Physician Cost Measures and Patient Relationship Codes (PCMP) Technical Expert Panel

February 6-7, 2020

Summary Report

May 2020

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

TABLE OF CONTENTS

1 Introduction ... 3
1.1 Project Context ... 3
1.2 TEP Panelists ... 5
2 Meeting Overview ... 9
2.1 Structure ... 9
2.2 Materials ... 10
2.3 Charter ... 10
3 Summary of Presentation and Discussion ... 11
3.1 Session 1-B: Chronic Episode-Based Cost Measure Framework ... 11
3.1.1 Summary of Presentation ... 11
3.1.2 Panelist Discussion ... 13
3.1.3 Key Findings ... 14
3.2 Session 1-C: Patient Relationship Categories and Codes ... 14
3.2.1 Summary of Presentation ... 15
3.2.2 Panelist Discussion ... 16
3.2.3 Key Findings ... 17
3.3 Session 1-D: Measure Maintenance and Re-Evaluation ... 17
3.3.1 Summary of Presentation ... 17
3.3.2 Panelist Discussion ... 19
3.3.3 Key Findings ... 21
3.4 Session 1-E: MSPB Hospital Measure Re-Evaluation ... 21
3.4.1 Summary of Presentation ... 21
3.4.2 Panelist Discussion ... 23
3.4.3 Key Findings ... 24
3.5 Session 2-B: Linking Cost and Quality to Achieve Value ... 24
3.5.1 Summary of Presentation ... 25
3.5.2 Panelist Discussion ... 26
3.5.3 Key Findings ... 29
3.6 Session 2-C: Measure Prioritization and Conceptualization for Future Cost Measure Development Waves ... 29
3.6.1 Summary of Presentation ... 30
3.6.2 Panelist Discussion ... 31
3.6.3 Key Findings ... 33
4 Next Steps ... 34
4.1 Session 1-B: Chronic Episode-Based Cost Measure Framework ... 34
4.2 Session 1-C: Patient Relationship Categories and Codes ... 34
4.3 Session 1-D: Measure Maintenance and Re-Evaluation ... 34
4.4 Session 1-E: MSPB Hospital Measure Re-Evaluation ... 35
4.5 Session 2-B: Linking Cost and Quality to Achieve Value ... 35
4.6 Session 2-C: Measure Prioritization and Conceptualization for Future Cost Measure Development Waves ... 35
5 Appendix A: PCMP Cost Measure Project Team ... 36

1 INTRODUCTION

The Medicare Access and Children's Health Insurance Program (CHIP) Reauthorization Act (MACRA) of 2015 established the Quality Payment Program, which rewards the delivery of high-quality patient care through Advanced Alternative Payment Models (Advanced APMs) or the Merit-based Incentive Payment System (MIPS). MIPS eligible clinicians are assessed in four performance categories – quality, promoting interoperability, improvement activities, and cost. MACRA requires that cost measures implemented in MIPS include consideration of care episode groups and patient condition groups (referred to as "episode groups"). The Centers for Medicare & Medicaid Services (CMS) has contracted with Acumen, LLC (referred to as "Acumen") to develop, maintain, and re-evaluate cost measures for use in the MIPS cost performance category through the *Physician Cost Measures and Patient Relationship Codes (PCMP)* contract (75FCMC18D0015/Task Order 75FCMC19F0004). This project is a continuation of the work performed under the *MACRA Episode Groups and Cost Measures* contract (HHSM-500-2013-13002I/ HHSM-500-T0002), which lasted from 2016 to 2019. Under the PCMP project, Acumen also maintains the Medicare Spending Per Beneficiary (MSPB) Hospital measure used in the Hospital Value-Based Purchasing (VBP) program.

This report provides a summary of the feedback shared by panelists during the February 6 and 7, 2020 Technical Expert Panel (TEP) meeting. Acumen has a standing TEP of 20 panelists, with diverse experience and perspectives, including representatives from physician and nursing specialty societies, academia, health administration, and person and family organizations. The remainder of this section provides an introduction of the PCMP project. Section 2 outlines the structure, materials, and composition of the panel. Section 3 presents a summary of the presentation, panelist discussion, and key findings for each session. The discussion summaries presented are not meant to represent a consensus view shared by all TEP panelists but rather to consolidate related feedback made by one or more panelist. Finally, Section 4 outlines the next steps for this project that take into account the feedback obtained from the TEP.

1.1 Project Context

Under this project, Acumen develops, maintains, and re-evaluates episode-based cost measures that meet the statutory mandate of MACRA for consideration for potential use in the Quality Payment Program. As a central part of constructing clinically valid cost measures, Acumen engages and works directly with stakeholders through stakeholder engagement activities such as TEPs, Clinical Subcommittees (CS), Clinician Expert Workgroups, and public comment periods. Acumen also collects input on the patient and family engagement (PFE) perspective. Acumen also hosts educational and outreach webinars to inform stakeholders on the measure development process.

Under the previous MACRA contract, Acumen's standing TEP met seven times between August 2016 and December 2018 to provide high-level guidance on the overall direction of measure development. The TEP meeting dates, locations, and topics for consideration are included in Table 1 below.

Table 1. MACRA TEP Meetings To Date

TEP Meeting Date	Meeting Location	Topics for TEP Consideration
August 15, 2016	Washington, DC	• Concepts of cost measure development • Alignment of cost and quality • Prioritization of measures for development
December 19, 2016	Washington, DC	• Discussion of cost measure development of procedural and acute inpatient medical condition episode groups including: o Defining an episode group o Attribution of the episode group to clinicians o Assignment of costs to the episode group
March 23, 2017	Virtual	• Structuring clinical input on the components of episode-based cost measures and prioritization of clinical areas • Approaches to alignment of cost and quality (led by Yale-New Haven Health Services Corporation, Center for Outcomes Research and Evaluation (CORE))
August 3, 2017	Washington, DC	• Risk adjustment to inform the measure development process • Re-evaluation of the Medicare Spending Per Beneficiary (MSPB) and Total Per Capita Cost (TPCC) measures
May 11, 2018	Washington, DC	• Refinement of MSPB and TPCC measures • Improving the Field Test Report for episode-based cost measures • Approaches for incorporating patient and family perspectives in cost measure development
November 14, 2018	Virtual	• Review summary statistics of the results of refinements discussed by the May 2018 TEP • Review stakeholder feedback from during the October 2018 field test of the MSPB and TPCC measures and discussed further refinements
December 14, 2018	Washington, DC	• Development and prioritization of chronic condition episode groups

In addition to the TEP, Acumen has assembled groups of clinical stakeholders to provide detailed input throughout the measure development process. Acumen uses a "wave" approach wherein sets of Clinical Subcommittees, each focused on a particular clinical area, convene to select episode groups to develop into cost measures and to provide input on the measures' specifications. Members of Clinical Subcommittees also provide direction for the selection of Clinician Expert Workgroups, which are smaller groups meant to facilitate focused discussions

to provide detailed input on each component of the cost measures. Table 2 presents a summary of CS convened in previous waves.

Table 2. Summary of Episode-based Measure Development

Development Cycle	# Clinical Subcommittees	# Clinical Subcommittee Members	# Affiliated Professional Societies with Clinical Subcommittee Members	# Workgroups	# Workgroup Members	# Affiliated Professional Societies with Workgroup Members	# Measures Developed
Wave 1 (2017 – 2018)	7	148	98	-	-	-	8
Wave 2 (2018 – 2019)	10	267	120	11	138	79	11
Wave 3 (2019 – 2020)	4	137	100	5	85	68	5

The Wave 1 measures were finalized for use in MIPS for the 2019 performance period and future years, and ten Wave 2 measures were finalized for use in MIPS for the 2020 performance period and future years. The Wave 3 measures will undergo field testing in summer-fall 2020.

Acumen also gathers input on the patient and family engagement perspective through discussions with beneficiaries and caregiver/family members of a Medicare patient who have experience with health care and/or patient advocacy, health care delivery, concepts of value, and outcomes that are important to patients across care delivery and trajectory and disease management. Finally, as part of the effort to involve and educate stakeholders on the measure development process, CMS and Acumen have hosted additional education and outreach activities, including listening sessions and national provider calls.

1.2 TEP Panelists

The PCMP TEP comprises 20 stakeholders with diverse perspectives and areas of expertise, as listed in Table 3, below. The panelists include expert stakeholders representing specialty societies, academia, health care and hospital administration, and patient and family member organizations. Eleven of the panelists are returning panelists from the standing TEP convened under the previous MACRA contract. Sixteen panelists attended the meetings in person, while three panelists attended the meetings virtually, and one panelist was unable to attend the meeting.

