligible for publication on Physician Compare under the first year of the QPP. This includes measures in the MIPS Quality and ACI categories, as those had been collected under the PQRS and Meaningful Use programs, respectively. Further, the TEP members recommended not publicly reporting the new IA category data on Physician Compare under the first year of the QPP, as this data was being collected for the first time.

5.2 Preliminary Data Analyses

The Physician Compare team presents preliminary quality measure data analyses to the TEP members on an annual basis. This provides the TEP an opportunity to weigh in on public reporting recommendations for measures that pass all public reporting standards and are eligible for publication on Physician Compare. Further, the Physician Compare team utilizes this opportunity to present processes for handling measures with unique circumstances, such as those that underwent specification changes throughout the program year. The Physician Compare team presented on numerous measures that were impacted by changes in clinical guidelines (technology updates, changes in treatment protocol, etc.) during the performance period and therefore had validity issues. The TEP members approved of the outlined data analysis approach.

6 FINAL PERFORMANCE YEAR 2017 PUBLIC REPORTING PLAN

Based on discussions from the May and August 2018 TEP meetings, CMS has decided to proceed with the following:

6.1 Measures Recommended for Public Reporting

See Appendix B for a list of star rated measures recommended for public reporting on the Physician Compare live-site. The following information will be eligible for publication on live-site profile pages and in the Physician Compare downloadable database:

- Group- and ACO-level CAHPS for MIPS displayed as a percentage
- Group- and individual-level QCDR measures displayed as a percentage
- Group- and ACO-level MIPS Quality measures displayed as a star rating
- Checkmark indicating successful reporting of the MIPS ACI category

In addition to all items listed above, the following data will be available exclusively in the Physician Compare downloadable database:

- MIPS final score
- Group-level measure benchmarks displayed as a percentage
- Group- and individual-level performance rates displayed as a percentage
- Clinician utilization data
- MIPS ACI attestations
- MIPS ACI measure performance rates displayed as a percentage

New measures will not be published on Physician Compare for PY 2017. Measures that were reported under legacy programs are eligible for public reporting. This includes measures in the Quality and ACI categories. Cost measures will not be publicly reported for PY 2017, given that these data were not used to determine scoring or payment adjustments. Measure retirement and specification changes were assessed on a case-by-case basis and CMS has ultimately recommended suppressing Measures 5, 112, and 118 from public reporting under the first year of the Quality Payment Program due to substantive specifications applied during the program year of interest (i.e., PY 2017).

6.1.1 ACI Category

ACI measures that were available under legacy CMS programs will be eligible for publication on Physician Compare. ACI performance will be available in the downloadable database as measure performance rates (%) and attestations (yes/blank). Reporters who achieved a score greater than zero in the ACI category will receive a successful reporting checkmark on their respective profile pages.

6.1.2 APM Measure Submissions

APM performance information will be publicly reported at the level at which data was collected, given that data submissions meet all public reporting criteria. APM affiliations will be publicly reported on individual clinician profile pages. CMS decided to not publish performance information submitted by TINs used for MIPS APM participation. However, if a clinician also submitted performance information under different TINs that were not used for MIPS APM participation, those submissions are available for public reporting for PY 2017. Qualifying APM Participants (QPs) will not have performance information on Physician Compare. Partial QPs will have MIPS performance information available for public reporting, if submitted to CMS.

6.1.3 MIPS Final and Performance Category Scores

Because the information does not resonate with Medicare beneficiaries and their caregivers, MIPS performance category and final scores will be published in the individual-level and group-level downloadable database. If an individual has multiple MIPS composite scores, CMS will publicly report the highest final score and respective category scores.

6.1.4 ACOs

ACO quality measure and CAHPS for ACO performance rates (%) will be publicly reported on ACO live-site profile pages. ACO participants' MIPS final and performance category scores will be publicly reported in the downloadable database.

APPENDIX A – PY 2016 STAR RATED MEASURES

The following data table presents the total number of reporters, number of reporters by star value, and star cut-off values for each measure that met all PQRS public-facing profile display public reporting requirements for publication in late 2017.

Table A.1 – PY 2016 Star Rated Measure Cut-offs and Number of Reporters

| #	Measure Information Title	# of Groups	Star Rating Cut-offs			
			2	3	4	5
6	Coronary Artery Disease (CAD): Antiplatelet Therapy	309	61%	74%	87%	100%
47	Advance Care Plan	607	25%	50%	75%	100%
48	Urinary Incontinence: Assessment of Presence or Absence of Urinary Incontinence in Women Aged 65 Years and Older	114	25%	50%	75%	100%
51	Chronic Obstructive Pulmonary Disease (COPD): Spirometry Evaluation	42	24%	48%	72%	96%
110	Preventive Care and Screening: Influenza Immunization	247	24%	45%	66%	87%
111	Pneumonia Vaccination Status for Older Adults	246	25%	48%	71%	93%
113	Colorectal Cancer Screening	241	21%	42%	63%	84%
117	Diabetes: Eye Exam - National Quality Strategy	224	21%	40%	60%	80%
128	Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow-Up Plan	251	35%	54%	74%	93%
134	Preventive Care and Screening: Screening for Clinical Depression and Follow-Up Plan	236	23%	47%	70%	93%
226	Preventive Care and Screening: Tobacco Use: Screening and Cessation Intervention	1,551	63%	75%	87%	100%
238	Use of High-Risk Medications in the Elderly	605	60%	74%	87%	100%
318	Falls: Screening for Future Fall Risk	242	25%	50%	74%	99%

APPENDIX B – PY 2017 STAR RATED MEASURES

The following data table presents the total number of reporters, number of reporters by star value, and star cut-off values for each measure that met all MIPS public-facing profile display public reporting requirements for publication in 2019.

Table B.1 – PY 2017 Star Rated Measure Cut-offs and Number of Reporters

#	Measure Information Title	# of Groups	Star Rating Cut-offs			
			2	3	4	5
6	Coronary Artery Disease (CAD): Antiplatelet Therapy	432	51%	68%	84%	100%
47	Advance Care Plan	1,179	25%	50%	75%	100%
48	Urinary Incontinence: Assessment of Presence or Absence of Urinary Incontinence in Women Aged 65 Years and Older	207	25%	50%	75%	100%
51	Chronic Obstructive Pulmonary Disease (COPD): Spirometry Evaluation	77	25%	50%	75%	100%
110	Preventive Care and Screening: Influenza Immunization	172	36%	56%	75%	94%
111	Pneumonia Vaccination Status for Older Adults	171	32%	53%	73%	94%
113	Colorectal Cancer Screening	164	23%	43%	64%	84%
117	Diabetes: Eye Exam - National Quality Strategy	191	79%	86%	93%	100%
128	Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow-Up Plan	469	25%	50%	75%	100%
134	Preventive Care and Screening: Screening for Clinical Depression and Follow-Up Plan	164	23%	46%	68%	91%
226	Preventive Care and Screening: Tobacco Use: Screening and Cessation Intervention	172	75%	83%	92%	100%
318	Falls: Screening for Future Fall Risk	171	34%	56%	77%	99%

APPENDIX C – PY 2017 STAR RATED MEASURE HISTOGRAMS

The following histograms present the performance rate and star rating distributions for each measure that met all public-facing profile display public reporting requirements for publication in 2019.

