ACUMEN

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

August 29 and 30, 2021

Summary Report

September 2022

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

TABLE OF CONTENTS

1 INTRODUCTION

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). Acumen also maintains the Medicare Spending Per Beneficiary (MSPB) Hospital measure used in the Hospital Value-Based Purchasing (VBP) program. The PCMP project is a continuation of a previous contract, *MACRA Episode Groups and Cost Measures* (2016 to 2019).

As part of this work, we convene a standing Technical Expert Panel (TEP) to provide input on overarching issues across all activities. This report provides a summary of the TEP meetings on August 29 and 30, 2022. Section 1 outlines the structure and composition of the panel. Section 2 summarizes the presentation, member discussion, and key findings for each session. The discussion summaries presented are not meant to represent consensus but rather to consolidate related feedback. Finally, Section 3 outlines the next steps for this project.

1.1 Project Context

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"). Acumen constructs clinically valid cost measures for MIPS using extensive engagement, including a TEP, measure-specific panels of clinician experts (referred to as Clinical Expert Workgroups), person and family engagement (PF) representatives, and the general public via field testing and comment periods.

1.2 Standing TEP

The PCMP TEP comprises 20 members with diverse perspectives and areas of expertise. The panel includes experts in health care and payment policy, payment models, and quality and cost measurement; clinicians across many specialties; and patient advisors who share their perspective from lived experiences. Please see Appendix A for the full list of TEP members or the TEP Composition List, posted alongside this report. Table 1, below, lists the TEP meetings and their discussion topics.

Table 1. PCMP TEP Meetings

TEP Meeting Date	Meeting Location	Topics
February 5-6, 2020	Washington, DC *(with virtual option)*	• Chronic episode-based cost measure framework • Patient Relationship Categories (PRC) and Codes reporting limitations • Measure maintenance and reevaluation • MSPB Hospital measure reevaluation • Alignment of cost and quality • Measure prioritization and conceptualization for future development
July 20, 2021	Virtual	• Refining service assignment • Cost measurement gaps • Approach to cost measure calculation
August 29, 2022	Virtual	• Risk adjustment and social risk factors • Cost measurement gaps
August 30, 2022	Virtual	• Accounting for mortality in cost measures • Comprehensive reevaluation

Most recently, the TEP met via webinar on August 29 and 30, 2022. On August 29, sixteen of the 20 members attended from 1:00 to 4:30pm ET. On August 30, eleven of the 20 members attended from 2:00 to 5:30pm ET. In preparation for the meetings, Acumen provided TEP members with an agenda and presentation slides. The *Physician Cost Measures and Patient Relationship Codes TEP Charter* was also distributed to the members for review and was approved by the 16 members who were in attendance at the start of the meetings.

The TEP meetings began with an introductory session to provide an update about project activities since the last TEP meeting in July 2021. The rest of the meeting consisted of three sessions focusing on different aspects of the project: refining service assignment rules to balance clinical specificity with comprehensibility; identifying and filling cost measurement gaps; and exploring an alternative approach to cost measure calculation. A moderator from Acumen presented discussion questions for the panel.

2 DISCUSSION SUMMARY

This section summarizes feedback shared by TEP members during the August 29 and 30, 2022 meetings. Each section focuses on a session of the meeting: Section 2.1 focuses on risk adjustment and social risk factors, Section 2.2 discusses cost measurement gaps, Section 2.3 summarizes whether to account for mortality in cost measures, and Section 2.4 discusses comprehensive measure reevaluation.

2.1 Risk Adjustment and Social Risk Factors

Acumen walked the TEP through a review of risk adjustment and social risk factors, including two conceptual models, before giving them questions to discuss. Section 2.1.1 provides a summary of Acumen's presentation, Section 2.1.2 summaries the TEP members' discussion, and Section 2.1.3 contains key takeaways.

2.1.1 Summary of Presentation

This session discussed Acumen's approach to testing whether to include Social Risk Factors (SRFs) in cost measure risk adjustment. The overarching goal of risk adjustment is to account for factors outside of clinicians' control so that performance can be fairly compared; however there is a risk that adjusting for SRFs can also mask poor provider performance and exacerbate disparities in care. Acumen recapped the risk adjustment framework for cost measures, as SRFs would need to also satisfy these criteria, and presented a conceptual model to conceptualize how SRFs relate to performance and cost. In particular, SRFs should be considered against the criteria requiring that risk adjustors must: have a clinical/conceptual relationship with the outcome of interest, not be redundant with other variables, and not be an indicator of care provided. These criteria have also been noted in various guidance, including from National Quality Forum (NQF) and Assistant Secretary for Planning and Evaluation's (ASPE). ASPE's 2020 report provided guidance for whether it is appropriate to adjust for SRFs, noting that this is a continuum.

