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

January 2017

Acumen, LLC
500 Airport Blvd., Suite 365
Burlingame, CA 94010

[This page is intentionally left blank.]

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1 ABOUT THE TEP

Physician Compare operates as the public reporting home of quality initiative and related measurement data for individual EPs and group practices under the authority of the Patient Protection and Affordable Care Act (ACA) of 2010. The legislation requires that data reported publicly on Physician Compare are accurate, valid, reliable, and comparable across the available submission methods. Further, any measures published to profile pages must resonate with and be correctly interpreted by website users. CMS has contracted with Acumen to ensure the information published on Physician Compare fulfills these public reporting standards.

To ensure that the data reported on Physician Compare portray physician performance accurately and robustly, the Physician Compare support team consulted with the Physician Compare Quality Measurement Technical Expert Panel (henceforth "TEP"). Per the CMS Measures Management Blueprint criteria, the TEP consists of members who represent the perspectives of the patient/caregiver dynamic, purchasers, and technical experts with a broad range of experience in publicly reporting performance measures, improving health care quality, and developing and testing quality measures. The Physician Compare support team and the TEP convened on August 24, 2016 to discuss the selection of candidate 2015 quality measures to publish on Physician Compare in late 2016. Table 1.1 lists the 12 individuals who comprise the TEP, seven of whom were present on the teleconference.

Table 1.1: TEP Members

TEP Member	Position(s),Organization	Location
David Baker, MD, MPH*	Michael A. Gertz Professor in Medicine, Chief of the Division of General Internal Medicine and Geriatrics, and Deputy Director of Institute for Public Health and Medicine at Feinberg School of Medicine, Northwestern University	Chicago, IL
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
Ted von Glahn, MS	Consultant	San Francisco, CA
Eric Holmboe, MD*	Internist, Senior Vice President, Milestones Development and Evaluation of the Accreditation Council for Graduate Medical Education (ACGME)	Philadelphia, PA

TEP Member	Position(s),Organization	Location
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
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
Robert Krughoff, JD*	Founder and President, Center for the Study of Services/Consumers' Checkbook	Washington, DC
Michael Mihlbauer, MS*	Practice Administrator, Anesthesiology Associates of Wisconsin	Milwaukee, WI
Sara Schoelle, DrPH*	Assistant Vice President, Research & Analysis/National Committee for Quality Assurance	Washington, DC
Dale Shaller, MPA* **(TEP Chair)**	Principal, Shaller Consulting Group	Stillwater, MN
Thomas Smith, MD, MS	Medical Director, Division of Managed Care, NYS Office of Mental Health/New York State Psychiatric Institute	New York, NY
A.J. Yates, MD	Associate Professor, Department of Orthopedic Surgery/University of Pittsburgh School of Medicine	Pittsburgh, PA

**TEP member was unable to participate in the teleconference, but received and reviewed all meeting materials and was invited to provide written feedback.*

The remainder of this report summarizes the discussions and conclusions from this meeting. Section 2 provides an overview of the 2015 Public Reporting Plan and reporting mechanisms available to practitioners. Section 3 describes the analyses performed to address the statistical public reporting standards and determine which measures are suitable for publication in the downloadable database. Section 4 reviews considerations for the testing undertaken to ensure adequate user understanding of measures posted to live site profile pages. Section 5 walks through the Consumer Assessment of Healthcare Providers and Systems (CAHPS) for PQRS measures recommendations. Section 6 recounts summary statistics describing the representation of quality measures and the specialties of eligible professionals (EPs) for the set of measures recommended for the live site pages. Any TEP input provided during the meeting is contained in the relevant section.

2 OVERVIEW OF 2015 PUBLIC REPORTING PLAN

Through Physician Fee Schedule (PFS) rulemaking, CMS phased in an annually increasing number of quality measures available for Physician Compare. The 2012 PFS final rule (76 FR 73417-73422) outlined a plan to publish measures reported by group practices through the Group Practice Reporting Option (GPRO) Web Interface in 2012, as well as those submitted by Shared Savings Program and Pioneer Accountable Care Organizations (ACOs). After identifying the candidate measures meeting public reporting standards, CMS made the first set of measures, consisting of five diabetes mellitus and coronary artery disease measures, available in February 2014. For the second phase, four 2013 PQRS Web Interface measures were posted to Physician Compare profile pages in December 2014 for group practices and ACOs.

The 2014 PFS final rule (78 FR 74446-74454) greatly expanded the amount of data available for public reporting by: (1) increasing the number of quality measures available for publication on Physician Compare (including CAHPS patient experience summary survey measures); (2) making data submitted by individual eligible professionals (EPs) publicly available; and (3) expanding the types of reporting mechanisms available for public reporting. Data submitted via Medicare Part B Fee-for-Service claims (for individuals), the GPRO Web Interface (for groups with at least 25 EPs), and qualified PQRS registries and electronic health records, or EHR, (for both individuals and groups of any size) became available to be publicly reported. CMS could not determine the accuracy of the 2014 EHR data, so these data were not considered for public reporting. In December 2015, CMS published the third and most recent set of quality measures to Physician Compare, including 14 PQRS Web Interface measures and 8 CAHPS measures submitted by groups and 6 PQRS measures submitted through claims by individual EPs. Since 2012, Physician Compare has increased the number of reporters represented from 66 group practices and 141 ACOs in phase one to 345 group practices, 353 ACOs, and roughly 37,000 individual EPs in the most recent release in December 2015.

As stipulated by the 2015 PFS final rule (79 FR 67761-67766), all PQRS measures[1] collected through any of the aforementioned submission methods by groups or individuals, 12 CAHPS for PQRS measures, established non-PQRS measures[2] submitted by Qualified Clinical Data Registries (QCDRs), and all Shared Savings Program and Pioneer ACO quality measures submitted in 2015 are available for public reporting in late 2016. As no first-year measures are currently publicly reported on Physician Compare and this is the first year CMS is using EHR-reported performance scores across programs, Physician Compare is not reporting these measures this year. CMS plans to revisit these measures next year for consideration for future

[1] Excluding any measures that are only available through a measures group.
[2] Measures that have been in use by CMS for longer than one year

public reporting. Table 2.1 presents the specifications for each reporting mechanism with 2015 data available for public reporting. The first column details the submission method. The next two columns describe the representation of eligible patients and number of measures required to meet CMS's criteria for satisfactory reporting. The final column reports whether the reporting mechanism is available to groups, individuals, or both.