Table 3. PCMP TEP Composition

Name, Credentials, Professional Role	Organizational Affiliation, City, State	Consumer/ Patient/ Family/ Caregiver Perspective	Performance Measurement	Coding and Informatics	Clinical Content	Clinician or Hospital Payment Policy	Conflict of Interest Disclosure
***Anita Bemis-Dougherty**, **PT, DPT, MAS**, Vice President, Clinical Practice	American Physical Therapy Association, Alexandria, VA		X		X		N
Kathleen Blake, **MD, MPH,** Vice President, Healthcare Quality	American Medical Association, Washington, DC		X	X	X		N
Akinluwa (Akin) Demehin, MPH, Director of Policy	American Hospital Association, Washington, DC		X			X	N
Kurtis Hoppe, MD, Physician	American Academy of Physical Medicine and Rehabilitation, Rochester, MN		X	X	X	X	N
Caroll Koscheski, **MD, FACG,** Gastroenterologist	American College of Gastroenterology, Hickory, NC		X		X		N
Alan Lazaroff, **MD,** Physician	American Geriatrics Society, Centennial, CO		X	X	X		N
***Shirley Levenson**, **PhD, FNP-BC, PMHNP-BC,** Nurse Practitioner	American Academy of Nurse Practitioners, Caldwell, TX		X			X	N
Robert Leviton, **MD, MPH, FACEP, FAMIA, ABPM-CI** Associate Chief of Emergency Medicine	American Medical Informatics Association, Mamaroneck, NY			X	X		N
Edison Machado, **MD, MBA,** Chief Strategy Officer	American Health Quality Association, Lake Success, NY		X		X		N

Name, Credentials, Professional Role	Organizational Affiliation, City, State	Consumer/ Patient/ Family/ Caregiver Perspective	Performance Measurement	Coding and Informatics	Clinical Content	Clinician or Hospital Payment Policy	Conflict of Interest Disclosure
James Naessens, **MPH, ScD,** Professor of Health Services Research	Mayo Clinic, Rochester, MN		X				N
Shelly Nash, **DO, FACOOG,** Chief Medical Information Officer Ambulatory and Chief of Quality	Adventist Health System, Altamonte Springs, FL		X	X	X	X	N
Diane Padden, **PhD, CRNP, FAANP,** Nurse Practitioner	American Association of Nurse Practitioners, Austin, TX		X		X		N
***Parag Parekh**, **MD, MPA,** Ophthalmologist	American Society of Cataract and Refractive Surgery, Dubois, PA		X		X		N
David Seidenwurm, **MD, FACR,** Neuroradiologist and Quality Director	American College of Radiology, Sacramento, CA		X		X		Shareholder: Sutter Medical Group, RASMG Medical Group; ACR MRI Accreditation Program Fees, ACR consulting fees, medical legal consulting fees, NQF, HSAF, CMS travel, food, and lodging
Mary Fran Tracy, **PhD, RN, APRN, CNS, FCNS, FAAN,** Associate Professor	National Association of Clinical Nurse Specialists, Minneapolis, MN		X		X	X	N
Janice Tufte, Patient Advisor	Society for Participatory Medicine, Seattle, WA	X					N
****Ugochukwu (Ugo) Uwaoma**, **MD, MPH, MBA, FACP,** President of the Medical Group and Provider Services	Trinity Health of New England, Hartford, CT		X		X	X	N

Name, Credentials, Professional Role	Organizational Affiliation, City, State	Consumer/ Patient/ Family/ Caregiver Perspective	Performance Measurement	Coding and Informatics	Clinical Content	Clinician or Hospital Payment Policy	Conflict of Interest Disclosure
Danny van Leeuwen, **RN, MPH,** Patient Advisor	Health Hats, Arlington, MA	X					N
Michael Wasserman, **MD, CMD,** Geriatrician	California Association of Long Term Care Medicine, Newbury Park, CA	X	X		X		N
Adolph Yates Jr., **MD,** Academic Orthopaedic Surgeon	American Association of Hip and Knee Surgeons, Pittsburgh, PA	X	X		X		N

* Denotes panelists who joined the meeting virtually.

** Denotes panelists who were unable to join the meeting, either virtually or in-person.

2 MEETING OVERVIEW

This section provides an overview of the TEP meeting held on February 6 and 7, 2020. The TEP met from 9:00 a.m. to 5:30 p.m. on February 6 and from 9:00 a.m. to 12:30 p.m. on February 7 at the Acumen Washington, D.C. office.

2.1 Structure

The TEP meeting began with an introductory session that included a background of the measure development work completed under the previous contract followed by six topic-driven sessions across the two days. Table 4 below provides the agenda for both days of the meeting. In these sessions, Acumen sought specific feedback on outstanding methodological questions for chronic condition measure framework, evaluation and testing of patient relationship categories (PRC) and codes, the maintenance and re-evaluation of episode-based cost measures and the MSPB Hospital measure, measure prioritization for future cost measure development waves, and the alignment of cost and quality.

Table 4. TEP Meeting Agenda

Session	Topic
	Day 1
1-A	Introductions and Project Overview
1-B	Chronic Episode-Based Cost Measure Framework
1-C	Patient Relationship Categories and Codes
1-D	Measure Maintenance and Re-Evaluation
1-E	MSPB Hospital Measure Re-Evaluation
	Day 2
2-A	Day 2 Introduction
2-B	Linking Cost and Quality to Achieve Value
2-C	Measure Prioritization and Conceptualization for Future Cost Measure Development Waves
2-D	Open Discussion

Acumen presented targeted questions to inform the discussion and to solicit recommendations that can be operationalized in the service of the overarching project goals. While no formal recommendations were made, bulleted highlights of those discussions are presented at the end of each section in this report.

2.2 Materials

Prior to the TEP, Acumen provided panelists with the following materials: meeting agenda, presentation slides, background materials, and supplemental meeting materials.

2.3 Charter

Prior to the TEP, the *Physician Cost Measures and Patient Relationship Codes TEP Charter* was distributed to the TEP panelists for review. The CMS Measure Development Blueprint requires that each TEP have a Charter to outline the purpose of the TEP along with the level of commitment expected of the panelists. The Charter was approved by the 18 panelists of the TEP who attended the meeting in-person or virtually on the first day.

3 SUMMARY OF PRESENTATION AND DISCUSSION

This section summarizes feedback shared by TEP panelists during the February 6-7, 2020 in-person meeting. The section is organized into the six subsections representing the six main sessions of the meeting. Within each subsection, the discussion questions that were posed to the TEP in the session are listed in italics, with TEP panelists' discussion summarized below. In certain instances where TEP panelists discussed issues or asked questions requiring clarification, a summary of the response from Acumen is included to provide context and accurately reflect the flow of the discussion.

3.1 Session 1-B: Chronic Episode-Based Cost Measure Framework

During this session, Acumen provided an overview of the framework of chronic condition episode-based cost measures and presented additional methodological decisions necessary to address challenges with assessing chronic care.

3.1.1 Summary of Presentation

During the December 2018 TEP meeting, panelists provided input on the framework for chronic condition cost measures and on chronic condition episode groups to prioritize for initial development. As part of Wave 3 of measure development, Acumen convened a Chronic Condition and Disease Management Clinical Subcommittee, which selected the Asthma/COPD and Diabetes episode groups for initial chronic measure development and provided recommendations for the composition of measure-specific workgroups, which have subsequently provided detailed input during development of the measures.

The framework for chronic condition cost measures includes four steps:

(1) Identify beneficiaries with the chronic condition and attribute their care to the clinician group (identified by the Taxpayer Identifier Number [TIN]) managing their treatment. Under this step, a "triggering event" of two services billed by a TIN within 180 days is used to identify the start or continuation of the management of a patient's chronic disease. The triggering event must contain a primary care evaluation and management (E&M) code with a chronic diagnosis and a subsequent "confirming claim." The confirming claim can be another primary care E&M code with a chronic diagnosis or a Current Procedural Terminology / Healthcare Common Procedure Coding System (CPT/HCPCS) code related to the management of the chronic disease.

(2) Define the episode window for which cost will be assessed in the performance period. The triggering event opens a one-year "attribution window," which can be extended if a "reaffirming claim" or a service identified during an attribution window reaffirms and extends a TIN's responsibility for managing a patient's chronic disease. The total attribution

window can span multiple years and vary in length between patients and TINs. Episodes are defined as one year segments of the total attribution period that overlap with the performance period. All clinically related costs occurring during the episode are assigned to the managing clinician.

(3) Assign services and calculate the observed cost. The costs of all assigned services are summed and averaged across the number of days in the episode to obtain the observed cost, which is then annualized by multiplying by 365.