Acumen's SRF testing used dual Medicare and Medicaid enrollment status as a proxy as analyses have shown that this is a good predictor of cost variation, and is uniformly available for beneficiaries in enrollment data. Patient-level SRFs can be reported through ICD-10 Z codes for SRF; however, the reporting rate remains low (0.11% of all FFS claims) so cannot yet be reliably used for risk adjustment. Other proxy variables considered are neighborhood-level variables from the American Community Survey (e.g., income, education, unemployment).

We presented each step of testing for two measures that illustrate how results can indicate whether or not a measure should adjust for SRFs: Diabetes, a chronic condition measure, and Sepsis, an acute inpatient measure. This SRF testing is conducted for cost measures so that any

decisions about risk adjustment are informed by empirical testing. The testing for most cost measures do not support adjusting for SRFs, as was the case for Sepsis in these analyses.[1] However, where the results indicate that adjusting for SRFs would be appropriate, such as for Diabetes, a risk adjustor for dual status is included in the model. The steps of the SRF analyses are detailed below.

- Question 1: Is dual status associated with higher episode cost?
    - We did this by examining whether the coefficient of an additional dual status risk adjustor is both positive and statistically significant, indicating that episodes with dual status have higher costs.
- Question 2: Are higher costs potentially due to the provider's poor performance?
    - We compared the results from the first question to another model that adds a risk adjustor for a provider's share of dual episodes in addition to patient dual status. If the coefficients are similar across both models, that suggests that the additional costliness of dual episodes is more influenced by patient factors than provider performance. Conversely, if the magnitude of the coefficient when both patient dual status and provider share of dual episodes decreases compared to the coefficient just for dual status, this indicates that the provider's factors are more influencing costs.
    - We also compared provider performance on dual vs non-dual episodes. If poor performance is observed on both dual and non-dual episodes, this suggests that the higher costs of dual episodes is more likely to be related to performance.
- Question 3: Can providers perform as well on dual as non-dual episodes?
    - If a large share of providers can perform as well or better on dual episodes as compared to non-dual episodes, this indicates that it is possible to improve performance.
- Question 4: Is there much impact on provider scores with a dual risk adjustor?
    - We examined the percentage of providers that would experience a change in their score if there was a dual risk adjustor added. This includes shifts in both directions (i.e., providers who perform better and providers who perform worse) with the added adjustor.

Acumen presented the following questions for discussion:

- *Does the Conceptual Model represent the salient points about relationship between SRFs, patient, and provider performance?*

[1] Sandhu, A.T., J. Bhattacharya, J. Lam, S. Bounds, B. Luo, D. Moran, A. Uwilingiyimana, D. Fenson, N. Choradia, R. Do, L. Feinberg, T. MaCurdy, and S. Nagavarapu, (2020) "Adjustment For Social Risk Factors Does Not Meaningfully Affect Performance On Medicare's MIPS Clinician Cost Measures," Health Affairs, 2020 September; 39(9): 1495-1503

- *Are there additional tests or results that would be useful in evaluating the need to risk adjust for SRFs?*
- *What kind of information would be useful to present to clinicians in feedback reports to shine a light on disparities in care?*
- *What other mechanisms can help support providers with high caseloads of dual eligible beneficiaries beyond the cost measures?*

2.1.2 TEP Member Discussion

The TEP generally supported Acumen's SRF testing approach, emphasizing that it should be considered in light of the goal of accountability measures to improve care outcomes and value. A few members acknowledged that dual status is readily available but that it has limitations as a variable for testing due to the variation across states in Medicaid programs. The TEP was overall in favor of augmenting the SRF testing with community-level variables (e.g., AHRQ variables). Members highlighted particular advantages: SRFs often operate on a community level, there are significant differences in care based on zip code (e.g., rural vs urban), and the community-level variables can provide additional nuance that dual status alone may not. Acumen clarified that community-level variables have had little impact in testing after dual status is added. Reasons for this include that these may co-occur with clinical conditions that are already part of risk adjustment and the nature of episode-based measures to only include clinically related services (e.g., costly health outcomes that are more prevalent for patients with higher SRF may not be included in particular measures). Other members noted that there are limitations in interpreting testing results using community-level variables, such as heterogeneity within zip codes (e.g., wealth disparity within a small geographic area) and patient mobility across neighborhoods. A panelist also suggested that testing should consider the interaction between SRF variables and overutilization.

Members agreed that there was a need to improve SRF data collection, such as increasing the use of ICD-10 Z codes. Panelists noted that the uptake of Z codes currently varies across settings, including due to differences in claim forms. Electronic systems could be leveraged to make it easier for clinicians to understand and report Z codes.