Table 2.1: The submission methods available for reporting year 2015 and their respective requirements for satisfactory reporting

Submission Method	Reporting Requirement	Number of Measures Required	Groups or Individuals?
Web Interface	At least 248 sampled patients*	17	Groups
Claims	50% of eligible patients	9**	Individuals
Registry	50% of eligible patients	9**	Both

*If the group practice does not have 248 eligible patients, it must submit data on 100% of its eligible patients.

**The measures must span 3 National Quality Strategy (NQS) domains; one must be a cross-cutting measure.

Individual EPs can also coordinate with a Qualified Clinical Data Registry (QCDR) to submit data on both PQRS and non-PQRS measures. PQRS measures submitted through QCDRs must either use registry or EHR specifications. Because EHR-reported measures are not available for public reporting this year, any QCDR-submitted measures using EHR specifications were excluded from candidacy. Eligible non-PQRS measures include those used by boards, specialty societies, and regional quality collaborations, among other CMS-approved measures. QCDRs can opt to publish their non-PQRS measures either on the Physician Compare live site profile pages or on their own websites. For those that choose to report their own measures, we provide a link on Physician Compare to the QCDR websites containing their non-PQRS measures, as long as the posted measures adhere to public reporting standards.

Two sets of measures are published on Physician Compare. The first consists of measures meeting statistical standards; all measures found to be accurate, reliable, valid, and comparable across submission methods are posted on the website in a downloadable database. The other set contains the measures included on EP and group practice profile pages; in addition to meeting the statistical standards, these measures must be understood by and resonate with website users. There are three types of measures discussed herein: PQRS, non-PQRS, and CAHPS for PQRS measures. All measures available for public reporting must have been in use for at least one year.

Sections 3-5 discuss the analyses and considerations factoring into our measures recommendations.

3 MEASURES RECOMMENDATIONS: DOWNLOADABLE DATABASE

The downloadable database serves as an exhaustive compilation of all measures meeting the mandated statistical standards, including non-PQRS measures published to QCDR websites (and linked to from Physician Compare profile pages). It details the performance scores received by each group and individual EP reporting quality data for each measure, as well as an indication of the reporting mechanism used to submit the measure. This section explains the analyses performed to inform our measures recommendations for the downloadable database. We describe our assessment of the comparability, reliability, and validity of the submitted measures data. Then we detail the additional considerations factoring into the recommendations for measures data to include in the downloadable database.

3.1 Comparability across Submission Methods

Existing regulation and mandates require that any quality data published to Physician Compare are comparable across the submission methods through which they were collected. To address this concern, we compared the performance rate distributions across the reporting mechanisms used for each individual measure and were unable to confirm comparability of data submitted from distinct methods. Thus, we conservatively recommended selecting data from a single submission method for each measure to publish on Physician Compare. To make as much information available as possible, we recommended publishing the data from the most represented submission method for each measure that meet the other statistical standards for public reporting. For individuals, the submission method with the highest number of EPs reporting data is chosen; for groups, representation is determined by tabulating the number of EPs in the group practices that have data for the measure.

3.2 Reliability

To satisfy the reliability standard, we assessed the reliability of the data submitted through each individual submission method available for each measure. Reliability describes the confidence with which one EP's performance can be distinguished from another and accounts for the precision with which the performance rates are measured. We calculated reliability using two methods: (1) the beta-binomial model and (2) a split-half reliability test.

Reliability scores calculated using a beta binomial model range from 0 to 1, where scores closer to one indicate better reliability. We used the 25th percentile of the reliability scores for a given measure to determine whether its data are reliable. Literature suggests that groups and individuals should have different thresholds for adequate reliability (0.7-0.8 for groups and 0.9

for individuals).[3] Because groups using a registry could technically be comprised of only two EPs, we used the individual standard (0.9) for these reporters. Web Interface users must be groups with at least 25 EPs, so we adopted the upper limit of the accepted range for sufficient group reliability (0.8) as the cutoff. Thus, for all individual EP submissions and for groups submitting data through a registry, if the 25th percentile of the reliability scores determined by the beta binomial model for a given measure was lower than 0.9, the measure was considered unreliable and is recommended against; for groups using the Web Interface, measures with a 25th percentile value under 0.8 were designated as unreliable and were recommended against.

The split-half reliability test randomly divides the population into halves and compares performance between the two halves. If the resultant intraclass correlation coefficient was under 0.75, we designated the measure as unreliable. A measure must pass both the beta binomial test and the split-half reliability test to maintain candidacy. Overall, 28 PQRS and 72 non-PQRS measures failed our reliability check (as shown in Table 5.2).

3.3 Validity

We addressed the validity of measure data in two ways: (1) a case-mix adjustment for relevant measures and (2) an investigation into the possibility of selective reporting.

3.3.1 Case Mix Adjustment

Performance on certain measures can be susceptible to influences from the patient population served by a reporting entity. Unlike process measures, which portray the frequency with which a clinical standard of care was performed by a reporter, outcome measures reflect patient health (e.g. the percentage of diabetic patients with controlled blood sugar), which is meant to illuminate whether the care provided by reporters translates into positive health outcomes. However, since PQRS outcome measures do not include risk adjustment strategies, performance on these measures could be associated with the types of patients treated, given that certain patient populations will be healthier than others (e.g. a reporting entity serving an affluent suburban client base will be expected to have patients with better health metrics than one treating patients in a more disadvantaged socioeconomic environment). If characteristics of the patient population impact the observed performance rates on outcome measures, those rates would be an invalid representation of the quality of care provided by the reporting entity.