(4) Calculate the expected spending for the episode and the measure score. A regression model with risk adjustment variables as covariates is ran to calculate the expected cost of each episode. The measure is calculated as a weighted average of the episode ratios (i.e., ratio of annualized standardized observed cost to annualized expected cost for each episode) across a TIN's attributed episodes. The average is weighted by the number of attributed days for each episode.

Acumen outlined four framework considerations to address the challenges of assessing chronic care. The first is the implication of attributed clinicians' effect on risk adjustment for future performance periods. Due to the continuous nature of managing the treatment of a chronic condition, TINs can be measured on the same patient in adjacent performance periods. Since the risk adjustment model uses a lookback period from the start of an episode, when a TIN is attributed a patient in multiple performance periods, risk adjustors captured in the lookback period for the subsequent performance period are defined during the preceding episode, a period in which the attributed clinician has an opportunity to influence the patient's condition. In the Asthma/COPD and Diabetes measures, nearly all TINs have at least one episode for which the same patient was attributed to the TIN in a previous performance period and a risk adjustor emerged during that time; however, these episodes represent a small portion of a TIN's total attributed episodes.

The second consideration is the impact of high-cost acute events and extended post-acute care (PAC) utilization. These services represent opportunities for cost savings through improved care practices however, they should not have an outsized effect on the overall measure performance. High-cost events have the potential to drive up a clinician's measure score, even if they provided cost effective care to most of their patients. Episodes assigned acute inpatient admissions or PAC services had much higher risk adjusted costs than the population of episodes. However, TINs with episodes assigned costs for acute inpatient admissions and/or home health do not have substantially larger measure scores than the overall population, mostly due to the distribution of these events across TINs. TINs with episodes for which skilled nursing facility (SNF), long-term care hospital (LTCH), and inpatient rehabilitation facility (IRF) services are assigned do show an increase in measure score.

The third consideration is the evaluation of beneficiaries that transition between managing clinicians. The chronic condition measure attribution methodology intends to reflect the continuous, team-based nature of managing chronic conditions and encourages care coordination. Under the current construction, episode windows can hold clinicians accountable for costs after they may have stopped managing a patient's chronic condition and could be truncated when there is evidence that a patient has transitioned to another provider and more accurately reflect clinician responsibility. Patient transition between TINs is infrequent but affects almost all TINs. Truncating episodes where a transition occurs is shown to have a small impact on most provider's scores.

The fourth and final consideration is whether to include episodes that are shorter than one year in length. Episodes can be truncated if an ending event (e.g., a lung transplant for a COPD patient) concludes the patient's chronic disease. Currently, these episodes are excluded but chronic measures average costs over the number of days in the episode, so it would be possible to include shorter episodes. This change would increase measure coverage but would result in higher cost variation when comparing the shorter episodes to episodes that are one year or longer.

Panelists were presented with the following questions to inform discussion for each framework consideration:

- *How can we account for a clinician's effect on future risk adjusted cost?*
- *How should we include condition-related acute inpatient and PAC services in the measure?*
- *How should we account for transition of care?*
- *How should we account for episodes shorter than one year?*

3.1.2 Panelist Discussion

Panelists discussed the challenges in determining the impact that a previously attributed clinician could have made on the course of the disease captured by a chronic cost measure, as there are some patients and conditions that will get progressively worse regardless of the care they receive. Ultimately, panelists advised not to make changes to the chronic measures and the conditions should be risk adjusted in full regardless of the clinician's previous involvement with the patient. Panelists also agreed that all comorbidities should be included in the risk model since we cannot determine when they arise due to the clinician or not. Discussion of the inclusion of costly PAC services in the measures, though related to the condition being treated, was primarily driven by the differences between PAC settings and the care provided in each. Broad considerations of the resources available at the community-and organization-level where clinicians practice and geography were raised as important factors that may influence decisions about PAC services and introduce variation in PAC use. The concern was raised that PAC

decisions about should be made for clinical reasons, not cost considerations. A few panelists made the argument for excluding LTCH altogether, or at least building in special considerations based on PAC setting such as for LTCH and IRF, as they consider patients in those settings to be separate populations. Panelists were less concerned with inpatient costs because, even though those episodes may have high risk-adjusted costs, clinician-level measure scores are not greatly affected by their inclusion since they occur frequently and the costs are distributed across clinicians.

Panelist discussion about truncating episodes for beneficiaries that transition between managing clinicians included the possible explanations for the transitions and the effects of truncating those episodes. There was agreement amongst panelists that a provider's responsibility for a patient does not end at the point of transition to a new provider. The responsibility of a patient's costs in these episodes should overlap with the subsequent provider's care for a period of time, though different lengths of time were suggested including between three and six months. Panelist support for including a period of overlap was predicated on certain factors, including the relative infrequency of transitions and the potential for shorter episodes to be excluded.

This session concluded with a brief discussion about the possibility of including truncated episodes that are less than one year in length. The presentation noted that this change would increase measure coverage since truncated episodes are currently excluded, however the average daily cost would have higher variation when compared to episodes lasting one year. Panelists generally agreed with the current approach of excluding truncated episodes.

3.1.3 Key Findings

- Panelists agreed that chronic episode-based cost measures should include comorbidities that arise from complications or the chronic condition worsening while the attributed clinician was caring for the patient in the past.
- Panelists highlighted that some PAC settings should be considered for exclusion, such as LTCH and IRF, based on the severity of patients treated in those settings.
- Panelists agreed that when a patient transitions to a new clinician, it is appropriate to attribute the costs of an episode to the original clinician for some period of time.
- There was agreement among panelists that truncated episodes, due to changes in a patient's chronic condition, should continue to be excluded from the measure.

3.2 Session 1-C: Patient Relationship Categories and Codes

During this session, Acumen provided an overview of the PRCs and their reporting status in addition to considerations for further engagement to increase reporting and potential

approaches to testing the validity and accuracy of PRCs prior to their potential use in cost measure attribution methodology.

3.2.1 Summary of Presentation

PRCs define a clinician-patient relationship at the time of furnishing a service, and were finalized as HCPCS Level II modifier codes (as shown in Table 5 below) in the CY 2018 Physician Fee Schedule final rule. CMS implemented voluntary reporting of PRCs starting in January 2018 to allow time for clinicians to become familiar with PRCs and for CMS to gather information and data to test their validity and accuracy.

Table 5. PRC Codes and Descriptions

Code	Category
X1	Continuous/Broad Services
X2	Continuous/Focused Services
X3	Episodic/Broad Services
X4	Episodic/Focused Services
X5	Only as Ordered by Another Clinician

To increase clinicians' familiarity with PRCs, CMS and Acumen have hosted education webinars, posted an FAQ and MLN Connect articles, and supported the QPP help desk to answer stakeholder's questions about how to report PRCs. Most recently in 2019, the PRCs were finalized as a high-weighted Improvement Activity to incentivize reporting.

CMS's plans for testing the PRCs include assessing the consistency and accuracy of PRC reporting before explore how PRCs can be used in the attribution of cost measures. PRCs have the highest potential for use in the attribution methodology for chronic condition episode-based cost measures and the Total Per Capita Cost (TPCC) measure. Some approaches for incorporating PRCs include relying solely on the presence of a PRC code to attribute an episode in place of the current attribution methodology, using PRCs to confirm episodes identified by the current methodology, or using PRCs as an additional method of attribution. However, the low rate of PRC reporting thus far presents obstacles to any testing plans. Acumen's analyses indicate that only 0.01% of Medicare Part B Physician/Supplier claim lines billed in 2018 had a PRC code and only 0.07 % of clinicians reported a PRC. Clinicians reporting PRC codes most frequently reported categories reflecting short-term and focused care (e.g., X4, X5) while few reported broad care. For most specialties, no clinicians reported PRCs.

Given the need for increased reporting before testing can be begin, Acumen suggested activities that might help engage clinicians around PRCs, such as building PRC discussion in to the measure development process, incorporating PRC information in Field Test Reports used during beta testing for the cost measures, and coordinating with the Medicare Administrative Contractors to disseminate PRC information through existing communication channels. Acumen sought TEP feedback on these suggestions, ways to identify barriers to PRC reporting, the

outlined approaches to testing PRCs and incorporating them into cost measure attribution methodology.