Separately from risk adjustment, the TEP agreed that there is a need to shine a light on disparities in care by providing more information to clinicians to increase their understanding of patient circumstances. For example, an unhoused patient with diabetes cannot keep insulin in a cooled environment for storage, which should be factored into a treatment plan. Mechanisms to provide this information could be via confidential reporting so that clinicians can see performance differences across different subsets of patients (e.g., by demographics) or by leveraging electronic systems to identify patients who have had SRFs previously coded. This information could prompt clinicians to consider care options.

Members discussed additional factors that could enhance the overall understanding of how SRFs impact costs. These include the degree of Medicare Advantage penetration as areas with high shares of MA patients tend to have low-cost, upfront care offers that pull poorer patients out of Medicare FFS and into MA. Another factor affecting costs would be FFS beneficiaries who have concurrent commercial coverage and compare cost with other FFS populations. A suggestion was also to account for differences in state Medicaid coverage for a more accurate assessment of costs for dual eligible beneficiaries.

2.1.3 Key Takeaways

- The TEP agreed on the importance of improving outcomes for vulnerable populations but there was some minor disagreement about SRF testing and the inclusion of community-level factors.
- The TEP acknowledged a need to improve data collection, such as ICD-10 Z codes.
- The TEP called for the distribution of information to clinicians about disparities in health outcomes, separate from risk adjustment. This would focus on the need to consider patient circumstances when providing care.
- The TEP agreed that an enhanced understanding of SRFs may compliment cost measure development and testing, and that mechanisms beyond risk adjustment may be helpful for providers with lots of dual-eligible patients.

2.2 Measurement Gaps

This session focused on how to identify, prioritize, and fill gaps in cost measurement. It also included a new approach to measuring anesthesia-related costs. Section 2.2.1 provides a summary of Acumen's presentation, Section 2.2.2 summaries the TEP members' discussion, and Section 2.2.3 contains key takeaways.

2.2.1 Summary of Presentation

This session covered two topics: first, reviewing the current coverage of episode-based cost measures to identify priority areas for future development, and second, exploring a specific case study of applying an alternative approach to measure construction for anesthesia care. There are currently 23 episode-based cost measures in the MIPS Cost performance category in 2022, with another 10 in development which CMS could add to MIPS in the future. With many areas of care involving high costs and clear opportunities for improvement already covered by cost measures, Acumen presented analyses into different ways to identify the remaining high-cost clinical areas: (i) inpatient and institutional PAC stays with the highest total standard allowed amounts, as these are high-cost care settings, (ii) diagnosis codes present on claims with the highest total Part A and B spending, as these can indicate which conditions are the most prevalent and associated with the highest costs, (iii) specialties with the highest total Part B physician/supplier billing, and (iv) conditions and types of care being assessed in MIPS Value

Pathways (MVPs). For each of these, we highlighted which areas have an applicable episode-based cost measure; for example, of the 20 specialties with the highest total standard allowed amount, 15 have an applicable episode-based cost measure leaving hematology/oncology, diagnostic radiology, medical oncology, podiatry, and anesthesiology as specialties without an episode-based cost measure that applies to the care that they provide.

In the second part of the presentation, Acumen presented an adapted approach to cost measurement for anesthesia care as this involves large specialties (i.e., anesthesiologists and CRNAs) and there is an MVP that assesses pre- and post-operative anesthesia care. This has previously been discussed by the TEP and in public comment as a challenging area for cost assessment as it is difficult to identify what is within the reasonable influence of the clinician providing anesthesia versus that of the surgeon. Our team presented analyses that looked for services before and after anesthesia that appear to be directly related to anesthesia care vs surgical care. We found a very low incidence of anesthesia complications (e.g., ED visit with diagnoses that could be related to anesthesia, like retention of urine) and of services for monitoring anesthesia (e.g., blood counts after anesthesia) in claims data.

To address these challenges, Acumen tested an alternative approach to creating measures for anesthesia care that does not try to parse out the anesthesia role separately from the surgical role and instead assigns some portion of cost from procedural episode-based cost measures to the clinician providing anesthesia. For each procedural episode-based cost measure, we attributed a "shadow" episode to the clinician billing anesthesia on the same day as the trigger surgical procedure. The costs for the "shadow" anesthesia episodes were a subset of the assigned services for the procedural measure, limited to the grouped inpatient, outpatient, and Part B physician/supplier assigned costs in 7 days following the anesthesia trigger. Initial testing confirmed that anesthesiologists and CRNAs were identified for the majority of procedural episodes and that there was some variation in observed costs.

