



**ACUMEN**

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

**November 2015**

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

[This page is intentionally left blank.]

## TABLE OF CONTENTS

<b>1</b>	<b>About the TEP</b> .....	<b>1</b>
<b>2</b>	<b>Overview of 2014 Public Reporting Plan</b> .....	<b>3</b>
2.1	Mechanism-Specific Considerations .....	3
2.1.1	Web Interface.....	4
2.1.2	Claims-Based Reporting .....	4
2.1.3	Registry.....	4
2.1.4	Electronic Health Record.....	4
2.1.5	TEP Input on Mechanism-Specific Considerations.....	4
2.2	Measures Available for Public Reporting.....	5
<b>3</b>	<b>Issues to Consider: Group Practices</b> .....	<b>7</b>
3.1	Measure Performance Rate Reliability .....	7
3.1.1	Reliability Results.....	7
3.1.2	TEP Input on Reliability .....	8
3.2	Comparability of Performance Rates across Reporting Mechanisms .....	8
3.2.1	TEP Input on the Comparability of Performance Rates .....	9
3.3	Impact of Case-Mix on Performance Rates.....	9
3.3.1	TEP Input on the Impact of Case-Mix.....	10
<b>4</b>	<b>Issues to Consider: Individual Eligible Professionals</b> .....	<b>11</b>
4.1	Measure Performance Rate Reliability Results .....	11
4.1.1	TEP Input on Cardiovascular Measures Group Reliability .....	11
4.2	Comparability of Performance Rates across Reporting Mechanisms .....	11
4.3	Measures Group Selection Bias .....	12
4.3.1	TEP Input of Measures Group Selection Bias .....	12
4.4	Impact of Case-Mix on Performance Rates.....	12
4.4.1	TEP Input on Case-Mix .....	12
<b>5</b>	<b>CAHPS for pQRS Results</b> .....	<b>13</b>
5.1	TEP Input on CAHPS .....	13
<b>6</b>	<b>Physician Compare Moving Forward</b> .....	<b>14</b>
6.1	TEP Input on Physician Compare Moving Forward .....	14
<b>7</b>	<b>Measures Recommended for Public Reporting: PQRS 2014</b> .....	<b>15</b>

## LIST OF TABLES AND FIGURES

Table 1.1:	TEP Members .....	1
Table 2.1:	Candidate PQRS Measures Available for Public Reporting .....	5
Table 5.1:	Candidate CAHPS for PQRS Measures Available for Public Reporting.....	13
Table 8.1:	Measures Recommended for Public Reporting .....	15

## 1 ABOUT THE TEP

---

Physician Compare serves as the public reporting home of quality initiative and related measurement data for individual health care practitioners and group practices. The Physician Compare website operates under the mandate of the Affordable Care Act (ACA). Section 10331 of the ACA stipulates that all measures data published to Physician Compare must be comparable, valid, reliable, and accurate. Existing regulation requires that the data displayed on the website resonates with and is accurately interpreted by consumers. Collectively, these conditions comprise the public reporting standards, which must be met for all data presented on Physician Compare. To uphold these standards, CMS has contracted Acumen, LLC (henceforth “Acumen”) to assist in the selection of physician quality measures for public reporting by assessing the comparability, validity, reliability, and accuracy of the data.

To ensure that the data reported on Physician Compare portray physician performance accurately and robustly, Acumen 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. Acumen and the TEP convened on August 20, 2015 to discuss the selection of PQRS 2014 candidate quality measures to publish on Physician Compare in late 2015 and future considerations for publicly reporting measures data in the context of new legislation and an expanded set of available candidate measures for PQRS 2015. Table 1.1 lists the 12 individuals who comprise the TEP, nine 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 Muhlbauer, 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* <b>(TEP Chair)</b>	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 2014 Public Reporting Plan and reporting mechanisms available to practitioners. Sections 3, 4 and 5 summarize results and feedback from the TEP for group practices, individual eligible professionals (EPs), and Consumer Assessment of Healthcare Providers and Systems (CAHPS) for PQRS, respectively. Section 6 discusses Physician Compare moving forward. Finally, Section 7 presents Acumen's recommendations for public reporting.