Acumen presented the following questions to inform the TEP's discussion:

- *What additional strategies can increase reporting rates?*
- *Are there particular types of materials or information that are needed to increase familiarity with PRCs (e.g., a rubric outlining the expected code to bill for a certain clinician type and service provided)?*
- *What questions should be posed to the public to better understand clinicians' PRC reporting experience?*
- *Are the proposed testing methods (e.g., inter-rater reliability, test-retest reliability) sufficient for assessing PRCs? How should we define the expected, or appropriate, usage of PRCs for validity testing?*
- *What mechanisms for incorporating PRCs into the current attribution methodology for chronic measures and TPCC should be tested?*
- *Is there any benefit of using PRCs in attribution of acute inpatient medical and procedural episodes?*

3.2.2 Panelist Discussion

Panelists discussed ways to reduce administrative burden associated with PRC reporting. One common suggestion was to leverage electronic health record (EHR) systems as an effective method to automate reporting of PRCs. There was agreement that without EHR systems, reporting PRC codes would be cumbersome for both clinicians and billing vendors. However, panelists advised that the timing be taken into consideration as EHR systems are currently implementing many billing changes that vendors and clinicians will need to learn, such as the high priority changes to the E&M codes. Panelists also pointed to the lengthy process of incorporating new codes in EHR systems (up to two years) and the significant expenses associated with this process as factors for Acumen to consider in its recommendations to CMS. Additionally, panelists noted that reporting a PRC code on each claim line could be burdensome to clinicians, and suggested a system for designating a default PRC code within the EHR system based on specialty type or an algorithm that maps to other billing codes. The default code would require clinicians to only make changes as needed, rather than select a code for each claim line.

Furthermore, panelists considered whether existing codes or attribution algorithms could sufficiently convey information for which PRCs were created. Existing codes, such as the 90-day global code and chronic care management codes, could be used to capture the relationships that PRCs aim to identify. Similarly, in terms of whether PRCs could improve current attribution methodology, panelists suggested the PRCs would only add value if they attribute differently from the current methodology, and highlighted TPCC as one of the few measures for which

PRCs have potential to improve attribution. One caveat that was raised was the potential for PRC codes to be used as a kind of defensive strategy for clinicians to report limited patient-clinician relationships. In general, panelists concluded that PRCs should be pursued only if they provide additional information that cannot be achieved by leveraging existing codes and attribution algorithms.

Finally, panelists discussed the potential barriers to reporting and engagement with PRCs. Generally, panelists felt the voluntary reporting does not provide enough incentive for clinicians to report PRC codes, and some panelists pointed to mandatory reporting or payment incentives as more effective approaches if CMS would like to make PRC reporting a priority. However, many panelists expressed concern about mandatory reporting and the additional burden it would place on clinicians to manually report PRC codes on each claim line if PRCs are not incorporated into EHR systems. Mandatory PRC reporting would also require clinicians to become familiar with PRCs. In terms of increased engagement, panelists expressed that clinicians may require guidance from CMS on when to report certain codes, especially when there is room for interpretation or in cases of an evolving relationship. Other panelists noted that would it be challenging to explain the value of PRCs to patients, suggesting that Acumen and CMS should consider ways to translate the technicality of PRCs into language that patients and the general public can understand.

3.2.3 Key Findings

- EHR systems should be leveraged to automate PRC reporting and reduce administrative burden.
- Identifying ways to reduce burden, such developing an algorithm for designating a default code that clinicians only need to update when necessary may help increase reporting.
- Rather than seeking further public comment, Acumen should work with CMS to determine whether PRCs present additional value not available through existing codes and attribution algorithms and explore approaches to minimize burden in the implementation of PRCs.

3.3 Session 1-D: Measure Maintenance and Re-Evaluation

During this session, Acumen presented the proposed process for maintenance of MIPS episode-based cost measures, and outlined two potential substantive updates for existing cost measures: inclusion of Part D costs and how to address inpatient outlier payments (IOPs).

3.3.1 Summary of Presentation

Acumen maintains 21 cost measures, 20 of which are currently in use in MIPS. This measure set includes 18 episode-based cost measures and two population-based measures. Measure maintenance is the process by which we ensure the measures are functioning properly, accurate, and are up-to-date. The CMS measure maintenance process includes three distinct

forms: routine annual maintenance, ad hoc review, and comprehensive measure re-evaluation. Routine annual maintenance ensures that a measure is functioning as intended and that procedure and diagnostic codes used in the measure are current. Ad hoc review includes significantly more information gathering and can occur at any time in a measure's lifecycle if there is persuasive new information or empirical evidence. Information gathering may include environmental scans, empirical analyses of updated data sources, and stakeholder input such as targeted conversations with experts or convening the measure's Clinician Expert Workgroup. Based on the outcome of the ad hoc review, a measure may be revised, left as-is, or suspended or removed from the program. Comprehensive re-evaluation occurs every three years and revisits all aspects of a measure. This process includes in-depth information gathering, analyzing measure performance, and, typically, convening clinical experts to assess the measures.

Acumen outlined considerations for incorporating Medicare Part D drug costs in episode-based cost measures developed in Waves 1 and 2, and how to account for IOPs in for inpatient hospitalization services assigned to episodes to gather input on whether to pursue ad hoc reviews or comprehensive re-evaluation.

The TEP previously recommended against including Part D costs until a price standardization methodology was developed. Now that a standardization methodology is nearly complete, it is important to consider if drug costs should be added, and, if so, which measures should be prioritized. If Part D drug costs are added retroactively to Wave 1 and 2 measures, criteria for prioritization include the following considerations: (i) the impact of including drug costs on variation in clinician cost performance, (ii) the importance of including drug costs to clinical face validity, (iii) the degree to which the measure dis-incentivizes Part D substitutes, and (iv) if other measure specifications will need to be changed as a result of the inclusion. Acumen's analyses indicated that 69.8% of Medicare fee-for-service beneficiaries are enrolled in Part D, which raises the question of how to account for beneficiaries that do not have prescription drug coverage. Acumen suggested sub-grouping by Part D enrolling as one approach to ensure separate risk-adjustment models are estimated for Part D enrollees and non-enrollees, such that the measure only compares observed to expected spending within each enrollment type.

IOPs are supplemental payments, beyond the standard diagnosis-related group (DRG) payments, that compensate hospitals for especially high-cost inpatient stays. IOPs are rare and occur in 2-3% of all Medicare fee for service inpatient hospital stays. Acumen typically includes IOPs in cost measures, unless the overall episode costs are so high that episodes are deemed to be statistical outliers and excluded from the measure calculation. For the episode-based cost measures developed so far, IOPs affect a small fraction of episodes (2.1% of all acute inpatient episodes). IOPs are significantly more frequent in the Acute Kidney Injury Requiring New

Inpatient Dialysis measure however (36.4%). Some considerations for including IOPs are that they are a meaningful source of cost variation and omitting them could incentivize long hospital stays to reduce PAC costs. Acumen found that excluding IOPs in the Acute Kidney Injury Requiring New Inpatient Dialysis resulted in a 20% reduction in the measure reliability score demonstrating that including the costs improves measure reliability. Alternatively, the degree to which IOPs are under the influence of the attributed clinician must be considered. Acumen proposed including IOPs but counting only a portion of the outlier costs, referred to as "discounting," and using risk adjustment to adjust for factors such as length of stay. Potential criteria for determining if IOPs should be discounted include the prevalence of IOPs within a measure, the degree to which IOPs are under the influence of the attributed clinician, whether other measure specifications address concerns about clinician influence, the impact of IOPs on measure reliability, and the ease of ensuring predictive accuracy through revisions of the risk adjustment model.

Acumen presented the following questions to inform the TEP's discussion:

- *How can the maintenance process be refined?*
- *What factors should be considered when determining if updates require stakeholder input?*
- *How can we strike a balance between keeping measures updated through frequent changes and clinicians' ability to understand how their performance is being evaluated?*
- *Are there other factors we should consider in prioritizing Wave 1 and 2 measures for inclusion of Part D costs?*
- *Which, if any, Wave 1 and 2 EBCMs should be re-specified to include Part D costs first?*
- *What are limitations of the proposed approach to account for Part D enrollment differences across beneficiaries?*
- *Are there criteria that should be added or removed when determining if inpatient outlier costs should be discounted?*

3.3.2 Panelist Discussion

Panelist comments were largely supportive of the measure maintenance framework and process outlined by Acumen and included suggestions for improving the collection of input from stakeholders when determining what form of maintenance is required for a given measure. The process could be expanded to be more accessible to non-clinical stakeholders that may not be aware of the opportunities to provide input. A dissenting view shared was that the process does not function well in practice and additional outreach from CMS and Acumen could improve the process. Panelists debated the benefits of evaluating cost measures at regular intervals with the need to incorporate changes mid-cycle, such as in the case of new or disruptive technology emerging. Ultimately, most panelists agreed that the process should be predictable and regular, except in outstanding cases. An alternative viewpoint, that measure updates could drive change

quickly, was also made in support of having a mechanism for early review in cases where disruptive technology may have been introduced or significant changes are enacted through rulemaking.