Figure C.1 – Performance Rate Distribution, Coronary Artery Disease (CAD): Antiplatelet Therapy (MIPS #6 – Group Registry)
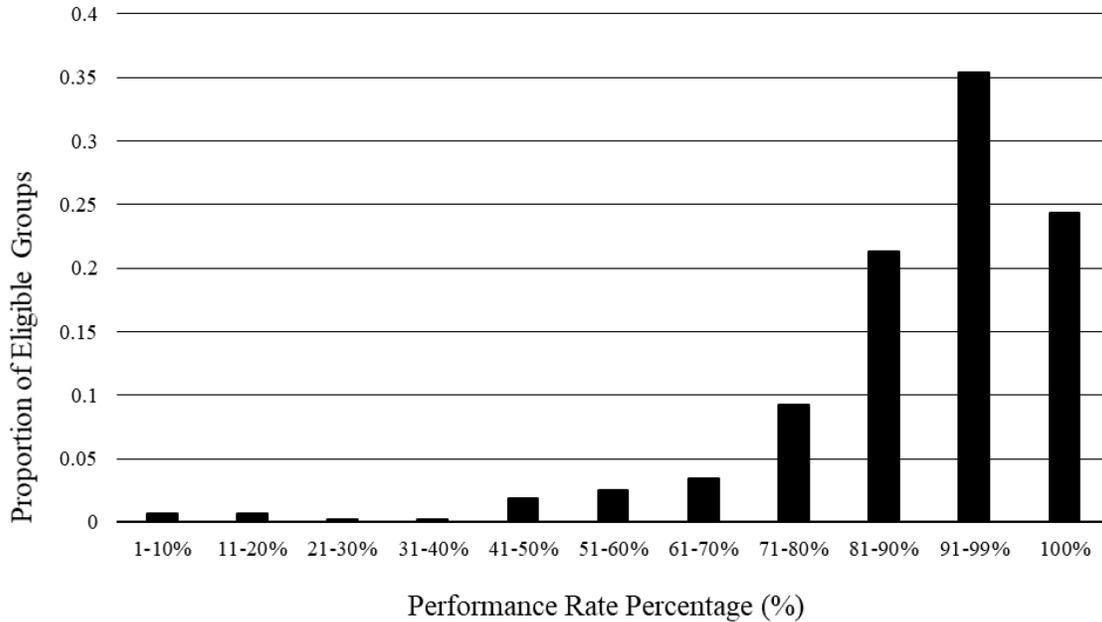
Figure C.2 – Star Rating Distribution, Coronary Artery Disease (CAD): Antiplatelet Therapy (MIPS #6 – Group Registry)
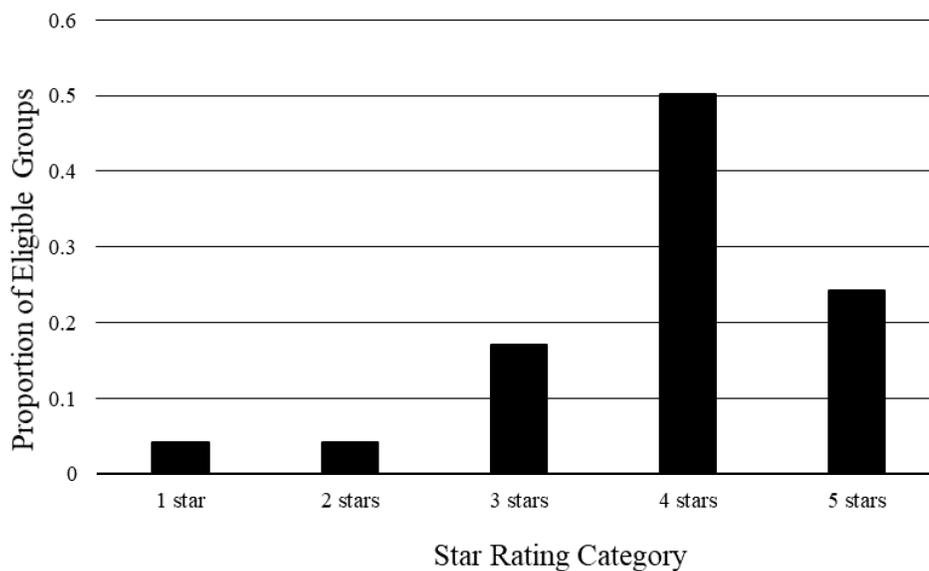
Figure C.3 – Performance Rate Distribution, Advance Care Plan (MIPS #47 – Group Registry)
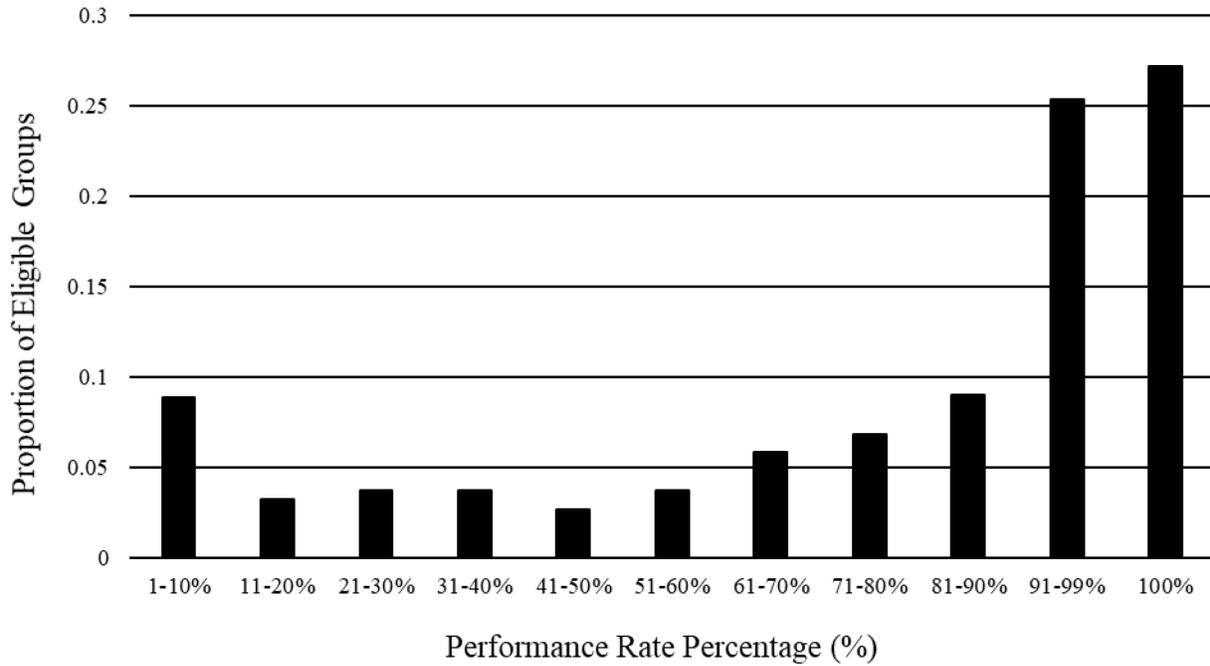
Figure C.4 – Star Rating Distribution, Advance Care Plan (MIPS #47 – Group Registry)
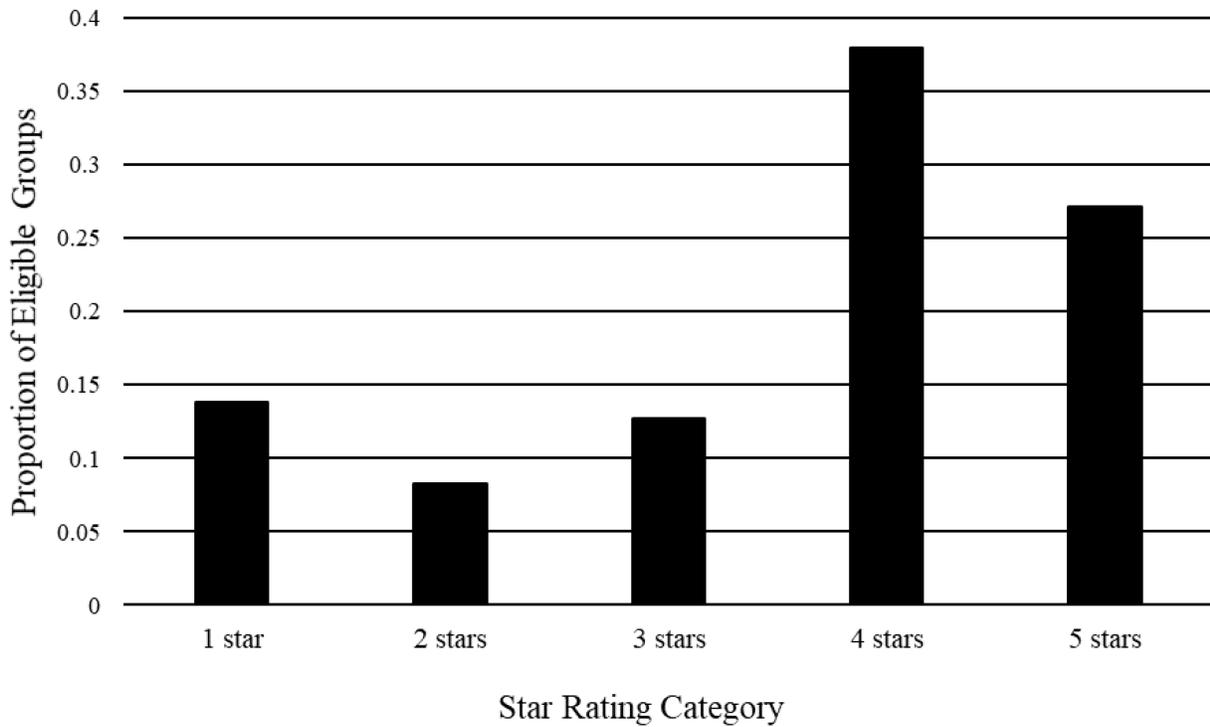
Figure C.5 – Performance Rate Distribution, Urinary Incontinence: Assessment of Presence or Absence of Urinary Incontinence in Women Aged 65 Years or Older (MIPS #48 – Group Registry)
