NATIONAL QUALITY FORUM
Driving measurable health improvements together

MAP Clinician 2020-2021 Preliminary Analysis Worksheet

MIPS

Cost Measures

- MUC20-0015: Asthma-Chronic Obstructive Pulmonary Disease (COPD) Episode-Based Cost Measure
  - Measure Specification
  - Preliminary Analysis
  - Public Comment
- MUC20-0016: Colon and Rectal Resection Episode-Based Cost Measure
  - Measure Specification
  - Preliminary Analysis
  - Public Comment
- MUC20-0017: Diabetes Episode-Based Cost Measure
  - Measure Specification
  - Preliminary Analysis
  - Public Comment
- MUC20-0018: Melanoma Resection Episode-Based Cost Measure
  - Measure Specification
  - Preliminary Analysis
  - Public Comment
- MUC20-0019: Sepsis Episode-Based Cost Measure
  - Measure Specification
  - Preliminary Analysis
  - Public Comment

Quality Measures

- MUC20-0034: Risk-Standardized Acute Unplanned Cardiovascular-Related Admission Rates for Patients with Heart Failure for the Merit-based Incentive Payment System
  - Measure Specification
  - Preliminary Analysis
  - Public Comment
- MUC20-0040: Intervention for Prediabetes
  - Measure Specification
  - Preliminary Analysis
  - Public Comment
- MUC20-0042 Person-Centered Primary Care Measure Patient Reported Outcome Performance Measure
  - Measure Specification
  - Preliminary Analysis
  - Public Comment

- MUC20-0043: Preventive Care and Wellness (composite)
  - Measure Specification
  - Preliminary Analysis
  - Public Comment
- MUC20-0045: CoV-2 Vaccination by Eligible Clinicians
  - Measure Specification
  - Preliminary Analysis
  - Public Comment

SSP

- MUC20-0033 ACO-Level Days at Home for Patients with Complex, Chronic Conditions
  - Measure Specification
  - Preliminary Analysis
  - Public Comment

General Public Comments

- General Public Comments

MUC20-0015: Asthma-Chronic Obstructive Pulmonary Disease (COPD) Episode-Based Cost Measure

Measure Information

Characteristic	Submitted Information
MUCID	MUC20-0015
Other Measure Identification Numbers	N/A
Title	Asthma/Chronic Obstructive Pulmonary Disease (COPD) Episode-Based Cost Measure
Program	Merit-based Incentive Payment System-Cost
Workgroup	MAP Clinician Workgroup
In what state of development is the measure?	Fully Developed
State of Development Details	The measure has been developed, field tested, and is now refined based on feedback received from field testing in summer 2020 (August 17 to September 18, 2020). As background, a list of draft episode groups and trigger codes, developed with input from a Clinical Committee convened in 2016, was posted by CMS in December 2016 to meet MACRA requirements. Building off this work, episode-based cost measures are developed using a "wave" approach wherein sets of Clinical Subcommittees (CS) are convened to select episode groups to develop into cost measures, while smaller, measure-specific Clinician Expert Workgroups provide detailed input on each component of the measures. The current wave of measure development began in May 2019 and includes 4 CS with a total of 166 members affiliated with 110 professional societies. The Chronic Conditions and Disease Management CS selected this episode group from the December 2016 draft list of episode groups for development. This CS comprises members representing clinician specialty societies in this clinical area. A Clinician Expert Workgroup met in August 2019 to discuss measure specifications for all components of the measure, followed by a webinar in January 2020 for follow-up discussions on service assignment and risk adjustment. The measure was field tested in the summer of 2020, during which clinicians and stakeholders learned about the measure and provide input on the draft cost measure specifications. After field testing, the Clinician Expert Workgroups revisited and refined the draft measure specifications based on the stakeholder feedback received. We conducted reliability testing of measures for clinicians (TIN-NPIs) and clinician groups (TINs), constructed using episodes ending between January 1, 2019 and December 31, 2019. Reliability refers to the extent to which a

	measure reflects true variation between clinicians' risk-adjusted episode spending, as opposed to random variation. The reliability metric specifically captures how much of the variance in a measure is due to systematic differences in episode spending between clinicians, rather than differences in episode spending within a clinician's set of episodes. A measure with high reliability suggests that comparisons of performance across clinicians can be expected to better reflect systematic differences in actual performance. Our testing results indicate that this measure is reliable for clinicians and clinician groups across a range of case minimums. • For TINs at a 20-episode case minimum, the mean reliability was 0.64. For TINs at a 30-episode case minimum, the mean reliability was 0.70. For TINs at a 40-episode case minimum, the mean reliability was 0.74. • For TIN-NPIs at a 20-episode case minimum, the mean reliability was 0.57. For TIN-NPIs at a 30-episode case minimum, the mean reliability was 0.63. For TIN-NPIs at a 40-episode case minimum, the mean reliability was 0.67.
Measure Description	The Asthma/COPD cost measure evaluates a clinician group's risk-adjusted cost to Medicare for patients receiving medical care to manage asthma or COPD. The measure score is a clinician group's weighted average of risk-adjusted cost for each episode attributed to the clinician group, where each episode is weighted by the number of assigned days during the episode. This chronic measure includes services that are clinically related and under the reasonable influence of the attributed clinician group. Services are assigned during an Asthma/COPD episode, which is a portion of the overall time period of a clinician group's responsibility for managing a patient's asthma or COPD. Medicare beneficiaries enrolled in Medicare Parts A and B during the performance period are eligible for the measure. This measure addresses the Patient-Focused Episode of Care goal of CMS's Meaningful Measures Initiative, the CMS high priority area of Efficiency/Cost Reduction and MACRA statutory requirements.
Numerator	The numerator for the Asthma/COPD measure is the sum of the ratio of observed to expected payment-standardized cost to Medicare for all episodes attributed to a clinician. This sum is then multiplied by the national average observed episode cost to generate a dollar figure. Mathematically, this is represented as: sum of (observed episode cost/expected episode cost) * national average observed cost.
Denominator	The denominator for the Asthma/COPD measure is the total number of episodes from this episode group attributed to a clinician.
Exclusions	The following episode-level exclusions apply: (a) The beneficiary has a primary payer other than Medicare for any amount of time overlapping the episode window or in the lookback period. (b) No attributed clinician is found for the episode. (c) The beneficiary's date of birth is missing. (d) The beneficiary's death date occurred before the episode ended. (e) The

	beneficiary was not enrolled in Medicare Parts A and B for the entirety of the lookback period plus episode window, or is enrolled in Part C for any part of the lookback period plus episode window. (f) The episode trigger claim was not performed in an office, IP, OP, or ASC setting based on its place of service. Exclusions specific to the Asthma/COPD measure are developed with input from the Asthma/COPD Clinician Expert Workgroup.
Measure type	Cost/Resource Use
What is the NQF status of the measure?	Never submitted
NQF ID number	0000
Year of next anticipated NQF CDP endorsement review	
Year of most recent NQF Consensus Development Process (CDP) endorsement	
Is the measure being submitted exactly as endorsed by NQF?	
If not exactly as endorsed, describe the nature of the differences	N/A
What data sources are used for the measure?	Claims
If EHR or Administrative Claims or Chart-Abstracted Data, description of parts related to	N/A

these sources.	
At what level of analysis was the measure tested?	The measure has been tested at the Clinician: Individual and Clinician: Group levels.
In which setting was this measure tested?	Ambulatory/office-based care, Hospital outpatient department (HOD)
What NQS priority applies to this measure?	
What one primary meaningful measure area applies to this measure?	Patient-focused episode of care
What secondary meaningful measure area applies to this measure?	
What one primary healthcare priority applies to this measure?	Make care affordable
What secondary healthcare priority applies to this measure?	
What area of specialty best fits the measure?	Family Medicine
What is the target population of the measure?	Medicare Fee for Service
Is this measure an eCQM?	No

If eCQM, enter Measure Authoring Tool (MAT) number	No
If eCQM, does the measure have a Health Quality Measures Format (HQMF) specification?	No
Comments	N/A
Measure steward	Centers for Medicare & Medicaid Services
Long-Term Measure Steward (if different)	N/A
Measure Steward Contact Information	Evans, Ronique; Center for Clinical Standards and Quality; 410-786-3966; Ronique.Evans1@cms.hhs.gov
Primary Submitter Contact Information	Jensen, Ross; Acumen, LLC; 650-558-8882; macra-episode-based-cost-measures-info@acumenllc.com
Long-Term Measure Steward Contact Information	N/A
Secondary Submitter Contact Information	N/A
Was this measure proposed for a previous year's MUC list?	No
In what prior year(s) was this measure	

proposed?	
What were the programs that NQF MAP reviewed the measure for in each year?	N/A
Why was the measure not recommended in those year(s)?	N/A
What were the MUC IDs for the measure in each year?	N/A
NQF MAP report page number being referenced for each year	N/A
What was the NQF MAP recommendation in each year?	N/A
List the NQF MAP workgroup(s) in each year	N/A
What is the history or background for including this measure on the new MUC list?	New measure never reviewed by MAP Workgroup or used in a CMS program
Range of years(s) this measure has been used by CMS Program(s)	N/A
What other federal	

programs are currently using this measure?	
Evidence that the measure can be operationalized	This is a claims-based measure and will not require any additional submission of data.
How is the measure expected to be reported to the program?	Claims
Is this measure similar to and/or competing with measure(s) already in a program?	No
Which existing measure(s) is your measure similar to and/or competing with?	N/A
How will this measure be distinguished from other similar and/or competing measures?	N/A
Rationale for how this measure will add to the CMS program	N/A
If this measure is being proposed to meet a statutory requirement, please list the corresponding statute.	Section 101(f) of MACRA

Evidence of performance gap	Evidence of Performance Gap response: A recent study indicates that clinician beliefs about treatment and the efficacy of particular therapies may be the most important factors explaining the variation in health care expenditures. However, clinicians are often unaware of how their care decisions influence the overall costs of care. Cost measures are intended to help inform clinicians on the costs attributed to their decision-making and to incentivize cost-effective, high-quality care. A cost measure offers opportunity for improvement if clinicians can exercise influence on the intensity or frequency of a significant share of costs during the episode, or if clinicians can achieve lower spending and better care quality through changes in clinical practice. According to the literature and previous feedback received through stakeholder input activities to date, this measure represents an area where there are opportunities for improvement. Various educational programs and interventions have been associated with reduced readmissions, hospitalizations, and complications among patients with asthma and COPD. Opportunities to reduce costs and improve the chronic care and clinical outcomes of asthma and COPD exist primarily in maintenance pharmacotherapy, proper use of inhalers, pulmonary rehabilitation, and smoking cessation. Advances in pharmacotherapy have led to the development of guidelines to improve the management and outcomes of patients with COPD. However, it is estimated that 71 percent of Medicare patients with COPD are not prescribed long-term maintenance pharmacotherapy. Research has also shown other measures of under-treatment of COPD patients in the Medicare population, with suboptimal treatment for smoking cessation (behavioral therapy or prescriptions for medications), bronchodilator therapy post hospitalization, and pneumococcal and influenza vaccinations. In addition to potential under-prescription, medication adherence has also been documented as suboptimal, with only 50 percent of Medicare patients adhering to medications, signaling that patients may not be benefiting from prescribed therapies. This highlights an important opportunity for clinicians to prescribe treatment, such as appropriate inhaler devices, and encourage medication adherence during the management of COPD patients. Current guidelines suggest that inhaled bronchodilators are the mainstay of COPD management and therapy, and patients with either asthma or COPD can benefit from them. However, research has shown that over 50 percent of patients with asthma and COPD do not handle inhaler devices as prescribed or instructed, and up to 92 percent of patients experience critical errors that may impact the drug's effectiveness. This has important implications as poor inhaler techniques and non-adherence to inhaled therapy limit the therapeutic benefit of medication for patients with asthma and COPD. Existing literature suggests that the primary care physician has an important role in selecting appropriate inhaler devices for patients with asthma or COPD to optimize outcomes, while also encouraging patients to be involved in the decision-making process to improve patient education. Promoting medication adherence and instructing patients on proper inhaler techniques through educational and training methods could facilitate a successful relationship between clinicians and patients and optimize health outcomes.

Treatments that promote physical activity and exercise have been shown to improve patient outcomes for individuals with asthma and COPD. Various studies have looked at different components of pulmonary rehabilitation treatments (i.e., intensity) and patient selection (i.e., weight or disease severity) among COPD patients, and have indicated the benefits of pulmonary rehabilitation in improving exercise capacity and muscle function. One study showed that comprehensive pulmonary rehabilitation programs are beneficial in both early and late stages of COPD. For asthmatic patients, one study found that pulmonary rehabilitation can reduce the number of exacerbations and clinical visits while improving symptoms and pulmonary function. A clinician's role in prescribing pulmonary rehabilitation has potential implications for cost savings and improved performance given the benefits of pulmonary rehabilitation. Smoking is a main causative factor for COPD. Despite evidence showing the benefits of interventions promoting smoking cessation, it is estimated that 30 to 40 percent of COPD patients continue to smoke. This is concerning given that COPD patients who smoke have a higher prevalence of respiratory symptoms and higher death rates compared to non-smokers. Clinicians have an opportunity to promote smoking cessation among their patients in an effort to improve clinical outcomes and reduce cost of care. Existing literature suggests that smoking cessation among COPD patients is an important therapeutic intervention that "slows the accelerated rate of lung function decline and improves survival compared with continued smoking," even in severe COPD cases. For asthmatic patients, smoking cessation improves asthma symptoms and lung function, particularly when coupled with other therapies. One study found that subjects with asthma who quit smoking saw improvements in lung function compared to those with asthma who continued smoking. To optimize the management of asthma and COPD, clinicians should approach smoking cessation interventions by utilizing both behavioral (patient counseling and support) and pharmacological therapy for comprehensive treatment of asthma and COPD and improved outcomes. Additionally, patients with asthma and COPD and who smoke are at a higher risk of pneumococcal disease and influenza, and as such, clinicians should target these individuals for pneumococcal and influenza vaccinations to prevent asthma and COPD exacerbations. Overall, currently available research identifies areas of intervention primarily under the influence of clinicians, where evidence-based action can be taken to achieve better long-term health outcomes in the Medicare population.

This measure aims to address these example areas of opportunities for improvement. Research has shown that both asthma and COPD are highly prevalent, costly conditions within the US population, and their overall disease burden and financial impact continue to rise. As such, the use of this cost measure can provide clinicians with information to improve care outcomes and reduce future health care costs. This is supported by our analysis into the variation in spending per episode that may be reduced through better clinical practices. We analyzed the measure's performance gap for clinicians (TIN-NPIs) and clinician groups (TINs) using episodes

	ending between January 1, 2019 and December 31, 2019. With no case minimum applied, there were 3,102,288 Asthma/COPD episodes for 2,507,655 beneficiaries. The TIN-NPI and TIN-level measure scores were calculated for 35,976 clinicians and 20,568 clinician groups who met the 20-episode case minimum. • The mean risk-adjusted cost per episode was $5,089 at the TIN level. The risk-adjusted cost per episode at the 10th percentile was $3,272, compared to $6,893 at the 90th percentile at the TIN level. • The mean risk-adjusted cost per episode was $5,047 at the TIN-NPI level. The risk-adjusted cost per episode at the 10th percentile was $3,155, compared to $7,042 at the 90th percentile at the TIN-NPI level. • For TINs, the mean measure score was 0.95. The score at the 10th percentile was 0.61, compared to 1.29 at the 90th percentile. For TIN-NPIs, the mean measure score was 0.94. The score at the 10th percentile was 0.59, compared to 1.32 at the 90th percentile.
Unintended consequences	Stakeholders have expressed concerns that cost measures could potentially have risks or unintended consequences. During past stakeholder input activities, including Technical Expert Panels (TEPs) and public comment periods, stakeholders have highlighted the importance of considering potential incentives created by cost measures that could adversely affect patients and clinicians. For patients, these potential adverse effects may include: (i) reducing access to care, and (ii) stinting care. For clinicians, these adverse effects could include: (i) making unfair comparisons among certain types of clinicians in assessing cost measure performance, and (ii) rewarding lower costs without appropriately balancing performance on quality. Overall, by promoting lower spending, stakeholders have expressed concerns that cost measures could potentially create barriers to patients receiving clinically appropriate care. To mitigate the potential risks of such incentives, stakeholders and TEP members have highlighted the importance of the following: Devising an appropriate risk adjustment model for episode-based cost measures; Aligning cost measures with indicators of quality; Identifying benchmarks for comparison that are fair and appropriate for all eligible clinicians; Potentially excluding certain types of patients from measure calculations.
Which clinical guideline(s)?	N/A
Briefly describe the peer reviewed evidence justifying this measure	Research has shown that both asthma and COPD are highly prevalent, costly conditions within the US population, and their overall disease burden and financial impact continue to rise. COPD is the third leading cause of death in the United States. In 2014, 15.7 million Americans were diagnosed with COPD, yet this number could be an underestimation since many people with low lung function are not aware they have COPD. The Centers for Disease Control and Prevention estimated that COPD-related costs grew by nearly $17 billion in the past decade in the United States, equating to an overall increase of 53 percent. Specifically, Medicare paid 51 percent of

	these COPD-related costs. One study found that the mean total health care costs were $20,500 higher among Medicare patients with COPD compared to those without COPD. Among the many factors that contribute to rising health care costs associated with COPD, increasing hospitalization and readmission rates are among the highest cost drivers. COPD is the fourth leading cause of 30-day readmissions, where nearly one-fifth of patients hospitalized for an acute exacerbation of COPD were readmitted within 30 days of discharge. More than 25 million Americans live with asthma, and it has been estimated that five percent of all Medicare patients have an asthma diagnosis. The total cost incurred for treatment of asthma was $81.9 billion in 2013. Recent estimates attribute more than 10 million lost work days among employed adults and nearly two million emergency department visits over one year to asthma. Much like COPD, the burden of asthma falls heavily on adults aged 65 years and older, who have the highest mortality rate for the condition compared to any other age group. Despite the differences in etiology, symptoms, and responses to therapy between asthma and COPD, these diseases overlap in disease presentation and pathophysiologic characteristics. There is also a substantial 15 to 20 percent overlap in the reported prevalence of comorbid cases of asthma and COPD. This overlapping relationship places an important role on clinicians to follow appropriate guidelines and utilize proper management strategies to classify and treat patients accurately. Given the high impact in terms of patient population and Medicare spending, the Asthma/COPD measure represents an opportunity for improvement on overall cost performance.

Preliminary Analysis – MUC ID: MUC20-15 Asthma-Chronic Obstructive Pulmonary Disease (COPD) Episode-Based Cost Measure

Criteria	Yes/No	Justification and Notes
Does the measure address a critical quality objective not currently adequately addressed by the measures in the program set?	Yes	MUC20-0015 addresses the Patient-Focused Episode of Care goal of CMS's Meaningful Measures initiative, the MIPS high priority area of Efficiency/Cost Reduction and the 2015 Medicare Access and Children's Health Insurance Program (CHIP) Reauthorization Act (MACRA) statutory requirements (section 101(f)). Better asthma and COPD management can significantly lower costs and improve patient outcomes. Currently there are 3 MIPS quality measures related to management of asthma and COPD. Of the current measures, none look at episode-based costs.