## **2 OVERVIEW OF 2014 PUBLIC REPORTING PLAN**

---

Per the 2012 (76 FR 73025) and 2013 (77 FR 68891) Medicare Physician Fee Schedule (PFS) final rules, CMS publicly reported a select set of Physician Quality Reporting System (PQRS) Group Practice Reporting Option (GPRO) quality measures (i.e., Diabetes Mellitus and Coronary Artery Disease measures) that were collected via the Web Interface. The 2012 PQRS GPRO measures were publicly reported on Physician Compare in February 2014 and the 2013 PQRS GPRO measures were publicly reported on Physician Compare in December 2014.

With the 2014 PFS final rule (78 FR 74229), CMS expanded the measures available for public reporting from six Diabetes Mellitus (DM) measures and two Coronary Artery Disease (CAD) measures to all group-level measures under the 2014 PQRS collected through the Web Interface and a subset of quality measures that group practices report via registries and EHRs. Group practices participating under the 2014 PQRS are identified through their Tax Identification Number (TIN) and consist of two or more individual EPs. Group Practices could report 2014 PQRS quality data through the following mechanisms: (i) electronic health records (EHR) using certified EHR technology (CEHRT), (ii) qualified registry, and (iii) Web Interface for groups of 25 EPs or more. Group practices can supplement these submissions with the Consumer Assessment of Healthcare Providers and Systems (CAHPS) for PQRS, which are surveys that measure patients' experience with their care at group practices. As CAHPS measures are distinct from the PQRS measures submitted via EHR, registry, and the Web Interface, they are discussed separately in Section 5.

Furthermore, the 2014 PFS final rule announced CMS's decision to publicly report measures for individual EPs that are in line with measures collected through the Web Interface. Individual EPs are identified by their unique TIN/National Provider Identifier (NPI) and could submit quality data through: (i) Medicare Part B claims, (ii) qualified registry, and (iii) EHR using CEHRT. Individuals reporting via registry can participate in a measures group, which combines a set of related measures. For PQRS 2014, registry data submitted as part of the Cardiovascular Prevention Measures Group are available for public reporting. The remainder of this section will cover the considerations surrounding the reporting mechanisms and the candidate measures available for public reporting.

### **2.1 Mechanism-Specific Considerations**

This section outlines the material presented to the TEP regarding the operational parameters and specifications for satisfactory reporting that must be considered when evaluating each reporting mechanism.

### **2.1.1 Web Interface**

Only group practices of 25 or more EPs can submit measures through the Web Interface. This reporting mechanism is unique in that CMS selects a random sample of Medicare beneficiaries on whom groups are to report patient-level data. Additionally, when group practices elect to report via the Web Interface, they are agreeing to report data on a minimum number of eligible patients for each of the 22 available measures.

### **2.1.2 Claims-Based Reporting**

The claims reporting option is available to individual EPs. When submitting Medicare Part B fee-for-service claims for reimbursement, EPs include the Current Procedural Terminology (CPT) code corresponding to the intended measure for eligible beneficiaries. The data for these beneficiaries are then submitted to PQRS. For CMS to consider EPs satisfactory claims reporters, they must report data for at least 50% of their measure-eligible beneficiaries.

### **2.1.3 Registry**

When opting for the registry submission method, group practices and individual EPs contract with a qualified PQRS registry that accesses their health records to report measures data to CMS. Registries identify beneficiaries eligible for PQRS measures and report the proportion of eligible beneficiaries who met the measures. Group practices can only submit individual items while EPs can report items individually or as a measures group. To report individual items satisfactorily, groups and individual EPs must report on at least half of their eligible patients. Individual EPs can also choose to submit quality data via registry as part of a measures group. Each measures group contains roughly six related measures. Measures group reporters must submit data on at least 20 beneficiaries who qualify for each measure in the measures group to be considered a satisfactory reporter. For 2014 PQRS data, only data from the Cardiovascular Prevention Measures Group are available for public reporting. The measures comprising the Cardiovascular Prevention Measures Group are: Use of Aspirin/Another Antithrombotic (PQRS 204), Tobacco Use Screening/Cessation Intervention (PQRS 226), Controlling High Blood Pressure (PQRS 236), and High Blood Pressure Screening (PQRS 317).

### **2.1.4 Electronic Health Record**

Acumen did not present data on the EHR reporting option to the TEP because CMS could not determine the accuracy of the EHR data.