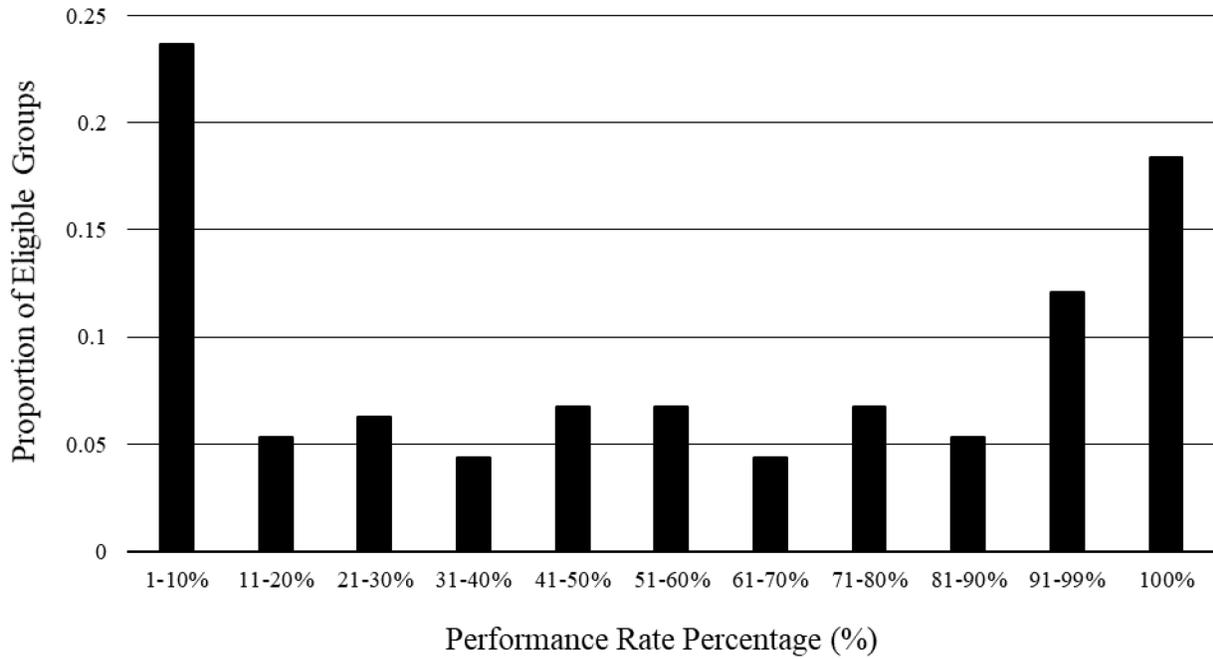
Figure C.6 – Star Rating Distribution, Urinary Incontinence: Assessment of Presence or Absence of Urinary Incontinence in Women Aged 65 Years or Older (MIPS #48 – Group Registry)
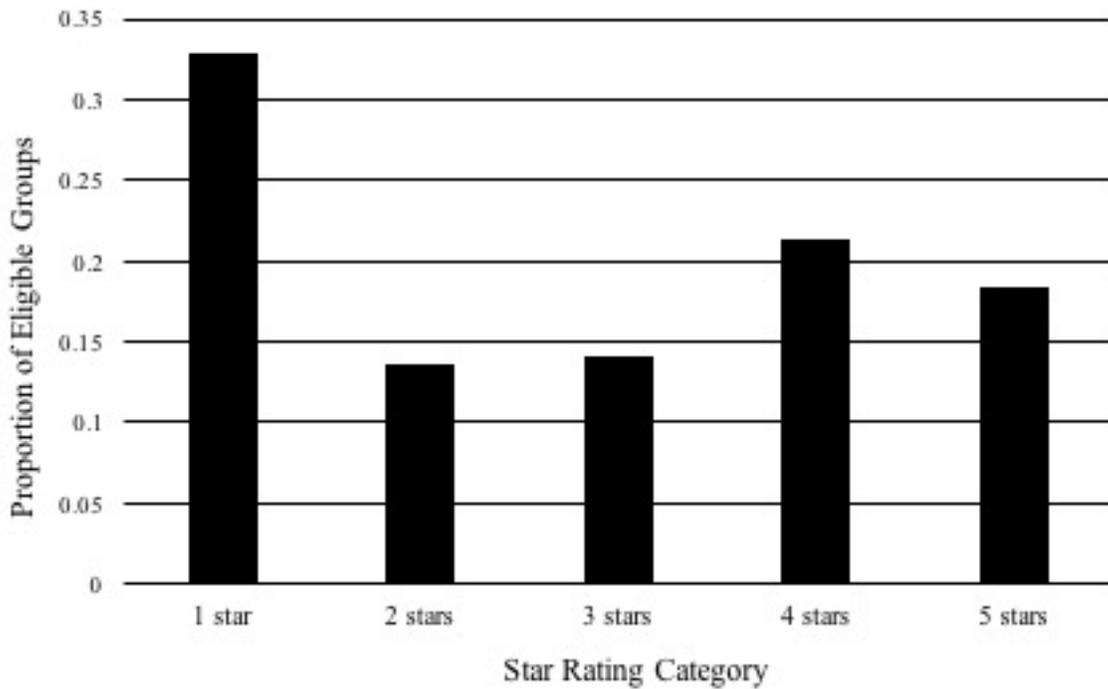
Figure C.7 – Performance Rate Distribution, Chronic Obstructive Pulmonary Disease (COPD): Spirometry Evaluation (MIPS #51 – Group Registry)
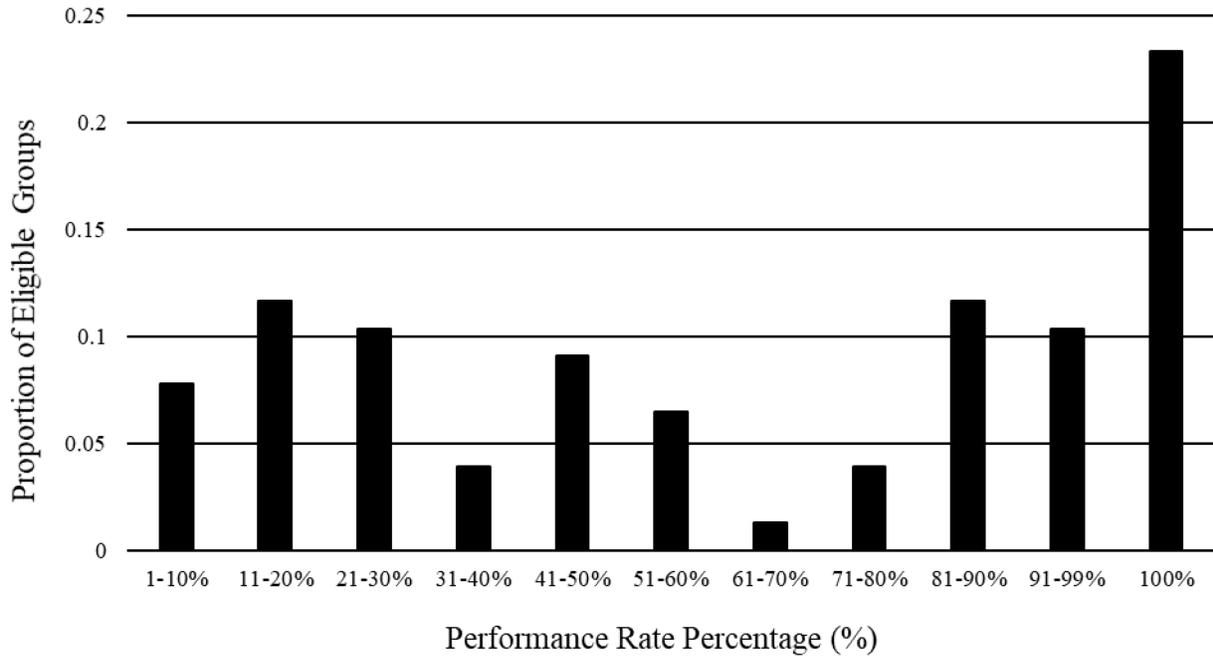
Figure C.8 – Star Rating Distribution, Chronic Obstructive Pulmonary Disease (COPD): Spirometry Evaluation (MIPS #51 – Group Registry)
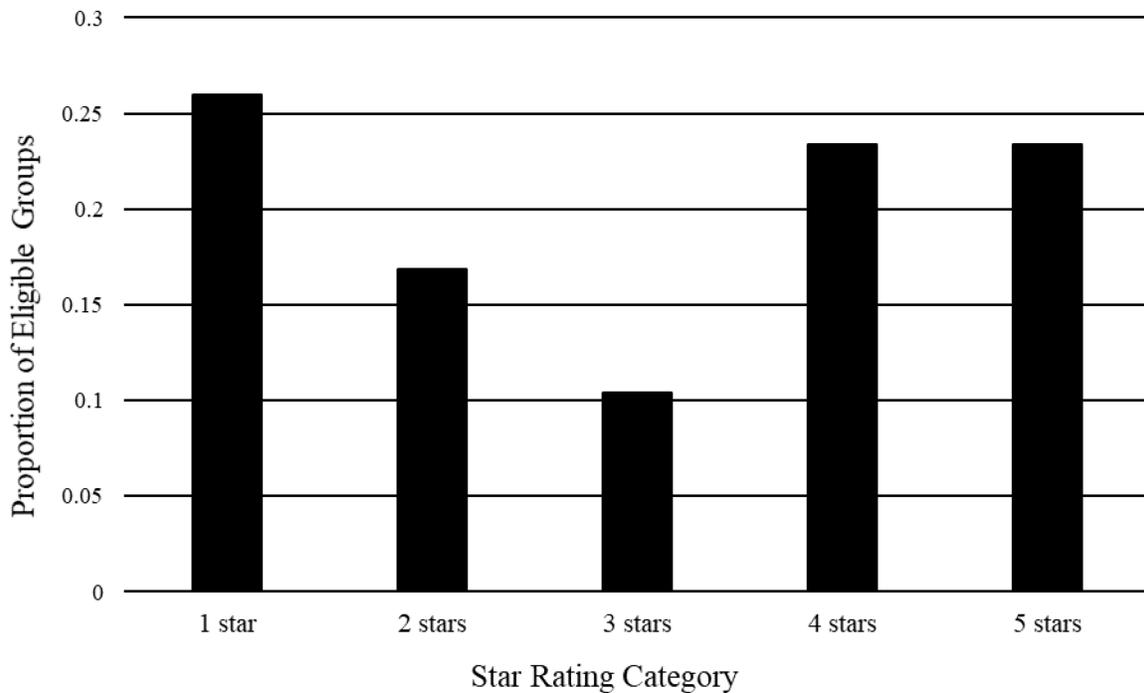
Figure C.9 – Performance Rate Distribution, Preventive Care and Screening: Influenza Immunization (MIPS #110 – CMS Web Interface)