Acumen presented the following discussion questions regarding the identification and prioritization of measurement gaps:

- *Are there hospitalizations and institutional PAC care that should be prioritized for development?*
- *Are there any chronic conditions with high costs and opportunities for improvement?*
- *Are there types of care provided by specialties that indicate potential measurement gaps?*
- *Are any of the MVPs without potential episode-based cost measures strong candidates for cost measure development?*

Acumen presented separate discussion questions about the alternative approach to creating measures for anesthesia care:

- *To what extent does this approach of creating "shadow" measures address the challenges of assessing anesthesia care?*
- *Could this approach apply to all procedural measures or only particular types of procedures?*
- *Are there other types of care where a similar approach could be applied?*
- *How should service assignment be defined so that they can be automatically calculated for measures? E.g., setting, timing*
- *Are there anesthesia-specific service assignment rules that apply across different types of procedures to supplement existing procedural rules? E.g., ED visits with nausea/vomiting*

2.2.2 TEP Member Discussion

The TEP identified conditions and procedures that appeared to have good potential for improving cost performance. A member noted that hip fractures were among the costly inpatient stays that are excluded from episode-based cost measures (e.g., from the Elective Primary Hip Arthroplasty measure). Similarly, osteoporosis and the management of care after a fracture was noted as a potential candidate as there were efforts underway to assess this type of care in an MVP. To address the specialty gap for podiatry, a member suggested diabetic foot care (e.g., care for foot ulcers), osteomyelitis, and partial and complete trans metatarsal amputation could be areas for cost management. Imaging was also identified as an area with cost savings and patient health improvement opportunities: there are options for less expensive types of tests and patients should not have frequent radiology tests when they are not necessary, suggesting room for improvement. A few members noted that dementia and Alzheimer's disease would have high importance as it has a high cost burden and would fill in a large gap for care provided by neurologists; however, the TEP also agreed that there would be many challenges. For example, a member noted many costs related to dementia and Alzheimer's disease are outside of Medicare billing, both out-of-pocket costs and Medicaid.

The TEP also discussed two potential measure concepts based on setting and a combination of conditions. Some members noted that PAC should be high priority for a cost measure because of the amount of costs involved and the need to change incentives for PAC usage which currently is driven more by financial considerations rather than patient needs. Others expressed concerns about a cost measure focused on PAC as there is not yet enough understanding of how PAC is structured to avoid unintended consequences and that prior attempts have not been successful. A suggestion was to instead focus on hospitalization as inpatient measures would capture PAC costs as downstream services. Members also discussed

that patients with multiple complex or chronic conditions were a cohort that should be considered as managing a chronic condition affects the costs for many other chronic conditions.

The TEP also suggested other criteria to identify measurement gaps outside of the methods Acumen presented. One criterion was to examine Part D pharmaceutical drug use as prescription drugs are extremely expensive and make up a substantial share of Medicare spending. Additionally, there is a lot of evidence for inappropriate drug use spiking total treatment cost. Examining deprescription efforts can help identify gaps in measurement. A few members suggested looking at the American College of Radiology's (ACR) Appropriateness Criteria to discover radiology topics with opportunities to reduce costs; this tool is seeing increasing usage among providers.

The TEP generally did not support creating "shadow" measures for anesthesia based on existing procedural measures. They agreed that there was a lack of variation reflecting the role of the clinician providing anesthesia compared to the surgeon leading the operation. In the presented measure, events within the 7-day episode window were deemed to be much more likely attributable to the surgeon than the clinician providing anesthesia. Instead of that approach, the TEP recommended examining specific drivers of anesthesia cost, such as the specific drugs used or the mode/method of administration. The TEP said that due to the plethora of scoring variables for anesthesiologists, and because they all must report two different scores, perhaps that could the basis for future versions of the "shadow" model concept. Additionally, Acumen could look at separating the specialty into two groups: clinicians who perform spinal injections and clinicians who administer anesthesia.

2.2.3 Key Takeaways

- The TEP also identified several areas with cost improvement opportunities: hip fractures, osteoporosis and secondary fractures, podiatric care, and radiology.
- The TEP noted several other areas with cost improvement opportunities but would be a significant challenge to measure: PAC, Dementia & Alzheimer's, and Multiple Chronic Conditions
- The TEP suggested two other ways to identify measurement gaps: prescription drug usage and the ACR Appropriateness Criteria.
- The TEP did not support the anesthesia "Shadow" measure in its current form and discussed areas for further exploration, such as clinicians who primarily provide injections compared to anesthesia, as well as specific drivers of cost (e.g. drugs, method/mode of anesthesia).

2.3 Accounting for Mortality

This session explored the current mortality exclusion against criteria for when to exclude episodes, as well as discussed potential approaches to account for mortality if the exclusion does not meet the criteria. Section 2.3.1 provides a summary of Acumen's presentation, Section 2.3.2 summarizes the TEP's discussion, and Section 2.3.3 highlights key takeaways.