Panelist discussion of the inclusion of Part D costs in cost measures recognized that the issue is complex but an important potential aspect of measuring the cost of care in many instances. There was agreement that the decision to include Part D costs should be considered on a case-by-case basis, as there are many unique factors in nearly every case. Retroactively adding Part D costs to existing cost measures, such as to procedural episode-based cost measures from Wave 1, is not likely to improve the measures as panelists expected that drugs are uniformly prescribed in the pre- and post-trigger periods for procedural episode groups. There was more support for adding Part D costs to chronic episode-based cost measures, where the management of the condition could be more cost effective and prescription drugs likely play a larger role in the overall episode cost. This support extended to other measures where drug costs are a large factor. Examining Part D data could provide an opportunity to learn more about cost measure episodes, for example, if maintenance medications are prescribed that might signal an ongoing chronic episode. The issue of clinical nuance was ubiquitous throughout the discussion of when it is appropriate to include Part D costs.

Discussion of standardized drug costs and the potential to sub-group within measures helped to allay panelist concern about making fair comparisons among providers and between patients that do and do not have Part D coverage. Panelists inquired about the mechanics of the standardization process and if it can address clinician uncertainty of the actual drug prices for patients. On the topic of sub-groups, some panelists questioned if sub-grouping would create groups too granular for accurate comparison and discussed what types of groups would be appropriate to create. Acumen agreed, noting that factors of cell size and the size of the patient cohort captured by the measure would be taken into consideration to determine appropriate sub-groups and that this would be determined in consultation with members of the Clinician Expert Workgroups. Panelists expressed reservations about how considering the costs of prescription drugs could affect clinician decisions and the negative impacts patients face by receiving cheaper, but less effective drugs. Some panelists indicated that new technology that informs a provider's selection of prescription drugs by factoring in the costs to patients in real-time could impact the inclusion of Part D costs in measures. Other panelists noted that these decisions could impact quality measurement of the care provided and may produce unintended consequences such as providers prescribing drugs that initially cost less but are detrimental to patients down the line. These types of examples could be an argument for not including Part D costs until there is the capability to capture long term benefits of provider choices.

Discussion of IOPs was grounded in the Acute Kidney Injury Requiring New Inpatient Dialysis measure though the concepts are potentially applicable to all measures. The Acute Kidney Injury Requiring New Inpatient Dialysis measure is unique with its high number of IOPs so any decision on the topic should not be based on that measure alone. There was some support among panelists for the "discounting" approach, as clinicians do have a role in IOPs. An opposing panelist view was that IOPs should not remain in cost measures across the board, based on the example of the Acute Kidney Injury Requiring New Inpatient Dialysis measure. There may be value in keeping the costs, or part of them, in order to see what can be learned from the episodes where they are present.

3.3.3 Key Findings

- Panelists were largely supportive of the proposed measure maintenance framework and thought it could be improved by expanding efforts to collect stakeholder feedback.
- Panelists agreed that Part D costs should be included on a case by case basis when their inclusion will make a difference and improve the measure. It is also necessary to evaluate the drugs and how they are grouped to ensure there is enough variation in cost to determine if they should be included.
- Panelists supported prioritizing the inclusion of Part D costs in chronic and acute inpatient medical condition episode groups ahead of procedural episode groups that may have less variability in prescription drugs costs.
- Panelists supported discounting costs for episode groups where the prevalence of IOPs is high.

3.4 Session 1-E: MSPB Hospital Measure Re-Evaluation

During this session, Acumen reviewed the measure construction steps for the MSPB Hospital measure and outlined proposed refinements to incorporate to the measure as part of its re-evaluation.

3.4.1 Summary of Presentation

The MSPB Hospital measure is the only measure within the efficiency and cost reduction domain of the Hospital VBP Program and began to affect payment in fiscal year 2015. The measure was endorsed by the National Quality Forum (NQF) in 2013 and re-endorsed in 2017. It is due for comprehensive re-evaluation in fall 2020.

The MSPB Hospital measure includes all Medicare Part A and B costs during an MSPB hospital episode, which spans the three days before a hospitalization, the duration of the hospitalization, and 30 days after discharge. These costs are risk-adjusted and attributed to acute care hospitals paid under the Inpatient Prospective Payment System where the hospitalization occurs. The measure calculation steps include: i) define the population of index admissions, ii)

standardize costs and calculate observed episode costs, iii) calculate expected episode costs, iv) winsorize expected costs and exclude statistical outliers, v) calculate the "MSPB Hospital Amount", and vi) calculate the MSPB Hospital measure. The MSPB Hospital measure centers the score around 1. If a hospital's MSPB measure score is greater than 1, then they are greater than the national median.

Acumen outlined four potential refinements to the MSPB Hospital measure. The first potential refinement is to narrow the all-cost approach of the measure in response to concerns expressed in some public comments about the inclusion of costs that may not be within the hospital's influence. These concerns are weighed against the ability of the all-cost approach to promote a broad incentive for care coordination. A refinement of the set of Medicare Part A and B services included in the MSPB Hospital measure can strike a meaningful balance between these perspectives and may improve measure reliability and could leverage the service exclusion rules that were incorporated into the MSPB Clinician measure during its re-evaluation. The second potential refinement is to allow readmissions to trigger a new episode to account for episodes and costs that are currently not included in the measure but that could be within the hospital's reasonable influence. Allowing a readmission to trigger a new episode would increase the number of episodes for which a clinician can be scored and align the incentives of the cost measure during readmissions and capture potentially high-cost services that are otherwise excluded. Preliminary analyses indicate that allowing readmissions to trigger episodes results in a 14% increase in triggered episodes and a small change (less than +/-0.3) in the measure score for 97% of hospitals.

The third potential refinement is to change the measure calculation from the sum of observed costs divided by the sum of expected costs (i.e., ratio of sums) to the mean of observed costs divided by expected costs (i.e., sum of ratios). The original format allows more costly episodes to be weighted proportionately, which can make the measure slightly more sensitive to outlier episodes, while the refinement would weight episodes equally (and thus reduce the impact of outlier episodes) and align with other MPSB cost measures and the episode-based cost measures. Preliminary results indicate that this refinement results in minimal changes to the measure score and a slight increase in reliability. The fourth potential refinement involved considerations of how to test social risk factors for potential inclusion in the risk adjustment model, building on past testing. Previous Acumen analyses of certain social risk factors (e.g., race, income-to-poverty ratio, dual eligibility) identified minimal impact on MSPB Hospital scores, though NQF panels in the past disputed the precision of the data that were used for these analyses.

Acumen posed the following questions for TEP panelists:

- *Should the current all-cost methodology of MSPB Hospital be narrowed? If so, how should MSPB Hospital adopt existing MSPB Clinician service exclusion rules – as is, or with further refinement?*
- *Should the MSPB Hospital methodology be refined to allow readmissions to trigger a new episode? What are the potential drawbacks from the increased capture of Medicare services?*
- *Should the MSPB Hospital measure calculation methodology be refined to be less sensitive to outlier episodes? What are the potential drawbacks from the decreased impact of outlier episodes?*
- *What other SRF testing can be done for the MSPB Hospital measure? Which other SRF/SES indicators or indices may be considered? And, at which levels of granularity (e.g., hospital and patient level)?*

3.4.2 Panelist Discussion

There was general agreement that the all-cost approach of the measure should not be narrowed to align with the MSPB Clinician measure as hospitals have a larger influence over patients and the cost of their care compared to clinicians measured under MSPB Clinician. Panelists also noted that hospitals have more episodes attributed to them compared to clinicians due to the large volume of hospital admissions. As a result, panelists felt that hospitals are less vulnerable to statistical outliers that may result from an all-cost measure and that for this reason, it was unnecessary to curtail the measure's scope. Panelists suggested exploring whether hospital performance varies when there is a bias towards medical or surgical episodes, and if there are potential opportunities to align with the two-pronged attribution methodology for medical versus surgical episodes in the MPSB Clinician measure, an approach that could also prove useful for identifying outliers.

For the second refinement, panelists agreed that readmissions should trigger MSPB Hospital episodes to capture costs in the subsequent 30 days post-discharge for the readmission, as it is clinically appropriate to hold the hospital responsible for these costs. This 30 day timeframe aligns with the 30-day post discharge window in the hospital readmissions reduction program. Panelists also discussed the potential for observation stays to be included in trigger methodology, but noted the hourly payment may present complications.

For the third refinement, there was general support for changing the measure calculation to the sum of ratios to reduce the impact of outlier episodes. Based on evidence Acumen presented, adjusting for outliers does not make a large difference, most likely because the outlier episodes are uniformly distributed across hospitals.