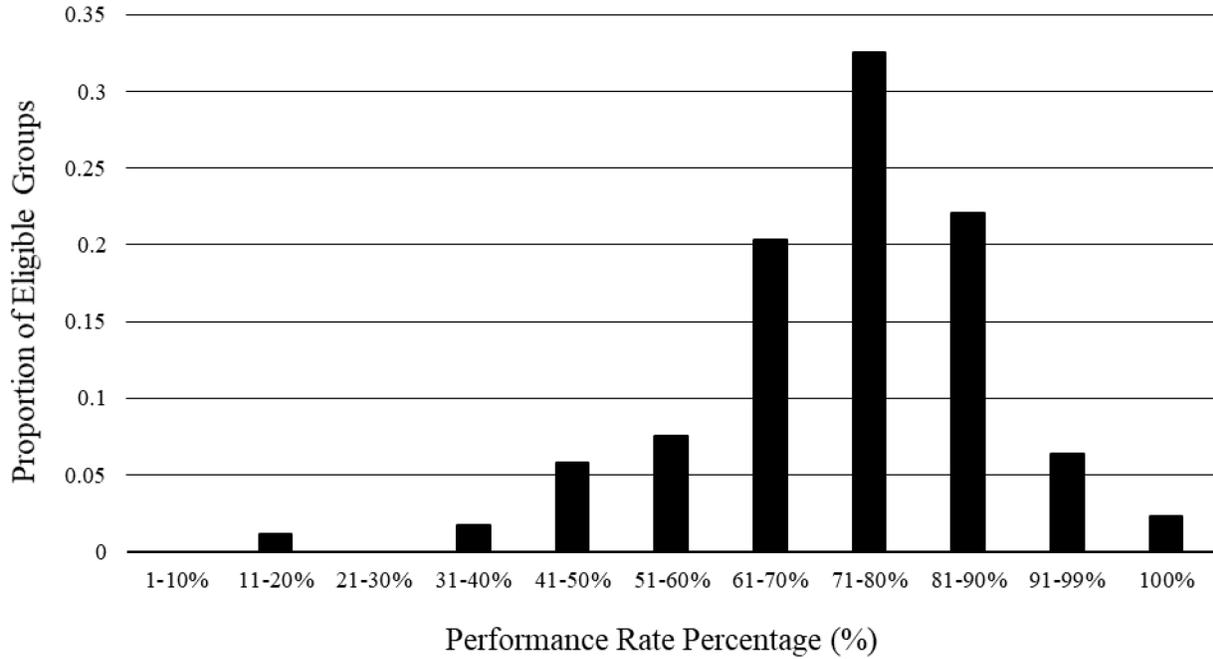
Figure C.10 – Star Rating Distribution, Preventive Care and Screening: Influenza Immunization (MIPS #110 – CMS Web Interface)
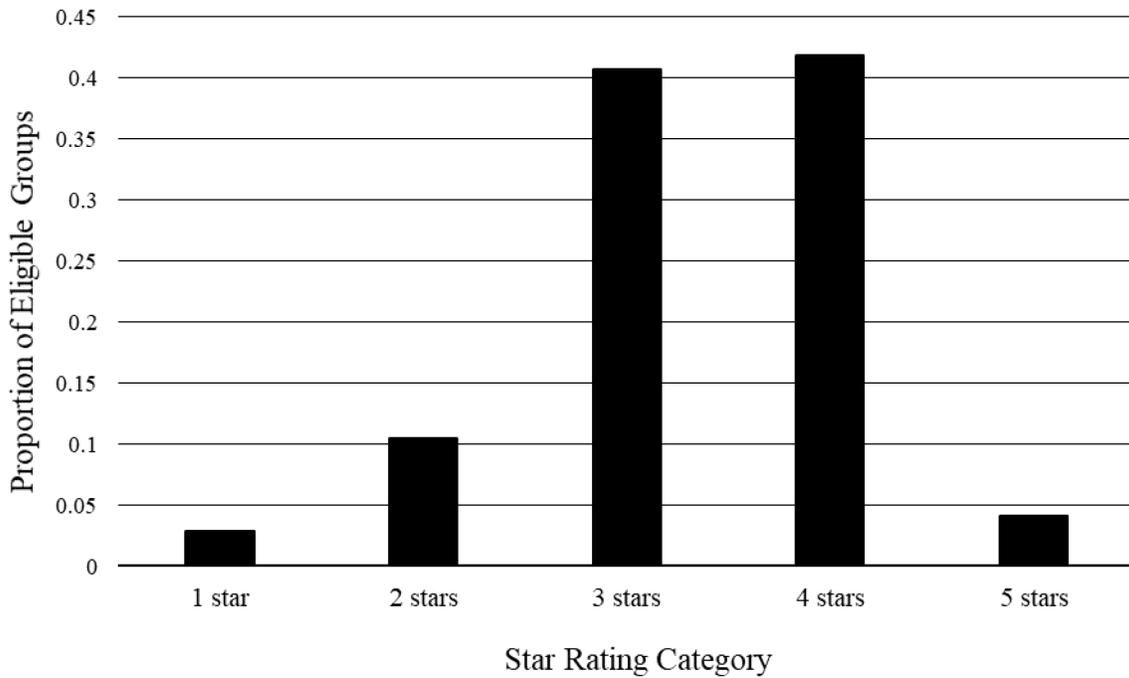
Figure C.11 – Performance Rate Distribution, Pneumococcal Vaccination Status for Older Adults (MIPS #111 – CMS Web Interface)
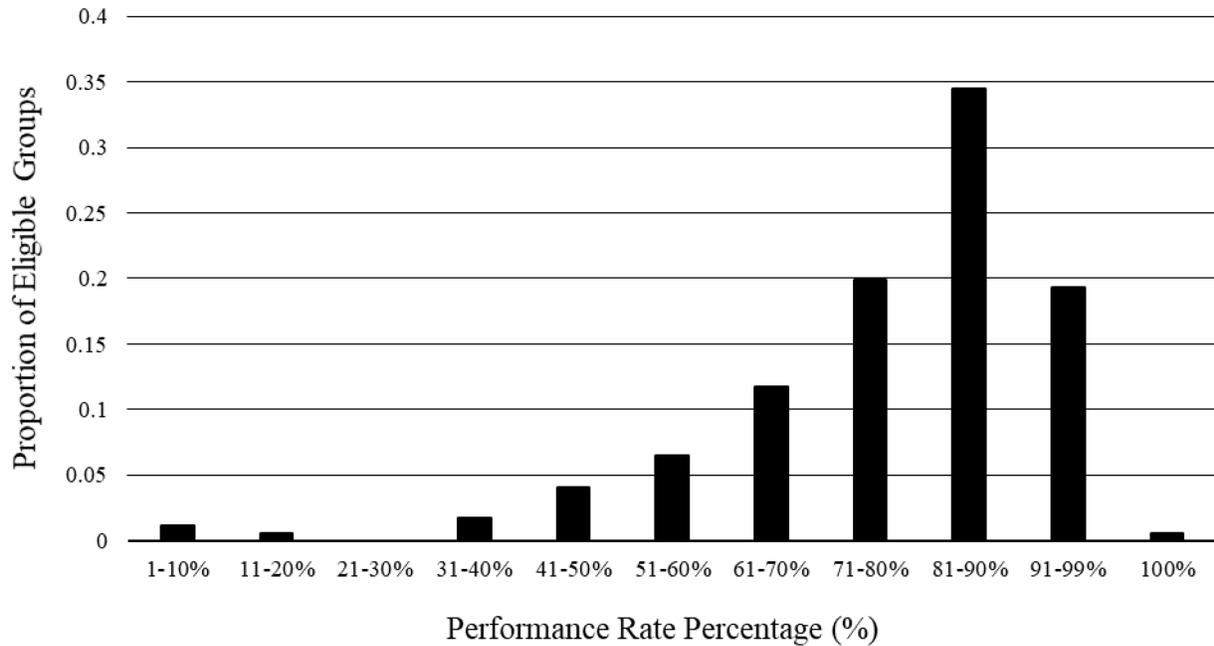
Figure C.12 – Star Rating Distribution, Pneumococcal Vaccination Status for Older Adults (MIPS #111 – CMS Web Interface)
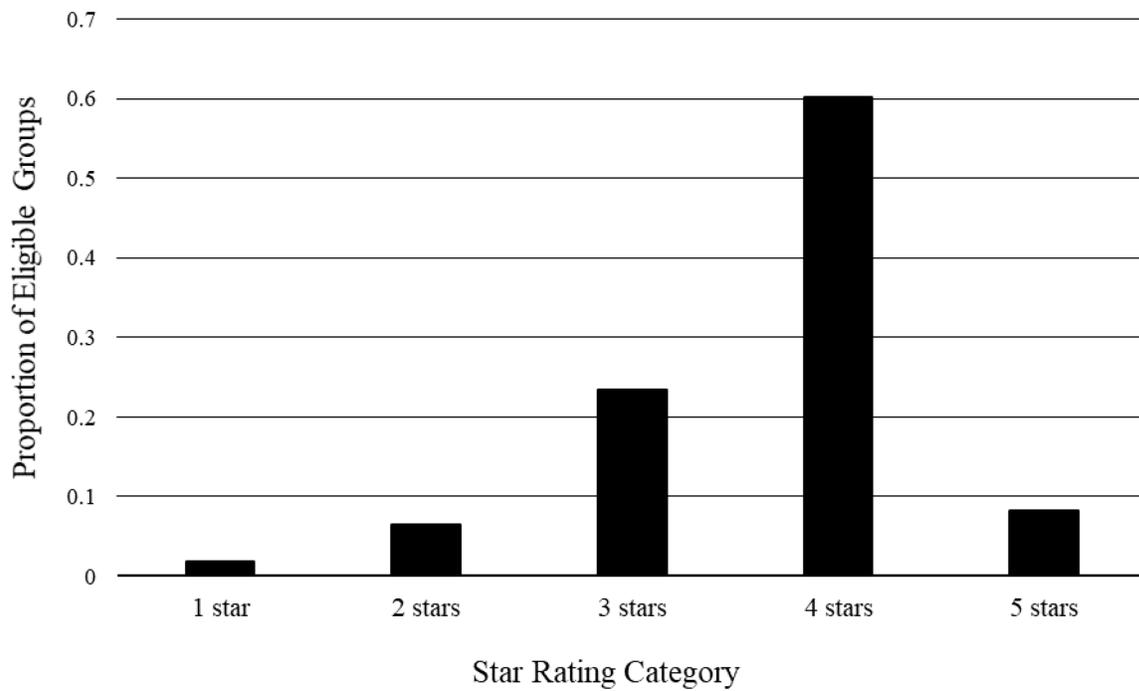
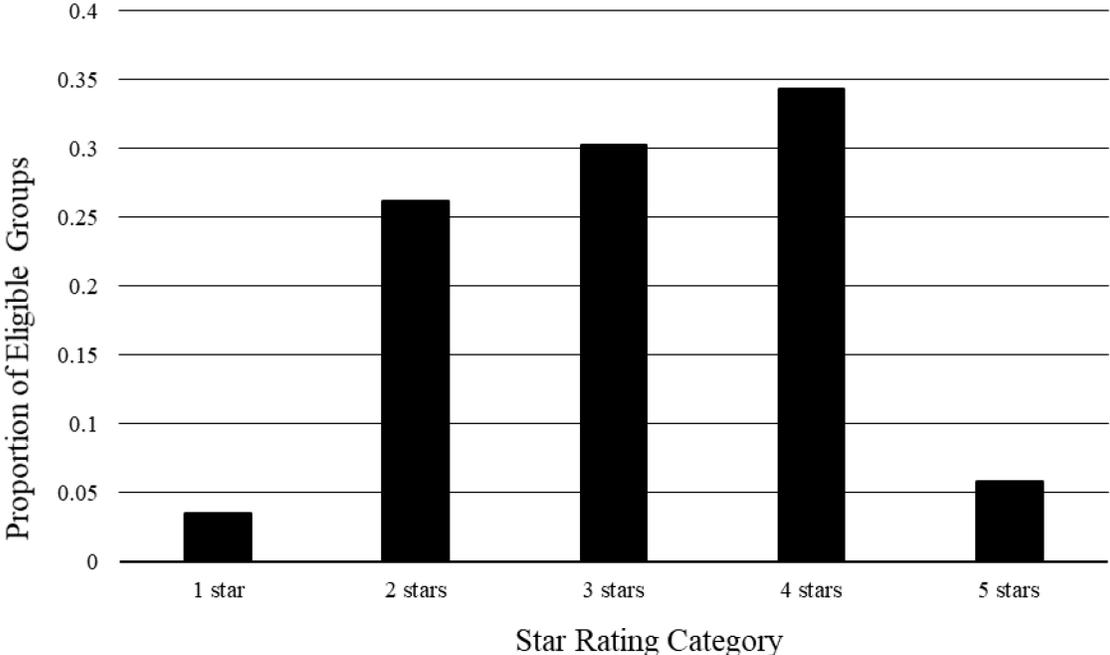**Figure C.13 – Performance Rate Distribution, Colorectal Cancer Screening (MIPS #113 – CMS Web Interface)**