Is the measure evidence-based and either strongly linked to outcomes or an outcome measure?	Yes	This cost/resource use measure aims to inform clinical decision-making related to asthma and COPD by reflecting cost of an episode of care and incentivizing cost-effective interventions. Studies suggest that knowledge and awareness of evidence-based practices and treatment risks can influence decision-making and can lead to lower costs (Cutler, et al., 2016).
Does the measure address a quality challenge?	Yes	Developer notes that COPD is the third leading cause of death in the United States. In 2014, 15.7M Americans were diagnosed with COPD. The CDC estimated that US COPD-related costs grew by nearly $17B in the past decade, equating to a 53% increase, with increasing hospitalization and readmission rates among the highest cost drivers. Medicare paid 51% of these COPD-related costs. More than 5% of all Medicare patients have an asthma diagnosis. The total cost incurred for treatment of asthma was $81.9 billion in 2013.
Does the measure contribute to efficient use of measurement resources and/or support alignment of measurement across programs?	Yes	This measure is not duplicative of other measures currently within the MIPS program. This and other 2020 measures under consideration were created in response to MACRA requirements to develop measures for potential implementation in the cost performance category of MIPS. Additionally, this measure is an episode-based measure; current cost-based measures in MIPs are all at the population level. Within MIPS there are two quality measures related to asthma control and one measure related to medication management for COPD. Across CMS programs, there are COPD and asthma quality measures within the Hospital Readmission Reduction Program, the Marketplace Quality Rating System, the Medicaid Child Core Set, Hospital Value-Based Purchasing, but no other measures of cost.
Can the measure be feasibly reported?	Yes	The measure uses Medicare claims data which is feasibility reported and a low-burden data source.
Is the measure applicable to and appropriately specified for the program's intended care setting(s), level(s) of analysis, and population(s)?	Yes	The measure is specified for Clinician: Individual and Clinician: Group/Practice levels, which aligns with MIPS reporting categories. The measure has not been reviewed for endorsement by an NQF Standing Committee.

If the measure is in current use, have negative unintended issues to the patient been identified? Have implementation challenges outweighing the benefits of the measure have been identified?	N/A	The submission identified a potential unintended consequence of reduction in access to care or stinting care. This measure could result in rewarding cost reductions with no net quality of care gain if it is not appropriately balanced with measures addressing performance on quality.
PAC/LTC Core Concept?		N/A
Impact Act Domain?		N/A
Hospice High Priority Areas?		N/A

Rural Workgroup Input	[redacted]	Relative priority/utility: • Access to pulmonary specialty providers or smoking cessation programs may be challenging for rural settings Data collection issues: • None identified. Calculation issues: • Definition of episodes was not clear. • Average cost for rural providers may exceed the national average cost and therefore may unduly impact rural providers Unintended consequences: • Cost savings focus may result in over-rationing of care resulting in lower quality Votes: Range is 1 – 5, where higher is more relevant to rural. Average: 3.9 • 1 – 1 vote • 2 – 5 votes • 3 – 5 vote • 4 – 7 vote • 5 – 1 votes
Preliminary Analysis Recommendation	Do Not Support with Potential for Mitigation	Mitigation is contingent on further evaluation of the correlation with clinical quality measures, as well as NQF endorsement. MAP noted a tension between expenses associated with good care that may result in reductions in overall cost of care but raise condition-specific care costs. MAP urged CMS to balance these cost measures with appropriate quality measures and to demonstrate the connection between them. MAP further noted that cost measures associated with upstream preventions should result in lowered downstream costs and expressed concerns that this is not the case for the measure, impacting its overall actionability. The developer noted that there is no specific measure that dictates that there will be lower cost, but rather that studies and other sources of evidence suggest that a given action will result in lower costs of care.

Summary: What is the potential value to the program measure set?	[redacted]	MUC20-0015 addresses the Patient-Focused Episode of Care goal of CMS's Meaningful Measures initiative, the MIPS high priority area of Efficiency/Cost Reduction and MACRA statutory requirements. US COPD-related costs grew by nearly $17B in the past decade, equating to a 53% increase, with increasing hospitalization and readmission rates among the highest cost drivers. Medicare paid 51% of these COPD-related costs. More than 5% of all Medicare patients have an asthma diagnosis. The total cost incurred for treatment of asthma was $81.9 billion in 2013. Currently, there are no MIPS measures that assess episode-based cost related to asthma and COPD.
Summary: What is the potential impact of this measure on quality of care for patients?	[redacted]	MAP noted that this measure was devised to reduce costs to Medicare claimants who experienced episodes of asthma and COPD events. While there are suggestions that effective interventions for asthma and COPD that result in lowered overall cost of care for beneficiaries and better patient outcomes, MAP suggested that these should be explicitly connected with MIPS asthma and COPD measure prior to implementation. Should testing data show that the measure appropriately measures episode-based cost and can be used to improve value of care, this measure would be valuable to add to the program measure set.

Measure Comments

Author	Submitted Comment
American Physical Therapy Association	We encourage the measure steward to include physical therapists as eligible clinicians that can be scored under this cost measure.
University of Colorado Medicine	The ability to measure cost per episode in a meaningful and useful manner relies on the ability to define an episode. It is not clear from the information provided how an episode will be defined- and what marks the start and end of an episode. Is an episode defined by an inpatient or observation hospitalization where asthma (or COPD) is the primary diagnosis, an ambulatory diagnosis for 'exacerbation' or what. What marks the end of an episode? This measure could result in increased use of exacaberation codes in the outpatient setting to increase the number of simple (and less costly) episodes or a decrease in valuable follow-up visits. It also might limit necessary testing. How will "episodes" that span providers/practices be judged, where one entity is not responsible for entire care episode? The final consideration is that HOPD-sites of care will never perform well on cost-based measures, and these arrangements are not clinician-driven decisions.

American Medical Association	The AMA continues to have significant concerns with MUC 20-0015 and believes additional field testing is necessary, especially considering the COVID-19 PHE and delayed implementation of the measure until the following issues can be addressed. Therefore, we urge the MAP to recommend "Do Not Support with Potential for Mitigation. More work must be done to achieve the desired results, which is why we are unable to support the measure at this time. • Many patients are incorrectly diagnosed with asthma by non-specialists and receive inappropriate and often expensive treatment before referral to an asthma specialist who determines the patient does not have asthma. • Ensuring that the risk adjustment methodology takes into consideration the fact that specialists typically provide care for patients with the highest disease severity. • Adequately incorporating social determinants of health, which play a large role in asthma care, into the risk algorithm.
Association of American Medical Colleges (AAMC)	The AAMC remains concerned that such cost measures are adjusted to account for social risk factors (SRFs). In addition to patient clinical complexity, SRFs can drive differences in average costs. In particular, physicians at academic medical centers (AMCs) care for vulnerable populations of patients who are sicker, poorer, and more complex than patients treated elsewhere. In regard to attribution – AAMC has previously commented that attribution methods used should be clear and transparent to clinicians and that it is critical that there be an accurate determination of the relationship between a patient and a clinician to ensure that the correct clinician is held responsible for the patient's outcomes and costs. Attribution is complicated, given that most patients receive care from numerous clinicians across several facilities, and AAMC has urged CMS to explore better data sources and analytic techniques to support more accurate attribution. The AAMC recommends that: (1) cost measures include risk-adjustment for SRFs, (2) the attribution methodology is transparent, and (3) the appropriate clinician is held responsible for the patient's outcomes and costs. Additionally, this measure has not been submitted for NQF endorsement. The AAMC has long held that measures should not be proposed for addition to public reporting programs unless vetted and endorsed by the NQF. The AAMC recommends that the highest level of MAP recommendation be "Do Not Support With Potential For Mitigation."
AdvaMed	AdvaMed strongly supports this measure, as it would provide useful information to support the development of new algorithms for early diagnosis and therapeutic guidance for COPD.
American College of Allergy, Asthma & Immunology	January 4, 2021 RE: Comments on the List of Measures Under Consideration Dear National Quality Forum Measure Applications Partnership:

The American College of Allergy, Asthma, and Immunology is pleased to provide comments on the List of Measures Under Consideration for December 21, 2020. Our members are all physicians who are board-certified in allergy and immunology and who are specialists in the care of asthma and allergic disease. Our comments are focused on the Asthma/COPD Episode-Based Cost Measure (MUC20-0015).

We appreciate that CMS is working to develop a cost measure for care of COPD and Asthma. However, as expressed in the feedback we provided to CMS on the Asthma/COPD Cost Measure Field Test Report, we have several concerns about the development of this measure and the abbreviated time for providing feedback. For this reason, we strongly urge that it undergo at least another year of testing before implementation. Providing one year of sample test reports to practices would allow allergists to get more familiar with the measure and would permit further refinement before the measure impacts payment. Rolling out new measures amid a pandemic when many physicians and their staff are already under enormous pressure as well as facing financial struggles is imprudent and threatens to overwhelm allergy practices, especially smaller groups which account for the majority of allergy practices.

As we explained in detail in our feedback on the Field Test Report, there are a number of areas where refinement of the Asthma/COPD measure is needed. These include:

• Addressing the fact that many patients are incorrectly diagnosed with asthma by non-specialists and receive inappropriate and often expensive treatment before referral to an asthma specialist who determines the patient does not have asthma.
• Ensuring that the risk adjustment methodology takes into consideration the fact that specialists typically provide care for patients with the highest disease severity.
• Adequately incorporating social determinants of health, which play a large role in asthma care, into the risk algorithm.

Moreover, allergists reported considerable difficulty understanding the Field Test Report and accompanying CSV files. In our comments on the Field Test Report, we emphasized the need for more granular data to make the Report meaningful and actionable.

For the above reasons, we urge that CMS delay implementation and allow for another year of testing before the Asthma/COPD measure is implemented.

Sincerely,

Luz S. Fonacier, MD, FACAAI
President, American College of Allergy, Asthma & Immunology

James Tracy, DO, FACAAI

	Chair, Advocacy Council of ACAAI
Roji Health Intelligence	For each cost measure under consideration, physician practices or groups must have the ability to examine data on episodes for procedures and conditions that reveal how these episodes relate to the cost target. In all cases involving multi-physician and facility care, this will require data and transparency from CMS, rather than simply providing aggregate final results. With CMS proposed measures, physician groups can only react to measurement but lack the ability to improve because (a) they don't have the data to replicate the episode-based measure data, since much of the costs are indirectly generated by other providers yet invisible to the attributed group; and (b) all of the scores are retrospective and there is no actionable data to help them improve. Adoption of cost measures should be contingent on CMS provision of claims data on a regular ongoing basis (for example monthly) so that the group can create their own applications for CMS cost measure-related episodes, and evaluate the specific reasons for cost excess. Because CMS cost measures involve medical services beyond an individual practice or group's own EMR or billing systems, physician groups do not have the necessary data without receipt of claims details from CMS. Receipt of that data would allow organizations the ability to make modifications to improve their cost-effectiveness. This is analogous to what is currently happening with Accountable Care Organizations. Roji Health Intelligence is a qualified CMS reporting registry. We have created Episode-Based cost measures for our physician practice and group clients based on the data present in their practice management and electronic health records. This information is presented in an on-line interactive format for illumination and comparison of episodes, cost drivers, and outcomes associated with episodes. The purpose is to engage clinicians and practices in understanding and acting on cost and outcome results for patients in episodes of care. For our cost measures to be most valuable, we suggest that our clients' receipt of CMS claims data would ensure that we could assist them in a more complete understanding of variation in costs and interventions that would lower health care expenditures.

Measure Comments (Post-Workgroup Meeting)

Author	Submitted Comment
Federation of American Hospitals	Support
American Medical Association	The American Medical Association (AMA) continues to have significant concerns with this measure and believes that revisions to the specifications and additional field testing are necessary. We ask that additional conditions be provided with this recommendation around ensuring that the minimum reliability rate be 0.7 or higher, evaluating costs within the context of the quality of care provided is completed, and removing Part D prescription drug costs.

Association of American Medical Colleges (AAMC)	The Clinician MAP conditionally supported two of the episode-based costs measures (Colon and Rectal Resection [MUC2020-0016] and Melanoma Resection [MUC2020-0018]) for future rulemaking for the Merit-based Incentive Payment System (MIPS) program subject to NQF endorsement. The Clinician MAP did not support the other three episode-based costs measures (Asthma/Chronic Obstructive Pulmonary Disease [MUC20-0015], Diabetes [MUC2020-0017], and Sepsis [MUC2020-0019]) for future rulemaking for the MIPS program with potential for mitigation. Mitigation for the three measures focused on evaluation of the actionability and connection between upstream medical interventions and downstream costs, in addition to NQF endorsement. The AAMC agrees with concerns about episode-based cost measures relying on the suggestion that providing certain upstream preventions will result in lower costs of care, and that lower costs will result in better patient outcomes. Furthermore, the AAMC remains concerned that none of the 13 cost measures are adjusted to account for social risk factors (SRFs). In addition to patient clinical complexity, SRFs can drive differences in average costs. In particular, physicians at academic medical centers (AMCs) care for vulnerable populations of patients who are sicker, poorer, and more complex than patients treated elsewhere. In regard to attribution – AAMC has previously commented that attribution methods used should be clear and transparent to clinicians and that it is critical that there be an accurate determination of the relationship between a patient and a clinician to ensure that the correct clinician is held responsible for the patient's outcomes and costs. Attribution is complicated, given that most patients receive care from numerous clinicians across several facilities, and AAMC has urged CMS to explore better data sources and analytic techniques to support more accurate attribution. In addition, the movement in medicine has been to team-based care, further complicating appropriate attribution to a single clinician. The MAP, through its recommendations, and CMS should be careful not to incent patterns of care that are outdated. The AAMC recommends that the MAP recommendation be "do not support with potential for mitigation" for each of the episode-based cost measures.
American Academy of Otolaryngic Allergy	On behalf of the American Academy of Otolaryngic Allergy, we support the American College of Allergy, Asthma, and Immunology (ACAAI) in recommending a delay in implementation for the Asthma/COPD Cost Measure to allow for more test data to be considered.
AdvaMed	AdvaMed strongly supports this measure, as it would provide useful information to support the development of new algorithms for early diagnosis and therapeutic guidance for COPD.

MUC20-0016: Colon and Rectal Resection Episode-Based Cost Measure

Measure Information

Characteristic	Submitted Information
MUCID	MUC20-0016
Other Measure Identification Numbers	N/A
Title	Colon and Rectal Resection Episode-Based Cost Measure
Program	Merit-based Incentive Payment System-Cost
Workgroup	Clinician
In what state of development is the measure?	Early Development, Field Testing
State of Development Details	The measure has been developed, field tested, and is now refined based on feedback received from field testing in summer 2020 (August 17 to September 18, 2020). As background, a list of draft episode groups and trigger codes, developed with input from a Clinical Committee convened in 2016, was posted by CMS in December 2016 to meet MACRA requirements. Building off this work, episode-based cost measures are developed using a “wave” approach wherein sets of Clinical Subcommittees (CS) are convened to select episode groups to develop into cost measures, while smaller, measure-specific Clinician Expert Workgroups provide detailed input on each component of the measures. The current wave of measure development began in May 2019 and includes 4 CS with a total of 166 members affiliated with 110 professional societies. The General and Colorectal Surgery CS selected this episode group from the December 2016 draft list of episode groups for development. This CS comprises members representing clinician specialty societies in this clinical area. A Clinician Expert Workgroup met in August 2019 to discuss measure specifications for all components of the measure, followed by a webinar in January 2020 for follow-up discussions on service assignment and risk adjustment. The measure was field tested in the summer of 2020, during which clinicians and stakeholders learned about the measure and provide input on the draft cost measure specifications. After field testing, the Clinician Expert Workgroups revisited and refine the draft measure specifications based on the stakeholder feedback received. We conducted reliability testing of measures for clinicians (TIN-NPIs) and for clinician groups (TINs), constructed using episodes ending between January 1, 2019 and December 31, 2019. Reliability refers to the extent to which a measure reflects true variation between clinicians’ risk-adjusted episode spending, as opposed to random variation. The reliability metric specifically captures how much of the variance in a measure is due to systematic

	differences in episode spending between clinicians, rather than differences in episode spending within a clinician's set of episodes. A measure with high reliability suggests that the comparisons of performance across clinicians can be expected to better reflect systematic differences in actual performance. Our testing results indicate that this measure is reliable for clinicians and clinician groups across a range of case minimums. • For TINs at a 10-episode case minimum, the mean reliability was 0.44. For TINs at a 20-episode case minimum, the mean reliability was 0.56. For TINs at a 30-episode case minimum, the mean reliability was 0.63. For TIN-NPIs at a 10-episode case minimum, the mean reliability was 0.33. For TIN-NPIs at a 20-episode case minimum, the mean reliability was 0.45. For TIN-NPIs at a 30-episode case minimum, the mean reliability was 0.54.
Measure Description	The Colon and Rectal Resection cost measure evaluates clinician or clinician group's risk-adjusted cost to Medicare for patients who receive colon or rectal resections for either benign or malignant indications. The measure score is a clinician or clinician group's average risk-adjusted cost for the episode group across all attributed episodes. This inpatient procedural measure includes services that are clinically related and under the reasonable influence of the attributed clinician or clinician group during the 15 days prior to the clinical event that opens or "triggers" the episode through 90 days after. Medicare beneficiaries enrolled in Medicare Parts A and B during the performance period are eligible for the measure.
Numerator	The numerator for the Colon and Rectal Resection measure is the sum of the ratio of observed to expected payment-standardized cost to Medicare for all episodes attributed to a clinician or clinician group. This sum is then multiplied by the national average observed episode cost to generate a dollar figure. Mathematically, this is represented as: sum of (observed episode cost/expected episode cost) * national average observed cost.
Denominator	The denominator for the Colon and Rectal Resection measure is the total number of episodes from this episode group attributed to a clinician.
Exclusions	The following episode-level exclusions apply: (a) The beneficiary has a primary payer other than Medicare for any amount of time overlapping the episode window or in the lookback period. (b) No attributed clinician is found for the episode. (c) The beneficiary's date of birth is missing. (d) The beneficiary's death date occurred before the episode ended. (e) The beneficiary was not enrolled in Medicare Parts A and B for the entirety of the lookback period plus episode window, or is enrolled in Part C for any part of the lookback period plus episode window. (f) The episode trigger claim was not performed in an office, IP, OP, or ASC setting based on its place of service. Exclusions specific to the Colon and Rectal Resection measure are developed with input from the Colon and Rectal Resection Clinician Expert Workgroup.

Measure type	Cost/Resource Use
What is the NQF status of the measure?	Never submitted
NQF ID number	0000
Year of next anticipated NQF CDP endorsement review	NA
Year of most recent NQF Consensus Development Process (CDP) endorsement	NA
Is the measure being submitted exactly as endorsed by NQF?	NA
If not exactly as endorsed, describe the nature of the differences	N/A
What data sources are used for the measure?	Claims
If EHR or Administrative Claims or Chart-Abstracted Data, description of parts related to these sources.	N/A
At what level of analysis was the measure tested?	The measure was testesd at both the Clinician: Individual and Clincian: Group levels.
In which setting was this	Ambulatory surgery center, Hospital outpatient department (HOD), Hospital

measure tested?	inpatient acute care facility
What NQS priority applies to this measure?	NA
What one primary meaningful measure area applies to this measure?	Patient-focused episode of care
What secondary meaningful measure area applies to this measure?	NA
What one primary healthcare priority applies to this measure?	Make care affordable
What secondary healthcare priority applies to this measure?	NA
What area of specialty best fits the measure?	Colorectal surgery
What is the target population of the measure?	Medicare Fee for Service
Is this measure an eCQM?	No
If eCQM, enter Measure Authoring Tool (MAT) number	No
If eCQM, does the measure have a Health Quality	No

Measures Format (HQMF) specification?	
Comments	N/A
Measure steward	Centers for Medicare & Medicaid Services
Long-Term Measure Steward (if different)	N/A
Measure Steward Contact Information	Evans, Ronique; Center for Clinical Standards and Quality; 410-786-3966; Ronique.Evans1@cms.hhs.gov
Primary Submitter Contact Information	Jensen, Ross; Acumen, LLC; 650-558-8882; macra-episode-based-cost-measures-info@acumenllc.com
Long-Term Measure Steward Contact Information	N/A
Secondary Submitter Contact Information	N/A
Was this measure proposed for a previous year's MUC list?	No
In what prior year(s) was this measure proposed?	NA
What were the programs that NQF MAP reviewed the measure for in each year?	N/A

Why was the measure not recommended in those year(s)?	N/A
What were the MUC IDs for the measure in each year?	N/A
NQF MAP report page number being referenced for each year	N/A
What was the NQF MAP recommendation in each year?	N/A
List the NQF MAP workgroup(s) in each year	N/A
What is the history or background for including this measure on the new MUC list?	New measure never reviewed by MAP Workgroup or used in a CMS program
Range of years(s) this measure has been used by CMS Program(s)	N/A
What other federal programs are currently using this measure?	
Evidence that the measure can be operationalized	This is a claims-based measure and will not require any additional submission of data.