For the fourth and final refinement, panelists discussed the limitations of current research on social risk factors and agreed additional research is necessary to make an informed decision on the types of indicators to include in measures. Panelists specifically discussed the use of dual-

eligibility as a proxy for social risk factors. While it is easily captured from Medicare data and can show valuable comparisons, it is limited as a proxy for social risk factors because Medicaid patients and programs vary greatly between states. Some suggestions around this included exploring whether there are underlying clinical risk factors for dual-eligible populations or using dual-eligibility as a proxy for functional impairment that is not being measured currently. Panelists also noted the limited ability to capture diversity with one binary stratification for race (i.e., black and non-black). Additionally, panelists noted that the five digit zip-code level data used to assess income level in previous analyses are not sufficiently precise as areas captured by the five digit zip-code level can include a diverse range of socioeconomic status, and suggested that census block regions or nine digit zip-codes might allow for more granularity. Panelists mentioned that social risk factors are more likely to have a compounding impact at the hospital level than the individual level. Acumen agreed, noting that a social risk factors such as dual eligibility might have minimal impact at the individual level, compared to a hospital with resources that are strained due to a large population of dual eligible patients. In this scenario, the care and outcomes for all patients, regardless of dual-eligibility status, could suffer. While this compounding effect might be difficult to capture in data, Acumen agreed that studying social risk factors at the hospital level in addition to the individual level is warranted.

3.4.3 Key Findings

- The all-cost approach for including Medicare Parts A and B services in the MSPB Hospital measure should be retained.
- Readmissions occurring during another episode should be allowed to trigger distinct episodes.
- There were no objections to calculating the measure score as the average of observed costs divided by expected costs.
- More empirical analyses on social risk factors should be provided before TEP panelists can provide recommendations on specific indicators to include in the MSPB Hospital measure.
- More precise and nuanced indicators should be used for social risk factors, and data with more precise mapping of social risk factors to individuals, should be used in further analyses.
- Acumen should assess whether there are compounding effects of social risk factors at the hospital level as well as the individual level.

3.5 Session 2-B: Linking Cost and Quality to Achieve Value

During this session Acumen discussed ways to assess and achieve value of clinical care through linking cost and quality measures and presented 12 principles for systematic value assessment in MIPS through the MIPS Value Pathways (MVPs).

3.5.1 Summary of Presentation

Recent legislation has introduced value-based purchasing and quality reporting programs to assess the value of care across Medicare settings. CMS initiated the Meaningful Measures Framework to identify priority domains for measuring value, which can help select a parsimonious set of measure required to measure value. Value is defined as patient health outcomes for a given cost, where cost (medical expenditures associated with improving health outcomes) and quality (health outcomes produced by medical intervention) are assessed together to determine value. Currently, MIPS faces several challenges in assessing value. These include (i) siloed measurement of cost and quality that reflects disparate aspects of performance rather than health outcomes for a given cost, (ii) measure and activity sets that do not capture cost and quality of care for the full range of clinical practice, (iii) voluntary reporting and submission requirements that limit comparability across clinicians, (iv) clinician burden in understanding, selecting, and reporting measures and complexity of publicly available information intended for patients to use to make care decisions, and (v) weighting/reweighting of performance categories, which could complicate comparisons of value across clinicians.

MVPs present an opportunity to address these limitations by creating connections across the Cost, Quality, and Improvement Activities performance categories for different specialties and conditions. MVPs include a foundational layer comprising measures from the Promoting Interoperability performance category and administrative claims-based quality measures for population health. The overarching purpose of MVPs is to make MIPS more meaningful to clinicians and patients by moving away from the compartmentalization of measures and activities and assessing clinicians on aligned measures relevant to clinicians' scope of practice. This is intended to help achieve five key goals: (i) simplify MIPS by addressing the issues around complexity and burden, (ii) improve value by creating connections across performance categories, (iii) reduce burden of understanding, selecting, and reporting measures, (iv) help patients compare clinician performance, and (v) better inform patient choice in selecting clinicians.

Acumen presented 12 principles for constructing value-based purchasing systems that fall within four categories (specific principles are listed in the discussion summary alongside panelist input):

- Clinical Scope (Principles 1-3): Assesses the clinical scope of the value framework, and the area of care where there are opportunities to improve value.
- Measures (Principles 4-10): Describes measures and activities that should be used in each performance category to capture value in a way that clinicians and patients can act upon to improve outcomes and reduce cost.

- Value-Centric Scoring (Principle 11): Assesses the method of scoring measures and activities to ensure that they create incentives for high-quality, cost effective care decisions.
- Stakeholder Engagement (Principle 12): Discusses the processes that CMS should use to engage with stakeholders.

Panelists were asked to consider the following questions:

- *When operationalizing the principles requires tradeoffs, which principles should be prioritized?*
- *Should any of the 12 principles be changed or removed?*
- *Are there decisions that CMS must make that should be integrated into these principles?*

3.5.2 Panelist Discussion

Panelists discussed the term 'value,' and agreed that the patient voice needs to be accounted for, as the term may have different meanings among patients and clinicians. From the purchaser or payer perspective, value (defined as the ratio of quality of outcome to cost) may lead to cherry-picking as healthier patients tend to have a higher value score.

Panelists provided general feedback on the principles. Overall, panelists thought the principles were a good articulation of what MVPs should include, but may be insufficient in capturing an MVP as a whole. There was a suggestion to include a principle that accounts for potential sources of bias to ensure that the design of MVPs does not place certain specialties at a disadvantage. Acumen clarified that having a direct and transparent stakeholder engagement process that ensures consistency among MVPs can help address this concern of bias. Panelists agreed that patient engagement and satisfaction are critical and should be included in this set of principles.

Panelists also commented on the limitations of claims data and the associated increased burden. While claims data can help for certain time-limited and focused interventions, it needs to be supplemented by medical documentation in a lot of other instances. Reporting within specialties could be a viable solution to capture clinicians within the same group to assess cost and quality measures in a meaningful way and reduce burden. Additionally, there was concern that population-based measures do not reflect specialty care and should therefore be excluded from MVPs. There was some expressed preference for streamlining the existing MIPS program by changing the scoring methodologies for example, instead of requiring clinicians to report on MVPs that may not reflect their clinical practice.

Panelists also provided specific feedback on most of the principles individually.

Principle # 1: Clinicians should be evaluated by MVPs that best reflect the scope of their practice as identified and assigned based on clinician billing patterns. There was general disagreement to use billing patterns to identify a clinician's scope of practice, as billing patterns vary by specialty and may not capture some aspects of care due to the complexity of care. While it was recommended to use both International Classification of Diseases [10th revision] (ICD-10) codes and CPT codes to identify the type of care provided by a clinician to enhance the reliability of this principle, it was also noted that some ICD-10 codes may be underutilized, inaccurate, and are less reliable compared to CPT codes. Panelists suggested that this principle does not promote team-based care and instead evaluates individual clinicians. They suggested using sub-TIN reporting to bring relevance to different sub-specialties in large group settings and promote multidisciplinary practice. Panelists suggested that the implementation of this principle would increase reporting burden. Acumen clarified that this principle could be rephrased to remove the mention of billing patterns, as this principle is intended to focus on clinicians' performance on clinical activities that represent their typical practice.

Principle #2: MVPs should move towards more in-depth assessments of a specialty's care for disease processes. There was some concern about this principle since the scope of a specialty's care can vary greatly. Panelists indicated that it would be difficult to evaluate a primary care clinician, for example, who sees complex or comorbid patients based on individual diseases. They noted the importance of prioritization and ensuring that the principle provides a balance between being clinically meaningful to capture performance on something important while also retaining a focus on the broader issues. Acumen clarified that this principle aims to focus on what clinicians' care patterns look like. Panelists also mentioned the potential for burden under this principle, particularly for clinicians who are part of an organization with large multispecialty clinical practices.

Principle #3: MVPs must assess the value of care using the same standards for clinicians participating as individuals and clinicians participating in groups. *No specific comments were provided for this individual principle.*

Principle #4: MVPs must include at least one cost measure, with preference for measures that count services relevant to clinicians' role in care. There was agreement with the list of criteria Acumen presented on what cost measures must satisfy to be used in value assessment. Some of these criteria include: (i) the assignment of costs should accurately capture the role of attributed clinicians, (ii) clinicians should be held accountable for the cost they can reasonably influence, and (iii) measures must convey concrete guidance indicating how clinicians can alter practice to improve measured performance. This last criteria could be revised or elaborated on, as it suggests the need for clinical support tools that allow clinicians to fix a less optimal decision.