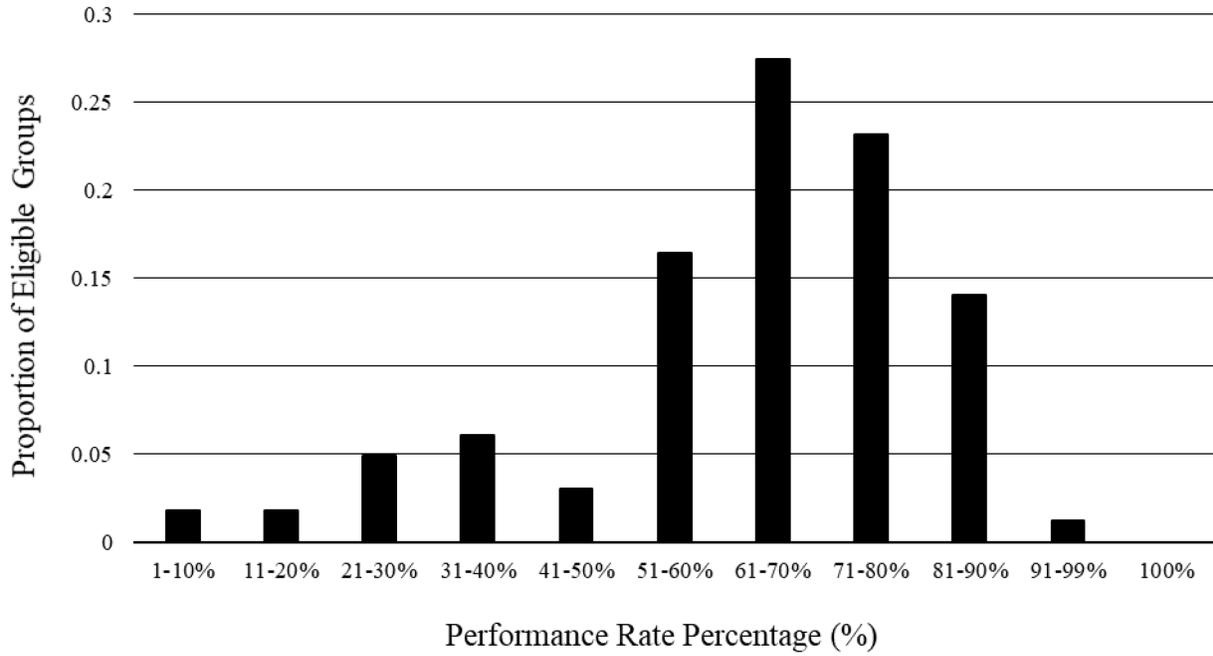
Figure C.14 – Star Rating Distribution, Colorectal Cancer Screening (MIPS #113 – CMS Web Interface)
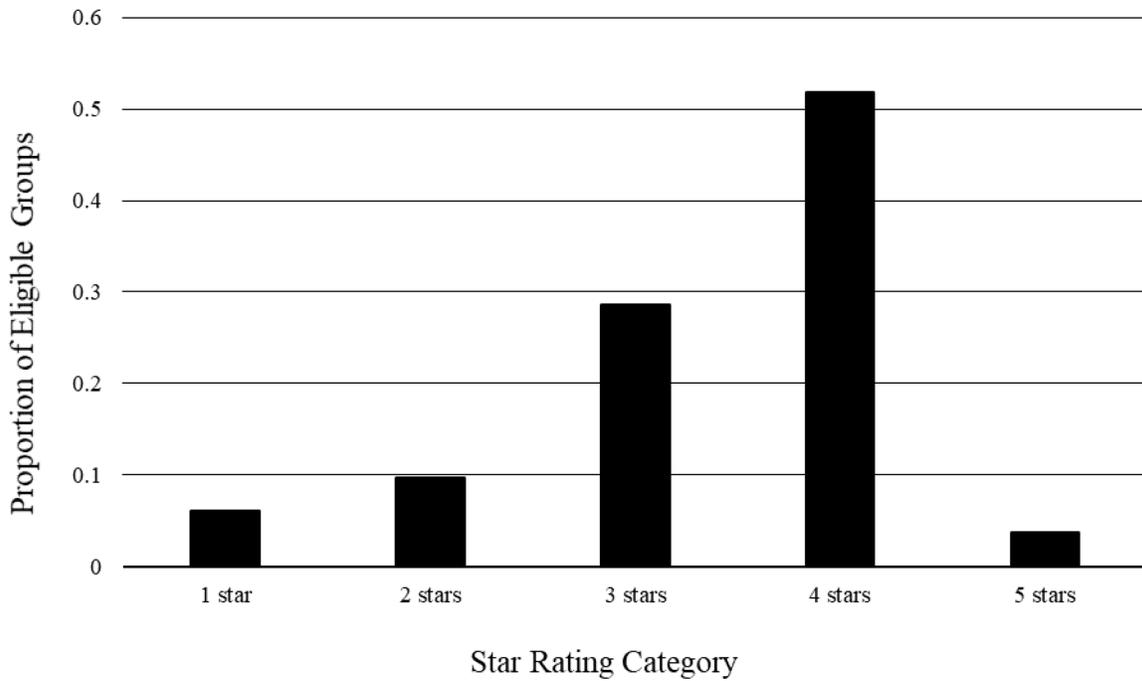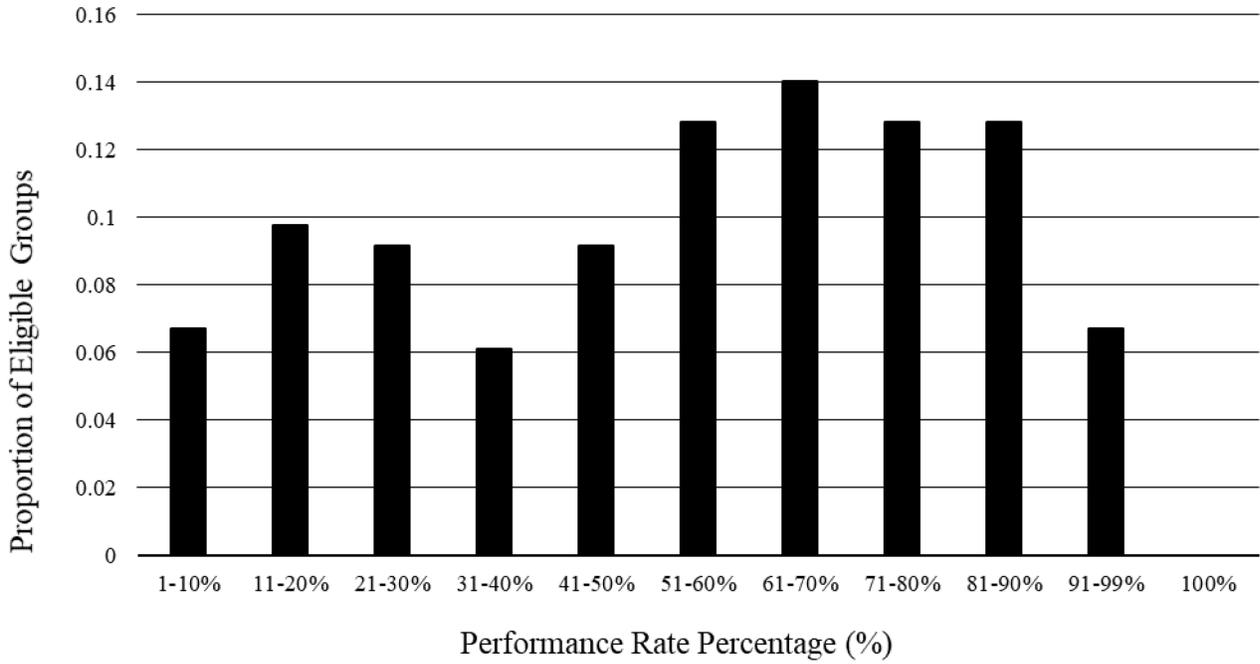
Figure C.15 – Performance Rate Distribution, Colorectal Cancer Screening (MIPS #117 – CMS Web Interface)
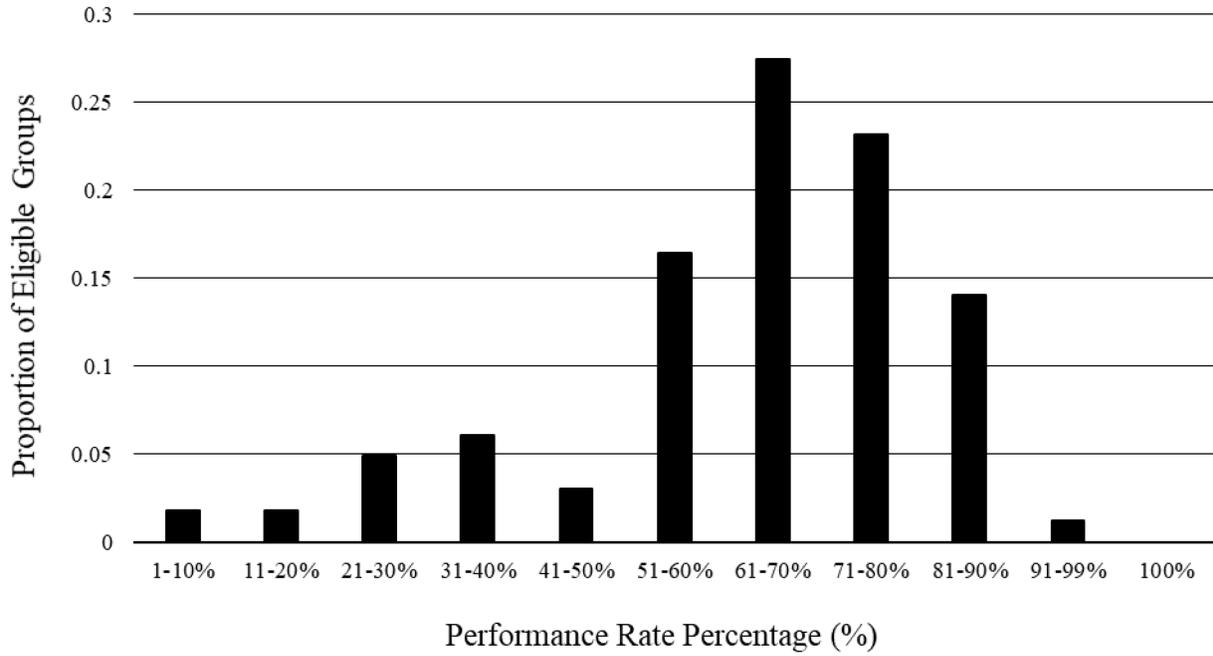
Figure C.16 – Star Rating Distribution, Colorectal Cancer Screening (MIPS #117 – CMS Web Interface)
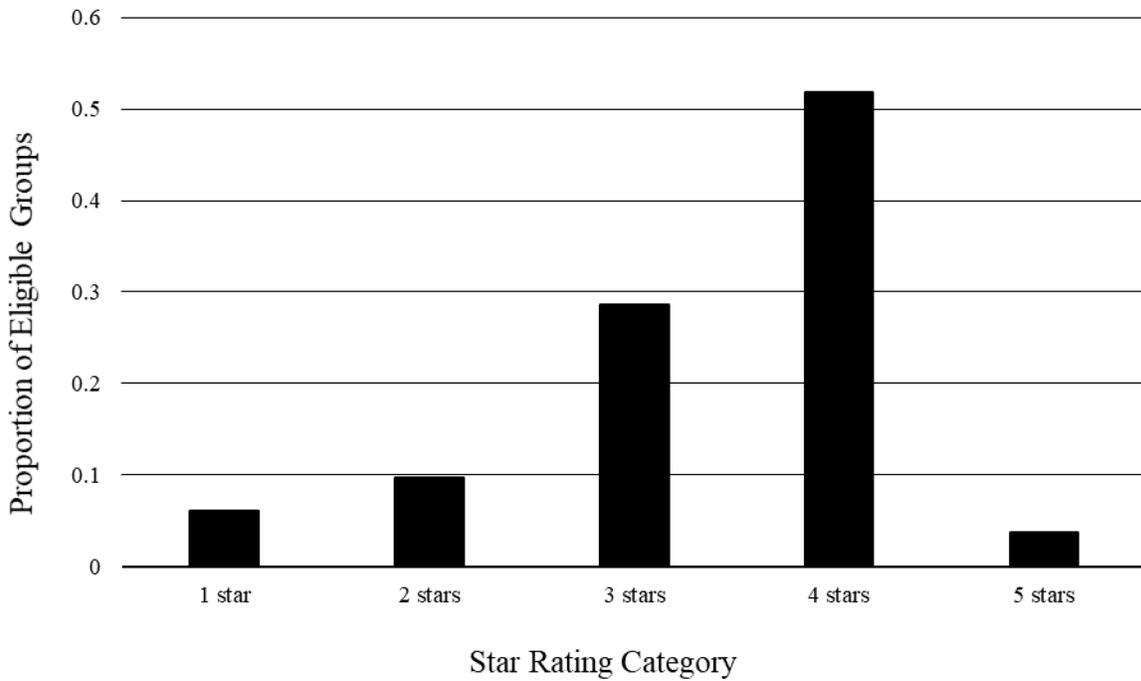
Figure C.17 – Performance Rate Distribution, Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow-Up Plan (MIPS #128 – CMS Web Interface)
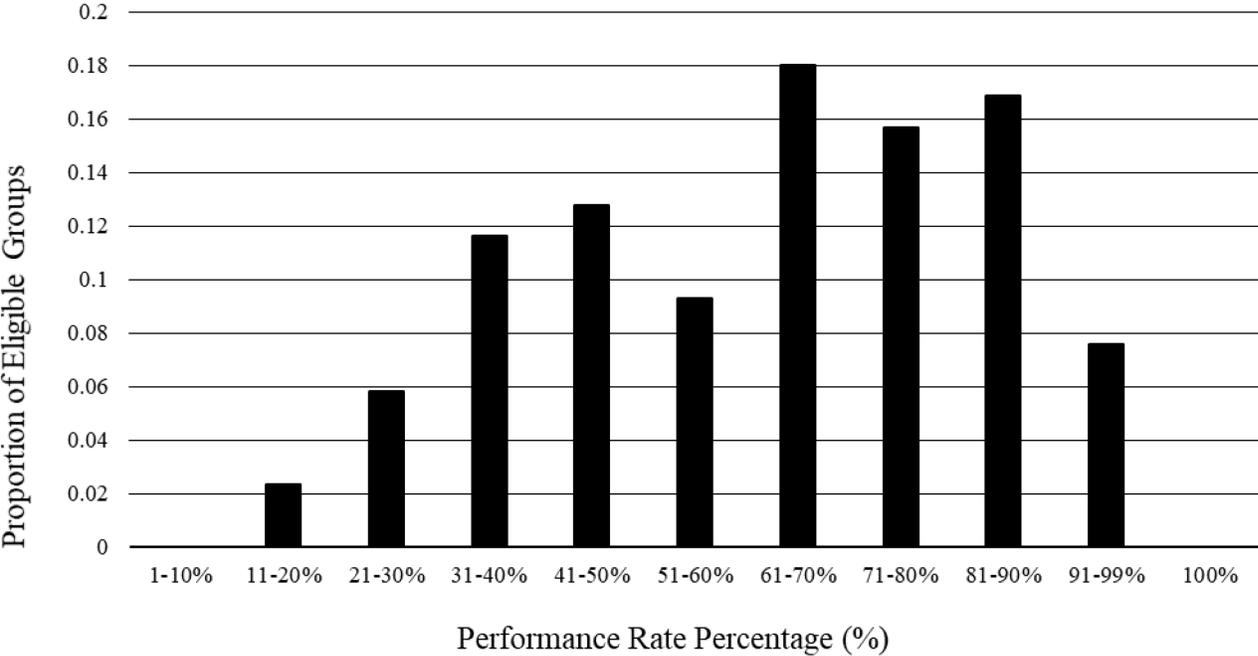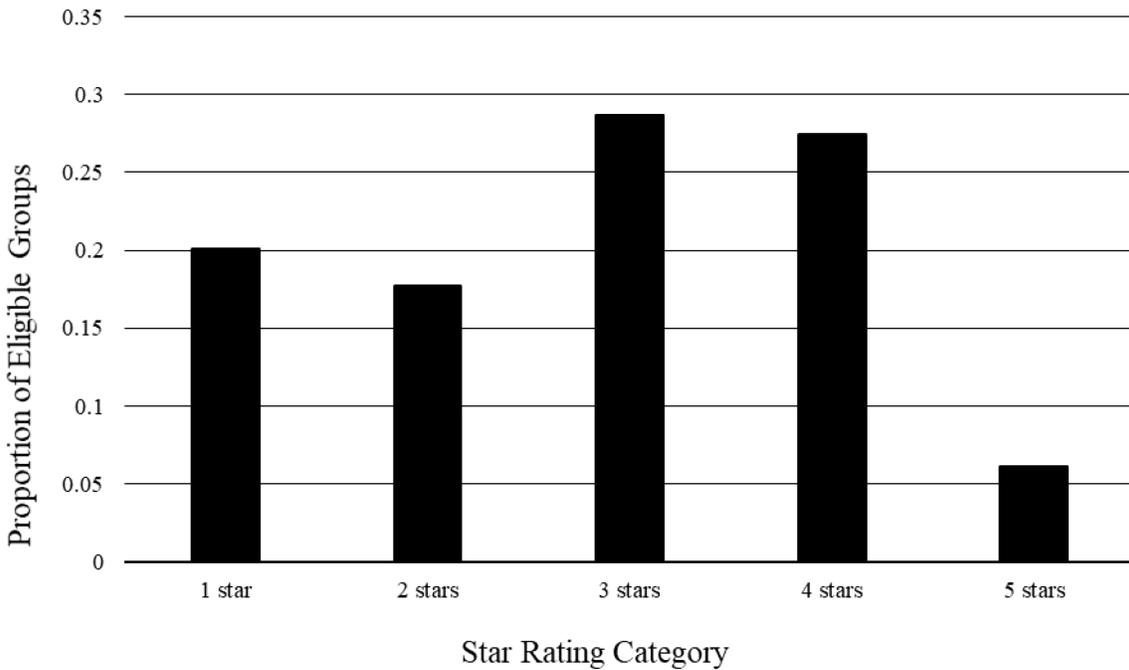