To investigate this issue, we built two models to adjust for case mix. The first controlled for basic patient demographic and clinical conditions (constructed from AHRQ CCS codes), and the second model added socioeconomic information (built from 2015 American Community Survey data) to the first. Patient-level data are only available through the Web Interface and

[3] Hays RD, Revicki D. Reliability and validity (including responsiveness). In: Fayers P, Hays R, eds. Assessing Quality of Life In Clinical Trials. New York: Oxford University Press; 2005

claims reporting mechanisms, so we could only perform adjustments using those data for groups and individual EPs, respectively. In both Web Interface and claims data, both models predicted performance values that differed from the observed performance rates for both of the relevant measures (PQRS 1: Diabetes – Poor Hemoglobin A1c Control and PQRS 236: Hypertension – Blood Pressure Control). Therefore, we recommended against using these measures.

TEP Input on Case-Mix Adjustments

The Physician Compare support team and a TEP member discussed how we perform the case-mix adjustments and make decisions from the results. We explained that our measures are not risk adjusted and that we cannot publish adjusted rates; we can only recommend against measures where we discover evidence of case-mix effects. A TEP member asked how we react to situations where we only observe evidence of case-mix effects when adding the socioeconomic variables in the second model (and fail to see a result with the first model). We noted that this situation has not occurred in analyses thus far. If there was an effect when factoring in the SES covariates, it was also evident when those variables were not included.

Additionally, in the past, we have found case-mix effects when controlling for basic demographics and clinical conditions. Because the observed performance appears to be influenced by the patient population served by reporters, publishing the performance scores without adjusting for basic demographics and clinical conditions would then be an invalid reflection of clinical quality. Therefore, we continually recommend against using these measures, even though users find these measures useful. We asked TEP members whether they agreed with our approach for handling these types of measures. TEP members agreed with our decision to exclude measures with evidence of case-mix effects and echoed our sentiment that the lack of risk adjustment is unfortunate because these are measures users find particularly helpful.

3.3.2 Selective Reporting

Groups or individual EPs using either a registry or Medicare claims are only required to submit quality data on half of their eligible patients to be considered satisfactory reporters. This creates an opportunity for reporters to select their best patients on whom to submit data in an attempt to inflate their performance rates (referred to as cherry-picking). If, in fact, a subset of reporters cherry-picked patients for measures data submissions, their artificially elevated performance rates would be an invalid comparator to those reporters choosing a more inclusive representation of their patient population.

To address this concern, we studied the relationship between reporting volume and performance in claims- and registry-submitted measures. We evaluated these measures by separating the EPs into reporting rate buckets of 10% from 50-100% reporting volume (50-60%, 60-70%, 70-80%, 80-90%, 90-100%) and determining the performance rate distribution inside each bucket. If, on average, EPs reporting half of their eligible patients were performing better

than their counterparts with higher reporting volume, we considered those measures data to have potential evidence of selective reporting and conservatively recommended against publication of their data on Physician Compare. Five claims-submitted measures had trends suggestive of selective reporting, four of which also had reliable data available through the registry reporting mechanism, so the registry data were recommended instead for these measures. The fifth measure (PQRS Measure 41) did not have reliable data through registry for individual EPs, so we recommended against publishing PQRS Measure 41. Because most registry measures appear to be well reported (97% of measures had a 100% median reporting rate), we only assessed the two candidate measures with more variable reporting rate distributions that passed our reliability assessment, both of which did not have sufficient evidence of selective reporting to recommend against the measures.

3.4 Zero Percent Performance from Registry Submissions

As mentioned in Section 3.1, reporters submitting data through a registry depend on a qualified vendor to handle their measures data. Therefore, it is possible that the EP or group practice may not be tracking to the entire set of measures submitted to CMS by the registry. These EPs or group practices may have measures reported with zero-percent performance rates for measures where they technically had eligible patients on whom they did not perform the care being measured because it was irrelevant to the reason for the patient visit. For instance, a neurologist may have a 55-year-old female patient whom he did not screen for breast cancer because it was not within his scope of care. Thus, CMS recommended suppressing zero-percent, registry-reported performance rates from public reporting. A counterpoint to this recommendation is that some zero-percent scores are real reflections of performance, and one goal of Physician Compare is to incentivize better clinical performance.

3.4.1 TEP Input on Zero-Percent Performance from Registry Submissions

Two TEP members debated the advantages and disadvantages of eliminating the zero-percent scores. The first member agreed with removal of the zero-percent rates because we are still in a transitional period of EPs or group practices getting up to speed on public reporting requirements and logistics. A zero-percent score may not mean the EP or group practice is doing a poor job but rather that they were not tracking to the measure or there was a data collection issue. Thus, it seems fair to give EPs and group practices a pass for a year or two until the processes become more refined. Another member argued the opposite position: if there is no consequence for zero-percent rates, there is no incentive to improve the process. One or two years could be stretched into an undesirably longer amount of time. The first member reinforced his original stance by pointing out that there are many small groups who still do not have electronic medical records, containing physicians who are facing the last few years of their careers and others who may end up being hired by hospitals and major groups. It seems

reasonable, as a courtesy to these smaller groups, that we err on the side of caution at this stage in public reporting.

3.5 Recommendations for the Downloadable Database

Measures with data analytically determined to be valid and reliable remain eligible for public reporting through the reporting mechanism with the highest representation. For groups, representation is based on the number of EPs comprising the group practices submitting data for the given measure. Furthermore, CMS decided to exclude measures only available through measures groups or EHR and to suppress true zero-percent performance rates from registry submissions. A true zero-percent performance score means that the group or EP did not report successful completion of the measure for a single patient; it is possible for a performance rate to be displayed as zero percent on Physician Compare if a reporter completed the measure for a small enough fraction of eligible patients that the score rounds down to zero. The measures recommended for public reporting in the downloadable database are reported in Appendix A.

4 MEASURES RECOMMENDATIONS: LIVE SITE PROFILE PAGES

This section discusses the testing planned to ensure that measures published to the profile pages are understood by and resonate with website users, as well as further considerations for selecting the measures to include on the live site.