2.3.1 Summary of Presentation

This session discussed how mortality affects costs in episode-based cost measures and the extent to which the current exclusion for episodes ending in death meets established criteria for when it is appropriate to exclude episodes. While mortality is a major health outcome, its relationship with cost is less clear than other serious complications. For example, episodes ending in death may have low costs if patients die early in an episode, or have high costs reflecting very complex patients.

Acumen provided an overview of four main criteria used in measure construction for when exclusions are appropriate (rather than being addressed through risk adjustment or service assignment), and presented analyses from acute and procedural measures for whether episodes ending in death meet those criteria. Overall, the results varied across measures depending on the clinical focus of the measure where the most pronounced differences in trends were between on routine, outpatient procedures (e.g., cataract removal and screening colonoscopy) and acute hospitalizations (e.g., sepsis). Results for other measures tended to fall along a continuum between these. The following summarize the exclusion criteria and high-level testing results:

(1) The cohort is a small set of patients.
   - We found that the cohort size for episodes ending in death varied depending on the type of care being measured and the total number of episodes.

(2) There is a clinical or conceptual relationship with the outcome of interest.
   - We compared all-cost metrics to test whether episodes ending in death are different in a way that affects costs. Almost all measures have higher total Parts A/B spending for episodes ending in death compared to other episodes.

(3) The costs for the cohort have extreme variability, and the variability cannot be addressed through service assignment or risk adjustment.
   - Applying service assignment rules to only include clinically related services generally reduced the cost differences between episodes ending in death and other episodes. There were some inpatient cost measures where there were greater cost differences after applying service assignment rules.

(4) The exclusion is not an indicator of the care provided.

- We examined this criterion through three sub-criteria to assess the relatedness of death to attributed clinician's care: (i) prior hospice which could indicate pre-existing risk of death, (ii) the timing of death relative to the episode trigger, and (iii) the share of grouped costs which could indicate the intensity of related and unrelated services (i.e., a high share of grouped costs can indicate that the bulk of care that a patient receives is clinically related to the procedure or condition being assessed, while a low share can indicate that the patient has high, unrelated healthcare needs).
- We found that the percentage of episodes ending in death with prior hospice is very low across measures. Regarding the timing of death, procedural measures tend to have more days between trigger and death than acute and inpatient-only procedural measures. Most cost measures have a lower mean share of grouped costs for episodes ending in death compared to other episodes, while some inpatient cost measures displayed the opposite trend.

Acumen outlined potential approaches for cost measures to account for mortality where exclusion does not meet these four main criteria, and the advantages and limitations for each approach. The current approach is to exclude episodes from measure calculation and allow mortality to be assessed through quality measures or program-level interventions. Although MIPS is designed to evaluate different dimensions of care, there are only two quality measures that assess mortality in MIPS for limited conditions (i.e., after total hip/knee arthroplasty and coronary artery bypass surgery). Another option is to continue to exclude episodes ending in death from the cost measure score but provide information on them in confidential beneficiary level reports for clinicians. This approach provides information about providers' performance for episodes ending in death without impacting scores so that they can use it to make practice changes. However, providers would not be held accountable for costs associated with mortality if a death is related to the care being assessed.

Another option is to change the current approach for excluding episodes ending in death and include them in measure score calculation. This could be done by itself (i.e., make no other adjustments to the measure) or with other methodological changes to address challenges with assessing the costs of care prior to death. For example, some measures with sufficiently large episode counts could create sub-groups for episodes ending in death and episodes not ending in death. It could also be possible to include episodes ending in death but assign a fixed cost for these episodes to prevent episodes where death occurs soon after the trigger from appearing to be lower cost. This approach would hold providers accountable for costs associated with mortality. However, there could be concerns regarding potential incentives for care stinting, difficulty in setting a fixed cost, and other unintended consequences.

Acumen presented the following discussion questions:

- *Should there be changes to this approach for analyzing mortality as measure exclusions? What other analyses should be done?*
- *How should cost measures account for mortality (e.g., continue to exclude or include in the measure with or without adjustments?*
- *Should different cost measures address mortality differently, given the variation in testing results? If so, what principles can be applied to determine which approach would be appropriate?*
- *If death is accounted for in a cost measure score, what conditions would define episodes ending in death?*
- *For the option of assigning a fixed cost to episodes ending in death, how should this cost be determined?*

2.3.2 TEP Member Discussion

The TEP agreed that mortality is an important topic to be addressed in measuring performance, but mostly believed that this should be assessed through quality metrics rather than cost measures. Members noted that mortality is influenced by many factors, such as the inherent risk of the medical procedures, patients' underlying health conditions and particularly frailty, family influence, and patient preference. These and other factors contribute to the difficulty of identifying the relationship between mortality and cost. Some panelists noted that assessing clinician performance on mortality metrics is necessary from a patient perspective, especially in major procedures and acute conditions, but would require considering how high risk are the clinician's patient-case mix.