Figure C.18 – Star Rating Distribution, Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow-Up Plan (MIPS #128 – CMS Web Interface)
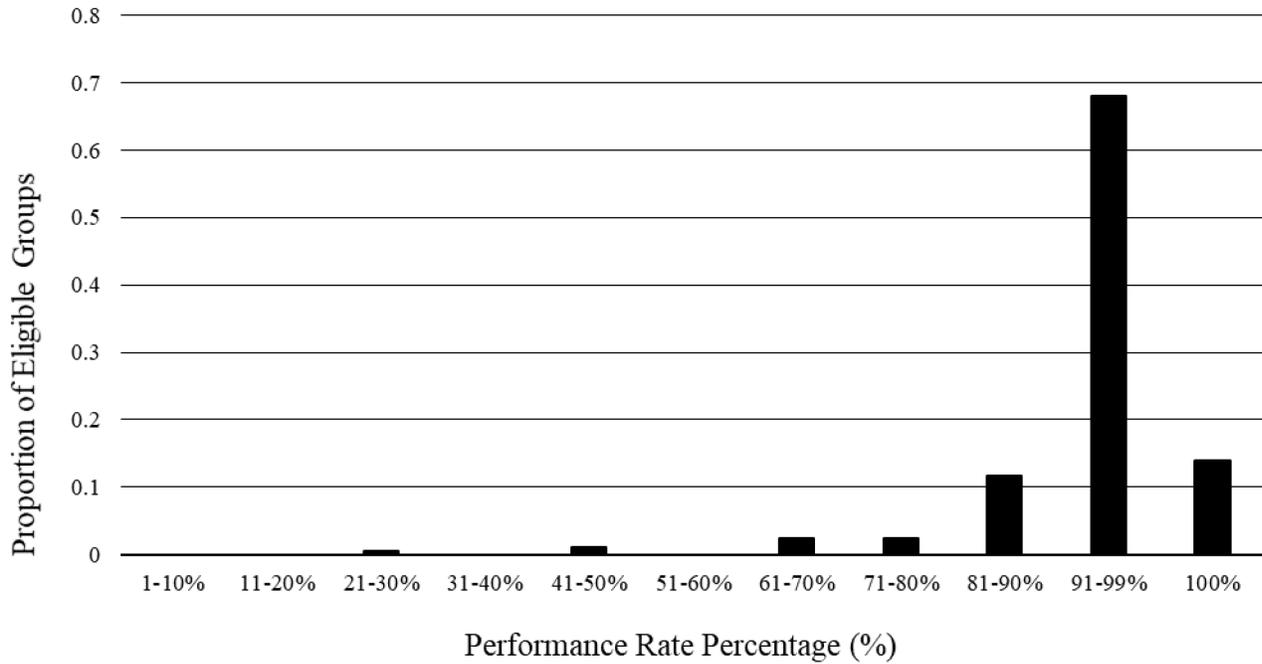
Figure C.19 – Performance Rate Distribution, Preventive Care and Screening: Screening for Clinical Depression and Follow-Up Plan (MIPS #134 – CMS Web Interface)
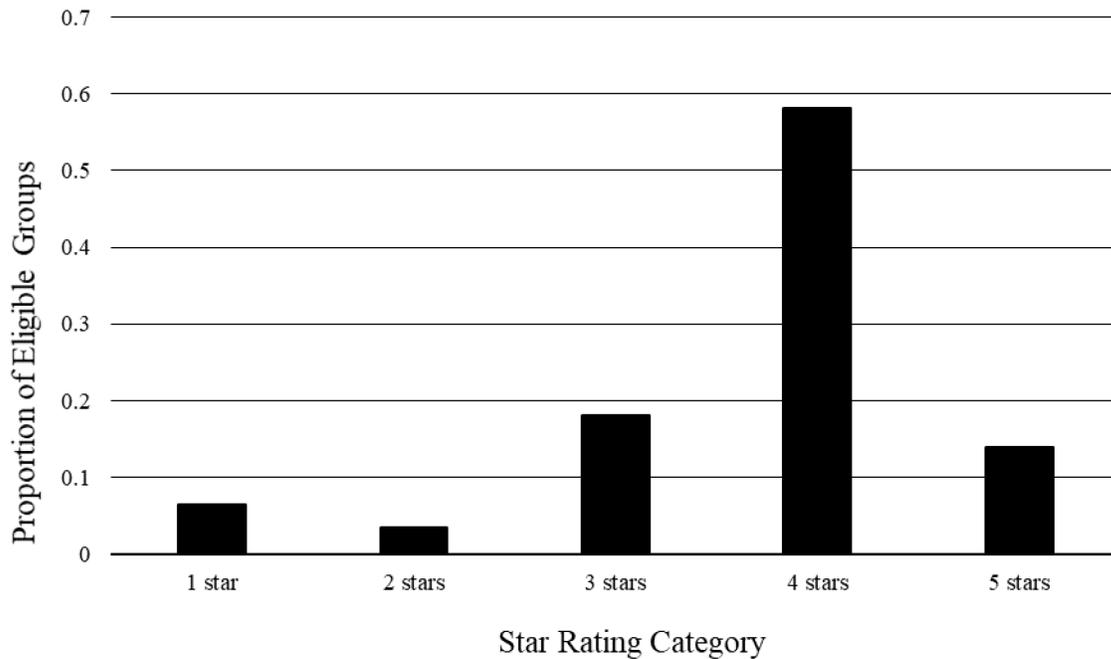