### **2.1.5 TEP Input on Mechanism-Specific Considerations**

The following bullet points provide comments from some TEP members regarding the specifications of individual reporting mechanisms:

- One TEP member wondered if requiring practitioners to report measures on only half of their eligible patients could allow them only to report on patients who would positively contribute to their

performance rates. To address this concern, Acumen investigated if practitioners' performance rates were related to the number of beneficiaries on whom they submitted data. Most measures showed no relationship between the number of beneficiaries reported and performance rate. The exceptions were PQRS 110-113, which, as described in Section 8, were recommended against for public reporting.

- Another member noticed that the number of patients required for satisfactory measures group reporting (20 beneficiaries) is quite low and mentioned that selection bias in this context could be especially pronounced. This issue and other related TEP comments are described more thoroughly in Sections 4.2 and 4.3.

## 2.2 Measures Available for Public Reporting

Table 2.1 summarizes candidate measures that have been evaluated for public reporting along with the reporting mechanisms through which the measures are collected.<sup>1</sup> The first and second columns list the PQRS measure number and Web Interface number, respectively. The third column provides the title of each measure. The final column indicates the reporting mechanism through which the measures may be submitted. As previously decided by CMS and past TEPs, lipid control measures are inconsistent with clinical guidelines and therefore have been excluded from the analyses.

**Table 2.1: Candidate PQRS Measures Available for Public Reporting**

PQRS Measure Number	Web Interface Number	Measure Title	Reporting Mechanism					
			Individual EPs			Group Practice		
			EHR	Registry	Claims	EHR	Registry	Web Interface
46	CARE-1	Medication Reconciliation		X	X		X	X
318	CARE-2	Falls: Screening for Future Fall Risk	X					X
197	CAD-2	Coronary Artery Disease (CAD): Lipid Control*		X			X	X
118	CAD-7	Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%)		X			X	X
1	DM-2	Diabetes: Hemoglobin A1c Poor Control	X	X	X	X	X	X
319	DM-13	Diabetes Composite (All or Nothing Scoring): Diabetes Mellitus: High Blood Pressure Control						X
319	DM-14	Diabetes Composite (All or Nothing Scoring): Diabetes Mellitus: Low Density Lipoprotein (LDL-C) Control*						X
319	DM-15	Diabetes Composite (All or Nothing Scoring): Diabetes Mellitus: Hemoglobin A1c Control (< 8%)						X
319	DM-16	Diabetes Composite (All or Nothing Scoring): Diabetes Mellitus: Daily Aspirin or Antiplatelet Medication Use for Patients with Diabetes and Ischemic Vascular Disease						X

<sup>1</sup> The measure specifications can be downloaded from: [https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/2014\\_Physician\\_Quality\\_Reporting\\_System.html](https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/2014_Physician_Quality_Reporting_System.html)

PQRS Measure Number	Web Interface Number	Measure Title	Reporting Mechanism					
			Individual EPs			Group Practice		
			EHR	Registry	Claims	EHR	Registry	Web Interface
319	DM-17	Diabetes Composite (All or Nothing Scoring): Diabetes Mellitus: Tobacco Non-Use						X
2	-	Diabetes Mellitus: Low Density Lipoprotein (LDL-C) Control*†	X	X	X			
8	HF-6	Heart Failure (HF): Beta-Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD)	X	X		X	X	X
236	HTN-2	Controlling High Blood Pressure†	X	X	X	X	X	X
241	IVD-1	Ischemic Vascular Disease (IVD): Complete Lipid Panel and LDL Control*†	X	X	X	X	X	X
204	IVD-2	Ischemic Vascular Disease (IVD): Use of Aspirin or Another Antithrombotic†	X	X	X	X	X	X
112	PREV-5	Breast Cancer Screening	X	X	X	X	X	X
113	PREV-6	Colorectal Cancer Screening	X	X	X	X	X	X
110	PREV-7	Preventive Care and Screening: Influenza Immunization	X	X	X	X	X	X
111	PREV-8	Pneumonia Vaccination Status for Older Adults	X	X	X	X	X	X
128	PREV-9	Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow-Up†	X	X	X	X	X	X
226	PREV-10	Preventive Care and Screening: Tobacco Use: Screening and Cessation Intervention	X	X	X	X	X	X
317	PREV-11	Preventive Care and Screening: Screening for High Blood Pressure and Follow-Up Documented†	X	X	X	X	X	X
134	PREV-12	Preventive Care and Screening: Screening for Clinical Depression and Follow-Up Plan	X	X	X		X	X