How is the measure expected to be reported to the program?	Claims
Is this measure similar to and/or competing with measure(s) already in a program?	No
Which existing measure(s) is your measure similar to and/or competing with?	N/A
How will this measure be distinguished from other similar and/or competing measures?	N/A
Rationale for how this measure will add to the CMS program	N/A
If this measure is being proposed to meet a statutory requirement, please list the corresponding statute.	Section 101(f) of MACRA
Evidence of performance gap	Evidence of Performance Gap response: A recent study indicates that clinician beliefs about treatment may be the most important factors explaining the variation in health care expenditures. However, clinicians are often unaware of how their care decisions influence the overall costs of care. Cost measures are intended to help inform clinicians on the costs attributed to their decision-making and to incentivize cost-effective, high-quality care. A cost measure offers opportunity for improvement if clinicians can exercise influence on the intensity or frequency of a significant share of

costs during the episode, or if clinicians can achieve lower spending and better care quality through changes in clinical practice. According to the literature and feedback received through stakeholder input activities to date, this measure's focus represents an area where there are opportunities for improvement. These include minimizing risks associated with clinicians' approach to performing a colon or rectal resection and the adoption of prevention strategies to mitigate the risk of common postoperative complications. A clinician's approach to performing a colorectal surgery has a significant impact on patient outcomes. Colorectal surgery can be performed using three different modalities: open, laparoscopic, and robotic. The benefits of performing colon resections laparoscopically or robotically are well established. These minimally invasive approaches are associated with reduced lengths of stay, reduced utilization of post-acute care, lower postoperative readmission rates, and lower mortality rates, especially among the older adult population. Although the use of such techniques may be more limited in scope for rectal resections due to added technical complexity, recent studies indicate that these techniques may also have a role in reducing postoperative complications following surgery for rectal cancer treatment. Specifically, studies and reviews of meta-analyses have demonstrated that robotic or laparoscopic surgery for rectal cancer treatment may reduce the incidence of postoperative complications when compared to open surgery. There remains wide variation in the utilization of open surgery, laparoscopic, and robotic approaches for different diagnoses. In 2012, open surgeries constituted 65.4% of all colorectal surgeries nationwide, while 31.2% and 3.4% were performed using laparoscopic or robotic techniques, respectively. Efforts to increase adoption of minimally invasive techniques, when appropriate, through surgeon education and training could be effective strategies to curb costs associated with prolonged lengths of stay and readmission. Colorectal resection accounts for a substantial share of postoperative readmissions among inpatient procedures, with one study approximating a 30-day postoperative readmission rate of 13.7%. Estimates of the inpatient cost for readmission following colorectal surgery range from $9,000 to $12,000 across studies. One study estimates that readmissions associated with colorectal surgery account for approximately $300 million in costs annually across the nation. Postoperative readmission is strongly associated with the occurrence of common complications such as surgical site infection (SSI), ileus, and urinary tract infections. Occurrence of SSI alone is estimated to contribute an additional estimated cost of $40,500 per patient and an estimated national total of $3 billion per year. Applying prevention strategies to emergency colorectal surgeries based on clinical guidelines for an "Enhanced Recovery After Surgery" (ERAS) protocol can decrease these post-operative complications and reduce morbidity. ERAS is a standard of perioperative care for elective colorectal surgeries; however, there appears to be low implementation of an ERAS protocol in emergent settings. This may be due to the fact that patients undergoing emergent surgeries have more risk factors and comorbidities that must be managed. Expanding the implementation of ERAS protocols has the potential to improve overall

quality of care and reduce related services and their associated costs. A diverting stoma, in which a surgeon “diverts the flow of the feces externally”, may be another avenue to mitigate common complications such as anastomotic leaks and the associated costs. Although there are benefits and tradeoffs to fecal diversion to protect an anastomosis, certain factors may indicate cases in which a diverting stoma may be the preferred surgical approach. For example, there is generally a consensus among researchers that the presence of a diverting stoma lowers the risk of anastomotic leak and can lower the risk of developing pelvic sepsis for patients who undergo a low anterior resection. Since the risks associated with diverting stomas are well documented, preventative pathways have been developed to address the potential for dehydration and other common causes of readmission due to colorectal surgeries. For example, one study reported reducing the rate of hospital readmissions and entirely eliminating readmissions related to dehydration by employing an educational intervention for patients with new, temporary or permanent ileostomies. This suggests that coupling diverting stomas with robust patient education may result in improved outcomes following colorectal surgery. Fecal diversion is also demonstrated to have a protective effect in terms of decreased mortality and morbidity for other high-risk cases. For example, recent studies have identified primary anastomosis with diversion as the preferred option for cases with active infections, such as peritonitis from diverticular disease, compared to Hartmann’s procedure. As such, diversions may play an important role in improving outcomes and reducing associated downstream costs for select high-risk colorectal cases.

This measure aims to address these example areas of opportunities for improvement. Research has shown that a single colectomy is estimated to cost $25,000 which can increase to nearly $50,000 with post-operative complications and complications related to diverticular disease with colectomy procedures account for more than $2 billion in treatment costs annually. [3,4,6,7] While diverticular disease is usually an asymptomatic condition, the incidence of complications such as colonic diverticulitis increases with age As such, the use of this cost measure can provide clinicians with information to improve care outcomes and reduce future health care costs. This is supported by our analysis into the variation in spending per episode that may be reduced through better clinical practices. We analyzed the measure’s performance gap for clinicians (TIN-NPIs) and clinician groups (TINs) using episodes ending between January 1, 2019 and December 31, 2019. With no case minimum applied, there were 54,626 Colon and Rectal Resection episodes for 54,414 beneficiaries. The TIN-NPI and TIN-level measure scores were calculated for 1,921 clinicians and 1,398 clinician groups who met the 10-episode case minimum.

- The mean risk-adjusted cost per episode at the TIN level was $25,281. The risk-adjusted cost per episode at the 10th percentile was $22,874, compared to $28,008 at the 90th percentile at the TIN level.
- The mean risk-adjusted cost per episode at the TIN-NPI level was $25,025. The risk-adjusted cost per episode at the 10th percentile

	was $22,268, compared to $28,132 at the 90th percentile at the TIN-NPI level. • For TINs, the mean measure score was 1.02. The score at the 10th percentile was 0.92, compared to 1.13 at the 90th percentile. For TIN-NPIs, the mean measure score was 1.01. The score at the 10th percentile was 0.90, compared to 1.13 at the 90th percentile.
Unintended consequences	Stakeholders have expressed concerns that cost measures could potentially have risks or unintended consequences. During past stakeholder input activities, including Technical Expert Panels (TEPs) and public comment periods, stakeholders have highlighted the importance of considering potential incentives created by cost measures that could adversely affect patients and clinicians. For patients, these potential adverse effects may include: (i) reducing access to care, and (ii) stinting care. For clinicians, these adverse effects could include: (i) making unfair comparisons among certain types of clinicians in assessing cost measure performance, and (ii) rewarding lower costs without appropriately balancing performance on quality. Overall, by promoting lower spending, stakeholders have expressed concerns that cost measures could potentially create barriers to patients receiving clinically appropriate care. To mitigate the potential risks of such incentives, stakeholders and TEP members have highlighted the importance of the following: Devising an appropriate risk adjustment model for episode-based cost measures; Aligning cost measures with indicators of quality; Identifying benchmarks for comparison that are fair and appropriate for all eligible clinicians; Potentially excluding certain types of patients from measure calculations.
Which clinical guideline(s)?	N/A
Briefly describe the peer reviewed evidence justifying this measure	Colorectal resection, or colectomy, is a common treatment for colorectal cancer and complications related to diverticular disease. According to the Agency for Healthcare Research and Quality, about 320,000 colorectal resection procedures were performed annually between 2001 and 2011. Colorectal cancer is the second leading cause of cancer-related deaths and the third most common cancer in both men and women in the United States. Colorectal cancer is especially common in the 85 and older adult population, with an incidence of 237 per 100,000 persons in 2016. Similarly, diverticular disease primarily affects older adults, occurring in 50-70% of those aged 80 or older. Diverticular disease accounts for more than $2 billion in treatment costs annually. While diverticular disease is usually an asymptomatic condition, the incidence of complications such as colonic diverticulitis increases with age. Morbidity and the risk of postoperative complications following colorectal resection also increase significantly for patients above age 65. According to the literature, a single colectomy is estimated to cost $25,000, and this cost can increase to nearly $50,000 with post-operative complications. Estimates of index hospitalization costs for colorectal surgery are similar and have been shown to range between about $18,000 to $21,000 among a cohort of Medicare beneficiaries, with variation in the cost

of care provided within a year of the surgery largely driven by readmissions and post-acute care utilization. Given the costs and frequency of treating colorectal cancer and complications related to diverticular disease with colectomy procedures in Medicare beneficiaries, the Colon and Rectal Resection cost measure represents an opportunity for improvement on overall cost performance.

Preliminary Analysis – MUC20-0016: Colon and Rectal Resection Episode-Based Cost Measure

Criteria	Yes/No	Justification and Notes
Does the measure address a critical quality objective not currently adequately addressed by the measures in the program set?	Yes	MUC20-0016 addresses the Patient-Focused Episode of Care goal of CMS's Meaningful Measures initiative, the MIPS high priority area of Efficiency/Cost Reduction and the 2015 Medicare Access and Children's Health Insurance Program (CHIP) Reauthorization Act (MACRA) statutory requirements (section 101(f)). Better colon and rectal resection decision-making can impact a patient's recovery time and can decrease risks of postoperative complications. Currently there are 8 MIPS quality measures related to the identification of or procedures for colorectal cancer. Of the current measures, none look at episode-based cost related to colectomy decision making.
Is the measure evidence-based and either strongly linked to outcomes or an outcome measure?	Yes	This cost/resource use measure aims to inform clinical decision-making related to colon and rectal resection by reflecting cost of an episode of care and minimizing risks with approaches to colorectal surgery. Studies suggest that knowledge and awareness of evidence-based practices and treatment risks can influence decision-making and can lead to lower costs (Sacks, et al., 2016; Cutler, et al., 2016).

Does the measure address a quality challenge?	Yes	Colorectal cancer represents 8.2% of all cancer diagnoses and impacts nearly 150,000 patients per year (NIH, 2020). Colectoral resection accounts for 14.7% of in-patient readmissions (Bliss, et al., 2015) and average $300 million in readmission costs annually (Wick, et al, 2011). Research suggests that a clinician's decision on modality of surgery can result in post-operative complications and higher costs. Unlike open colectomy, laparoscopic colectomy is associated with shorter hospital stays, lower risk of post-operative complications, and lower cost due to lower complication rates (Fitch, et al., 2017; Flynn, et al, 2014, Sheetz, et al, 2017) as well as quicker bowel function post-surgery, lower morbidity rates, reduced pain and overall better quality of life for patients (Alsowaina, et al, 2019).
Does the measure contribute to efficient use of measurement resources and/or support alignment of measurement across programs?	Yes	This measure is not duplicative of other measures currently within the MIPS program. This and other 2020 MIPS cost measures under consideration were created in response to MACRA requirements to develop measures for potential implementation in the cost performance category of MIPS. Additionally, this measure is an episode-based measure; current cost-based measures in MIPS are all at the population level. While no colorectal surgery cost measures were identified outside of MIPS, colorectal screening measures were found in multiple CMS quality programs including the Marketplace Quality Ratings System, Part C & D Stars, Shared Savings Program, MIPS and Promoting Interoperability.
Can the measure can be feasibly reported?	Yes	The measure uses Medicare claims data which is feasibly reported and a low-burden data source.
Is the measure applicable to and appropriately specified for the program's intended care setting(s), level(s) of analysis, and population(s)?	Yes	The measure is specified for Clinician: Individual and Clinician: Group/Practice levels, which aligns with MIPS reporting categories. The measure has not been reviewed for endorsement by an NQF Standing Committee.

If the measure is in current use, have negative unintended issues to the patient been identified? Have implementation challenges outweighing the benefits of the measure been identified?	N/A	The measure is new and not currently in use. The submission identified a potential unintended consequence of reduction in access to care or stinting care. This measure could result in rewarding cost reductions with no net quality of care gain if it is not appropriately balanced with measures addressing performance on quality.
PAC/LTC Core Concept?		N/A
Impact Act Domain		N/A
Hospice High Priority Areas		N/A

Rural Workgroup Input	[redacted]	Relative priority/utility: • This measure was suggested to be less applicable in rural settings with potential for low volume concerns. Data collection issues: • None identified Calculation issues: • Average cost for rural providers may exceed the national average cost and therefore may unduly impact rural providers. • Combining of benign and malignant conditions may not be appropriate. • 30 day length pre-surgery period was suggested as potentially better than a 15 day. • There were concerns expressed related to measure reliability due to low volume of events. Unintended consequences: • Concerns were noted for higher potential to penalize rural providers due to tendency to catch cancers in later stages. • It was also noted that cost measures may result in restriction of care if not paired with quality measures. Votes: Range is 1 – 5, where higher is more relevant to rural. Average: 3.2 • 1 – 0 vote • 2 – 6 votes • 3 – 5 votes • 4 – 7 vote • 5 – 1 vote
Preliminary Analysis Recommendation	Conditional Support for Rulemaking	Conditional support for rulemaking is contingent on NQF endorsement.
Summary: What is the potential value to the program measure set?	[redacted]	MUC20-0016 addresses the Patient-Focused Episode of Care goal of CMS's Meaningful Measures initiative, the CMS high priority area of Efficiency/Cost Reduction and MACRA statutory requirements. Currently, there are no measures that assess episode-based cost related to colectomy.

Summary: What is the potential impact of this measure on quality of care for patients?	[redacted]	Colorectal cancer represents 8.2% of all cancer diagnoses, impacting nearly 150,000 patients per year (NIH, 2020). Evidence suggest that surgical decision making and treatment course related to colon and rectal resection can reduce length of hospital stay, risk of major post-operative complications, and cost. Should testing data show that the measure appropriately measures episode-based cost while maintaining quality, this measure would be valuable to add to the program measure set. Conditional support for rulemaking is contingent on NQF endorsement.

Measure Comments

Author	Submitted Comment
University of Colorado Medicine	More clarity needed for the verbiage located in the measure description that states "This inpatient procedural measure includes services that are...'under the reasonable influence of the attributed clinician'..."
American Medical Association	Overall, the AMA supports the episode-based measure development process and movement to episode-based measures over broad cost measures. However, we continue to have significant concerns with the lower than desirable mean reliability rate. The AMA believes that the minimum acceptable thresholds should be 0.7. In addition, we strongly support the tenet that cost must be assessed within the context of the quality of care provided; yet, none of these assessment have been provided for the existing MIPS cost measures or the ones currently under MAP consideration. We request further refinements to the process to ensure the cost measures are accurate and fair when used in MIPS and, ultimately, publicly reported. The AMA strongly opposes including Part D prescription drug costs in the Medicare episode-based cost measures. To hold physicians accountable for costs that are negotiated between CMS and Medicare Prescription Drug Plans is fundamentally problematic, and physicians and patients do not always have information about coverage, formularies, out-of-pocket costs and list prices at the point of care. This also assumes there is a viable, evidence-based, less expense alternative option for patients. In general, we urge the MAP to make “Conditional Support” its highest level of recommendation and to recommend “Do Not Support with Potential for Mitigation” for any measures where relevant specialties have raised serious concerns or the average reliability rate is less than 0.7 as is the case with several of these measures.
Association of American Medical Colleges (AAMC)	The AAMC remains concerned that such cost measures are adjusted to account for social risk factors (SRFs). In addition to patient clinical complexity, SRFs can drive differences in average costs. In particular, physicians at academic medical centers (AMCs) care for vulnerable populations of patients who are sicker, poorer, and more complex than patients treated elsewhere. In regard to attribution – AAMC has previously commented that attribution methods used should be clear and transparent to clinicians and that it is critical that there be

	an accurate determination of the relationship between a patient and a clinician to ensure that the correct clinician is held responsible for the patient's outcomes and costs. Attribution is complicated, given that most patients receive care from numerous clinicians across several facilities, and AAMC has urged CMS to explore better data sources and analytic techniques to support more accurate attribution. The AAMC recommends that: (1) cost measures include risk-adjustment for SRFs, (2) the attribution methodology is transparent, and (3) the appropriate clinician is held responsible for the patient's outcomes and costs. Additionally, this measure has not been submitted for NQF endorsement. The AAMC has long held that measures should not be proposed for addition to public reporting programs unless vetted and endorsed by the NQF. The AAMC recommends that the highest level of MAP recommendation be "Do Not Support With Potential For Mitigation."
AdvaMed	AdvaMed strongly supports this measure and urges that total cost include addressing anastomotic leaks ((a major complication of colorectal surgery). This measure as worded appears to capture the cost of a leak, as these are usually discovered quickly post-operatively, but should be clarified accordingly. This is an important clarification as the cost of managing a leak is usually 5-10x more than the initial surgery itself.
American College of Surgeons	On behalf of the over 80,000 members of the American College of Surgeons (ACS), we appreciate the opportunity to submit comments to the Measure Applications Partnership (MAP). The ACS is a scientific and education association of surgeons founded in 1913 to improve the quality of care for the surgical patient by setting high standards for surgical education and practice. ACS has a vested interest in CMS' MAP and the CMS Measures Under Consideration (MUC) list because of our dedication to improving the value of care for surgical patients. With our 100-year history in developing quality programs to optimize the delivery of surgical services, we believe that we can offer valuable insight to the MAPs deliberations. Acumen Episode-Based Cost Measures ACS has advocated extensively in prior letters to CMS on the need to measure cost and quality over the same episode of care in order to achieve higher value. Furthermore, for the Cost category of MIPS to be meaningful, the measures used must be not only reliable, but also actionable. That is, they should provide information on how a physician or care team currently uses resources and allow for comparisons with others who may be more efficient. ACS does not believe that the Acumen process for developing episode-based cost measures is structured in a way to truly measure cost and does not result in measures that are actionable. CMS will unlikely be able to reach their goals for increased cost accountability with the current Acumen measures. For both Wave 1 and Wave 2, Acumen presented a single, basic framework that all clinical subcommittees were required to follow to develop cost measures. While

	members of the clinical subcommittees appreciated the opportunity to define the length of the episode and decide on trigger codes, exclusions and risk adjusters, they found it concerning that Acumen and CMS had already determined the general framework for measuring physician costs and felt that they had little say over whether this was the most appropriate strategy. Physician cost, quality, and overall value cannot be evaluated using a one-size-fits-all approach. Procedures and patient populations are vastly different and cannot always be evaluated for appropriateness in the same manner. Another ongoing problem with the Acumen cost measure development process is that it relies exclusively on claims data. The limitations of administrative data interfere with the critical tasks of risk stratification, subgrouping, and defining accurate and appropriate inclusion and exclusion criteria. Throughout the Acumen process there have also been concerns about applying the CMS-Hierarchical Condition Categories (HCCs) risk adjustment methodology to these episode-based cost measures. The CMS-HCCs were not designed to risk adjust narrowly defined patient cohorts such as episode groups and is a poor performer even for its intended purpose of determining payment to Medicare Advantage plans based on traditional Medicare cost benchmarks. Another drawback of the Acumen cost measure development approach is that the process fails to account for the impact that cost reduction may have on patient outcomes or other measures of quality. It is virtually impossible for clinicians to evaluate, provide meaningful feedback on, and find significance in cost performance data when it is presented in a vacuum with no consideration of quality. Even if relevant MIPS quality metrics do not exist, some of the services included in the post-trigger period of some of the surgical episodes are tied to complications (e.g., need for re-operation, treatment of infection, DVT, etc.) so it should be relatively easy for CMS to separately provide data on the rates of these complications for a given surgeon and to benchmark that data against peers. Finally, although some improvements have been made, the timeline for cost measure development and field testing is still rushed and driven by arbitrary timelines that focus exclusively on the goal of measure implementation (i.e., the MUC/rulemaking process), rather than actually getting the measure right. Clinicians still find the field test reports challenging to access and cumbersome to navigate/digest. In summary, we do not believe this process yields measure that are designed to be an incentive to optimize the cost of healthcare.
Roji Health Intelligence	For each cost measure under consideration, physician practices or groups must have the ability to examine data on episodes for procedures and conditions that reveal how these episodes relate to the cost target. In all cases involving multi-physician and facility care, this will require data and transparency from CMS, rather than simply providing aggregate final results. With CMS proposed measures, physician groups can only react to measurement but lack the ability to improve because (a) they don't have the data to replicate the episode-based measure data, since much of the costs are indirectly generated by other providers yet invisible to the attributed group; and (b) all of the scores are retrospective and there is no actionable data to

help them improve.