Principle #5: Complementary quality measures should assess important healthcare outcomes not included in cost measures. Panelists suggested incorporating the patient voice in this principle, such as noting the preference of patient-reported outcomes that reflect the experience of patients. This is informed by the NQF's approach of defining cost and quality as efficiency and incorporating value when the patient's perspective is included. Patient satisfaction was noted as another important aspect to consider to identify quality care, and could be modeled from the diabetic prevention program that captures patient-entered self-data and focuses on lifestyle, behavioral health, and physical examination changes. It was also suggested to add information collected from the Hospital Consumer Assessment of Healthcare Providers and Systems Survey to understand patients' thoughts on treatments they receive in the hospital. There was disagreement to use readmissions, complications, and the variability of care provided as dimensions of quality as they are not adequately risk adjusted for.

Principle #6: MVP construction should consider what updates or additions to measures would move towards an optimal value construct and meet Meaningful Measure Framework priorities. The Meaningful Measure Framework was viewed as a good framework that provides the opportunity for stakeholders to come together to provide input on meaningful measures that could be used to shape the construction of MVPs.

Principle #7: Process measures should be used only if the associated health outcomes are difficult to assess meaningfully or if outcome measures are not available. As in other principles, panelists recommended including patient input, and suggested that patient education as a process measure could be viewed by patients as something important to be collected and reported on. Process measures can be a defensible proxy for outcomes, particularly among specialties that do not have immediate outcome measures available to differentiate quality of care. This principle could be rephrased to say that process measures should be used "preferably" because associated health outcome measures are difficult to assess meaningfully or are unavailable. Outcomes cannot always be measured reliably and outcome measures cannot adequately adjust for patient risk.

Principle #8: Measures (or overall assessments of value) should suitably risk adjust for factors outside of clinician influence. *No specific comments were provided for this principle.*

Principle #9: MVPs should minimize clinician costs of data collection and reporting for the included cost and quality measures. Panelists agreed with this principle, as minimizing costs of data collection is important. There is a need to have a system in place that promotes primary care practices to lighten the reporting burden on primary care providers.

Principle #10: Measures and activities within an MVP should be mandatory. There was concern and some opposition to the mandatory participation in MVPs, with the belief that

clinicians should have the ability to select measures and activities that reflect their practice. Panelists indicated that there needs to be a clear case made if measures and activities in MVPs are mandatory. Certain tradeoffs in clinical practice, including reduced burden on clinicians, improvement in care, and lower costs, should be met before mandating the reporting of measures and activities within MVPs. Ensuring that certain quality standards are met before receiving cost-sharing money or bonuses can lead to acceptance of MVPs because it would reflect clinicians and patients' values.

Principle #11: Scoring of cost, quality, and IA in MIPS final score should reflect agency priorities particular to each MVP. *No specific comments were provided for this principle.*

Principle #12: Direct and structured stakeholder engagement is necessary for MVPs to accurately reflect clinician value of care and to enhance understanding and buy-in. There was agreement with including the concept of consistency across MVPs in this principle, and ensuring there is consistency in capturing quality across the spectrum, including Medicare Advantage programs. Through this process, it is also important to account for access to care among vulnerable populations. There was also support to use patient-reported outcomes and have CMS engage with patients who can contribute to reporting on outcomes.

3.5.3 Key Findings

- Panelists agreed that the principles were a good articulation of what MVPs should include, but noted that they may be insufficient in capturing an MVP as a whole.
- Panelists strongly recommended prioritizing principles that incorporated the patient perspective and team-based care approach.
- Panelists supported the use of patient-reported outcomes and person-centered process measures.
- Panelists were concerned with the use of billing patterns to identify a clinician's scope of practice, as billing patterns vary by specialty and may not capture some aspects of care due to the complexity of care.
- There was some concern for reporting burden under some principles, specifically principles 1, 2, 9, and 10.
- Panelists emphasized balancing the level of detail and focus of broader issues when developing these principles.

3.6 Session 2-C: Measure Prioritization and Conceptualization for Future Cost Measure Development Waves

During this session, Acumen outlined how prior TEP input has helped guide strategic decisions in prior waves of development and discussed potential strategies for future measure prioritization.

3.6.1 Summary of Presentation

Acumen develops cost measures in "waves" of development. To date there have been three waves of development: Wave 1 (eight measures developed), Wave 2 (11 measures developed), and Wave 3 (five measures under development, including two chronic condition measures for the first time). As noted in earlier sections, the CS convened in each wave select measures for development and discuss composition of Workgroups for each selected measure; Workgroups provide detailed clinical input on measure specifications for each measure.

Principles reflecting TEP input help guide strategic decisions in each Wave, including which CS to convene and guidance for CS and Workgroups to ensure consistency and high-impact across measures for MIPS. These principles include clinical coherence, impact and importance to MIPS, the opportunity for cost performance improvement, and alignment with quality indicators to assess clinician value, and have informed CS considerations for measure prioritization, including a focus on novel areas for development in Wave 3.

Acumen presented three strategic priorities to inform the TEP's input on measure prioritization for Wave 4 and future waves of measure development. The first strategy is to develop measures for additional types of care or specialties not covered by the current set of measures. A comparison of episode groups developed thus far to those included in the December 2016 posting reveals a measurement gap for particular specialties. The current set of 19 episode-based cost measures captures low percentages of clinicians in various large specialties including nurse practitioner, family practice, physician assistant, emergency medicine, physical therapy, diagnostic radiology, certified registered nurse anesthetist, and anesthesiology. Potential novel measures for some of these additional specialties include Emergency Visit for Shortness of Breath, Physical Therapy for Major Lower Extremity Musculoskeletal Procedures, Mammography for Breast Cancer Screening, and Anesthesia for CABG and Valvular Surgery. A cross cutting issue for these specialties is the extent to which these measures can take a broader form rather than focusing on narrow conditions or procedures.

The second strategy is to focus on building out the patient care continuum for different conditions. Measures assessing clinical care at each point of on a patient's care continuum (through overlapping episodes) can align incentives for different types of clinicians and encourage coordination. This framework can include measures for different clinician types providing care during the same event (e.g., orthopedic surgery measure and anesthesiology measure for a knee replacement episode) or measures for different stages of patient care (e.g., chronic care measure and acute inpatient exacerbation for COPD). In either scenario, each episode would only be compared to other episodes of the same episode group and costs would be comparable to those providing the same type of care.

Since measure development to date has focused on acute inpatient and procedural episode groups, PAC and chronic condition measures represent portions of the care continuum for which there is a gap in measurement in the current measure set. Examples of chronic conditions for

which measures could be developed include Coronary Artery Disease (CAD), Chronic Heart Failure (CHF), and Osteoarthritis. Examples of conditions or procedures for which PAC measures could be developed include COPD, Coronary Artery Bypass Graft (CABG) procedure, CHF exacerbation, and Lower Extremity Musculoskeletal procedure. These measures can add to the care continuum for conditions or procedures for which measures already exist. For instance the COPD PAC measure would complement the existing inpatient COPD exacerbation measure and the chronic Asthma/COPD measure. For many of these conditions, measures that capture new specialties or clinical areas as discussed in the first strategy would also contribute to a care continuum measure set. For instance, emergency medicine measures for Shortness of Breath and Chest Pain would complement measures in the care continuum for COPD and CHF, and CAD, respectively. Additionally, a Lower Extremity Musculoskeletal measure for anesthesia would add to the measure set for Osteoarthritis.

The final strategy is to target conditions and procedures that compose a large component of care for a particular type of clinician. For specialties that have multiple existing measures such as orthopedic surgery (e.g., Hip Replacement and Knee Replacement), the question would be whether to develop additional procedural measures to capture other aspects of care within this specialty. There are also instances where multiple measures have been created in a broad clinical area (e.g., cardiology), but do not focus on one type of clinician. In the case of cardiology, there is one acute inpatient medical condition measure, STEMI with PCI, and two procedural measures, Non-Emergent CABG and Elective Outpatient PCI. In examining the candidate episode groups in the December 2016 posting, acute inpatient Heart Failure & Shock and chronic CHF could be developed in tandem to address heart failure for cardiologists.