Figure C.20 – Star Rating Distribution, Preventive Care and Screening: Screening for Clinical Depression and Follow-Up Plan (MIPS #134 – CMS Web Interface)

Figure C.21 – Performance Rate Distribution, Preventative Care and Screening: Tobacco Use: Screening and Cessation Intervention (MIPS #226 – CMS Web Interface)
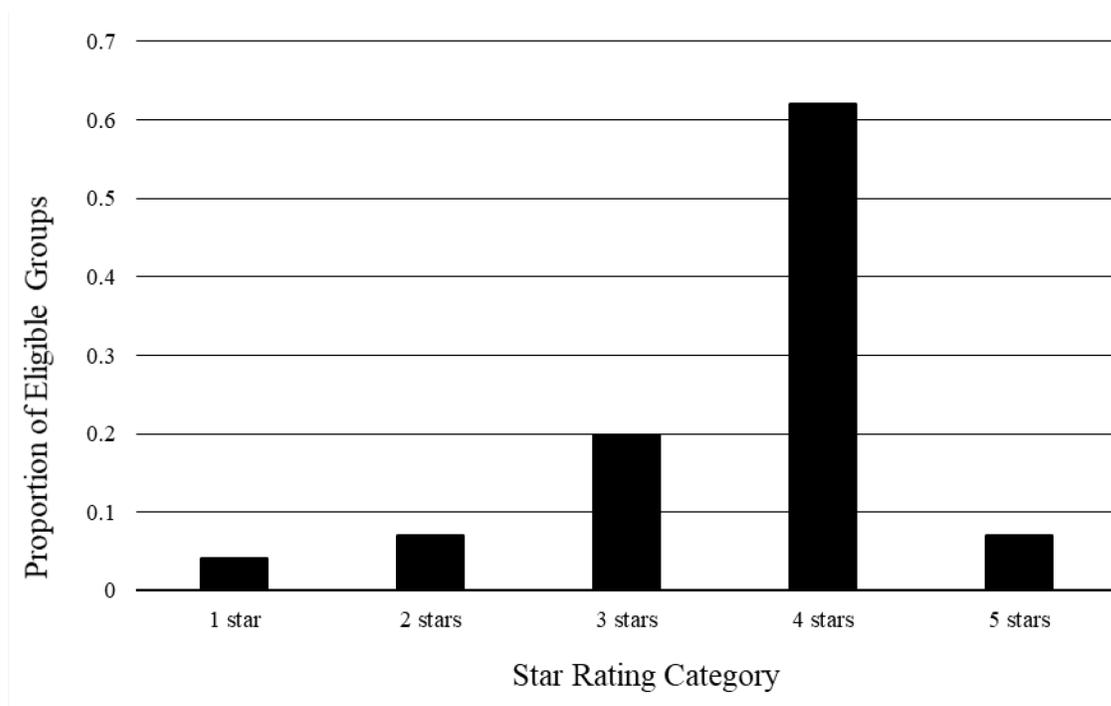