Members found Acumen's analyses thorough and suggested additional dimensions to build out analyses examining the extent to which episodes ending in death meet exclusion criteria. Members advised that Acumen should analyze risk-adjusted costs as the cost differentials could be attributable to risk factors associated with mortality. They suggested examining risk factors including age, number of hierarchical condition category (HCC) codes, and SRFs. To determine whether accounting for mortality in cost measures would be impactful, one member suggested concentrating on particular conditions with high mortality rates and assessing whether those conditions constitute a significant portion of Medicare spending. TEP members also noted limitations to analyses using claims data. They discussed missing data on relevant risk factors; while they could be incorporated into EHRs (e.g., Z-codes for SRFs), there were concerns that adding more fields would create additional burdens on clinicians to enter data into EHRs. For the analysis of prior hospice, members recommended analyzing hospice referrals rather than hospice service usage as a result of access to hospice services. Members suggested considering the role of advanced directives, such as Do Not Resuscitate (DNR) status, with one member noting that DNR orders are often coded inconsistently by providers.

In terms of whether episode-based cost measures should continue to exclude episodes ending in death or account for them in some way, TEP members generally agreed that mortality should continue to be excluded as the evidence did not strongly point to the need for a different approach. Members recommended that mortality should be addressed in the domain of quality improvement and encouraged more quality measures to account for mortality. A few panelists questioned what would indicate that cost measures should include episodes ending in death. Some suggestions were to consider whether something meaningful could be achieved, examine the share of Medicare expenditure incurred at end of life, and whether a substantial amount of cost information is lost through the exclusion. One member noted that it would be inevitable to consider costs associated with mortality and ways to address the increasing medical expenditures for patients at the end of life.

The TEP also highlighted the need to be cautious in changing the approach to excluding episodes ending in death due to the potential for unintended consequences. A few members were concerned that accounting for mortality in cost performance may undermine providers' ability to consistently deliver patient-centered care. For example, a physician may be ambivalent about a patient's decision to undergo a high-risk procedure when the patient presents higher risks for mortality.

2.3.3 Key Takeaways

- Members noted the many complexities affecting mortality across conditions and patients. The high degree of complexity and potential for unintended consequences necessitate a cautious approach in assessing mortality.
- The TEP generally agreed that episode-based cost measures should continue to exclude mortality. Mortality as an outcome is important to capture in some way, and for now, would be better assessed through quality measures.
- Members suggested focusing future analyses on clinical areas where there is a high mortality rate and where the excluded episodes mean that a substantial amount of cost performance information is lost. Metrics to examine include: risk-adjusted costs and mortality risk factors, such as age, SRFs, and number of HCC codes, and other diagnostic information.
- Members also noted limitations with available data which adds to the complexity of assessing mortality. For example, hospice referrals rather than hospice usage would be more accurate. Missing data from administrative claims data could be addressed through EHRs.

2.4 Comprehensive Reevaluation

This session reviewed recent comprehensive reevaluation activities and discussed principles for future reevaluation. Section 2.3.1 provides a summary of Acumen's presentation, Section 2.3.2 summarizes the TEP's discussion, and Section 2.3.3 highlights key takeaways.

2.4.1 Summary of Presentation

This session outlined the purpose of measure re-evaluation to ensure that cost measures remain current and clinically valid over time, and sought input on the process that aims to balance measure stability, regulatory processes, and the flexibility to make updates to measure specifications. Acumen has developed a measure maintenance framework to ensure that cost measures would continue to function as intended. The framework includes three components: (i) annual updates for non-substantive changes, (ii) early maintenance/ad hoc review for significant unforeseen problems, and (iii) comprehensive reevaluation occurring every three years for extensive changes. Wave 1 cost measures were added to MIPS in 2019 and are eligible for comprehensive reevaluation in 2022. We presented an overview of comprehensive reevaluation activities of these measures. There are four phases to comprehensive reevaluation: (i) initial information gathering, (ii) select measures for revisions based on information gathering findings, (iii) reconvene Wave 1 Clinical Expert 9Workgroups for selected measures, and (iv) public comment.