4.1 User Testing

For users to have a productive website experience, they must be able to correctly interpret measures and find them useful for health care decision-making. To accomplish this, we first produced plain-language versions of the measure titles and descriptions for measures meeting the statistical reporting standards. We then tested these plain language versions of the measures with people who receive Medicare and their caregivers. The plain-language measure titles and descriptions transform the technical descriptions and measure specifications into language that is understandable to users while maintaining the clinical meaning. User testing had yet to take place at the TEP meeting, but Section 4.2 details an issue discussed with TEP members regarding user understanding of inverse measures.

4.2 User Understanding of Inverse Measures

For inverse measures, lower performance rates indicate higher clinical quality. For instance, PQRS 1 (Diabetes: Hemoglobin A1c Poor Control) measures the percentage of a EP's diabetic patients whose blood sugar was uncontrolled during the measurement period (HbA1c > 9.0%). Here, a lower performance rate reflects superior management of blood glucose levels in diabetic patients and, hence, better clinical outcomes. In the past, inverse measures have not tested well with users. Users generally do not correctly interpret a lower score as a representation of better care, especially given that all of the other scores are reported in the more intuitive direction where a higher score is better.

One potential solution to aid user interpretation of inverse measures is to "flip" the performance scores and the plain language measure titles and descriptions so that a higher score is consistently better on profile pages across all measures. The fall 2016 user testing aimed to assess whether users respond well to the "flipped" presentation of inverse measure data.

Table 4.1 displays the examples of "flipped" plain language measure titles and descriptions presented to TEP members during the meeting. The two examples were meant to accentuate difficulties beyond just inverting the language so that a higher score is better. Even with successfully "flipped" measure language, users may still have difficulty interpreting the underlying pertinence of some measures to clinical quality. The aforementioned blood glucose control measure (PQRS 1) does not seem likely to cause issues with user understanding because proper control of blood sugar in patients with diabetes is a widely known area of importance.

However, concepts inlaid in other measures can be more difficult for users to grasp without further education, such as antibiotic prescriptions being inappropriate treatment for certain conditions.

Table 4.1: Examples of Flipped Plain Language Measure Titles and Descriptions Presented to TEP Members

Measure	Version	Title	Description
PQRS 1	Technical	Diabetes: Hemoglobin A1c Poor Control	Percentage of patients 18-75 years of age with diabetes who had hemoglobin A1c > 9.0% during the measurement period.
	Flipped Plain Language	Controlling Hemoglobin A1c in Patients with Diabetes	More stars are better because it means that more patients with diabetes have controlled blood sugar (blood glucose) levels. Diabetes is characterized by high blood glucose levels. Lower levels of hemoglobin A1c in patients with diabetes can reduce complications in your blood vessels and nerves. To give this health care professional a score, Medicare looked at the percentage of this health care professional's patients with diabetes whose recent hemoglobin A1c level was at a controlled level (less than or equal to 9 percent).
PQRS 331	Technical	Adult Sinusitis: Antibiotic Prescribed for Acute Sinusitis (Appropriate Use)	Percentage of patients, aged 18 years and older, with a diagnosis of acute sinusitis who were prescribed an antibiotic within 7 days of diagnosis or within 10 days after onset of symptoms.
	Flipped Plain Language	Appropriate Antibiotic Use for Acute Sinusitis	More stars are better because it means that fewer patients were got antibiotics for acute sinusitis. Overtreatment of acute sinusitis (inflamed and swollen sinuses) is common. Treatment with antibiotics may increase patient harm and can lead to antibiotic resistance. To give this health care professional a score, Medicare looked at the percentage of this health care professional's patients who got antibiotics within 7 days of acute sinusitis diagnosis or within 10 days after onset of symptoms.

4.2.1 TEP Input on User Understanding of Inverse Measures

During the discussion about inverse measures, one TEP member offered a potential drawback to "flipping" certain inverse measures. He presented an exaggeratory example: if the measure were measuring mortality after sinusitis, it would be important to be able to measure against zero-percent performance. Ideally, there would not be a single instance fof death from

sinusitis, so comparing it to zero is helpful. “Flipping” the measure might make this critical point less intuitive. Thus, he suggested the Physician Compare support team consider keeping the measurement inverse when it is imperative to report against zero or one hundred percent compliance. We agreed that this is a valid concern and will keep it in mind as we move forward. Another TEP member expressed agreement with the planned approach to inverse measures: adjusting the measure language to align with a “flipped” score and testing it with users to gauge understanding. Recommendations for the Live Site Profile Pages

Aside from concerns with user interpretability, we recommend against including a measure on profile pages when:

- The measure does not pertain to Physician Compare’s users (i.e. people with Medicare);
- The measure had multiple strata (i.e. measures with multiple performance rates) and we selected one stratum over others for live site publication;
- The measure had fewer than ten reporters.

Table 4.2 details the number of measures lost to each exclusion criterion for both the downloadable and live site measures sets, and the particular measures recommended for live site are detailed in Appendix A.

Table 4.2. The number of measures lost to each exclusion criterion

Reason for Exclusion	Number of Measures	
	PQRS	Non-PQRS
Measures Excluded from Downloadable and Live Site	**140**	**264**
No Reporters	20	182
Insufficient Sample for Reliability	6	10
Failed Reliability	28	72
Failed Case-Mix	2	0
EHR Only	36	0
Measures Group Only	48	0
Measures Excluded from Live Site Only	**23**	**0**
Population Off-Target	5	0
Recommend Other Stratum	5	0
Under 10 Reporters	10	0
Concerns with User Interpretability:		
Resource Use	1	0
Inverse (and Resource Use)	2	0

5 CAHPS FOR PQRS MEASURES RECOMMENDATIONS

In addition to the available PQRS measures, 12 CAHPS for PQRS summary survey measures are also available for public reporting per the 2015 PFS final rule (79 FR 67761-67766). The available measures and display recommendations are summarized in Table 6.1.

Of these 12 summary survey measures, Health Status/Functional Status was not reviewed for inclusion on the site because this measure is a context question in the survey and not a stand-alone set of items. Of the remaining 11 summary measures, three were assessed by RAND to have reliability too low for public reporting. These included the Access to Specialists, Shared Decision Making, and Helping You to Take Medication as Directed summary survey measures. Previous user testing completed by the Physician Compare support team also showed that both Access to Specialists and Shared Decision Making were consistently misinterpreted by users. As a result, the Physician Compare support team recommends not reporting these three summary survey measures and publicly reporting the remaining eight summary survey measures for group practices with 2015 CAHPS for PQRS data.