Figure C.22 – Star Rating Distribution, Preventative Care and Screening: Tobacco Use: Screening and Cessation Intervention (MIPS #226 – CMS Web Interface)

Figure C.23 – Performance Rate Distribution, Falls: Screening for Future Fall Risk (MIPS #318 – CMS Web Interface)
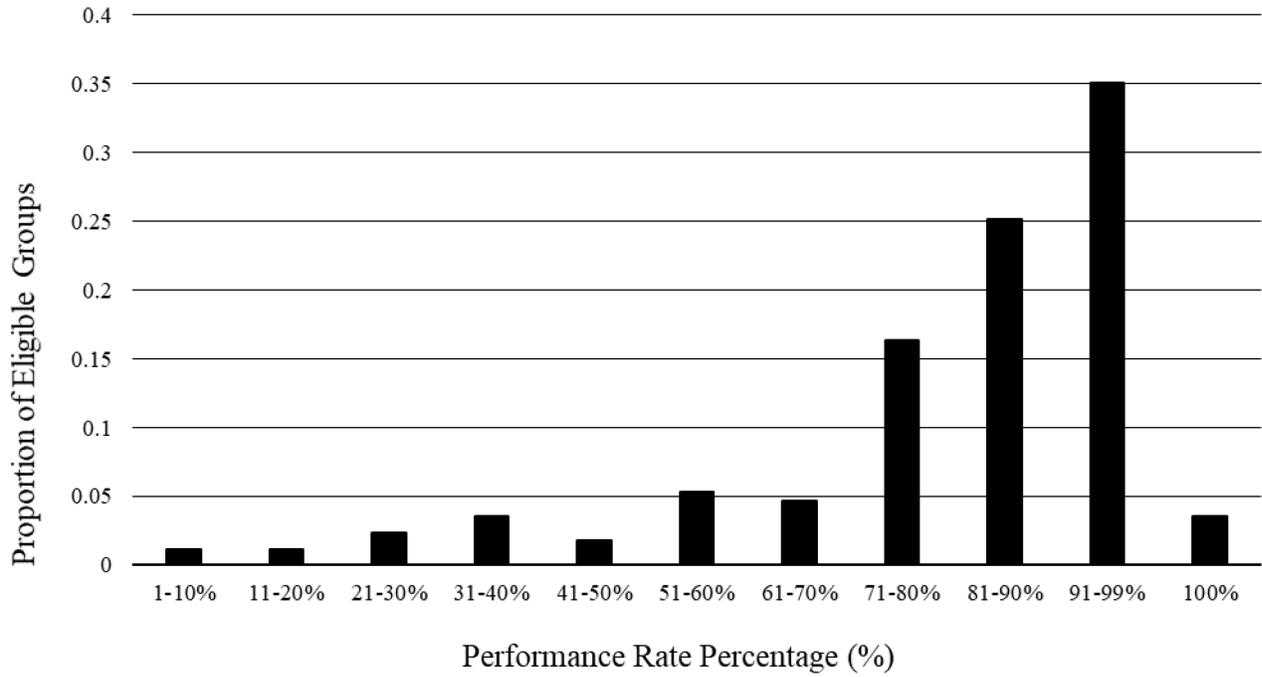
Figure C.24 – Star Rating Distribution, Falls: Screening for Future Fall Risk (MIPS #318 – CMS Web Interface)

APPENDIX D – PY 2017 APM AFFILIATIONS

Clinicians who participated in the following APMs will have an indicator on their profile page for QPP Year 1:

- Bundled Payments for Care Improvement (models 2, 3, & 4)
- Comprehensive Joint Replacement (tracks 1 & 2)
- Comprehensive ESRD Care (tracks 1, 2, & 3)
- Comprehensive Primary Care Plus (CPC+)
- Frontier Community Health Integration Project demonstration
- Independence at Home Demonstration
- Initiative to Reduce Avoidable Hospitalization
- Million Hearts: Cardiovascular Disease Risk Reduction
- Oncology Care Model
- Transforming Clinical Practice Initiative
- Medicare Shared Savings Program (tracks 1, 2, & 3)
- Next Generation ACO Model

APM affiliations will be indicated at the clinician level, with the exception of MSSP and NextGen APMs which will have affiliations at the group level.