Acumen has completed Phases 1 and 2 in the first half of 2022. The first phase was to identify new information that was not available at the time of measure development through stakeholder feedback, monitoring with empirical analyses, and environmental scan/literature review. The second stage involved evaluating changes in clinical practice and guidelines and evidence from measure use, such as unintended consequences, program gaps, and agency priorities, to select Wave 1 measures for revisions. For example, changes in the MIPS quality performance category can present new opportunities for alignment. Given the evidence from various sources, CMS has approved the Routine Cataract Removal with IOL Implantation, Simple Pneumonia with Hospitalization, and ST-Elevation Myocardial Infarction (STEMI) with Percutaneous Coronary Intervention (PCI) measures for comprehensive reevaluation so that their specifications can be considered for updating to reflect any new information. The Elective Outpatient PCI, Knee Arthroplasty, Revascularization for Lower Extremity Chronic Critical Limb Ischemia, Screening/Surveillance Colonoscopy, and Intracranial Hemorrhage or Cerebral Infarction measures will be retained, where CMS keeps the measures active with their current specifications and minor changes.

Phases 3 and 4 will be conducted over the coming months. The Wave 1 Clinical Expert Workgroups will reconvene in October 2022 to provide input on the selected measures. After the

convention of Clinical Expert Workgroups, there will be a public comment period. If substantive changes are made to the cost measures, those measures would go through the pre-rulemaking and rulemaking process to be used in MIPS.

Acumen presented the following discussion questions:

- *What principles should guide whether a measure should be reevaluated and if so, what the scope should be?*
- *How should changes under "revise" be distinguished from those under the "retain" disposition?*
- *Are there any steps that should be taken while measures are being revised and before they can be used in the program?*
- *What types of information are most useful to stakeholders when reviewing and providing feedback on measure specifications?*
- *What are other avenues for increasing engagement with reevaluation?*

2.4.2 TEP Member Discussion

The TEP expressed interest in analyses on Medicare spending and trends in care; that is, to see how the cost measures were functioning within MIPS and the overall goal of improving the value of care. Metrics of interest included quality of care, overall Medicare spending, length of hospital stay, and need for post-acute care. Acumen noted that analyses would be limited within one year of operation as the cost performance category was reweighted to 0% due to the impact of COVID-19. We are working on constructing a monitoring framework to assess the impact of cost measures and welcome feedback on outcomes and metrics to be included in the framework.

The TEP also highlighted the need to examine the extent to which the measures continue to meet criteria for use in MIPS, such as variation in cost performance and that case volume continues to be high enough to be impactful. Some members noted that the sources of cost variation could change over time: for example, clinical guidance might recommend certain procedures should not be performed as frequently. This would mean that the opportunity for cost improvement includes the appropriateness of when such procedures are performed and not just the cost effectiveness of the procedure. Another factor highlighted by the TEP was to consider place of service, such as the shift from inpatient to outpatient (e.g., removal from the Inpatient-Only List), and from outpatient to ambulatory surgery centers. In each of these cases, measure specifications would need to be examined. Acumen noted risk adjustment and sub-grouping accounts for segmenting inpatient and outpatient cases in many measures. Some members also noted that bundled payments and Medicare Advantage penetration in particular geographic areas could also affect cost measure trends.

The TEP expressed the need to assess unintended consequences of cost measures during reevaluation. A few members suggested considering the concept of medical necessity. Members noted that holding clinicians accountable for costs in medically necessary conditions would undermine providers' ability to make proper medical decisions. Additionally, there were concerns about access to care and whether providers are "cherry-picking" healthier patients. For example, patients who are overweight are often denied hip or knee arthroplasty because they might require more intensive care afterward, resulting in financial penalties on providers.

TEP members discussed approaches to increase stakeholder engagement with re-evaluation. Members noted that conducting outreach to experts and specialty societies is an important way to solicit feedback on measure specifications. It would be helpful to share such analyses with specialty societies to identify outstanding trends even for cost measures to be retained. One member asked whether the request for input on comprehensive reevaluation was included in CMS regulations. Acumen clarified that the request for input was conducted outside the rulemaking process. The public comment opportunity was posted on the Measures Management System (MMS) public comment page, and outreach was done via the MMS and the Quality Payment Program listservs.

2.4.3 Key Takeaways

- Members expressed interest in big-picture trends into the impact of cost measures. For instance, the impact on overall Medicare spending and how care outcomes have changed in relation to cost.
- The TEP highlighted changes in practice and policy that could affect cost measures that should be considered during reevaluation. For instance, practice changes in place of setting and sources of cost variation should be considered to ensure that measures continue to be impactful and clinically meaningful.
- Members expressed the need to assess unintended consequences of cost measures during reevaluation, such as access to care for complex or vulnerable patients, and cost considerations on medically necessary conditions.
- Members suggested increasing outreach to specialty societies to gather input on outstanding trends to be addressed through revisions for measure specifications and encouraged Acumen to keep abreast of changes in clinical practices and how that might fit in with cost measures.