Adoption of cost measures should be contingent on CMS provision of claims data on a regular ongoing basis (for example monthly) so that the group can create their own applications for CMS cost measure-related episodes, and evaluate the specific reasons for cost excess. Because CMS cost measures involve medical services beyond an individual practice or group's own EMR or billing systems, physician groups do not have the necessary data without receipt of claims details from CMS. Receipt of that data would allow organizations the ability to make modifications to improve their cost-effectiveness. This is analogous to what is currently happening with Accountable Care Organizations.

Roji Health Intelligence is a qualified CMS reporting registry. We have created Episode-Based cost measures for our physician practice and group clients based on the data present in their practice management and electronic health records. This information is presented in an on-line interactive format for illumination and comparison of episodes, cost drivers, and outcomes associated with episodes. The purpose is to engage clinicians and practices in understanding and acting on cost and outcome results for patients in episodes of care. For our cost measures to be most valuable, we suggest that our clients' receipt of CMS claims data would ensure that we could assist them in a more complete understanding of variation in costs and interventions that would lower health care expenditures.

Measure Comments (Post-Workgroup Meeting)

Author	Submitted Comment
American Association of Nurse Anesthetists	Providers need information about the applicability of specific measures to their practices. Would this measure be attributed to specific provider or clinician types, such as surgeon, attending physician or anesthesia professional?
Federation of American Hospitals	The Federation of American Hospitals (FAH) recommends that the minimum reliability threshold should 0.7 or higher both at the individual clinician and practice levels and believes that validity testing must demonstrate the presence or absence of correlations of this cost measure to one or more quality measures prior to use in MIPS. As a result, FAH requests that the highest level of MAP recommendation be "Do Not Support with Potential for Mitigation."
American Medical Association	The American Medical Association (AMA) continues to have significant concerns with this measure and believes that revisions to the specifications and additional field testing are necessary. We ask that additional conditions be provided with this recommendation around ensuring that the minimum reliability rate be 0.7 or higher, evaluating costs within the context of the quality of care provided is completed, and removing Part D prescription drug costs. The AMA requests that the highest level of MAP recommendation be "Do Not Support with Potential for Mitigation."
AdvaMed	AdvaMed strongly supports this measure and urges that total cost include addressing anastomotic leaks ((a major complication of colorectal surgery). This

	measure as worded appears to capture the cost of a leak, as these are usually discovered quickly post-operatively, but should be clarified accordingly. This is an important clarification as the cost of managing a leak is usually 5-10x more than the initial surgery itself.
Association of American Medical Colleges (AAMC)	The Clinician MAP conditionally supported two of the episode-based costs measures (Colon and Rectal Resection [MUC2020-0016] and Melanoma Resection [MUC2020-0018]) for future rulemaking for the Merit-based Incentive Payment System (MIPS) program subject to NQF endorsement. The Clinician MAP did not support the other three episode-based costs measures (Asthma/Chronic Obstructive Pulmonary Disease [MUC20-0015], Diabetes [MUC2020-0017], and Sepsis [MUC2020-0019]) for future rulemaking for the MIPS program with potential for mitigation. Mitigation for the three measures focused on evaluation of the actionability and connection between upstream medical interventions and downstream costs, in addition to NQF endorsement. The AAMC agrees with concerns about episode-based cost measures relying on the suggestion that providing certain upstream preventions will result in lower costs of care, and that lower costs will result in better patient outcomes. Furthermore, the AAMC remains concerned that none of the 13 cost measures are adjusted to account for social risk factors (SRFs). In addition to patient clinical complexity, SRFs can drive differences in average costs. In particular, physicians at academic medical centers (AMCs) care for vulnerable populations of patients who are sicker, poorer, and more complex than patients treated elsewhere. In regard to attribution – AAMC has previously commented that attribution methods used should be clear and transparent to clinicians and that it is critical that there be an accurate determination of the relationship between a patient and a clinician to ensure that the correct clinician is held responsible for the patient's outcomes and costs. Attribution is complicated, given that most patients receive care from numerous clinicians across several facilities, and AAMC has urged CMS to explore better data sources and analytic techniques to support more accurate attribution. In addition, the movement in medicine has been to team-based care, further complicating appropriate attribution to a single clinician. The MAP, through its recommendations, and CMS should be careful not to incent patterns of care that are outdated. The AAMC recommends that the MAP recommendation be "do not support with potential for mitigation" for each of the episode-based cost measures.

MUC20-0017: Diabetes Episode-Based Cost Measure

Measure Information

Characteristic	Submitted Information
MUCID	MUC20-0017
Other Measure Identification Numbers	N/A
Title	Diabetes Episode-Based Cost Measure
Program	Merit-based Incentive Payment System-Cost

Workgroup	Clinician
In what state of development is the measure?	Early Development, Field Testing
State of Development Details	The measure has been developed, field tested, and is now refined based on feedback received from field testing in summer 2020 (August 17 – September 18, 2020). As background, a list of draft episode groups and trigger codes, developed with input from a Clinical Committee convened in 2016, was posted by CMS in December 2016 to meet MACRA requirements. Building off this work, episode-based cost measures are developed using a "wave" approach wherein sets of Clinical Subcommittees (CS) are convened to select episode groups to develop into cost measures, while smaller, measure-specific Clinician Expert Workgroups provide detailed input on each component of the measures. The current wave of measure development began in May 2019 and includes 4 CS with a total of 166 members affiliated with 110 professional societies. The Chronic Conditions and Disease Management CS selected this episode group from the December 2016 draft list of episode groups for development. This CS comprises members representing clinician specialty societies in this clinical area. A Clinician Expert Workgroup met in August 2019 to discuss measure specifications for all components of the measure, followed by a webinar in January 2020 for follow-up discussions on service assignment and risk adjustment. The measure was field tested in the summer of 2020, during which clinicians and stakeholders learned about the measure and provided input on the draft cost measure specifications. After field testing, the Clinician Expert Workgroups revisited and refined the draft measure specifications based on the stakeholder feedback received. We conducted reliability testing of measures for clinicians (TIN-NPIs) and for clinician groups (TINs), constructed using episodes ending between January 1, 2019 and December 31, 2019. Reliability refers to the extent to which a measure reflects true variation between clinicians' risk-adjusted episode spending, as opposed to random variation. The reliability metric specifically captures how much of the variance in a measure is due to systematic differences in episode spending between clinicians, rather than differences in episode spending within a clinician's set of episodes. A measure with high reliability suggests that the comparisons of performance across clinicians can be expected to better reflect systematic differences in actual performance. Our testing results indicate that this measure is reliable for clinicians and clinician groups across a range of case minimums. • For TINs at a 20-episode case minimum, the mean reliability was 0.60. For TINs at a 30-episode case minimum, the mean reliability was 0.64. For TINs at a 40-episode case minimum, the mean reliability was 0.68. For TIN-NPIs at a 20-episode case minimum, the mean reliability was 0.55.

	For TIN-NPIs at a 30-episode case minimum, the mean reliability was 0.59. For TIN-NPIs at a 40-episode case minimum, the mean reliability was 0.62.
Measure Description	The Diabetes cost measure evaluates a clinician group's risk-adjusted cost to Medicare for patients receiving medical care to manage type 1 or type 2 diabetes. The measure score is a clinician group's weighted average of risk-adjusted cost for each episode attributed to the clinician group, where each episode is weighted by the number of assigned days during the episode. This chronic measure includes services that are clinically related and under the reasonable influence of the attributed clinician group. Services are assigned during a Diabetes episode, which is a portion of the overall time period of a clinician group's responsibility for managing a patient's diabetes. Medicare beneficiaries enrolled in Medicare Parts A and B during the performance period are eligible for the measure. This measure addresses the Patient-Focused Episode of Care goal of CMS's Meaningful Measures Initiative, the CMS high priority area of Efficiency/Cost Reduction and MACRA statutory requirements.
Numerator	The numerator for the Diabetes measure is the sum of the ratio of observed to expected payment-standardized cost to Medicare for all episodes attributed to a clinician or clinician group. This sum is then multiplied by the national average observed episode cost to generate a dollar figure. Mathematically, this is represented as: sum of (observed episode cost/expected episode cost) * national average observed cost.
Denominator	The denominator for the Diabetes measure is the total number of episodes from this episode group attributed to a clinician or clinician group.
Exclusions	The following episode-level exclusions apply: (a) The beneficiary has a primary payer other than Medicare for any amount of time overlapping the episode window or in the lookback period. (b) No attributed clinician is found for the episode. (c) The beneficiary's date of birth is missing. (d) The beneficiary's death date occurred before the episode ended. (e) The beneficiary was not enrolled in Medicare Parts A and B for the entirety of the lookback period plus episode window, or is enrolled in Part C for any part of the lookback period plus episode window. (f) The episode trigger claim was not performed in an office, IP, OP, or ASC setting based on its place of service. Exclusions specific to the Diabetes measure are developed with input from the Diabetes Clinician Expert Workgroup.
Measure type	Cost/Resource Use
What is the NQF status of the measure?	Never submitted
NQF ID number	N/A
Year of next anticipated NQF CDP	N/A

endorsement review	
Year of most recent NQF Consensus Development Process (CDP) endorsement	N/A
Is the measure being submitted exactly as endorsed by NQF?	N/A
If not exactly as endorsed, describe the nature of the differences	N/A
What data sources are used for the measure?	Claims
If EHR or Administrative Claims or Chart-Abstracted Data, description of parts related to these sources.	N/A
At what level of analysis was the measure tested?	The measure was testesds at both the Clinician: Individual and Clinician: Group levels.
In which setting was this measure tested?	Ambulatory/office-based care, Hospital outpatient department (HOD)
What NQS priority applies to this measure?	
What one primary meaningful measure area	Patient-focused episode of care

applies to this measure?	
What secondary meaningful measure area applies to this measure?	
What one primary healthcare priority applies to this measure?	Make care affordable
What secondary healthcare priority applies to this measure?	
What area of specialty best fits the measure?	Family Medicine
What is the target population of the measure?	Medicare Fee for Service
Is this measure an eCQM?	No
If eCQM, enter Measure Authoring Tool (MAT) number	No
If eCQM, does the measure have a Health Quality Measures Format (HQMF) specification?	No
Comments	N/A
Measure steward	Centers for Medicare & Medicaid Services

Long-Term Measure Steward (if different)	N/A
Measure Steward Contact Information	Evans, Ronique; Center for Clinical Standards and Quality; 410-786-3966; Ronique.Evans1@cms.hhs.gov
Primary Submitter Contact Information	Jensen, Ross; Acumen, LLC; 650-558-8882; macra-episode-based-cost-measures-info@acumenllc.com
Long-Term Measure Steward Contact Information	N/A
Secondary Submitter Contact Information	N/A
Was this measure proposed for a previous year's MUC list?	No
In what prior year(s) was this measure proposed?	N/A
What were the programs that NQF MAP reviewed the measure for in each year?	N/A
Why was the measure not recommended in those year(s)?	N/A
What were the MUC IDs for the measure in each year?	N/A

NQF MAP report page number being referenced for each year	N/A
What was the NQF MAP recommendation in each year?	N/A
List the NQF MAP workgroup(s) in each year	N/A
What is the history or background for including this measure on the new MUC list?	New measure never reviewed by MAP Workgroup or used in a CMS program
Range of years(s) this measure has been used by CMS Program(s)	N/A
What other federal programs are currently using this measure?	N/A
Evidence that the measure can be operationalized	This is a claims-based measure and will not require any additional submission of data.
How is the measure expected to be reported to the program?	Claims
Is this measure similar to and/or competing with measure(s)	No

already in a program?	
Which existing measure(s) is your measure similar to and/or competing with?	N/A
How will this measure be distinguished from other similar and/or competing measures?	N/A
Rationale for how this measure will add to the CMS program	N/A
If this measure is being proposed to meet a statutory requirement, please list the corresponding statute.	Section 101(f) of MACRA
Evidence of performance gap	Evidence of Performance Gap response: A recent study indicates that clinician beliefs about treatment and the efficacy of particular therapies may be the most important factors explaining the variation in health care expenditures. However, clinicians are often unaware of how their care decisions influence the overall costs of care. Cost measures are intended to help inform clinicians on the costs attributed to their decision-making and to incentivize cost-effective, high-quality care. A cost measure offers opportunity for improvement if clinicians can exercise influence on the intensity or frequency of a significant share of costs during the episode, or if clinicians can achieve lower spending and better care quality through changes in clinical practice. Diabetes mellitus is a group of metabolic disorders characterized by chronic hyperglycemia. The most common of these metabolic disorders in the Medicare population are type 1 and type 2 diabetes, both of which have their particular sets of causes, clinical manifestations, and management strategies, ranging from lifestyle changes to medication. Specifically, 7-12 percent of both the Medicare and broader United States diabetic population have type 1 diabetes, which is characterized by little to no insulin production by the insulin-producing beta

cells of the pancreatic islets. Conversely, 87-91 percent of the Medicare and broader United States diabetic population have type 2 diabetes, which is characterized by insulin resistance. According to the literature and previous feedback received through stakeholder input activities to date, the clinical focus of this measure represents an area where there are opportunities for improvement. These include mitigating the use of institutional post-acute care and inpatient stays, reducing overutilization, and increasing the use of preventive care to minimize downstream costs, which have the potential to be addressed by (i) promoting diabetes self-management education and support, (ii) increasing the use of appropriate medications, and (iii) encouraging adherence to correct preventive treatment guidelines. One way that clinicians may be able to contain costs associated with the management of diabetes is the promotion of diabetes self-management education and support (DSME/S). Given that diabetes is a chronic condition that requires patients to make several daily self-management decisions, DSME/S provides diabetes patients with a foundation to navigate these decisions and activities that are necessary to manage their condition (e.g., through medical nutrition therapy or other appropriate specialist referrals). For providers, there are national standards for DSME/S, which include but are not limited to developing an individualized DSME/S plan with diabetes patients, making diabetes patients aware of options and resources available for ongoing support of their initial education, and monitoring and communicating whether diabetes patients are achieving their self-management goals and other outcomes. Through promoting DSME/S, managing clinicians have an opportunity to reduce their patients' diabetes-related hospital admissions and readmissions, reduce their lifetime health care costs for diabetes-related complications, improve their glycated hemoglobin (HbA1C), an indicator of patient blood glucose levels, by as much as 1 percent, and reduce the onset or advancement of their diabetes-related complications, among other benefits. Increasing the use of appropriate medications offers another way for clinicians to contain costs associated with the management of diabetes. These pharmacological options, which are often supplemented by lifestyle changes, may vary depending on the type of diabetes. For patients with type 1 diabetes or poorly-controlled type 2 diabetes, insulin therapy helps to maintain normal blood glucose levels. In patients with type 1 diabetes, early and chronic exogenous insulin coverage, either through multiple daily injections or through use of an infusion pump, can reduce diabetes-related microvascular and macrovascular complications. In patients with type 2 diabetes, insulin therapy can reduce diabetes-related microvascular complications and in the long-term, can improve cardiovascular prognosis. Other diabetes management medications, such as metformin, aim to further regulate blood glucose levels by decreasing gluconeogenesis or increasing pancreatic insulin secretion. For most patients with type 2 diabetes, metformin is recommended as the preferred initial glucose lowering medication. This is due, in part, to its effectiveness in lowering blood glucose levels, its minimal hypoglycemia risk when used as monotherapy, and its weight loss benefits in some patients with type 2 diabetes. Through identifying these and other

appropriate medication(s) and promoting patient adherence to their medication regimes, managing clinicians have an opportunity to prevent the onset or progression of costly diabetes-related complications in their patients. Current literature also suggests that the managing clinician has an opportunity to contain diabetes-related costs by encouraging adherence to correct preventive treatment guidelines. It is well established that poor monitoring and control of blood glucose, lipid levels, and blood pressure can drastically increase the risk and severity of diabetes-related complications. This is especially salient for older adults whose diabetes treatment may be complicated by their clinical, cognitive, and functional heterogeneity. For example, higher rates of cognitive impairment in older adults have been associated with an increased risk of hypoglycemia, which can lead to falls, seizures, and loss of consciousness. One study showed that lower cognitive ability was associated with a twofold higher incidence of severe hypoglycemia. This study demonstrates that by screening older adults with diabetes for cognitive impairment during clinical visits, clinicians can better assess their patients' potential risk for worsening of their glycemic control, allowing clinicians to modify a patient's treatment plan to accommodate these cognitive changes and to continue to effectively manage their patient's diabetes care. Furthermore, diabetic patients also face an increased risk of cardiovascular disease and require close monitoring of lipid profiles and blood pressure to prevent stroke, coronary artery disease (CAD), and heart failure. One study found that improved control of HbA1C, lipid levels, and blood pressure predicted a 28-49 percent reduction in the probability of diabetes-related complications and a 7-10 percent decrease in total cost of care. To manage blood pressure, during each office visit, clinicians should measure their diabetic patients' blood pressure. If the readings on at least two of the visits are ≥ 130/80 mmHg, then clinicians should initiate medications (e.g., ACE inhibitors or angiotensin receptor blockers (ARB) and lifestyle changes (e.g., diet and exercise) for these patients. For lipid levels, it is recommended that clinicians screen patients with diabetes annually for their fasting serum lipid levels, and for those with dyslipidemia, clinicians should encourage lifestyle interventions (e.g., medical nutrition therapy or smoking cessation) and/or pharmacological interventions (e.g. statins) to control lipid levels. In following these and other preventive treatment guidelines, managing clinicians have another avenue to stem the onset or progression of diabetes-related complications in their patients. Literature suggests that given the high impact of diabetes within the Medicare patient population and consequential effect on Medicare spending, the Diabetes episode group represents an area with significant opportunity for improvement with respect to cost containment.