Table 5.1: Recommendations for CAHPS measures

CAHPS Measure	Recommendation for Program Year 2015
Getting timely care, appointments, and information	Yes
How well providers Communicate	Yes
Patient's Rating of Provider	Yes
Access to Specialists	No
Health Promotion & Education	Yes
Shared Decision Making	No
Health Status/Functional Status	No
Courteous and Helpful Office Staff	Yes
Care Coordination	Yes
Between Visit Communication	Yes
Helping You to Take Medication as Directed	No
Stewardship of Patient Resources	Yes

5.1 TEP Input on CAHPS Measures

A TEP member asked about moving CAHPS for PQRS to the individual EP level. We noted that one limitation is that care is coordinated in such a manner that only a subset of these CAHPS for PQRS measures would be applicable to each specific EP. The team explained that for the upcoming reporting year (2016), CAHPS for PQRS measures will only be available from groups. However, this is an important point of discussion. As the conversation advances, there is a contingent of stakeholders interested in looking at the opportunity of taking CAHPS for PQRS measures to the individual level, but there are concerns regarding sample size and that some of

the currently available measures relate more to the environment or team than the individual EP's performance.

Another TEP member asked if there is a push for rolling up the CAHPS for PQRS measure domains into a higher level of summary than they are currently. We responded by indicating that we are looking to AHRQ for their guidance about which summary level seems appropriate.

A third TEP member emphasized the importance of the Shared Decision Making measure and encouraged the continued work being done to make the measure more understandable for possible future publication.

6 LIVE SITE MEASURE AND EP REPRESENTATION

Users will make more edified health care decisions from the measures published on Physician Compare when there are sufficient measures available to represent as many EPs as possible. However, this sentiment must be balanced with caution about burdening Physician Compare users with too much information, as users who are given an excess of information tend to make worse decisions. To address this concern, we determined the distribution of the number of measures that will be included on live site profile pages. As shown in Table 6.1, on average, with the recommended set of measures, group practice profile pages will have four measures, and individual EP profile pages will have 2.5 measures. This number of measures should not threaten to overwhelm a user.

Table 6.1: The distribution of the number of measures per group practice and individual EP profile pages

Reporting Unit	Number of Reporters	Number of Measures/Reporter									
		Mean	Std	Min	25th	50th	75th	90th	95th	99th	Max
Group Practice	2,696	3.9	3.6	1	2	3	5	8	14	14	29
Individual EP	222,477	2.5	2.1	1	1	2	3	6	7	10	18

The recommended live site measures set also represents a considerable proportion of the EPs belonging to each specialty. To determine within-specialty representation, we calculated the proportion of EPs within each specialty with at least one measure for live site publication out of the total number of EPs belonging to the given specialty who submitted any data for Physician Compare. All specialties except pathology (57% representation) had at least 70% of available EPs with one or more measures recommended for publication. Table 6.2 shows the representation of EPs within the most searched specialties (as determined from past Google Analytics results for the Physician Compare website).

In conclusion, the recommended live site measures list both avoids presenting too many measures on one profile page accessed by users and provides substantial representation of the EPs who submitted measures data across all specialties.

Table 6.2: Within-specialty representation for the recommended live site measures set

Top Searched Specialties*	% All EPs Submitting Data
Family Practice/General Practice/Internal Medicine	81%
Psychiatry	86%
Orthopedic Surgery	81%
Clinical Psychologist	89%
Gastroenterology	95%
Dermatology	93%
Obstetrics/Gynecology	86%
Neurology	89%
Pain Management	87%
Ophthalmology	92%

*Results from Google Analytics (July 2016)

APPENDIX A

A.1 2015 PQRS Measures Recommendations for Public Reporting

Table A.1 lists the 2015 PQRS measures recommended for publication on Physician Compare and designates whether the measures are recommended for the downloadable database (D) only or for live site profile pages (L) for Group Practices and/or Individual EPs.

Table A.1: Recommended PQRS measures

Measure	Stratum	Measure Title	Downloadable / Live Site Recommendation	Reporting Unit (Group or Individual EP)
5	1	Heart Failure (HF): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy for Left Ventricular Systolic Dysfunction (LVSD)	D and L	Group Practice
6	1	Coronary Artery Disease (CAD): Antiplatelet Therapy	D and L	Group Practice
7	2	Coronary Artery Disease (CAD): Beta-Blocker Therapy – Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF < 40%)	D and L	Group Practice
7	1	Coronary Artery Disease (CAD): Beta-Blocker Therapy – Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF < 40%)	D and L	Group Practice
8	1	Heart Failure (HF): Beta-Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD)	D and L	Group Practice
12	1	Primary Open-Angle Glaucoma (POAG): Optic Nerve Evaluation:	D and L	Both
14	1	Age-Related Macular Degeneration (AMD): Dilated Macular Examination	D and L	Both
19	1	Diabetic Retinopathy: Communication with the Physician Managing Ongoing Diabetes Care	D and L	Both
21	1	Perioperative Care: Selection of Prophylactic Antibiotic – First OR Second Generation Cephalosporin	D and L	Both
22	1	Perioperative Care: Discontinuation of Prophylactic Parenteral Antibiotics (Non-Cardiac Procedures)	D and L	Both
23	1	Perioperative Care: Venous Thromboembolism (VTE) Prophylaxis (When Indicated in ALL Patients)	D and L	Both
24	1	Osteoporosis: Communication with the Physician Managing On-going Care Post-Fracture of Hip, Spine or Distal Radius for Men and Women Aged 50 Years and Older	D and L	Both
32	1	Stroke and Stroke Rehabilitation: Discharged on Antithrombotic Therapy	D and L	Both
33	1	Stroke and Stroke Rehabilitation: Anticoagulant Therapy Prescribed for Atrial Fibrillation (AF) at Discharge	D and L	Both
39	1	Screening or Therapy for Osteoporosis for Women Aged 65 Years and Older	D and L	Both