3 NEXT STEPS

The input provided by this TEP will serve as guidance for Acumen on new measure development and measure maintenance. After the meetings concluded, Acumen followed up with TEP members on their feedback and recommendations, gathering targeted input through a post-meeting survey. Based on the guidance received during the meetings and the survey responses, we will be taking the following next steps:

- **Risk Adjustment and Social Risk Factors:** We will continue with our approach to the risk adjustment conceptual model, and the structured approach to assessing whether social risk factors should be adjusted for in a measure. We will also continue to explore the potential for including community-level factors.
- **Measurement Gaps:** We will consider the TEP's suggested priority clinical areas for any future measure development, including hip fractures (and potentially pelvic, spinal, or femur fractures), post-acute care, osteoarthritis, dementia, and Alzheimer's disease. We will also consider whether the management of multiple chronic conditions - such as cardiac diseases (e.g., AMI, IHD, CHF, Arrhythmias) - could be a viable future measure concept.
- **Accounting for Mortality:** We will monitor the effect that mortality has on cost measures during measure development and maintenance efforts, while continuing to exclude any episodes where a patient death occurs.
- **Comprehensive Reevaluation:** We will explore ways to increase the outreach to specialty societies and other interested parties about the re-evaluation process. We will also consider ways that we can assess unintended consequences of cost measures and ensure that there have not been changes to clinical practices when reevaluating measures.

4 APPENDIX A: TEP MEMBER COMPOSITION

The table below includes the full list of TEP members, their professional roles, and their affiliated professional organizations.

Table A1. PCMP TEP Composition

Name, Credentials	Professional Role	Organizational Affiliation, City, State
****Anita Bemis-Dougherty, PT, DPT, MAS**	Vice President, Clinical Practice, APTA	American Physical Therapy Association, Alexandria, VA
Akinluwa (Akin) Demehin, MPH	Director of Policy	American Hospital Association, Washington, DC
Kurtis Hoppe, MD	Physician	American Academy of Physical Medicine and Rehabilitation, Rochester, MN
****Caroll Koscheski, MD, FACG**	Gastroenterologist	American College of Gastroenterology, Hickory, NC
Alan Lazaroff, MD	Physician	American Geriatrics Society, Centennial, CO
****Shirley Levenson, PhD, FNP-BC, PMHNP-BC**	Nurse Practitioner	American Academy of Nurse Practitioners, Caldwell, TX
Robert Leviton, MD, MPH, FACEP, FAMIA	Physician Advisor	American Medical Informatics Association, Mamaroneck, NY
***Edison Machado, MD, MBA**	Chief Strategy Officer	American Health Quality Association, Lake Success, NY
James Naessens, MPH, ScD	Professor of Health Services Research	Mayo Clinic, Rochester, MN
Shelly Nash, DO, FACOOG	Senior Vice President, Chief Medical Information Officer	Fresenius Healthcare North America, Altamonte Waltham, MA
***Diane Padden, PhD, CRNP, FAANP**	Nurse Practitioner	American Association of Nurse Practitioners, Austin, TX
***Parag Parekh, MD, MPA**	Ophthalmologist	American Society of Cataract and Refractive, Surgery Dubois, PA
****David Seidenwurm, MD, FACR**	Neuroradiologist and Quality Director	American College of Radiology, Sacramento, CA
***Mary Fran Tracy, PhD, RN, APRN, CNS, FCNS, FAAN**	Associate Professor	National Association of Clinical Nurse Specialists, Minneapolis, MN
Janice Tufte	Patient Advisor	Society for Participatory Medicine, Seattle, WA
Ugochukwu (Ugo) Uwaoma, MD, MBA, MPH, FACP	President of the Medical Group and Provider Services	Trinity Health of New England, Hartford, CT
Danny van Leeuwen, RN, MPH	Patient Advisor	Health Hats, Arlington, MA
Michael Wasserman, MD, CMD	Geriatrician	California Association of Long Term Care Medicine, Newbury Park, CA
Gregory Wozniak, PhD	Vice President, Health Outcome Analytics	American Medical Association, Washington, DC
Adolph Yates, Jr., MD	Academic Orthopaedic Surgeon	American Association of Hip and Knee Surgeons, Pittsburgh, PA

*Denotes members unable to join one meeting. **Denoted members unable to join both meetings.

5 APPENDIX B: 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:

- David Moore, Moderator
- Sri Nagavarapu, Co-Project Director
- Joyce Lam, Project Manager
- Laurie Feinberg, Clinical Lead
- Heather Litvinoff, Clinical Associate
- Sam Bounds, Associate Research Manager
- Binglie Luo, Associate Research Manager
- Allie Newsome, Policy Associate
- Ken Tran, Senior Policy Associate
- Nastia Biriuchinskaia, Policy Lead
- Henry Parkhurst, Data & Policy Analyst
- Elizabeth Peters, Data & Policy Analyst