This measure aims to address these example areas of opportunities for improvement. Research has shown that In the United States, there are approximately 13.5 million people ages 65 and older living with diabetes, and treatment of diabetes in the United States costs over $348 billion annually which warrants the exploration of potential cost measures which aim to achieve more cost-effective care for a given condition [1]. As such,

	the use of this cost measure can provide clinicians with information to improve care outcomes and reduce future health care costs. This is supported by our analysis into the variation in spending per episode that may be reduced through better clinical practices. We analyzed the measure's performance gap for clinicians (TIN-NPIs) and clinician groups (TINs) using episodes ending between January 1, 2019 and December 31, 2019. With no case minimum applied, there were 6,329,215 Diabetes episodes for 4,598,059 beneficiaries. The TIN-NPI and TIN-level measure scores were calculated for 83,271 clinicians and 39,321 clinician groups who met the 20-episode case minimum. • The mean risk-adjusted cost per episode was $7,031 at the TIN level. The risk-adjusted cost per episode at the 10th percentile was $4,738, compared to $9,464 at the 90th percentile at the TIN level. • The mean risk-adjusted cost per episode was $6,726 at the TIN-NPI level. The risk-adjusted cost per episode at the 10th percentile was $4,280, compared to $9,402 at the 90th percentile at the TIN level. • For TINs, the mean measure score was 0.99. The score at the 10th percentile was 0.67, compared to 1.33 at the 90th percentile. For TIN-NPIs, the mean measure score was 0.95. The score at the 10th percentile was 0.60, compared to 1.33 at the 90th percentile.
Unintended consequences	Stakeholders have expressed concerns that cost measures could potentially have risks or unintended consequences. During past stakeholder input activities, including Technical Expert Panels (TEPs) and public comment periods, stakeholders have highlighted the importance of considering potential incentives created by cost measures that could adversely affect patients and clinicians. For patients, these potential adverse effects may include: (i) reducing access to care, and (ii) stinting care. For clinicians, these adverse effects could include: (i) making unfair comparisons among certain types of clinicians in assessing cost measure performance, and (ii) rewarding lower costs without appropriately balancing performance on quality. Overall, by promoting lower spending, stakeholders have expressed concerns that cost measures could potentially create barriers to patients receiving clinically appropriate care. To mitigate the potential risks of such incentives, stakeholders and TEP members have highlighted the importance of the following: Devising an appropriate risk adjustment model for episode-based cost measures; Aligning cost measures with indicators of quality; Identifying benchmarks for comparison that are fair and appropriate for all eligible clinicians; Potentially excluding certain types of patients from measure calculations.
Which clinical guideline(s)?	N/A
Briefly describe the peer reviewed evidence justifying this	The high prevalence and cost of diabetes mellitus and its associated complications to the United States health care system warrants the exploration of potential cost measures which aim to achieve more cost-effective care for a given condition. In the United States, there are approximately 13.5 million people ages 65 and older living with diabetes,

measure	and treatment of diabetes in the United States costs over $348 billion annually. In 2012, 59 percent of healthcare costs related to diabetes were associated with patients over the age of 65. In 2017, approximately 57 percent ($9,600 out of $16,750) of annual medical expenditures incurred for patients diagnosed with diabetes were related to their diabetes diagnosis. Additionally, on average, patients with diabetes had medical expenditures 2.3 times higher than those for patients without a diabetes diagnosis. Significant cost drivers in the care of diabetes are the occurrence of acute complications such as acute hyperglycemic crises (diabetic ketoacidosis and hyperglycemic hyperosmolar nonketotic syndrome) and longer-term complications of diabetes such as retinopathy, neuropathy, diabetic foot ulcers, cardiovascular events, and amputations. For example, over $2.4 billion in costs from hospital treatment were attributed to acute hyperglycemic crises, and over $1.84 billion for acute hypoglycemia and related injuries. Overall, patients with multiple diabetes complications had a higher risk of readmissions for severe dysglycemia (hyperglycemia or hypoglycemia) as well as causes that are unrelated to diabetes. It was also estimated that the prevalence of diabetic retinopathy among diabetic patients 65 years and older was 29.5 percent. Similarly, in 2007, 8.1 percent of Medicare diabetic beneficiaries enrolled in Medicare Parts A and B had diabetic foot ulcers, incurring spending that was significantly higher than that for beneficiaries without chronic wounds ($31,363 vs. $11,692, respectively). Given the prevalence of diabetes in the Medicare population, and the high costs associated with the management of the disease and its complications, the Diabetes cost measure represents an opportunity for improvement on overall cost performance.

Preliminary Analysis – MUC ID: MUC20-0017 Diabetes Episode-Based Cost Measure

Criteria	Yes/No	Justification and Notes
Does the measure address a critical quality objective not currently adequately addressed by the measures in the program set?	Yes	MUC20-0017 addresses the Patient-Focused Episode of Care goal of CMS's Meaningful Measures initiative, the MIPS high priority area of Efficiency/Cost Reduction, and the 2015 Medicare Access and Children's Health Insurance Program (CHIP) Reauthorization Act (MACRA) statutory requirements (section 101(f)). Diabetes care in the US requires more healthcare resources than any other disease, with 80% of costs generated from medications and ambulatory care (Dielman, et al, 2016). Currently there are 8 MIPS quality measures related to diabetes care. Of the current measures, none look at episode-based cost related to diabetes.
Is the measure evidence-based and either strongly linked to outcomes or an outcome measure?	Yes	This cost/resource use measure aims to improve episodic costs associated with the management of diabetes. Studies suggest that knowledge and awareness of evidence-based practices and treatment risks can influence decision-making and can lead to lower costs (Cutler, et al., 2016). Early initiation of diabetes mellitus (DM) management has been shown to reduce costs and risks of more serious complications (Li, et al., 2010; CDC, 2002). Opportunities for improvement include reducing the use of institutional post-acute care and inpatient stays, reducing overutilization, and increasing the use of preventive care to minimize downstream costs, which have the potential to be addressed by (i) promoting diabetes self-management education and support, (ii) increasing the use of appropriate medications, and (iii) encouraging adherence to correct preventive treatment guidelines. The measure is still being tested and results are not available.
Does the measure address a quality challenge?	Yes	Approximately one third of Medicare patients have DM, with higher rates among minorities (Cubanski, et al., 2019; CDC, 2020). US total costs for DM exceed $348M annually (International Diabetes Federation, 2017). Recent studies indicate that DM prevalence and incidence has declined, but overall burden remains high (Benoit, et al., 2019).

Does the measure contribute to efficient use of measurement resources and/or support alignment of measurement across programs?	Yes	This measure is not duplicative of other measures currently within the MIPS program. This and other 2020 MIPS cost measures under consideration were created in response to MACRA requirements to develop measures for potential implementation in the cost performance category of MIPS. Additionally, this measure is an episode-based measure; current cost-based measures in MIPs are all at the population level. While diabetes measures are common throughout CMS quality programs, no additional episode-based cost measures were identified.
Can the measure can be feasibly reported?	Yes	This measure uses Medicare claim data which is feasibly reported and a low-burden data source.
Is the measure applicable to and appropriately specified for the program's intended care setting(s), level(s) of analysis, and population(s)?	Yes	The measure is specified for Clinician: Individual and Clinician: Group/Practice levels, which aligns with MIPS reporting categories. The measure has not been reviewed for endorsement by an NQF Standing Committee.
If the measure is in current use, have negative unintended issues to the patient been identified? Have implementation challenges outweighing the benefits of the measure have been identified?	N/A	The measure is not in current use. The submission identified a potential unintended consequence of reduction in access to care or stinting care. This measure could result in rewarding cost reductions with no net quality of care gain if it is not appropriately balanced with measures addressing performance on quality.
PAC/LTC Core Concept?		N/A
Impact Act Domain		N/A

Hospice High Priority Areas		N/A
Rural Workgroup Input	[redacted]	Relative priority/utility: • This cost measure was noted to be for an important common condition in rural settings. Data collection issues: • None identified Calculation issues: • Average cost for rural providers may exceed the national average cost and therefore may unduly impact rural providers Unintended consequences: • Cost measures may result in restriction of care if not paired with quality measures Votes: Range is 1 – 5, where higher is more relevant to rural. Average: 4.1 • 1 – 0 votes • 2 – 1 vote • 3 – 1 vote • 4 – 13 votes • 5 – 4 votes
Preliminary Analysis Recommendation	Do Not Support with Potential for Mitigation	Mitigation is contingent on further evaluation of the correlation with clinical quality measures, as well as NQF endorsement. MAP noted a tension between expenses associated with good care that may result in reductions in overall cost of care but raise condition-specific care. MAP urged CMS to balance these cost measures with appropriate quality measures that are connected with lower costs. MAP further noted that upstream preventions should result in reduced downstream costs and expressed concerns that this is not the case for the measure, impacting its overall actionability. MAP noted that this measure aims to improve care by optimizing resource use associated with diabetes management. While there are measures in MIPS related to individual treatments for diabetes, this measure would potentially focus care on the most cost-effective interventions, but these should be connected.

Summary: What is the potential value to the program measure set?	[redacted]	Diabetes is both prevalent and costly to treat. A third of Medicare beneficiaries have diabetes and the cost of diabetes care in the US exceeds $348 billion annually. This measure aims to improve care by optimizing resource use associated with diabetes management. While there are measures in MIPS related to individual treatments for diabetes, this measure would potentially focus care on the most cost-effective interventions.
Summary: What is the potential impact of this measure on quality of care for patients?	[redacted]	This measure could improve Medicare costs of diabetes by incentivizing risk reduction treatments that are cost effective. Should testing data show that the measure appropriately measures episode-based cost and can be used to improve value of care, this measure would be valuable to add to the program measure set.

Measure Comments

Author	Submitted Comment
American Physical Therapy Association	We encourage the measure steward to include physical therapists as eligible clinicians that can be scored under this cost measure.
University of Colorado Medicine	"Numerator does not mention "risk-adjusted"" standardized cost. We understand that this measure will be risk-adjusted due to that verbiage in the Description, but it may be useful to qualify this in the numerator as well.
Diabetes Advocacy Alliance	January 6, 2021 TO: Measure Application Partnership MAP MUC 2020 Comment Period FROM: Hannah Martin, Academy of Nutrition and Dietetics Diabetes Advocacy Alliance Co-Chair hmartin@eatright.org Kate Thomas, Association of Diabetes Care & Education Specialists Diabetes Advocacy Alliance Co-Chair kthomas@adces.org RE: Comments to MUC20-0017, Diabetes Episode-Based Cost Measure The Diabetes Advocacy Alliance (DAA) appreciates the opportunity to submit comments related to Measures Under Consideration: MUC20-0017, Diabetes Episode-Based Cost Measure.

The DAA is a coalition of 26 diverse member organizations, representing patient, professional and trade associations, other non-profit organizations, and corporations, all united in the desire to change the way diabetes is viewed and treated in America. Since 2010, the DAA has worked with legislators and policymakers to increase awareness of, and action on, the diabetes epidemic. The organizations that comprise the DAA share a common goal of elevating diabetes on the national agenda so we may ultimately defeat this potentially devastating chronic disease.

The DAA recognizes that health inequities have had, and continue to have, a tremendous negative impact on our society's ability to identify those at risk, prevent new cases of diabetes, and effectively treat people with diabetes and obesity to help prevent the many serious complications of the disease. The DAA believes that quality measures are important, but we urge CMS to recognize that inequities have an effect on the data that are collected and thus some quality measure data may not be representative of populations most affected by the proposed measures.

MUC20-0017, Diabetes Episode-Based Cost Measure

Description: The Diabetes cost measure evaluates a clinician group's risk-adjusted cost to Medicare for patients receiving medical care to manage type 1 or type 2 diabetes. The measure score is a clinician group's weighted average of risk-adjusted cost for each episode attributed to the clinician group, where each episode is weighted by the number of assigned days during the episode. This chronic measure includes services that are clinically related and under the reasonable influence of the attributed clinician group. Services are assigned during a Diabetes episode, which is a portion of the overall time period of a clinician group's responsibility for managing a patient's diabetes. Medicare beneficiaries enrolled in Medicare Parts A and B during the performance period are eligible for the measure.

Measure Type: Cost/Resource Use

Measure Steward: Centers for Medicare and Medicaid Services (CMS)

CMS Program: Merit-Based Incentive Payment System (MIPS)

This measure would add value for at least two reasons:

• First, it creates a path for recognizing and financially rewarding clinicians who are working with their patients to ensure that they are managing their diabetes effectively, as opposed to not managing it. Proactive management avoids complications such as emergency room visits as is documented in the literature cited in CMS' summary.
• Second, because the measure includes a risk adjustment feature, it theoretically should account for clinicians with more medically complex patients with diabetes.

We do want to state that before finalizing this measure, CMS and NQF should confirm that the risk-adjustment methodology it has proposed ensures that

	clinicians who primarily work with medically underserved population or Black or Latinx populations with higher rates of type 2 diabetes can take advantage of this measure and its incentive payments because of their work with these populations.
American Medical Association	Overall, the AMA supports the episode-based measure development process and movement to episode-based measures over broad cost measures. However, we continue to have significant concerns with the lower than desirable mean reliability rate. The AMA believes that the minimum acceptable thresholds should be 0.7. In addition, we strongly support the tenet that cost must be assessed within the context of the quality of care provided; yet, none of these assessment have been provided for the existing MIPS cost measures or the ones currently under MAP consideration. We request further refinements to the process to ensure the cost measures are accurate and fair when used in MIPS and, ultimately, publicly reported. The AMA strongly opposes including Part D prescription drug costs in the Medicare episode-based cost measures. To hold physicians accountable for costs that are negotiated between CMS and Medicare Prescription Drug Plans is fundamentally problematic, and physicians and patients do not always have information about coverage, formularies, out-of-pocket costs and list prices at the point of care. This also assumes there is a viable, evidence-based, less expense alternative option for patients. In general, we urge the MAP to make "Conditional Support" its highest level of recommendation and to recommend "Do Not Support with Potential for Mitigation" for any measures where relevant specialties have raised serious concerns or the average reliability rate is less than 0.7 as is the case with several of these measures.
Association of American Medical Colleges (AAMC)	The AAMC remains concerned that such cost measures are adjusted to account for social risk factors (SRFs). In addition to patient clinical complexity, SRFs can drive differences in average costs. In particular, physicians at academic medical centers (AMCs) care for vulnerable populations of patients who are sicker, poorer, and more complex than patients treated elsewhere. In regard to attribution – AAMC has previously commented that attribution methods used should be clear and transparent to clinicians and that it is critical that there be an accurate determination of the relationship between a patient and a clinician to ensure that the correct clinician is held responsible for the patient's outcomes and costs. Attribution is complicated, given that most patients receive care from numerous clinicians across several facilities, and AAMC has urged CMS to explore better data sources and analytic techniques to support more accurate attribution. The AAMC recommends that: (1) cost measures include risk-adjustment for SRFs, (2) the attribution methodology is transparent, and (3) the appropriate clinician is held responsible for the patient's outcomes and costs. Additionally, this measure has not been submitted for NQF endorsement. The AAMC has long held that measures should not be proposed for addition to public reporting programs unless vetted and endorsed by the NQF. The AAMC recommends that the highest level of MAP recommendation be "Do Not

	Support With Potential For Mitigation."
AdvaMed	AdvaMed strongly supports this measure and urge that an episode include the cost of management of complications resulting from diabetes, including chronic wounds. This measure should define "diabetes episode", given that diabetes is a chronic condition managed over a lifetime. Technologies that facilitate much greater control over diabetes are now available and are increasingly showing cost effectiveness relative to a plan of care that does not include these technologies. Specifically, continuous glucose monitors and insulin pumps permit people with diabetes to be much more aware of the nature of their condition and the impact of various lifestyle choices and habits. Unfortunately, while adoption is increasing, these technologies are not currently used as widely or effectively as they could be. Implementation of this measure would encourage providers to understand and appreciate the benefits these technologies and take steps needed to ensure their patients adopt and successfully use these critical tools.
Roji Health Intelligence	For each cost measure under consideration, physician practices or groups must have the ability to examine data on episodes for procedures and conditions that reveal how these episodes relate to the cost target. In all cases involving multi-physician and facility care, this will require data and transparency from CMS, rather than simply providing aggregate final results. With CMS proposed measures, physician groups can only react to measurement but lack the ability to improve because (a) they don't have the data to replicate the episode-based measure data, since much of the costs are indirectly generated by other providers yet invisible to the attributed group; and (b) all of the scores are retrospective and there is no actionable data to help them improve. Adoption of cost measures should be contingent on CMS provision of claims data on a regular ongoing basis (for example monthly) so that the group can create their own applications for CMS cost measure-related episodes, and evaluate the specific reasons for cost excess. Because CMS cost measures involve medical services beyond an individual practice or group's own EMR or billing systems, physician groups do not have the necessary data without receipt of claims details from CMS. Receipt of that data would allow organizations the ability to make modifications to improve their cost-effectiveness. This is analogous to what is currently happening with Accountable Care Organizations. Roji Health Intelligence is a qualified CMS reporting registry. We have created Episode-Based cost measures for our physician practice and group clients based on the data present in their practice management and electronic health records. This information is presented in an on-line interactive format for illumination and comparison of episodes, cost drivers, and outcomes associated with episodes. The purpose is to engage clinicians and practices in understanding and acting on cost and outcome results for patients in episodes of care. For our cost measures to be most valuable, we suggest that our clients' receipt of CMS claims data would ensure that we could assist them in a more complete understanding of variation in costs and interventions that would lower health care expenditures.

Measure Comments (Post-Workgroup Meeting)

Author	Submitted Comment
American Heart Association/American Stroke Association	The AHA does not support this measure for inclusion in the MIPS program. We question whether attributing this at the physician group level is appropriate. We are also concerned that it may create a disincentive to provide appropriate care and has the potential to worsen disparities. It is not clear from the documentation provided, but we assume that the risk-adjustment model relies solely on administrative claims data, which often cannot reflect the complexity and circumstances of each patient. It is critical that risk adjustment models incorporate clinical information related to severity of illness to accurately assess risk. At a minimum, we believe that MAP should not recommend this measure and CMS should not include it in a proposed rule until reviewed and endorsed by NQF.
Federation of American Hospitals	Support
American Medical Association	The American Medical Association (AMA) continues to have significant concerns with this measure and believes that revisions to the specifications and additional field testing are necessary. We ask that additional conditions be provided with this recommendation around ensuring that the minimum reliability rate be 0.7 or higher, evaluating costs within the context of the quality of care provided is completed, and removing Part D prescription drug costs.
Novo Nordisk	Novo Nordisk is a global healthcare company with more than 95 years of innovation and leadership in diabetes care. This heritage has given us the experience and capabilities necessary to help people defeat other serious chronic diseases such as hemophilia, growth disorders and obesity. As an organization, we are committed to ensuring patients have access to high-quality, affordable health care. While we support the idea of the Diabetes Episode-Based Cost Measure and the ongoing drive to curtail costs in the U.S. Health System, we align with the current MAP recommendation: "Do Not Support with Potential for Mitigation." Our specific concern is as follows: CMS should be encouraged to balance cost measures with appropriate quality measures that are connected with evidence-based care and have been shown to be effective for improving patient care and potentially lower costs. We agree with the MAP notation that while this measure aims to improve care by optimizing resource use associated with diabetes management, there are measures in MIPS related to individual treatments for diabetes that could be used in conjunction with a cost measure to assess the value of providing the most cost-effective and evidence-based interventions. Assessing care by cost alone could harm patients and result in decreased quality, which is not aligned with a value-based care system.

AdvaMed	AdvaMed strongly supports this measure and urge that an episode include the cost of management of complications resulting from diabetes, including chronic wounds. This measure should define "diabetes episode", given that diabetes is a chronic condition managed over a lifetime. Technologies that facilitate much greater control over diabetes are now available and are increasingly showing cost effectiveness relative to a plan of care that does not include these technologies. Specifically, continuous glucose monitors and insulin pumps permit people with diabetes to be much more aware of the nature of their condition and the impact of various lifestyle choices and habits. Unfortunately, while adoption is increasing, these technologies are not currently used as widely or effectively as they could be. Implementation of this measure would encourage providers to understand and appreciate the benefits these technologies and take steps needed to ensure their patients adopt and successfully use these critical tools.
Association of American Medical Colleges (AAMC)	The Clinician MAP conditionally supported two of the episode-based costs measures (Colon and Rectal Resection [MUC2020-0016] and Melanoma Resection [MUC2020-0018]) for future rulemaking for the Merit-based Incentive Payment System (MIPS) program subject to NQF endorsement. The Clinician MAP did not support the other three episode-based costs measures (Asthma/Chronic Obstructive Pulmonary Disease [MUC20-0015], Diabetes [MUC2020-0017], and Sepsis [MUC2020-0019]) for future rulemaking for the MIPS program with potential for mitigation. Mitigation for the three measures focused on evaluation of the actionability and connection between upstream medical interventions and downstream costs, in addition to NQF endorsement. The AAMC agrees with concerns about episode-based cost measures relying on the suggestion that providing certain upstream preventions will result in lower costs of care, and that lower costs will result in better patient outcomes. Furthermore, the AAMC remains concerned that none of the 13 cost measures are adjusted to account for social risk factors (SRFs). In addition to patient clinical complexity, SRFs can drive differences in average costs. In particular, physicians at academic medical centers (AMCs) care for vulnerable populations of patients who are sicker, poorer, and more complex than patients treated elsewhere. In regard to attribution – AAMC has previously commented that attribution methods used should be clear and transparent to clinicians and that it is critical that there be an accurate determination of the relationship between a patient and a clinician to ensure that the correct clinician is held responsible for the patient's outcomes and costs. Attribution is complicated, given that most patients receive care from numerous clinicians across several facilities, and AAMC has urged CMS to explore better data sources and analytic techniques to support more accurate attribution. In addition, the movement in medicine has been to team-based care, further complicating appropriate attribution to a single clinician. The MAP, through its recommendations, and CMS should be careful not to incent patterns of care that are outdated. The AAMC recommends that the MAP recommendation be "do not support with potential for mitigation" for each of the episode-based cost measures.