Measure	Stratum	Measure Title	Downloadable / Live Site Recommendation	Reporting Unit (Group or Individual EP)
40	1	Osteoporosis: Management Following Fracture of Hip, Spine or Distal Radius for Men and Women Aged 50 Years and Older	D and L	Both
41	1	Osteoporosis: Pharmacologic Therapy for Men and Women Aged 50 Years and Older	D and L	Group Practice
44	1	Coronary Artery Bypass Graft (CABG): Preoperative Beta-Blocker in Patients with Isolated CABG Surgery	D and L	Both
46	1	Medication Reconciliation	D only	Both
46	2	Medication Reconciliation	D and L	Both
47	1	Care Plan	D and L	Both
48	1	Urinary Incontinence: Assessment of Presence or Absence of Urinary Incontinence in Women Aged 65 Years and Older	D and L	Both
50	1	Urinary Incontinence: Plan of Care for Urinary Incontinence in Women Aged 65 Years and Older	D and L	Both
51	1	Chronic Obstructive Pulmonary Disease (COPD): Spirometry Evaluation	D and L	Both
52	1	Chronic Obstructive Pulmonary Disease (COPD): Inhaled Bronchodilator Therapy	D and L	Both
53	2	Asthma: Pharmacologic Therapy for Persistent Asthma - Ambulatory Care Setting	D only	Group Practice
53	1	Asthma: Pharmacologic Therapy for Persistent Asthma - Ambulatory Care Setting	D only	Group Practice
53	3	Asthma: Pharmacologic Therapy for Persistent Asthma - Ambulatory Care Setting	D and L	Group Practice
54	1	Emergency Medicine: 12-Lead Electrocardiogram (ECG) Performed for Non-Traumatic Chest Pain	D and L	Both
65	1	Appropriate Treatment for Children with Upper Respiratory Infection (URI)	D only	Group Practice
66	1	Appropriate Testing for Children with Pharyngitis	D only	Individual EP
67	1	Hematology: Myelodysplastic Syndrome (MDS) and Acute Leukemias: Baseline Cytogenetic Testing Performed on Bone Marrow	D only	Group Practice
68	1	Hematology: Myelodysplastic Syndrome (MDS): Documentation of Iron Stores in Patients Receiving Erythropoietin Therapy	D only	Group Practice
69	1	Hematology: Multiple Myeloma: Treatment with Bisphosphonates	D and L	Group Practice
70	1	Hematology: Chronic Lymphocytic Leukemia (CLL): Baseline Flow Cytometry	D and L	Group Practice
71	1	Breast Cancer: Hormonal Therapy for Stage IC - IIIC Estrogen Receptor/Progesterone Receptor (ER/PR) Positive Breast Cancer	D and L	Both
72	1	Colon Cancer: Chemotherapy for AJCC Stage III Colon Cancer Patients	D and L	Group Practice
76	1	Prevention of Central Venous Catheter (CVC)-Related Bloodstream Infections	D and L	Both

Measure	Stratum	Measure Title	Downloadable / Live Site Recommendation	Reporting Unit (Group or Individual EP)
82	1	Adult Kidney Disease: Peritoneal Dialysis Adequacy: Solute	D only	Group Practice
91	1	Acute Otitis Externa (AOE): Topical Therapy	D and L	Individual EP
93	1	Acute Otitis Externa (AOE): Systemic Antimicrobial Therapy – Avoidance of Inappropriate Use	D and L	Both
100	1	Colorectal Cancer Resection Pathology Reporting: pT Category (Primary Tumor) and pN Category (Regional Lymph Nodes) with Histologic Grade	D only	Both
102	1	Prostate Cancer: Avoidance of Overuse of Bone Scan for Staging Low Risk Prostate Cancer Patients	D and L	Both
109	1	Osteoarthritis (OA): Function and Pain Assessment	D and L	Both
110	1	Preventive Care and Screening: Influenza Immunization	D and L	Both
111	1	Pneumonia Vaccination Status for Older Adults	D and L	Both
112	1	Breast Cancer Screening	D and L	Both
113	1	Colorectal Cancer Screening	D and L	Both
116	1	Antibiotic Treatment for Adults with Acute Bronchitis: Avoidance of Inappropriate Use	D only	Both
117	1	Diabetes: Eye Exam	D and L	Both
118	1	Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy -- Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%)	D and L	Group Practice
119	1	Diabetes: Medical Attention for Nephropathy	D and L	Both
121	1	Adult Kidney Disease: Laboratory Testing (Lipid Profile)	D and L	Both
122	2	Adult Kidney Disease: Blood Pressure Management	D only	Individual EP
122	1	Adult Kidney Disease: Blood Pressure Management	D only	Both
122	3	Adult Kidney Disease: Blood Pressure Management	D and L	Both
126	1	Diabetes Mellitus: Diabetic Foot and Ankle Care, Peripheral Neuropathy – Neurological Evaluation	D and L	Both
127	1	Diabetes Mellitus: Diabetic Foot and Ankle Care, Ulcer Prevention – Evaluation of Footwear	D and L	Both
128	1	Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow-Up Plan	D and L	Both
130	1	Documentation of Current Medications in the Medical Record	D and L	Both
131	1	Pain Assessment and Follow-Up	D and L	Both
134	1	Preventive Care and Screening: Screening for Clinical Depression and Follow-Up Plan	D and L	Both
137	1	Melanoma: Continuity of Care – Recall System	D and L	Both
138	1	Melanoma: Coordination of Care	D and L	Both
140	1	Age-Related Macular Degeneration (AMD): Counseling on Antioxidant Supplement	D and L	Both
141	1	Primary Open-Angle Glaucoma (POAG): Reduction of Intraocular Pressure (IOP) by 15% OR Documentation of a Plan of Care	D and L	Both