*\*As previously decided by CMS and past TEPs, lipid control measures are inconsistent with clinical guidelines and therefore have been excluded from the analyses.*

*†Part of the Cardiovascular Prevention Measures Group*

### 3 ISSUES TO CONSIDER: GROUP PRACTICES

---

During the August teleconference, Acumen brought up key issues for the TEP's consideration while reviewing the analytic results for group practices. The first issue is reliability. For consumers to make meaningful comparisons across group practices, there must be sufficient variation in performance rates across different groups, and these rates must be reported with adequate precision. Second, performance rates generated from data collected across various submission methods must be comparable to assure consistency in the performance rates published to Physician Compare across reporting mechanisms. Lastly, to confirm the performance rates reflect a disparity in clinical quality between group practices rather than differences in a group practice's patient population that are outside of its control, the impact of case-mix must be considered for the four outcome measures.

#### 3.1 Measure Performance Rate Reliability

Measure reliability refers to the extent to which differences in performance rates for each quality measure are due to actual differences in group practice performance rates versus variation that arises from patient-level differences. Statistically, reliability depends on performance variation for a measure across group practices, the random variation in performance for a measure within a practitioner's panel of beneficiaries, and the number of beneficiaries attributed to the practitioner. High reliability for a measure suggests that comparisons of relative performance across group practices are likely to be stable over different performance periods and that the performance of one group practice on the quality measure can be confidently distinguished from another. Potential reliability values range from zero to one, where one (highest possible reliability) means that all variation in the measure's rates is the result of variation in differences in performance across group practices, while zero (lowest possible reliability) means that all variation is a result of patient-level differences.<sup>2</sup> Acumen calculated reliability using the beta binomial and test-retest methods for the Web Interface and registry reporting options.

##### 3.1.1 Reliability Results

Acumen concluded that reliability was high across all measures for the Web Interface reporting option; the 25th percentile ranged from 0.84 to 0.99, which is well above the range considered acceptable for drawing inferences about group practices (i.e., 0.70 – 0.80). Additionally, to measure between-group practice variation and within-group practice variation, Acumen calculated the test-retest reliability using the intra-class correlation coefficient (ICC).

---

<sup>2</sup> For more information about reliability testing for physician performance measurement, as well as the methodology for constructing the reliability score reported on Table 6, see "Reliability of Provider Profiling: A Tutorial" by John Adams, RAND. [http://www.rand.org/pubs/technical\\_reports/TR653.html](http://www.rand.org/pubs/technical_reports/TR653.html)

ICC values that approach 1 indicate that the fraction of the total variance due to between-group variation is high. The ICC values across all measures range from 0.77 to 0.99, indicating that most of the total variation is due to between-group practice variation. Similar to Web Interface results, reliability was also high across all measures for the registry reporting option; the 25th percentile ranged from 0.86 to 1.00, and the ICC values across all measures range from 0.80 to 1.0, indicating that reliability is high for the registry reporting option.

To further assess how much variation there was in the performance rates across the group practices, Acumen looked at the proportion of the group practices that had performance rates that were statistically different than the average performance rate across all group practices for any given measure. For all of the measures reported through the Web Interface and registry reporting options, at least 50% of group practices had performance rates that differed significantly from the mean.

### **3.1.2 TEP Input on Reliability**

Generally, the TEP did not have concerns about the reliability of the data. One member asked if CMS was planning on using distributional scoring (i.e., the proportion above or below an average mean in a population observed, where a set of practitioners will be at the top or bottom) and recommended using benchmark scoring. Acumen clarified that we are displaying the performance rate for a selected measure as a percent along with the graphical representations of the performance score as stars, and that while there will be no distributional or benchmark scoring for the 2014 data, this will be an option in future years per rulemaking.