MUC20-0018: Melanoma Resection Episode-Based Cost Measure

Measure Information

Characteristic	Submitted Information
MUCID	MUC20-0018
Other Measure Identification Numbers	N/A
Title	Melanoma Resection Episode-Based Cost Measure
Program	Merit-based Incentive Payment System-Cost
Workgroup	Clinician
In what state of development is the measure?	Fully Developed
State of Development Details	The measure has been developed, field tested, and is now refined based on feedback received from field testing in summer 2020 (August 17 – September 18, 2020). As background, a list of draft episode groups and trigger codes, developed with input from a Clinical Committee convened in 2016, was posted by CMS in December 2016 to meet MACRA requirements. Building off this work, episode-based cost measures are developed using a “wave” approach wherein sets of Clinical Subcommittees (CS) are convened to select episode groups to develop into cost measures, while smaller, measure-specific Clinician Expert Workgroups provide detailed input on each component of the measures. The current wave of measure development began in May 2019 and includes 4 CS with a total of 166 members affiliated with 110 professional societies. The Dermatologic Disease Management CS selected this episode group from the December 2016 draft list of episode groups for development. This CS comprises members representing clinician specialty societies in this clinical area. A Clinician Expert Workgroup met in August 2019 to discuss measure specifications for all components of the measure, followed by a webinar in January 2020 for follow-up discussions on service assignment and risk adjustment. The measure will be field tested in the summer of 2020, during which clinicians and stakeholders can learn about the measure and provide input on the draft cost measure specifications. After field testing, the Clinician Expert Workgroups will revisit and refine the draft measure specifications based on the stakeholder feedback received. We conducted reliability testing of measures for clinicians (TIN-NPIs) and for clinician groups (TINs), constructed using episodes ending between January 1, 2019 and December 31, 2019. Reliability refers to the extent to which a measure reflects true variation between clinicians’ risk-adjusted episode spending, as opposed to random variation. The reliability metric specifically

	captures how much of the variance in a measure is due to systematic differences in episode spending between clinicians, rather than differences in episode spending within a clinician's set of episodes. A measure with high reliability suggests that the comparisons of performance across clinicians can be expected to better reflect systematic differences in actual performance. Our testing results indicate that this measure is reliable for clinicians and clinician groups across a range of case minimums. • For TINs at a 10-episode case minimum, the mean reliability was 0.80. For TINs at a 20-episode case minimum, the mean reliability was 0.87. For TINs at a 30-episode case minimum, the mean reliability was 0.90. For TIN-NPIs at a 10-episode case minimum, the mean reliability was 0.79. For TIN-NPIs at a 20-episode case minimum, the mean reliability was 0.87. For TIN-NPIs at a 30-episode case minimum, the mean reliability was 0.90.
Measure Description	The Melanoma Resection cost measure evaluates clinician or clinician group'srisk-adjusted cost to Medicare for patients who undergo an excision procedure to remove a cutaneous melanoma. The measure score is a clinician's average risk-adjusted cost for the episode group across all episodes attributed to the clinician or clinician group. This procedural measure includes services that are clinically related and under the reasonable influence of the attributed clinician during the 30 days prior to the clinical event that opens or "triggers" the episode through 90 days after. Medicare beneficiaries enrolled in Medicare Parts A and B during the performance period are eligible for the measure.
Numerator	The numerator for the Melanoma Resection measure is the sum of the ratio of observed to expected payment-standardized cost to Medicare for all episodes attributed to a clinician or clinician group. This sum is then multiplied by the national average observed episode cost to generate a dollar figure. Mathematically, this is represented as: sum of (observed episode cost/expected episode cost) * national average observed cost.
Denominator	The denominator for the Melanoma Resection measure is the total number of episodes from this episode group attributed to a clinician or clinician group.
Exclusions	The following episode-level exclusions apply: (a) The beneficiary has a primary payer other than Medicare for any amount of time overlapping the episode window or in the lookback period. (b) No attributed clinician is found for the episode. (c) The beneficiary's date of birth is missing. (d) The beneficiary's death date occurred before the episode ended. (e) The beneficiary was not enrolled in Medicare Parts A and B for the entirety of the lookback period plus episode window, or is enrolled in Part C for any part of the lookback period plus episode window. (f) The episode trigger claim was not performed in an office, IP, OP, or ASC setting based on its place of service. Exclusions specific to the Melanoma Resection measure are developed with input from the Melanoma Resection Clinician Expert

	Workgroup.
Measure type	Cost/Resource Use
What is the NQF status of the measure?	Never submitted
NQF ID number	0000
Year of next anticipated NQF CDP endorsement review	NA
Year of most recent NQF Consensus Development Process (CDP) endorsement	NA
Is the measure being submitted exactly as endorsed by NQF?	NA
If not exactly as endorsed, describe the nature of the differences	N/A
What data sources are used for the measure?	Claims
If EHR or Administrative Claims or Chart-Abstracted Data, description of parts related to these sources.	N/A
At what level of analysis was the measure tested?	The measure has been tested at both the Clinician: Individual and Clinician: Group levels.

In which setting was this measure tested?	Ambulatory surgery center, Ambulatory/office-based care, Hospital outpatient department (HOD)
What NQS priority applies to this measure?	NA
What one primary meaningful measure area applies to this measure?	Patient-focused episode of care
What secondary meaningful measure area applies to this measure?	NA
What one primary healthcare priority applies to this measure?	Make care affordable
What secondary healthcare priority applies to this measure?	NA
What area of specialty best fits the measure?	Dermatology
What is the target population of the measure?	Medicare Fee for Service
Is this measure an eCQM?	No
If eCQM, enter Measure Authoring Tool (MAT) number	No
If eCQM, does the measure	No

have a Health Quality Measures Format (HQMF) specification?	
Comments	N/A
Measure steward	Centers for Medicare & Medicaid Services
Long-Term Measure Steward (if different)	N/A
Measure Steward Contact Information	Evans, Ronique; Center for Clinical Standards and Quality; 410-786-3966; Ronique.Evans1@cms.hhs.gov
Primary Submitter Contact Information	Jensen, Ross; Acumen, LLC; 650-558-8882; macra-episode-based-cost-measures-info@acumenllc.com
Long-Term Measure Steward Contact Information	N/A
Secondary Submitter Contact Information	N/A
Was this measure proposed for a previous year's MUC list?	No
In what prior year(s) was this measure proposed?	NA
What were the programs that NQF MAP reviewed the measure for in	N/A

each year?	
Why was the measure not recommended in those year(s)?	N/A
What were the MUC IDs for the measure in each year?	N/A
NQF MAP report page number being referenced for each year	N/A
What was the NQF MAP recommendation in each year?	N/A
List the NQF MAP workgroup(s) in each year	N/A
What is the history or background for including this measure on the new MUC list?	New measure never reviewed by MAP Workgroup or used in a CMS program
Range of years(s) this measure has been used by CMS Program(s)	N/A
What other federal programs are currently using this measure?	
Evidence that the measure can be	This is a claims-based measure and will not require any additional submission of data.

operationalized	
How is the measure expected to be reported to the program?	Claims
Is this measure similar to and/or competing with measure(s) already in a program?	No
Which existing measure(s) is your measure similar to and/or competing with?	N/A
How will this measure be distinguished from other similar and/or competing measures?	N/A
Rationale for how this measure will add to the CMS program	N/A
If this measure is being proposed to meet a statutory requirement, please list the corresponding statute.	Section 101(f) of MACRA
Evidence of performance gap	Evidence of Performance Gap response: A recent study indicates that clinician beliefs about treatment may be the most important factors explaining the variation in health care expenditures. However, clinicians are often unaware of how their care decisions influence the overall costs of care. Cost measures are intended to inform clinicians on the costs attributed to their decision-making and to incentivize cost-effective, high-quality care. A cost measure offers opportunity for improvement if clinicians can exercise

influence on the intensity or frequency of a significant share of costs during the episode, or if clinicians can achieve lower spending and better care quality through changes in clinical practice. According to the literature and feedback received through stakeholder input activities, the measure focus represents an area where there are opportunities for improvement. Primarily, opportunities for improving melanoma resection outcomes include operating under established clinical guidelines to reduce downstream complications and potential follow-up procedures to attain local disease control. Following established clinical guidelines for melanoma resection, the primary surgical procedure to excise a cutaneous malignant melanoma, can reduce complications and improve patient outcomes. Excisions for cutaneous melanoma are curative in 85-90% of cases. Melanomas curable by resection are predominantly localized, with localized melanomas seeing a 99% five-year survival rate. It is standard clinical practice to use different excision margins around the melanoma, depending on the depth and location of the melanoma. These margins are meant to balance surgical efficacy, patient impact, and burden on the healthcare system. Typically, wider margins are used for deeper melanomas, but these margins may be adjusted to account for any post-operative aesthetic concerns, such as to minimize scarring for resections performed on the head or face. Though the literature generally suggests that wide and narrow excision margins result in similar patient outcomes, one meta-analysis found that wider margins offered favorable melanoma-specific survival, recurrence free survival, and loco-regional recurrence rates compared to narrow margins. Variation in the treatment of melanoma can also exist in the timing of any necessary reconstructive procedures. Reconstructive procedures may be performed immediately after the excision of the melanoma or delayed to allow any pathological reports to indicate whether larger margins are needed. While performing a reconstructive procedure immediately potentially allows for a reconstructed wound to have residual disease, one study suggested that doing so immediately after excision is safe due to an acceptably low rate of residual tumors requiring operation. Additionally, the paper notes that not delaying the reconstructive procedure generates substantial savings, especially in the inpatient setting, noting a 38.5% decrease in cost. Melanomas that have spread beyond the local site can increase variability in patient treatment by adding downstream complications and/or additional procedures to address worsened disease severity. For example, a Sentinel Lymph Node (SLN) biopsy is a costly procedure commonly performed after a melanoma excision when the size of the melanoma indicates potential disease spread. A positive SLN biopsy result indicates non-localized and thus more severe disease (regional/distant). One study suggested that a patient's progression-free survival and overall survival decreases as the number of positive SLN nodes and maximum metastasis size (the size of the SLN being removed) increases. Negative SLN biopsy results have been shown to be strong predictors of melanoma-specific survival. One study found a significant increase in five-year melanoma-specific survival for patients with a negative SLN biopsy result compared to patients with a positive SLN biopsy result (88.9% compared to 64.8%). However, SLN

biopsies have higher complication rates. The overall complication rate for melanoma is estimated to be 4.2%, while the complication rate for SLN biopsies is estimated to be nearly triple that rate, at 11.3%. Though SLN biopsies can cause complications, they remain an important procedure following some melanoma resections to identify next steps in the treatment arc. Additional complications relevant to melanoma resection include surgical site infection (SSI) and delayed wound healing or wound dehiscence, requiring antibiotics, additional treatments, and skin grafts or skin substitutes, among others. While clinical characteristics may predispose certain patients to SSIs, the likelihood of an SSI can be reduced through evidence-based practices. These practices include but are not limited to proper administration of any necessary antibiotics and appropriate use of medical and sanitary equipment by the medical staff, including wearing proper surgical attire and disinfecting the surgical site prior to excision.

This measure aims to address these example areas of opportunities for improvement. Research shows that an estimated 196,060 cases of invasive and in situ melanoma will be newly diagnosed in 2020, with all melanoma cases costing the health care system an estimated $3.3 billion annually, a figure that is anticipated to continue to rise due to the increasing incidence of melanoma [2]. As such, the use of this cost measure can provide clinicians with information to improve care outcomes and reduce future health care costs. This is supported by our analysis into the variation in spending per episode that may be reduced through better clinical practices. We analyzed the measure's performance gap for clinicians (TIN-NPIs) and clinician groups (TINs) using episodes ending between January 1, 2019 and December 31, 2019. With no case minimum applied, there were 79,535 Melanoma Resection episodes for 67,094 beneficiaries. The TIN-NPI and TIN-level measure scores were calculated for 2,186 clinicians and 1,799 clinician groups who met the 10-episode case minimum.

- The mean risk-adjusted cost per episode was $1,760 at the TIN level. The risk-adjusted cost per episode at the 10th percentile was $1,280, compared to $2,298 at the 90th percentile at the TIN level.
- The mean risk-adjusted cost per episode was $1,799 at the TIN-NPI level. The risk-adjusted cost per episode at the 10th percentile was $1,280, compared to $2,433 at the 90th percentile at the TIN-NPI level.
- For TINs, the mean measure score was 0.95. The score at the 10th percentile was 0.69, compared to 1.24 at the 90th percentile.
- For TIN-NPIs, the mean measure score was 0.96. The score at the 10th percentile was 0.68, compared to 1.30 at the 90th percentile.

Unintended consequences	Stakeholders have expressed concerns that cost measures could potentially have risks or unintended consequences. During past stakeholder input activities, including Technical Expert Panels (TEPs) and public comment periods, stakeholders have highlighted the importance of considering potential incentives created by cost measures that could adversely affect

	patients and clinicians. For patients, these potential adverse effects may include: (i) reducing access to care, and (ii) stinting care. For clinicians, these adverse effects could include: (i) making unfair comparisons among certain types of clinicians in assessing cost measure performance, and (ii) rewarding lower costs without appropriately balancing performance on quality. Overall, by promoting lower spending, stakeholders have expressed concerns that cost measures could potentially create barriers to patients receiving clinically appropriate care. To mitigate the potential risks of such incentives, stakeholders and TEP members have highlighted the importance of the following: Devising an appropriate risk adjustment model for episode-based cost measures; Aligning cost measures with indicators of quality; Identifying benchmarks for comparison that are fair and appropriate for all eligible clinicians; Potentially excluding certain types of patients from measure calculations.
Which clinical guideline(s)?	N/A
Briefly describe the peer reviewed evidence justifying this measure	In the US, the average age when melanoma is diagnosed is 65, with incidence and melanoma-specific mortality increasing with age and peaking in the Medicare-aged population. Additionally, it is estimated that 196,060 cases of melanoma will be newly diagnosed in 2020, with all melanoma cases costing the health care system an estimated $3.3 billion annually, a figure that is anticipated to continue to rise due to the increasing incidence of melanoma. Opportunities for improvement for melanoma resection are primarily found within the variation in the timing of certain stages of post-excision treatment, procedure selection in context with patient characteristics to minimize complications, and adherence to established clinical excision margins.

Preliminary Analysis – MUC ID: MUC20-0018 Melanoma Resection Episode-Based Cost Measure

Criteria	Yes/No	Justification and Notes
Does the measure address a critical quality objective not currently adequately addressed by the measures in the program set?	Yes	MUC20-0018 addresses the Patient-Focused Episode of Care goal of CMS's Meaningful Measures initiative, the MIPS high priority area of Efficiency/Cost Reduction, and the 2015 Medicare Access and Children's Health Insurance Program (CHIP) Reauthorization Act (MACRA) statutory requirements (section 101(f)). Considering the high treatment cost for melanoma, this measure aims to reduce costs for a population that has a high incidence and mortality of melanoma. Currently there are 4 MIPS quality measures related to melanoma reporting and care continuity. Of the current measures, none look at episode-based cost related to melanoma.
Is the measure evidence-based and either strongly linked to outcomes or an outcome measure?	Yes	This cost/resource use measure aims to inform clinical decision-making related colon and rectal resection by reflecting cost of an episode of care and minimizing risks with approaches to melanoma resection. Studies suggest that knowledge and awareness of evidence-based practices and treatment risks can influence decision-making and can lead to lower costs (Sacks, et al., 2016; Cutler, et al., 2016). The evidence presented by the developer suggests that clinician approach significantly influences the cost and quality of care received.
Does the measure address a quality challenge?	Yes	Melanoma is especially common in the Medicare population. Nationwide estimates exceed 190,000 melanoma cases in 2020 (American Cancer Society, 2020), accounting for 5.6% of all cancer diagnoses (NIH, 2020). While five-year survival rates for melanoma are high (~99%), costs associated with treatment are also high. In one review of economic burden for all stages of melanoma, annual treatment costs were estimated to be $3.3B per year (American Cancer Society, 2020)

Does the measure contribute to efficient use of measurement resources and/or support alignment of measurement across programs?	Yes	This measure is not duplicative of other measures currently within the MIPs program. This and other 2020 MIPS cost measures under consideration were created in response to MACRA requirements to develop measures for potential implementation in the cost performance category of MIPS. Additionally, this measure is an episode-based measure; current cost-based measures in MIPS are all at the population level. No melanoma measures were identified in CMS programs outside of MIPS.
Can the measure can be feasibly reported?	Yes	This measure uses Medicare claims data which is feasibility reported and a low-burden data source.
Is the measure applicable to and appropriately specified for the program's intended care setting(s), level(s) of analysis, and population(s)?	Yes	The measure is specified for Clinician: Individual and Clinician: Group/Practice levels, which aligns with MIPS reporting categories. The measure has not been reviewed for endorsement by an NQF Standing Committee.
If the measure is in current use, have negative unintended issues to the patient been identified? Have implementation challenges outweighing the benefits of the measure have been identified?	N/A	The measure is not in current use. The submission identified a potential unintended consequence of reduction in access to care or stinting care. This measure could result in rewarding cost reductions with no net quality of care gain if it is not appropriately balanced with measures addressing performance on quality.
PAC/LTC Core Concept?		N/A
Impact Act Domain		N/A
Hospice High Priority Areas		N/A

Rural Workgroup Input	[redacted]	Relative priority/utility: • This measure was noted to be reliable at low case thresholds for a relatively common cancer type. Data collection issues: • None identified Calculation issues: • Average cost for rural providers may exceed the national average cost and therefore may unduly impact rural providers. Unintended consequences: • Cost measures may result in restriction of care if not paired with quality measures. Votes: Range is 1 – 5, where higher is more relevant to rural. Average: 3.8 • 1 – 0 votes • 2 – 0 votes • 3 – 4 votes • 4 – 12 votes • 5 – 1 vote
Preliminary Analysis Recommendation	Conditional Support for Rulemaking	Conditional support for rulemaking is contingent on NQF endorsement.
Summary: What is the potential value to the program measure set?	[redacted]	MUC20-0018 addresses the Patient-Focused Episode of Care goal of CMS's Meaningful Measures initiative, the MIPS high priority area of Efficiency/Cost Reduction and MACRA statutory requirements. Currently, there are no measures that assess episode-based cost related to melanoma. Melanoma is of growing concern to the Medicare population. The total annual treatment cost for melanoma is estimated at $3.3 billion, while melanoma resection is cited as curative in 85-90% of cases, with a 99% five-year survival rate. This measure aims to optimize resource use associated with melanoma resection. Clinician decision making is cited as being an important predictor of cost and an important pathway for risk reduction in melanoma care.

Summary: What is the potential impact of this measure on quality of care for patients?	[redacted]	Melanoma represents 5.6% of all cancer diagnoses, impacting over 190,000 patients per year. This measure could reduce costs of melanoma treatment and incentivize reduction of treatments that are not cost effective. Should testing data show that the measure appropriately measures episode-based cost while maintaining quality, this measure would be valuable to add to the program measure set. Conditional support for rulemaking is contingent on NQF endorsement.