Measure	Stratum	Measure Title	Downloadable / Live Site Recommendation	Reporting Unit (Group or Individual EP)
143	1	Oncology: Medical and Radiation – Pain Intensity Quantified	D and L	Both
144	1	Oncology: Medical and Radiation – Plan of Care for Pain	D and L	Both
145	1	Radiology: Exposure Time Reported for Procedures Using Fluoroscopy	D and L	Both
146	1	Radiology: Inappropriate Use of "Probably Benign" Assessment Category in Mammography Screening	D and L	Both
147	1	Nuclear Medicine: Correlation with Existing Imaging Studies for All Patients Undergoing Bone Scintigraphy	D and L	Both
154	1	Falls: Risk Assessment	D and L	Both
155	1	Falls: Plan of Care	D and L	Both
156	1	Oncology: Radiation Dose Limits to Normal Tissues	D and L	Both
163	1	Diabetes: Foot Exam	D and L	Both
172	1	Hemodialysis Vascular Access Decision-Making by Surgeon to Maximize Placement of Autogenous Arterial Venous (AV) Fistula	D and L	Individual EP
173	1	Preventive Care and Screening: Unhealthy Alcohol Use – Screening	D and L	Both
178	1	Rheumatoid Arthritis (RA): Functional Status Assessment	D and L	Both
181	1	Elder Maltreatment Screen and Follow-Up Plan	D and L	Both
182	1	Functional Outcome Assessment	D and L	Both
185	1	Colonoscopy Interval for Patients with a History of Adenomatous Polyps – Avoidance of Inappropriate Use	D and L	Both
187	1	Stroke and Stroke Rehabilitation: Thrombolytic Therapy	D and L	Individual EP
191	1	Cataracts: 20/40 or Better Visual Acuity within 90 Days Following Cataract Surgery	D and L	Both
193	1	Perioperative Temperature Management	D and L	Both
194	1	Oncology: Cancer Stage Documented	D and L	Both
195	1	Radiology: Stenosis Measurement in Carotid Imaging Reports	D and L	Both
204	1	Ischemic Vascular Disease (IVD): Use of Aspirin or Another Antithrombotic	D and L	Group Practice
205	1	HIV/AIDS: Sexually Transmitted Disease Screening for Chlamydia, Gonorrhea, and Syphilis	D only	Group Practice
217	1	Functional Deficit: Change in Risk-Adjusted Functional Status for Patients with Knee Impairments	D only	Group Practice
218	1	Functional Deficit: Change in Risk-Adjusted Functional Status for Patients with Hip Impairments	D and L	Individual EP
219	1	Functional Deficit: Change in Risk-Adjusted Functional Status for Patients with Lower Leg, Foot or Ankle Impairments	D and L	Individual EP

Measure	Stratum	Measure Title	Downloadable / Live Site Recommendation	Reporting Unit (Group or Individual EP)
220	1	Functional Deficit: Change in Risk-Adjusted Functional Status for Patients with Lumbar Spine Impairments	D Only	Group Practice
221	1	Functional Deficit: Change in Risk-Adjusted Functional Status for Patients with Shoulder Impairments	D Only	Group Practice
223	1	Functional Deficit: Change in Risk-Adjusted Functional Status for Patients with Neck, Cranium, Mandible, Thoracic Spine, Ribs, or Other General Orthopedic Impairments	D and L	Individual EP
224	1	Melanoma: Overutilization of Imaging Studies in Melanoma	D only	Both
225	1	Radiology: Reminder System for Screening Mammograms	D and L	Both
226	1	Preventive Care and Screening: Tobacco Use: Screening and Cessation Intervention	D and L	Both
238	2	Use of High-Risk Medications in the Elderly	D and L	Group Practice
238	1	Use of High-Risk Medications in the Elderly	D only	Both
249	1	Barrett's Esophagus	D and L	Individual EP
250	1	Radical Prostatectomy Pathology Reporting	D only	Both
261	1	Referral for Otologic Evaluation for Patients with Acute or Chronic Dizziness	D and L	Both
265	1	Biopsy Follow-Up	D and L	Both
303	1	Cataracts: Improvement in Patient's Visual Function within 90 Days Following Cataract Surgery	D and L	Individual EP
304	1	Cataracts: Patient Satisfaction within 90 Days Following Cataract Surgery	D and L	Individual EP
317	1	Preventive Care and Screening: Screening for High Blood Pressure and Follow-Up Documented	D and L	Both
318	1	Falls: Screening for Fall Risk	D and L	Group Practice
320	1	Appropriate Follow-Up Interval for Normal Colonoscopy in Average Risk Patients	D and L	Both
324	1	Cardiac Stress Imaging Not Meeting Appropriate Use Criteria: Testing in Asymptomatic, Low-Risk Patients	D only	Group Practice
325	1	Adult Major Depressive Disorder (MDD): Coordination of Care of Patients with Specific Comorbid Conditions	D and L	Individual EP
326	1	Atrial Fibrillation and Atrial Flutter: Chronic Anticoagulation Therapy	D and L	Both
331	1	Adult Sinusitis: Antibiotic Prescribed for Acute Sinusitis (Appropriate Use)	D and L	Both
332	1	Adult Sinusitis: Appropriate Choice of Antibiotic: Amoxicillin Prescribed for Patients with Acute Bacterial Sinusitis (Appropriate Use)	D and L	Group Practice
333	1	Adult Sinusitis: Computerized Tomography for Acute Sinusitis (Overuse)	D and L	Individual EP
334	1	Adult Sinusitis: More than One Computerized Tomography (CT) Scan Within 90 Days for Chronic Sinusitis (Overuse)	D only	Both

Measure	Stratum	Measure Title	Downloadable / Live Site Recommendation	Reporting Unit (Group or Individual EP)
337	1	Tuberculosis Prevention for Psoriasis,Psoriatic Arthritis and Rheumatoid Arthritis Patients on a Biological Immune Response Modifier	D and L	Both
343	1	Screening Colonoscopy Adenoma Detection Rate Measure	D and L	Individual EP
346	1	Rate of Postoperative Stroke or Death in Asymptomatic Patients undergoing Carotid Endarterectomy (CEA)	D and L	Individual EP
358	1	Patient-Centered Surgical Risk Assessment and Communication	D and L	Both
383	1	Adherence to Antipsychotic Medications for Individuals with Schizophrenia	D and L	Both
386	1	Amyotrophic Lateral Sclerosis (ALS) Patient Care Preferences	D only	Group Practice
389	1	Cataract Surgery: Difference Between Planned and Final Refraction	D and L	Individual EP
390	1	Discussion and Shared Decision Making Surrounding Treatment Options	D and L	Individual EP
391	2	Follow-up After Hospitalization for Mental Illness (FUH)	D only (groups) D and L (EPs)	Both
391	1	Follow-up After Hospitalization for Mental Illness (FUH)	D only (groups) D and L (EPs)	Both
395	1	Lung Cancer Reporting (Biopsy/Cytology Specimens)	D and L	Group Practice
397	1	Melanoma Reporting	D and L	Both
400	1	Hepatitis C: One-Time Screening for Hepatitis C Virus (HCV) for Patients at Risk	D and L	Both
402	1	Tobacco Use and Help with Quitting Among Adolescents	D only	Group Practice