## **3.2 Comparability of Performance Rates across Reporting Mechanisms**

Measures publicly reported on Physician Compare should be comparable across reporting mechanisms with respect to (i) the specifications and (ii) the impact the reporting mechanism has on performance. Steps are being taken to ensure measure specifications are appropriately aligned across reporting mechanisms. In addition, operational differences across mechanisms could culminate in incomparable performance rate distributions. To address this concern, Acumen compared the performance rate distributions between Web Interface and group registry submissions and found that Web Interface and registry rates were not comparable.

CMS cannot publish incomparable performance rates from multiple mechanisms to Physician Compare. During the TEP, Acumen presented performance rate distributions for a subset of the candidate measures, stratified by submission method, as examples of performance rate distributions that are dissimilar across mechanisms and requested TEP guidance on which mechanism would be most appropriate for public reporting.

### **3.2.1 TEP Input on the Comparability of Performance Rates**

Some TEP members had reservations about registry data and suggested that they may not meet public reporting standards.

### **3.3 Impact of Case-Mix on Performance Rates**

Case-mix adjustment refers to the statistical process of identifying and adjusting for differences in population characteristics (i.e., risk factors) before comparing outcomes of care. While case-mix adjustment is generally not applied to certain structure and process measures, it may be necessary for outcome measures that are not fully within a practitioners' control.

The DM and HTN outcome measures (i.e., DM-2: Hemoglobin A1c Poor Control, DM-13: High Blood Pressure Control, DM-15: Hemoglobin A1c Control (< 8%), and HTN-2: Controlling High Blood Pressure) do not include case-mix adjustment as part of their specifications. To determine the impact of case-mix on group practice performance rates across measures, Acumen adjusted the performance rates for certain patient characteristics that are outside the control of a group practice (e.g., demographic characteristics and pre-existing health conditions) and evaluated how group-level performance rates were affected.

To compare the impact of different sets of case-mix factors, Acumen constructed two predictive models; Model 1 included basic demographic characteristics and health status variables (i.e., age, sex, Medicare eligibility, Medicaid eligibility, and the presence of assorted health conditions) that Medicare commonly uses as part of case-mix adjustment for other publicly reported measures. However, group practice performance rates may vary systematically based on racial and regional attributes that Medicare does not typically use for case-mix adjustment; Model 2 was an expanded model that included these additional characteristics (i.e., race, region, region type, typical household income, and typical home value). Based on these models, Acumen reached the following conclusions about the impact of case-mix adjustment on the DM and HTN outcome measures:

(1) Adjusting for the demographic and clinical characteristics in Model 1 impacted the performance rates of the hemoglobin A1c measures DM-2 and DM-15 but minimally affected rates for the blood pressure measures DM-13 and HTN-2.

(2) Adjusting for the sociodemographic characteristics in Model 2 moderately influenced the performance rates of DM-13 and induced more pronounced effects on the rates of DM-2, DM-15, and HTN-2.

### **3.3.1 TEP Input on the Impact of Case-Mix**

Some TEP members were concerned about the impact of case-mix, and one member expressed concern that administrative claims data do not capture all of the factors important for case-mix adjustment.

## **4 ISSUES TO CONSIDER: INDIVIDUAL ELIGIBLE PROFESSIONALS**

---

This section summarizes Acumen's results and feedback from the TEP on reliability, comparability, and the impact of case-mix for individual eligible professionals. Detailed descriptions on methodology used to derive these results are provided in Section 3.

### **4.1 Measure Performance Rate Reliability Results**

As described in Section 3.1, measure reliability refers to the extent to which differences in performance rates for each quality measure are due to actual differences in practitioner performance versus variation that arises from patient-level differences. For the claims based reporting option, across all measures, more than 75% of individual EPs had high reliability scores ( $>0.80$ ). The ICC values across all measures are greater than 0.75 across all measures, indicating that most of the total variation is due to between-EP variation.

For the registry reporting option, reliability is high for all measures except for PQRS 118: ACEI/ARB Therapy, which had a borderline ICC of 0.70, and PQRS 8: Beta-Blocker Therapy, for which the median beta binomial score is 0.68 and the ICC is 0.57.

For the cardiovascular measures group registry option, across all measures, more than 75% of group practices had high reliability scores ( $>0.80$ ). The ICC values across all measures are greater than 0.75. However, high reliability in this reporting mechanism is in part due to many measures group reporters having 100% performance rates as opposed to variation in practitioner performance.