Measure Comments

Author	Submitted Comment
American Medical Association	Overall, the AMA supports the episode-based measure development process and movement to episode-based measures over broad cost measures. However, we continue to have significant concerns with the lower than desirable mean reliability rate. The AMA believes that the minimum acceptable thresholds should be 0.7. In addition, we strongly support the tenet that cost must be assessed within the context of the quality of care provided; yet, none of these assessment have been provided for the existing MIPS cost measures or the ones currently under MAP consideration. We request further refinements to the process to ensure the cost measures are accurate and fair when used in MIPS and, ultimately, publicly reported. The AMA strongly opposes including Part D prescription drug costs in the Medicare episode-based cost measures. To hold physicians accountable for costs that are negotiated between CMS and Medicare Prescription Drug Plans is fundamentally problematic, and physicians and patients do not always have information about coverage, formularies, out-of-pocket costs and list prices at the point of care. This also assumes there is a viable, evidence-based, less expense alternative option for patients. In general, we urge the MAP to make "Conditional Support" its highest level of recommendation and to recommend "Do Not Support with Potential for Mitigation" for any measures where relevant specialties have raised serious concerns or the average reliability rate is less than 0.7 as is the case with several of these measures.
Association of American Medical Colleges (AAMC)	The AAMC remains concerned that such cost measures are adjusted to account for social risk factors (SRFs). In addition to patient clinical complexity, SRFs can drive differences in average costs. In particular, physicians at academic medical centers (AMCs) care for vulnerable populations of patients who are sicker, poorer, and more complex than patients treated elsewhere. In regard to attribution – AAMC has previously commented that attribution methods used should be clear and transparent to clinicians and that it is critical that there be an accurate determination of the relationship between a patient and a clinician to ensure that the correct clinician is held responsible for the patient's outcomes and costs. Attribution is complicated, given that most patients receive care from numerous clinicians across several facilities, and AAMC has urged CMS to explore better data sources and analytic techniques to support more accurate attribution.

	The AAMC recommends that: (1) cost measures include risk-adjustment for SRFs, (2) the attribution methodology is transparent, and (3) the appropriate clinician is held responsible for the patient's outcomes and costs. Additionally, this measure has not been submitted for NQF endorsement. The AAMC has long held that measures should not be proposed for addition to public reporting programs unless vetted and endorsed by the NQF. The AAMC recommends that the highest level of MAP recommendation be "Do Not Support With Potential For Mitigation."
American Academy of Dermatology Association	January 5, 2021 Seema Verma Administrator Centers for Medicare & Medicaid Services Department of Health and Human Services Attention: CMS-1734-P Mail Stop C4-26-05, 7500 Security Boulevard, Baltimore, MD 21244-1850 Re: Melanoma Resection Episode based Cost Measure Submitted electronically via Qualityforum.org Dear Administrator Verma, The American Academy of Dermatology Association (AADA) represents close to 14,000 dermatologists nationwide. We are writing to provide feedback on the Melanoma Resection Episode based Cost measure. The AADA is committed to excellence in the medical and surgical treatment of skin disease; advocating for high standards in clinical practice, education, and research in dermatology and dermatopathology; and driving continuous improvement in patient care and outcomes while reducing the burden of disease We appreciate the opportunity to provide feedback on the Melanoma Resection Episode base cost measure for the Merit-based Incentive Payment System (MIPS) program. The Melanoma Resection cost measure is of significant interest to dermatologists. Currently, there are no cost measures available for Dermatologists under the MIPS program. The Melanoma Cost measure would give them the opportunity to fully participate in the MIPS program. Therefore, the AADA supports adopting the melanoma cost measure into the MIPS program. We believe that the measure specifications listed achieve the goal of accurately measuring the cost of melanoma resection that is within the control of the attributed clinician.

	Thank you for your consideration of our recommendations and comments on the Melanoma Cost measure. If you have any questions, please contact Helen Olkaba, Assistant Director, Healthcare Economics at holkaba@aad.org or 202-712-2612. Sincerely, Bruce H. Thiers, MD, FAAD President, American Academy of Dermatology Association
Roji Health Intelligence	For each cost measure under consideration, physician practices or groups must have the ability to examine data on episodes for procedures and conditions that reveal how these episodes relate to the cost target. In all cases involving multi-physician and facility care, this will require data and transparency from CMS, rather than simply providing aggregate final results. With CMS proposed measures, physician groups can only react to measurement but lack the ability to improve because (a) they don't have the data to replicate the episode-based measure data, since much of the costs are indirectly generated by other providers yet invisible to the attributed group; and (b) all of the scores are retrospective and there is no actionable data to help them improve. Adoption of cost measures should be contingent on CMS provision of claims data on a regular ongoing basis (for example monthly) so that the group can create their own applications for CMS cost measure-related episodes, and evaluate the specific reasons for cost excess. Because CMS cost measures involve medical services beyond an individual practice or group's own EMR or billing systems, physician groups do not have the necessary data without receipt of claims details from CMS. Receipt of that data would allow organizations the ability to make modifications to improve their cost-effectiveness. This is analogous to what is currently happening with Accountable Care Organizations. Roji Health Intelligence is a qualified CMS reporting registry. We have created Episode-Based cost measures for our physician practice and group clients based on the data present in their practice management and electronic health records. This information is presented in an on-line interactive format for illumination and comparison of episodes, cost drivers, and outcomes associated with episodes. The purpose is to engage clinicians and practices in understanding and acting on cost and outcome results for patients in episodes of care. For our cost measures to be most valuable, we suggest that our clients' receipt of CMS claims data would ensure that we could assist them in a more complete understanding of variation in costs and interventions that would lower health care expenditures.

Measure Comments (Post-Workgroup Meeting)

Author	Submitted Comment
American Association of Nurse Anesthetists	Providers need information about the applicability of specific measures to their practices. Would this measure be attributed to specific provider or clinician types, such as surgeon, attending physician or anesthesia professional?
Federation of American Hospitals	The Federation of American Hospitals (FAH) recommends that the minimum reliability threshold should 0.7 or higher both at the individual clinician and practice levels and believes that validity testing must demonstrate the presence or absence of correlations of this cost measure to one or more quality measures prior to use in MIPS. As a result, FAH requests that the highest level of MAP recommendation be "Do Not Support with Potential for Mitigation."
American Medical Association	The American Medical Association (AMA) continues to have significant concerns with this measure and believes that revisions to the specifications and additional field testing are necessary. We ask that additional conditions be provided with this recommendation around ensuring that the minimum reliability rate be 0.7 or higher, evaluating costs within the context of the quality of care provided is completed, and removing Part D prescription drug costs.
Association of American Medical Colleges (AAMC)	The Clinician MAP conditionally supported two of the episode-based costs measures (Colon and Rectal Resection [MUC2020-0016] and Melanoma Resection [MUC2020-0018]) for future rulemaking for the Merit-based Incentive Payment System (MIPS) program subject to NQF endorsement. The Clinician MAP did not support the other three episode-based costs measures (Asthma/Chronic Obstructive Pulmonary Disease [MUC20-0015], Diabetes [MUC2020-0017], and Sepsis [MUC2020-0019]) for future rulemaking for the MIPS program with potential for mitigation. Mitigation for the three measures focused on evaluation of the actionability and connection between upstream medical interventions and downstream costs, in addition to NQF endorsement. The AAMC agrees with concerns about episode-based cost measures relying on the suggestion that providing certain upstream preventions will result in lower costs of care, and that lower costs will result in better patient outcomes. Furthermore, the AAMC remains concerned that none of the 13 cost measures are adjusted to account for social risk factors (SRFs). In addition to patient clinical complexity, SRFs can drive differences in average costs. In particular, physicians at academic medical centers (AMCs) care for vulnerable populations of patients who are sicker, poorer, and more complex than patients treated elsewhere. In regard to attribution – AAMC has previously commented that attribution methods used should be clear and transparent to clinicians and that it is critical that there be an accurate determination of the relationship between a patient and a clinician to ensure that the correct clinician is held responsible for the patient's outcomes and costs. Attribution is complicated, given that most patients receive care from numerous clinicians across several facilities, and AAMC has urged CMS to explore better data sources and analytic techniques to support more accurate attribution. In addition, the movement in medicine has been to team-based care, further complicating appropriate attribution to a single clinician. The MAP, through its recommendations, and CMS should be careful not to incent patterns of care that are outdated. The AAMC recommends that the MAP recommendation be "do not support with potential for mitigation" for each of the episode-based cost measures.

MUC20-19 Sepsis Episode-Based Cost Measure

Measure Information

Characteristic	Submitted Information
MUCID	MUC20-0019
Other Measure Identification Numbers	N/A
Title	Sepsis Episode-Based Cost Measure
Program	Merit-based Incentive Payment System-Cost
Workgroup	MAP Clinician Workgroup
In what state of development is the measure?	Fully Developed
State of Development Details	The measure has been developed, field tested, and is now refined based on feedback received from field testing in summer 2020 (August 17 – September 18 2020). As background, a list of draft episode groups and trigger codes, developed with input from a Clinical Committee convened in 2016, was posted by CMS in December 2016 to meet MACRA requirements. Building off this work, episode-based cost measures are developed using a “wave” approach wherein sets of Clinical Subcommittees (CS) are convened to select episode groups to develop into cost measures, while smaller, measure-specific Clinician Expert Workgroups provide detailed input on each component of the measures. The current wave of measure development began in May 2019 and includes 4 CS with a total of 166 members affiliated with 110 professional societies. The Hospital Medicine CS selected this episode group from the December 2016 draft list of episode groups for development. This CS comprises members representing clinician specialty societies in this clinical area. A Clinician Expert Workgroup met in August 2019 to discuss measure specifications for all components of the measure, followed by a webinar in January 2020 for follow-up discussions on service assignment and risk adjustment. The measure will be field tested in the summer of 2020, during which clinicians and stakeholders can learn about the measure and provide input on the draft cost measure specifications. After field testing, the Clinician Expert Workgroups revisited and refined the draft measure specifications based on the stakeholder feedback received. We conducted reliability testing of measures for clinicians (TIN-NPIs) and clinician groups (TINs), constructed using episodes ending between January 1, 2019 and December 31, 2019. Reliability refers to the extent to which a measure reflects true variation between clinicians’ risk-adjusted episode spending, as opposed to random variation. The reliability metric specifically

	captures how much of the variance in a measure is due to systematic differences in episode spending between clinicians, rather than differences in episode spending within a clinician's set of episodes. A measure with high reliability suggests that comparisons of performance across clinicians can be expected to better reflect systematic differences in actual performance. Our testing results indicate that this measure is reliable for clinicians and clinician groups across a range of case minimums. • For TINs at a 10-episode case minimum, the mean reliability was 0.55. For TINs at a 20-episode case minimum, the mean reliability was 0.68. For TINs at a 30-episode case minimum, the mean reliability was 0.75. For TIN-NPIs at a 10-episode case minimum, the mean reliability was 0.37. For TIN-NPIs at a 20-episode case minimum, the mean reliability was 0.47. For TIN-NPIs at a 30-episode case minimum, the mean reliability was 0.55
Measure Description	The Sepsis cost measure evaluates clinicians' risk-adjusted cost to Medicare for patients who receive inpatient medical treatment for sepsis. The measure score is a clinician's average risk-adjusted cost for the episode group across all attributed episodes. This acute inpatient medical condition measure includes services that are clinically related and under the reasonable influence of the attributed clinician's role in managing care during each episode from the clinical event that opens or "triggers" the episode through 45 days after. Medicare beneficiaries enrolled in Medicare Parts A and B during the performance period are eligible for the measure. This measure addresses the Patient-Focused Episode of Care goal of CMS's Meaningful Measures Initiative, the CMS high priority area of Efficiency/Cost Reduction and MACRA statutory requirements.
Numerator	The numerator for the Sepsis measure is the sum of the ratio of observed to expected payment-standardized cost to Medicare for all episodes attributed to a clinician or clinician group. This sum is then multiplied by the national average observed episode cost to generate a dollar figure. Mathematically, this is represented as: sum of (observed episode cost/expected episode cost) * national average observed cost.
Denominator	The denominator for the Sepsis measure is the total number of episodes from this episode group attributed to a clinician or clinician group.
Exclusions	The following episode-level exclusions apply: (a) The beneficiary has a primary payer other than Medicare for any amount of time overlapping the episode window or in the lookback period. (b) No attributed clinician is found for the episode. (c) The beneficiary's date of birth is missing. (d) The beneficiary's death date occurred before the episode ended. (e) The beneficiary was not enrolled in Medicare Parts A and B for the entirety of the lookback period plus episode window, or is enrolled in Part C for any part of the lookback period plus episode window. (f) The episode trigger claim was not performed in an office, IP, OP, or ASC setting based on its place of service. Exclusions specific to the Sepsis measure are developed with input

	from the Sepsis Expert Workgroup.
Measure type	Cost/Resource Use
What is the NQF status of the measure?	Never submitted
NQF ID number	0000
Year of next anticipated NQF CDP endorsement review	
Year of most recent NQF Consensus Development Process (CDP) endorsement	
Is the measure being submitted exactly as endorsed by NQF?	
If not exactly as endorsed, describe the nature of the differences	N/A
What data sources are used for the measure?	Claims
If EHR or Administrative Claims or Chart-Abstracted Data, description of parts related to these sources.	N/A
At what level of analysis was the measure tested?	The measure was tested at both the Clinician: Individual and Clincian: Group levels.

In which setting was this measure tested?	Hospital inpatient acute care facility
What NQS priority applies to this measure?	
What one primary meaningful measure area applies to this measure?	Patient-focused episode of care
What secondary meaningful measure area applies to this measure?	
What one primary healthcare priority applies to this measure?	Make care affordable
What secondary healthcare priority applies to this measure?	
What area of specialty best fits the measure?	Internal Medicine
What is the target population of the measure?	Medicare Fee for Service
Is this measure an eCQM?	No
If eCQM, enter Measure Authoring Tool (MAT) number	No
If eCQM, does the measure	No

have a Health Quality Measures Format (HQMF) specification?	
Comments	N/A
Measure steward	Centers for Medicare & Medicaid Services
Long-Term Measure Steward (if different)	N/A
Measure Steward Contact Information	Evans, Ronique; Center for Clinical Standards and Quality; 410-786-3966; Ronique.Evans1@cms.hhs.gov
Primary Submitter Contact Information	Jensen, Ross; Acumen, LLC; 650-558-8882; macra-episode-based-cost-measures-info@acumenllc.com
Long-Term Measure Steward Contact Information	N/A
Secondary Submitter Contact Information	N/A
Was this measure proposed for a previous year's MUC list?	No
In what prior year(s) was this measure proposed?	
What were the programs that NQF MAP reviewed the measure for in	N/A

each year?	
Why was the measure not recommended in those year(s)?	N/A
What were the MUC IDs for the measure in each year?	N/A
NQF MAP report page number being referenced for each year	N/A
What was the NQF MAP recommendation in each year?	N/A
List the NQF MAP workgroup(s) in each year	N/A
What is the history or background for including this measure on the new MUC list?	New measure never reviewed by MAP Workgroup or used in a CMS program
Range of years(s) this measure has been used by CMS Program(s)	N/A
What other federal programs are currently using this measure?	
Evidence that the measure can be	This is a claims-based measure and will not require any additional submission of data.

operationalized	
How is the measure expected to be reported to the program?	Claims
Is this measure similar to and/or competing with measure(s) already in a program?	No
Which existing measure(s) is your measure similar to and/or competing with?	N/A
How will this measure be distinguished from other similar and/or competing measures?	N/A
Rationale for how this measure will add to the CMS program	N/A
If this measure is being proposed to meet a statutory requirement, please list the corresponding statute.	Section 101(f) of MACRA
Evidence of performance gap	Evidence of Performance Gap response: A recent study indicates that clinician beliefs about treatment may be the most important factors explaining the variation in health care expenditures. However, clinicians are often unaware of how their care decisions influence the overall costs of care. Cost measures are intended to help inform clinicians on the costs attributed to their decision-making and to incentivize cost-effective, high-

quality care. A cost measure offers opportunity for improvement if clinicians can exercise influence on the intensity or frequency of a significant share of costs during the episode, or if clinicians can achieve lower spending and better care quality through changes in clinical practice. According to the literature and feedback received through stakeholder input activities to date, this measure's focus represents an area where there are opportunities for improvement. The primary areas for improvement are early recognition, prompt and appropriate administration of antibiotics and provision of resuscitation, and improved post-discharge care coordination. These interventions may prevent progression of sepsis, thereby avoiding longer hospital stays, higher readmissions, and overall higher cost. There is an opportunity for improvement in preventing more severe forms of sepsis (and related complications) by improving early sepsis screening and recognition. The Surviving Sepsis Campaign's International Guidelines for Management of Sepsis and Septic Shock and other guidelines such as the sepsis 3-hour resuscitation bundle and the 6-hour septic shock bundle all stress the importance of early recognition for sepsis. Various studies have found that delayed sepsis diagnosis and treatment has an adverse effect on sepsis outcomes, including progression to severe sepsis and septic shock, which represents higher mortality and overall cost. A 2020 study found that among all Medicare sepsis hospitalizations in 2018, the average total payment for septic shock cases was over $9,000 more than the average for sepsis hospitalizations. The mean length of stay for septic shock is also substantially longer than for sepsis inpatient stays. Once sepsis is recognized, sepsis treatment may include fluid resuscitation, antimicrobial therapy, source control interventions, vasoactive medications, corticosteroids, blood products, and mechanical ventilation, when necessary. Early recognition of sepsis and adherence to treatment guidelines have been shown to be the primary means of improving sepsis outcomes. Several programs and emerging technologies focused on training clinical staff in early detection of sepsis and prompt administration of antibiotics have been associated with lower inpatient mortality rates and costs. For example, a 2015 study found that a sepsis intervention program yielded an over 8% reduction in the sepsis-associated mortality rate and a significant decrease in Medicare costs without a compensatory rise in post-acute care discharges. This intervention program had 4 components: (i) an intervention designed and refined by a multidisciplinary physician-chaired committee, (ii) a screening tool designed for integration with routine nursing care, (iii) data-driven revisions to screening and response protocols to target higher risk units and patients, and (iv) periodic education and training for all clinical staff on the epidemiology of sepsis along with the proper usage of the screening tool. Another 2016 study found that a sepsis intervention program yielded a lower mortality rate and a reduced length of stay for sepsis patients; its intervention program included parameters for emergent antibiotic therapy, intravenous antibiotics, antimicrobial treatment, source control, and periodic review of available information to appropriately modify the antibiotic treatment. In addition to staff training interventions, as technology progresses, there are improving software products and devices

	that can streamline patient monitoring, blood culture analysis, alerts, and communication. In tandem with training-based interventions, technology solutions may improve the timeliness and subsequent outcomes of sepsis treatments. As post-discharge mortality for sepsis hospitalizations has decreased in the past decade, there is an increasing number of patients surviving sepsis. These patients have an increased risk for new or worsened functional and cognitive impairment as well as worsening of chronic health conditions, leading to increasing risk of readmission. A 2018 literature review on enhancing recovery from sepsis concluded that post-discharge management should focus on the following: (i) screening for common and treatable post-sepsis impairments (e.g., functional disability, swallowing impairment, mental health impairment) and referring to appropriate treatment, (ii) reviewing and adjusting long-term medication for appropriateness, and (iii) evaluating for treatable conditions that commonly result in readmission (e.g., infection, heart failure, and renal failure). This measure aims to address these example areas of opportunities for improvement. Research has shown that sepsis represents a significant share of hospitalizations that have an average length of stay that is greater than other conditions and these hospitalizations are associated with a high hospital mortality rate. As such, the use of this cost measure can provide clinicians with information to improve care outcomes and reduce future health care costs. This is supported by our analysis into the variation in spending per episode that may be reduced through better clinical practices. We analyzed the measure's performance gap for clinicians (TIN-NPIs) and clinician groups (TINs) using episodes ending between January 1, 2019 and December 31, 2019. With no case minimum applied, there were 514,234 sepsis episodes for 448,430 beneficiaries. The TIN-NPI and TIN-level measure scores were calculated for 50,735 clinicians and 6,463 clinician groups who met the 10-episode case minimum. • The mean risk-adjusted cost per episode was $19,794 at the TIN level. The risk-adjusted cost per episode at the 10th percentile was $16,965 compared to $22,930 at the 90th percentile at the TIN level. • The mean risk-adjusted cost per episode was $22,832 at the TIN-NPI level. The risk-adjusted cost per episode at the 10th percentile was $19,216 compared to $26,719 at the 90th percentile at the TIN-NPI level. • For TINs, the mean measure score was 1.01. The score at the 10th percentile was 0.87, compared to 1.17 at the 90th percentile. For TIN-NPIs, the mean measure score was 1.04. The score at the 10th percentile was 0.88, compared to 1.22 at the 90th percentile.
Unintended consequences	Stakeholders have expressed concerns that cost measures could potentially have risks or unintended consequences. During past stakeholder input activities, including Technical Expert Panels (TEPs) and public comment periods, stakeholders have highlighted the importance of considering potential incentives created by cost measures that could adversely affect patients and clinicians. For patients, these potential adverse effects may include: (i) reducing access to care, and (ii) stinting care. For clinicians,