Panelists were asked to consider the following questions:

- *What additional guidance would be helpful for the CS or Workgroups to prioritize and develop measures?*
- *Given the number of EBCMs that can be created per year, how should we evaluate the tradeoffs between covering a broader array of clinicians versus covering a given type of clinician in more depth?*
- *Which chronic condition measures should be our next priority? What are ways to address concerns regarding the specific conditions discussed extensively in the Chronic CS?*
- *How broadly should PAC EBCMs be defined?*
- *Which additional specialties should we prioritize for Waves 4-6?*

3.6.2 Panelist Discussion

There was support for developing measures that capture specialties and areas of care not covered by the current measure set, particularly, mental/behavioral health, physical therapy, and diagnostic radiology. There was wide support for the development of measures focused on behavioral health, such as major depressive disorder, though panelists also expressed

reservations about the ability to properly capture this patient cohort among Medicare enrollees. One panelist pointed to the Mammography for Breast Cancer Screening as straightforward and a good first measure for diagnostic radiology noting that it is actionable, low burden and would have appropriate quality metrics for alignment. This panelist also suggested lung cancer screening for diagnostic radiology as well as peripheral vascular disease, hepatic cirrhosis, and stroke for interventional radiology, noting that these would be more complex but still good to consider. Additional suggestions for more broad types of care included pain management, low back pain, and spinal injection. Some panelists noted that some specialties such as Anesthesiology may not be conducive to cost assessment under the MIPS program as the wide range of outcomes that could occur due to a surgeon's care decisions would make it difficult to compare anesthesiologists to their peers.

In terms of high impact areas, panelists noted CHF as a condition that not only contributes to high costs but also affects a lot of specialties, including physical therapy and other specialties not covered under the current measure set. The Cardiovascular Disease Management CS discussed CHF in previous waves and there was a lot of interest because of the impact and cost. Acumen summarized the challenges around developing a CHF measure that the CS and previous TEP cited such as multiple etiologies in heart failure and the difficulty it would present to sub-grouping the patient sub-populations into clinically homogeneous patient cohorts for comparison. The development of a CHF measure should include considerations for addressing these challenges.

More generally, there was discussion regarding whether episode-based cost measures were appropriate for chronic conditions, which panelists noted occur in a continuum rather than in separate episodes of care with various specialists. Some panelists suggested person-centered care process measures as an appropriate way to measure clinicians and the management of chronic conditions. Panelists also mentioned looking into frailty evaluation process measures, noting that there is extensive literature on these measures and that simple measures of frailty are already available for clinician use.

Additionally, panelists pointed to the challenges in managing chronic conditions for patients with multiple morbidities and the challenges in determining the disease progression as there can be a period of time in which a patient is asymptomatic. Citing this aspect of chronic disease, one panelist emphasized the importance of taking into account the patient's disease stage in cost measurement. Relatedly, another challenge panelists raised was the lack of reporting of ICD-10 information that could be used to delineate severity for conditions such as chronic liver disease. Acumen noted that this challenge presents a tradeoff for whether to address this issue by developing measures that incentivize the reporting of this information or identifying other ways to incentivize this reporting. One panelist mentioned that EHR systems are often programmed to

provide clinicians the top diagnosis code based on previous billing patterns, but machine learning and artificial intelligence could help doctors more accurately diagnose patients in the future by suggesting appropriate diagnosis codes for conditions including chronic liver disease.

In terms of PAC measures, some panelists advised prioritizing home health given increased efforts to shift patient convalescence to the home, noting that this would be a good opportunity to align with patient reported outcomes that measure the family and caregiver burden as well. LTCH facilities were also suggested as another option because of the site-neutral payment. However, the panelist also cautioned that LTCH patients are sicker especially those with a vent weaning which could be problematic. Panelists noted that SNF might be most challenging as this setting can include a wide range of diagnoses and SNFs tend to have little guidance and few resources. Dissenting panelists advised that Acumen avoid efforts to develop measures for PAC settings citing industry pressure and minimal clinician influence as factors that might render measures for PAC setting non-viable.

In general, panelists agreed that there should be some effort to align priorities with measure development projects undertaken by other agencies, such as the Center for Medicare & Medicaid Innovation (CMMI) to ensure that measures developed complement each other. Panelists recommended that Acumen align with quality measures being developed by other contractors at CMS and focus on areas that were already identified as high priority.

3.6.3 Key Findings

- Panelists agreed with prioritizing measures capturing new specialties not covered under the current measure, particularly for following Clinical Social Work or Mental Health Therapy and Physical Therapy
- In addition to measures proposed by Acumen, panelists suggested the following additional measures to consider for future development:
  - Lung cancer screening
  - Low back pain
  - Spinal injection
  - Deep vein thrombosis
  - Peripheral vascular disease
  - Major depressive disorder
  - Hepatic cirrhosis
- Several panelists expressed reservations for the development of PAC measures, while others suggested home health and LTCH for initial development of PAC measure.
- Panelists advised that where possible, future measure development should be in sync with other measure development projects to ensure that resulting measures are complementary and not redundant.

4 NEXT STEPS

The input provided by this TEP, along with input from other stakeholders, will provide guidance to Acumen throughout the cost measure development process. The remainder of this section will discuss specific steps for how we aim to address and incorporate the feedback we received from this in-person TEP meeting.

4.1 Session 1-B: Chronic Episode-Based Cost Measure Framework

There was support from panelists for a period of overlap when a patient transitions to a new clinician where the costs of a chronic episode would be attributed to the original clinician. Panelists agreed that it is appropriate to include all identified health conditions in the risk adjustment model as clinicians cannot control the natural decline of a patient's health or the worsening of a chronic condition. Discussion of differences between PAC settings and the need to account for the severity of patients was informative and we will discuss the input with CMS when developing future chronic episode-based cost measures. We appreciate the panelist support for continuing to exclude truncated episodes that are shorter than one year from chronic condition measures and agree that a period of overlap when a patient transitions to a new clinician is appropriate as it reflects the ongoing and team-based nature of caring for patients with chronic conditions. We will take this input into consideration as we work with CMS to refine the chronic condition measures currently under development and in the development of future chronic condition measures.

4.2 Session 1-C: Patient Relationship Categories and Codes

Panelists expressed support for automating PRC reporting and suggested incorporating PRC codes into EHR systems to ease the burden of reporting and to increase reporting rates. We appreciate the panelists' suggestions and will discuss them with CMS and stakeholders through our ongoing PRC education and outreach efforts.

4.3 Session 1-D: Measure Maintenance and Re-Evaluation

Panelists were supportive of the proposed measure maintenance and re-evaluation framework, discounting costs for episodes with a high prevalence of IOPs, and they provided recommendations for which episode-based cost measure-types to prioritize when incorporating Part D costs in existing measures. We appreciate the support of the proposed framework and we will discuss the suggestions to include additional opportunities for incorporating stakeholder input in measure maintenance and re-evaluation through our work with CMS to refine the process. Panelist discussion in support of discounting costs for episodes with a high prevalence of IOPs will inform future measure maintenance and re-evaluation decisions.

4.4 Session 1-E: MSPB Hospital Measure Re-Evaluation

Panelists were supportive of refinements to the MSPB Hospital measure, including to align the measure calculation methodology with other MSPB measures and episode-based cost measures and allowing hospital readmissions to trigger new episodes. They supported the measure remaining an all-cost measure. This input will be taken into consideration as we work with CMS to refine the measure as part of the ongoing comprehensive measure re-evaluation. We will also explore additional analyses of social risk factors to inform the re-evaluation of the MSPB Hospital measure and the ongoing consideration to include social risk factors in CMS programs.

4.5 Session 2-B: Linking Cost and Quality to Achieve Value

We appreciate the feedback received from panelists regarding the challenges of the current approaches to assess value in MIPS. Panelist recommendations to prioritize the principles for systematic assessment of value that incorporate patient and team-based perspectives and to support patient-reported outcomes will be considered in our ongoing work with CMS to assess value in MIPS.

4.6 Session 2-C: Measure Prioritization and Conceptualization for Future Cost Measure Development Waves

Panelists recommended prioritizing the development of measures that capture clinical social workers, physical therapists, radiologists, and cardiologists. We appreciate the discussion of the challenges associated with PAC measurement and the opportunities identified in certain PAC settings. This input and the suggestion to align our work with measure developers that are developing quality measures, when possible, will be important factors in selecting episode groups to consider for future development into episode-based cost measures.

5 APPENDIX A: PCMP COST MEASURE PROJECT TEAM

The Acumen measure development team is multidisciplinary and includes individuals with knowledge and expertise in the areas of cost measure development, clinician payment policy, health economics, clinical practice, pay-for-performance, and value-based purchasing and quality improvement. The following 12 individuals from the project team attended the TEP:

- Sri Nagavarapu, Co-Project Director
- Jay Bhattacharya, Stanford Clinical Lead
- Rose Do, Clinical Lead
- Laurie Feinberg, Clinical Lead
- Nirmal Choradia, Clinical Lead
- Joyce Lam, Senior Policy Lead
- Sam Bounds, Senior Policy Lead
- Binglie Luo, Senior Policy Lead
- Aimée Uwilingiyimana, Senior Policy Lead
- Amanda Swygard, Policy Lead
- Taylore Fox, Data & Policy Analyst
- David Ruiz, Senior Research Associate