#### **4.1.1 TEP Input on Cardiovascular Measures Group Reliability**

A TEP member's response to lack of variation in performance rates for the Cardiovascular Prevention Measures Group was to remind the TEP that variation is an important characteristic for defining a good measure.

### **4.2 Comparability of Performance Rates across Reporting Mechanisms**

As explained in Section 3.2, measures publicly reported on Physician Compare should be comparable across reporting mechanisms with respect to (i) the specifications and (ii) the impact the reporting mechanism has on performance. Acumen examined the performance rate distributions across mechanisms to assess comparability, and, as observed for measures submitted by group practices, the distributions across mechanisms were incomparable. Acumen presented examples of dissimilar distributions across mechanisms to the TEP, including distributions showing better performance among EPs submitting measures as part of the Cardiovascular Prevention Measures Group.

### **4.3 Measures Group Selection Bias**

As covered in Section 2.1.1, measures group reporters are only required to report on 20 beneficiaries for all of the measures included in the measures group, regardless of the size of the total patient population served by the individual EP. Acumen's analyses showed that (1) EPs reporting via the Cardiovascular Prevention Measures Group (CPMG) tend to report on the minimum number of beneficiaries required (i.e. 20) and (2) in general, EPs reporting via the CPMG performed better than EPs reporting via claims or individual measures registry and achieved a higher proportion of 100% performance rates across relevant measures. These results suggest that the measures group could be more susceptible to selective reporting compared to other reporting options and that the performance rates are less likely to reflect EPs' true performance.

#### **4.3.1 TEP Input of Measures Group Selection Bias**

Multiple TEP members raised concerns about the measures group data. In addition to the comment voiced during the conference that is described in Section 2.1.5, members referred to the CPMG performance rates as being "the biggest outlier", "signaling gaming" (i.e. working a system to achieve a desired outcome), and suggesting "cherry-picking."

### **4.4 Impact of Case-Mix on Performance Rates**

Acumen investigated whether the case-mix of the patients served by individual EPs impacted their performance rates for outcome measures (PQRS 1: Hemoglobin A1c Poor Control and PQRS 236: Controlling High Blood Pressure) through adjustment with Model 1 and Model 2 (described in Section 3.3). Based on these models, Acumen reached the following conclusions about the impact of case-mix adjustment on the candidate outcome measures:

(1) Adjusting for the demographic and clinical characteristics in Model 1 impacted the performance rates for PQRS 1 moderately and the rates for PQRS 236 minimally.

(2) Adjusting for the sociodemographic characteristics in Model 2 produced similar effects on PQRS 1 and PQRS 236 performance rates as Model 1.

#### **4.4.1 TEP Input on Case-Mix**

Even though, in general, Acumen's case-mix adjustment did not appear to have a large impact on performance rates for these measures, the TEP was still concerned about the idea of posting performance rates for individual EPs without adjusting for case-mix.

## 5 CAHPS FOR PQRS RESULTS

---

In addition to the available PQRS measures, 12 CAHPS for PQRS summary survey measures are also available for public reporting per the 2014 PFS final rule. The candidate measures are summarized in Table 5.1:

**Table 5.1: Candidate CAHPS for PQRS Measures Available for Public Reporting**

CAHPS Measure Description
Getting timely care, appointments, and information
How well providers Communicate
Patient's Rating of Provider
Access to Specialists
Health Promotion & Education
Shared Decision Making
Health Status/Functional Status
Courteous and Helpful Office Staff
Care Coordination
Between Visit Communication
Helping You to Take Medication as Directed
Stewardship of Patient Resources

Of these 12, the Health Status/Functional Status measure was not reviewed for inclusion on the site because this is essentially a context question in the survey and not a stand-alone set of items. Of the remaining 11 measures, the following three were assessed by RAND to have reliability too low for public reporting: *Access to Specialists*, *Shared Decision Making*, and *Helping You to Take Medication as Directed*. Consumer testing completed by the Physician Compare support team also showed that both Access to Specialists and Shared Decision Making were consistently misinterpreted by consumers.

### 5.1 TEP Input on CAHPS

- TEP members generally thought CAHPS for PQRS data are important for public reporting.
- One TEP member thought it would be informative to report data for individual EPs.