A.2 2015 Non-PQRS Measures Recommendations for Public Reporting

Table A.2 lists the 2015 Non-PQRS measures recommended for publication on Physician Compare and designates whether the measures are recommended for the downloadable database (D) only or for live site profile pages (L).

Table A.2: Recommended Non-PQRS measures (reported only for Individual EPs)

Measure	Stratum	QCDR Name	Measure Title	Downloadable / Live Site Recommendation
AAAAI6	1	AAAAI Allergy, Asthma, & Immunology QCDR (CECity)	Documentation of Clinical Response to Allergen Immunotherapy within One Year	D and L
AAAAI8	1	AAAAI Allergy, Asthma, & Immunology QCDR (CECity)	Achievement of Projected Effective Dose of Standardized Allergens for Patient Treated With Allergen Immunotherapy for at Least One Year	D and L
AAAAI11	1	AAAAI Allergy, Asthma, & Immunology QCDR (CECity)	Asthma Assessment and Classification	D and L
AAAAI12	1	AAAAI Allergy, Asthma, & Immunology QCDR (CECity)	Lung Function/Spirometry Evaluation	D and L
AAAAI14	1	AAAAI Allergy, Asthma, & Immunology QCDR (CECity)	Patient Self-Management and Action Plan	D and L
AAAAI15	1	AAAAI Allergy, Asthma, & Immunology QCDR (CECity)	Body Mass Index	D and L
AAAAI16	1	AAAAI Allergy, Asthma, & Immunology QCDR (CECity)	Optimal Asthma Control	D and L
ACCPin1	1	American College of Cardiology - PINNACLE Registry and Diabetes Collaborative Registry	Hypertension (HTN): Blood Pressure (BP) Management	D and QCDR Site
ACR1	1	American College of Rheumatology, Rheumatology Informatics System for Effectiveness Registry	Disease Activity Measurement for Patients with Rheumatoid Arthritis (RA)	D and QCDR Site
ACR2	1	American College of Rheumatology, Rheumatology Informatics System for Effectiveness Registry	Functional Status Assessment for Patients with Rheumatoid Arthritis (RA)	D and QCDR Site
ACR4	1	American College of Rheumatology, Rheumatology Informatics System for Effectiveness Registry	Tuberculosis Test Prior to First Course Biologic Therapy	D and QCDR Site
ACR5	1	American College of Rheumatology, Rheumatology Informatics System for Effectiveness Registry	Glucocorticosteroids and Other Secondary Causes	D and QCDR Site
ACRad5	1	National Radiology Data Registry NRDR	Screening Mammography Abnormal Interpretation Rate (Recall Rate)	D and L
ACRad14	1	National Radiology Data Registry NRDR	Participation in a National Dose Index Registry	D and L

Measure	Stratum	QCDR Name	Measure Title	Downloadable / Live Site Recommendation
ASBS1	1	The American Society of Breast Surgeons Mastery of Breast Surgery Program	Surgeon assessment for hereditary cause of breast cancer	D and QCDR Site
ASBS2	1	The American Society of Breast Surgeons Mastery of Breast Surgery Program	Surgical Site Infection and Cellulitis After Breast and/or Axillary Surgery	D and QCDR Site
ASBS5	1	The American Society of Breast Surgeons Mastery of Breast Surgery Program	Perioperative Care: Selection of Prophylactic Antibiotic - First OR Second Generation Cephalosporin	D and QCDR Site
ASBS6	1	The American Society of Breast Surgeons Mastery of Breast Surgery Program	Perioperative Care: Discontinuation of Prophylactic Parenteral Antibiotics (Non-Cardiac Procedures)	D and QCDR Site
CDR1	1	US Wound Registry	Adequate Off-loading of Diabetic Foot Ulcers at each visit	D and QCDR Site
CDR4	1	US Wound Registry	Diabetic Foot & Ankle Care: Comprehensive Diabetic Foot Examination	D and QCDR Site
CDR5	1	US Wound Registry	Adequate Compression at each visit for Patients with Venous Leg Ulcers (VLU)	D and QCDR Site
CDR10	1	US Wound Registry	Vascular Assessment of patients with chronic leg ulcers	D and QCDR Site
CDR11	1	US Wound Registry	Wound Bed Preparation Through Debridement of Necrotic or Non-viable Tissue	D and QCDR Site
GIQIC2	1	GI Quality Improvement Consortium	Adequacy of bowel preparation	D and L
GIQIC3	1	GI Quality Improvement Consortium	Photodocumentation of the cecum (also known as cecal intubation rate) – All Colonoscopies	D and L
GIQIC4	1	GI Quality Improvement Consortium	Photodocumentation of the cecum (also known as cecal intubation rate) – Screening Colonoscopies	D and L
GIQIC9	1	GI Quality Improvement Consortium	Documentation of history and physical rate - Colonoscopy	D and L
GIQIC10	1	GI Quality Improvement Consortium	Appropriate management of anticoagulation in the peri-procedural period rate – EGD	D and L
GIQIC11	1	GI Quality Improvement Consortium	Helicobacter pylori (H. pylori) status rate	D and L
GIQIC12	1	GI Quality Improvement Consortium	Appropriate indication for colonoscopy	D and L
OBERD6	1	OBERD QCDR	Orthopedic Pain: Mental Health Assessment	D and QCDR Site