## **6 PHYSICIAN COMPARE MOVING FORWARD**

---

The last topic discussed during the TEP was the future direction of Physician Compare spanning from the implementation of the Medicare Access and CHIP Reauthorization Act (MACRA) and the 2016 Proposed Physician Fee Schedule Rule (PFS).

Per MACRA, the Physician Quality Reporting System (PQRS), EHR Incentive, and Value-Based Modifier programs are going to be replaced by the Merit-Based Incentive Payment System (MIPS). With MIPS, there will be an opportunity to address many of the issues discussed during the TEP from which reporting mechanisms are available for public reporting to requirements for satisfactory reporting.

In addition, with the proposed 2016 PFS rule, there could be close to three hundred PQRS measures available for public reporting next year. The TEP was encouraged to provide recommendations for how to parse through what could be up to 300 measures. From a feasibility stand point, it would be critical to narrow down the measures to a set that are going to be useful to consumers.

### **6.1 TEP Input on Physician Compare Moving Forward**

- A TEP member mentioned that MACRA will be an opportunity for more reporters and quality measures.
- Another TEP member proposed grouping measures as composites in the future.
- A TEP member suggested treating data displayed on the Physician Compare website and downloadable database differently as they should have different priorities. However, per existing PFS rulemaking, Physician Compare already has the flexibility to handle these separately.

## 7 MEASURES RECOMMENDED FOR PUBLIC REPORTING: PQRS 2014

Based on its analyses and the TEP's input on the results, Acumen recommends publishing measures from one submission method each for group practices and individual EPs: Web Interface and Medicare claims, respectively. Registry data did not resonate well with consumers and did not meet public reporting standards. The Cardiovascular Prevention Measures Group data seemed skewed by selection bias. The measures recommended for public reporting are listed in Table 8.1. Reasons for exclusion include poor reliability, being reported exclusively from a submission method determined to be unsuitable for public reporting, and concerns regarding the impact of case-mix. Furthermore, results from analyses suggested by the TEP suggested four of the claims-based measures could be prone to selective reporting (PQRS 110-113). For these measures, in general, the more beneficiaries on whom an EP reported data, the lower their resultant performance rates; this trend suggests that certain EPs may have chosen to submit data selectively for those beneficiaries who would increase their performance rates. None of the other claims-reported measures showed this relationship.

**Table 8.1: Recommendations for Public Reporting**

PQRS Number	WI Number	Description	Measure Type	Recommendation for Program Year 2014	
				GPRO	Individual EPs
46	CARE-1	Medication Reconciliation	Process	Yes	Yes
318	CARE-2	Screening for Future Fall Risk	Process	No	No
118	CAD-7	ACE Inhibitor/ARB Therapy – DM/ LVSD	Process	Yes	No
1	DM-2	Hemoglobin A1c Poor Control	Outcome	No	No
319	DM-13	High Blood Pressure Control	Outcome	Yes	-
319	DM-15	Hemoglobin A1c Control	Outcome	No	-
319	DM-16	Daily Aspirin/Antiplatelet Medication Use – DM/LVSD	Process	Yes	-
319	DM-17	Tobacco Non-Use	Process	No	-
8	HF-6	Beta-Blocker Therapy - LVSD	Process	Yes	No
236	HTN-2	Controlling High Blood Pressure	Outcome	No	No
204	IVD-2	Use of Aspirin/Another Antithrombotic	Process	Yes	Yes
112	PREV-5	Breast Cancer Screening	Process	Yes	No
113	PREV-6	Colorectal Cancer Screening	Process	Yes	No
110	PREV-7	Influenza Immunization	Process	Yes	No
111	PREV-8	Pneumonia Vaccination Status for Older Adults	Process	Yes	No
128	PREV-9	Body Mass Index (BMI) Screening	Process	Yes	Yes
226	PREV-10	Tobacco Use Screening and Cessation Intervention	Process	Yes	Yes
317	PREV-11	High Blood Pressure Screening (Documented Follow-Up)	Process	Yes	Yes

PQRS Number	WI Number	Description	Measure Type	Recommendation for Program Year 2014	
				GPRO	Individual EPs
134	PREV-12	Clinical Depression Screening (Documented Follow-Up)	Process	Yes	Yes

\* As previously decided by CMS and past TEPs, lipid control measures are inconsistent with clinical guidelines and therefore have been excluded from the analyses.