	these adverse effects could include: (i) making unfair comparisons among certain types of clinicians in assessing cost measure performance, and (ii) rewarding lower costs without appropriately balancing performance on quality. Overall, by promoting lower spending, stakeholders have expressed concerns that cost measures could potentially create barriers to patients receiving clinically appropriate care. To mitigate the potential risks of such incentives, stakeholders and TEP members have highlighted the importance of the following: Devising an appropriate risk adjustment model for episode-based cost measures; Aligning cost measures with indicators of quality; Identifying benchmarks for comparison that are fair and appropriate for all eligible clinicians; Potentially excluding certain types of patients from measure calculations.
Which clinical guideline(s)?	N/A
Briefly describe the peer reviewed evidence justifying this measure	Sepsis represents a significant share of hospitalizations and Medicare cost. A recent study indicated that from 2012 to 2018, the annual number of Medicare Parts A and B (fee-for-service) beneficiaries with a sepsis hospitalization (defined as having a sepsis diagnosis) rose from around 800,000 to over 1.1 million; annual total cost for these hospitalizations rose from $17.8 billion to over $22.4 billion. Additionally, the total cost of skilled nursing facility care in the 90 days after the sepsis hospitalization discharge rose from $3.9 billion to over $5.6 billion over that same interval. An earlier study using a 2013 sample estimated that sepsis hospitalizations represented over 8% of Medicare costs. Hospitalizations with sepsis have an average length of stay that is greater than other conditions, and it is longer for cases of septic shock. Sepsis hospitalizations also have a significant level of mortality. According to the Centers for Disease Control and Prevention, at least 1.7 million adults develop sepsis each year, and 1 in 3 patients who die in a hospital have sepsis (i.e., about 270,000 deaths annually). A 2020 study found that the one-week, six-month, and one-year mortality rates for Medicare beneficiaries admitted for sepsis hospitalizations range from 7.2 – 40.6%, 26.5 – 60.1%, and 32.9 – 64.6%, respectively, based on severity. Overall, hospital mortality rate is significantly higher for cases with septic shock. Given the high cost associated with providing care for sepsis and frequent use of post-acute care services following sepsis hospitalizations, sepsis cost measurement provides an opportunity for improvement on overall cost performance. According to the 2020 study of 2012-2018 Medicare sepsis hospitalizations, the average hospital cost in 2018 ranged from about $16,000 to over $29,000, based on severity, with significantly higher cost for cases where sepsis is not present on admission. There are also substantial downstream costs associated with sepsis; for example, patients hospitalized for sepsis are more likely to be discharged to either a short-term care facility or long-term care institution compared to patients hospitalized for other conditions. The 2020 study also found that, within 6 months of discharge, patients hospitalized for sepsis relative to patients hospitalized for other conditions had: (i) 22.6% fewer discharges to the home, (ii) a more than two-fold increase in mortality, and (iii) a larger

share of patients in skilled nursing facilities (or other nursing care), hospice care, or readmitted to an inpatient hospital.

Preliminary Analysis – MUC ID: MUC20-19 Sepsis Episode-Based Cost Measure

Criteria	Yes/No	Justification and Notes
Does the measure address a critical quality objective not currently adequately addressed by the measures in the program set?	Yes	MUC20-0019 addresses the Patient-Focused Episode of Care goal of CMS's Meaningful Measures initiative, the MIPS high priority area of Efficiency/Cost Reduction, and the 2015 Medicare Access and Children's Health Insurance Program (CHIP) Reauthorization Act (MACRA) statutory requirements (section 101(f)). Early recognition, prompt and appropriate administration of antibiotics and provision of resuscitation, and improved post-discharge care coordination have all been shown to lower costs and improve outcomes related to sepsis (Jones, et al., 2015). There are currently no MIPS quality measures related to sepsis and bacteremia.
Is the measure evidence-based and either strongly linked to outcomes or an outcome measure?	Yes	This cost/resource use measure aims to inform clinical decision-making related sepsis by reflecting cost of an episode of care and incentivizing cost-effective interventions. Studies suggest that knowledge and awareness of evidence-based practices and treatment risks can influence decision-making and can lead to lower costs (Cutler, et al., 2016).
Does the measure address a quality challenge?	Yes	Sepsis represents a significant share of hospitalizations and Medicare cost. A recent study indicated that from 2012 to 2018, the annual number of Medicare Parts A and B beneficiaries with a sepsis hospitalization rose from around 800,000 to over 1.1M; annual total cost for these hospitalizations rose from $17.8B to over $22.4B. Additionally, the total cost of skilled nursing facility care in the 90 days after the sepsis hospitalization discharge rose from $3.9B to over $5.6B over that same interval. An earlier study using a 2013 sample estimated that sepsis hospitalizations represented over 8% of Medicare costs (T.G. Buchman, et al., 2020)

Does the measure contribute to efficient use of measurement resources and/or support alignment of measurement across programs?	Yes	This measure is not duplicative of other measures currently within the MIPS program. This and other 2020 measures under consideration were created in response to MACRA requirements to develop measures for potential implementation in the cost performance category of MIPS. Additionally, this measure is an episode-based measure; current cost-based measures in MIPS are all at the population level. Currently there are no sepsis nor bacteremia measures in MIPS. Hospital Inpatient Quality Reporting has a sepsis management composite measure. A methicillin resistant staphylococcus bacteremia measure is used in Hospital Value-Based Payment, Hospital Acquired Condition Reduction Program, and Prospective Payment Exempt Cancer Hospital Quality Reporting Program.
Can the measure can be feasibly reported?	Yes	The measure uses Medicare claims data which is feasibility reported and a low-burden data source.
Is the measure applicable to and appropriately specified for the program's intended care setting(s), level(s) of analysis, and population(s)?	Yes	The measure is specified for Clinician: Individual and Clinician: Group/Practice levels, which aligns with MIPS reporting categories. The measure has not been reviewed for endorsement by an NQF Standing Committee.
If the measure is in current use, have negative unintended issues to the patient been identified? Have implementation challenges outweighing the benefits of the measure have been identified?	N/A	The measure is not in current use. The submission identified a potential unintended consequence of reduction in access to care or stinting care. This measure could result in rewarding cost reductions with no net quality of care gain if it is not appropriately balanced with measures addressing performance on quality.
PAC/LTC Core Concept?		N/A
Impact Act Domain		N/A

Hospice High Priority Areas		N/A
Rural Workgroup Input	[redacted]	Relative priority/utility: • This measure was noted to be relevant to rural clinicians and hospitals, especially internal medicine. Data collection issues: • None identified Calculation issues: • Measure was suggested to be potentially more reliable for clinician groups over individual clinicians. • Average cost for rural providers may exceed the national average cost and therefore may unduly impact rural providers. Unintended consequences: • Cost measures may result in restriction of care if not paired with quality measures. Votes: Range is 1 – 5, where higher is more relevant to rural. Average: 3.5 • 1 – 0 votes • 2 – 2 votes • 3 – 5 votes • 4 – 12 votes • 5 – 0 votes
Preliminary Analysis Recommendation	Do Not Support with Potential for Mitigation	MAP did not support the measure with potential for mitigation, with the mitigation points being NQF endorsement, an analysis of the potential for gaming associated with overdiagnosis of sepsis, and further evaluation of the correlation with clinical quality measures.
Summary: What is the potential value to the program measure set?	[redacted]	MUC20-0019 addresses the Patient-Focused Episode of Care goal of CMS's Meaningful Measures initiative, the MIPS high priority area of Efficiency/Cost Reduction and MACRA statutory requirements. Currently, there are no MIPS measures that assess episode-based cost related to sepsis.

Summary: What is the potential impact of this measure on quality of care for patients?	[redacted]	This measure was devised to reduce costs to Medicare septicemia-related events which represent a significant share of hospitalizations and Medicare cost. MAP noted that the annual number of Medicare beneficiaries with a sepsis hospitalization exceeds 1.1M, with over $22B in costs. Should testing data show that the measure appropriately measures episode-based cost while maintaining quality and a clear indication that there is not gaming through overdiagnosis, this measure would be valuable to add to the program measure set.

Measure Comments

Author	Submitted Comment
American Physical Therapy Association	We encourage the measure steward to include physical therapists as eligible clinicians that can be scored under this cost measure. Physical function remains below population norms for survivors and often does not return to pre-sepsis levels. Patients with sepsis may acquire neurological impairments such as delirium and impaired consciousness during hospitalization through a variety of conditions, including cerebral ischemia, metabolic derangements, and neuroinflammation. Among survivors, long-term impairments are seen in memory, attention, verbal fluency, and executive functioning.
University of Colorado Medicine	Questioning the ability to control this measure and/or the start/end of the episode is vague.
American Medical Association	Overall, the AMA supports the episode-based measure development process and movement to episode-based measures over broad cost measures. However, we continue to have significant concerns with the lower than desirable mean reliability rate. The AMA believes that the minimum acceptable thresholds should be 0.7. In addition, we strongly support the tenet that cost must be assessed within the context of the quality of care provided; yet, none of these assessment have been provided for the existing MIPS cost measures or the ones currently under MAP consideration. We request further refinements to the process to ensure the cost measures are accurate and fair when used in MIPS and, ultimately, publicly reported. The AMA strongly opposes including Part D prescription drug costs in the Medicare episode-based cost measures. To hold physicians accountable for costs that are negotiated between CMS and Medicare Prescription Drug Plans is fundamentally problematic, and physicians and patients do not always have information about coverage, formularies, out-of-pocket costs and list prices at the point of care. This also assumes there is a viable, evidence-based, less expense alternative option for patients. In general, we urge the MAP to make “Conditional Support” its highest level of recommendation and to recommend “Do Not Support with Potential for Mitigation” for any measures where relevant specialties have raised serious concerns or the average reliability rate is less than 0.7 as is the case with several of

	these measures.
Association of American Medical Colleges (AAMC)	The AAMC remains concerned that such cost measures are adjusted to account for social risk factors (SRFs). In addition to patient clinical complexity, SRFs can drive differences in average costs. In particular, physicians at academic medical centers (AMCs) care for vulnerable populations of patients who are sicker, poorer, and more complex than patients treated elsewhere. In regard to attribution – AAMC has previously commented that attribution methods used should be clear and transparent to clinicians and that it is critical that there be an accurate determination of the relationship between a patient and a clinician to ensure that the correct clinician is held responsible for the patient's outcomes and costs. Attribution is complicated, given that most patients receive care from numerous clinicians across several facilities, and AAMC has urged CMS to explore better data sources and analytic techniques to support more accurate attribution. The AAMC recommends that: (1) cost measures include risk-adjustment for SRFs, (2) the attribution methodology is transparent, and (3) the appropriate clinician is held responsible for the patient's outcomes and costs. Additionally, this measure has not been submitted for NQF endorsement. The AAMC has long held that measures should not be proposed for addition to public reporting programs unless vetted and endorsed by the NQF. The AAMC recommends that the highest level of MAP recommendation be "Do Not Support With Potential For Mitigation."
AdvaMed	AdvaMed strongly supports this measure, as it would provide useful information to support the development of new algorithms for early diagnosis and therapeutic guidance for Sepsis.
Roji Health Intelligence	For each cost measure under consideration, physician practices or groups must have the ability to examine data on episodes for procedures and conditions that reveal how these episodes relate to the cost target. In all cases involving multi-physician and facility care, this will require data and transparency from CMS, rather than simply providing aggregate final results. With CMS proposed measures, physician groups can only react to measurement but lack the ability to improve because (a) they don't have the data to replicate the episode-based measure data, since much of the costs are indirectly generated by other providers yet invisible to the attributed group; and (b) all of the scores are retrospective and there is no actionable data to help them improve. Adoption of cost measures should be contingent on CMS provision of claims data on a regular ongoing basis (for example monthly) so that the group can create their own applications for CMS cost measure-related episodes, and evaluate the specific reasons for cost excess. Because CMS cost measures involve medical services beyond an individual practice or group's own EMR or billing systems, physician groups do not have the necessary data without receipt of claims details from CMS. Receipt of that data would allow organizations the ability to make modifications to improve their cost-effectiveness. This is analogous to what is currently happening

with Accountable Care Organizations.

Roji Health Intelligence is a qualified CMS reporting registry. We have created Episode-Based cost measures for our physician practice and group clients based on the data present in their practice management and electronic health records. This information is presented in an on-line interactive format for illumination and comparison of episodes, cost drivers, and outcomes associated with episodes. The purpose is to engage clinicians and practices in understanding and acting on cost and outcome results for patients in episodes of care. For our cost measures to be most valuable, we suggest that our clients' receipt of CMS claims data would ensure that we could assist them in a more complete understanding of variation in costs and interventions that would lower health care expenditures.

Measure Comments (Post-Workgroup Meeting)

Author	Submitted Comment
Federation of American Hospitals	Support
American Medical Association	The American Medical Association (AMA) continues to have significant concerns with this measure and believes that revisions to the specifications and additional field testing are necessary. We ask that additional conditions be provided with this recommendation around ensuring that the minimum reliability rate be 0.7 or higher, evaluating costs within the context of the quality of care provided is completed, and removing Part D prescription drug costs.
AdvaMed	AdvaMed strongly supports this measure, as it would provide useful information to support the development of new algorithms for early diagnosis and therapeutic guidance for Sepsis.

MUC20-0034 Risk-Standardized Acute Unplanned Cardiovascular-Related Admission Rates for Patients with Heart Failure for the Merit-based Incentive Payment System Title

Measure Information

Characteristic	Submitted Information
MUCID	MUC20-0034
Other Measure Identification Numbers	N/A
Title	Risk-Standardized Acute Unplanned Cardiovascular-Related Admission

	Rates for Patients with Heart Failure for the Merit-based Incentive Payment System
Program	Merit-based Incentive Payment System-Quality
Workgroup	Clinician
In what state of development is the measure?	Fully Developed
State of Development Details	The measure is fully developed. We tested the measure using October 1, 2017 – September 30, 2018 as the measurement year in a MIPS Medicare FFS population (e.g. we only include patients who are assigned to a MIPS-eligible provider in 2017 or 2018). 2,444,341 patients met the inclusion/exclusion criteria. The initial individual-level attribution algorithm assigned 49.5% of patients to PCPs and 47.2% to cardiologists; it left unassigned 3.3% of patients who did not visit in the measurement year or whose pattern of visits did not allow us to identify the clinician most responsible for the patients' care. We further excluded patients who were attributed to a non-MIPS eligible clinician and who were not at risk for admission. Thus, the final HF cohort used for model building and testing included 1,855,941 patients. We calculated measure scores for all TINS; of note, the TIN-level analysis includes all clinicians – those who report as individuals and those reporting through MIPS groups. As expected, the results showed wide variation in the number of patients per TIN, ranging from 1 to 6,275 patients, with a median of 7 and an interquartile range (IQR) of 2 to 19. The measure scores also showed wide variation at the TIN level. When calculated for TINs (n=45,093), risk-standardized acute cardiovascular-related admission rates (RSCAR) measure scores, ranged from 9.6 to 62.4 per 100 person-years, with a median of 24.8 and an IQR of 24.0 to 25.9 per 100 person-years. The minimum sample sizes needed to achieve provider-level measure score reliabilities of 0.5 and 0.4 or greater are ≥32 and ≥21 patients per TIN, respectively. At these thresholds, reliability scores ranged from 0.50 to nearly 1.0, with a median value of 0.70 and an IQR of 0.58-0.84 and 0.40 to nearly 1, with a median value of 0.60 and an IQR of 0.48-0.78, respectively. If CMS established these cutoffs of ≥32 or ≥21 patients per TIN for public reporting, 83.2% or 76.1% of the TINs would be excluded, respectively; however, the measure would include 84.5% or 88.9% of patients, respectively, and 64.8% or 69.8% of clinicians, respectively. We developed the measure with input from a national TEP, Clinician Committee, and CMS. CMS decided to not adjust the measure for Medicare-Medicaid dual-eligibility status in June 2019; this form therefore was updated in July 2019 to reflect this update and include results of testing. We also polled both the TEP and Clinician Committee on face validity. Below are the results: TEPOf the 17 TEP members who were active through the end of the project, 12 responded. The majority of the respondents, 10/12 or 83%, moderately or somewhat agreed that the MIPS HF measure can be used to distinguish good from poor quality of care. Similarly, the majority of

	the respondents, 9/12 or 75%, moderately or somewhat agreed that the MIPS HF measure scores (RSCARs) will provide MIPS TINs with information that could be used to improve the quality of care for HF patients. Clinician Committee Of 13 Clinician Committee members who were active through the end of the project, 13 responded. The majority of the respondents, 11/13 or 85%, strongly, moderately, or somewhat agreed that the MIPS HF measure can be used to distinguish good from poor quality of care. Similarly, the majority of the respondents, 11/13 or 85%, strongly, moderately, or somewhat agreed that the MIPS HF measure scores will provide MIPS TINs with information that could be used to improve the quality of care for HF patients.
Measure Description	Annual risk-standardized rate of acute, unplanned cardiovascular-related admissions among Medicare Fee-for-Service (FFS) patients aged 65 years and older with heart failure (HF) or cardiomyopathy.
Numerator	The outcome for this measure is the number of acute cardiovascular-related admissions per 100 person-years at risk for admission during the measurement year. Time at risk is calculated as the number of days a patient is alive, from the start of the measurement period or first visit, until heart transplantation, LVAD implantation, or home inotropic therapy; enrollment in hospice; death; or the end of the measurement period. Time not considered at risk and excluded: Days spent in a hospital, SNF, or acute rehabilitation facility; 10 days following discharge from a hospital, SNF, or acute rehabilitation facility; and Time during and after LVAD implantation, home inotropic therapy, or heart transplantation. Acute cardiovascular-related admissions are defined using individual ICD-10-CM codes and the Agency for Healthcare Research and Quality's (AHRQ) Clinical Classification Software (CCS) diagnosis categories, which group clinically similar codes together. AHRQ CCS diagnosis categories used to define outcome: 55: Fluid and electrolyte disorders; 96: Heart valve disorders; 97: Peri-; endo-; and myocarditis; cardiomyopathy (except that caused by tuberculosis or sexually transmitted disease); 98: Essential hypertension; 100: Acute myocardial infarction; 102: Nonspecific chest pain; 104: Other and ill-defined heart disease; 105: Conduction disorders; 106: Cardiac dysrhythmias; 107: Cardiac arrest and ventricular fibrillation; 108: Congestive heart failure; non-hypertensive; 110: Occlusion or stenosis of precerebral arteries; 112: Transient cerebral ischemia; 115: Aortic; peripheral; and visceral artery aneurysms; 116: Aortic and peripheral arterial embolism or thrombosis; 157: Acute and unspecified renal failure; 245: Syncope. Subsets of the following AHRQ CCS diagnosis categories used to define outcome: 99: Hypertension with complications and secondary hypertension; 101: Coronary atherosclerosis and other heart disease; 103: Pulmonary heart disease; 109: Acute cerebrovascular disease; 114: Peripheral and visceral atherosclerosis; 117: Other circulatory disease; 130: Pleurisy; pneumothorax; pulmonary collapse; 131: Respiratory failure; insufficiency; arrest (adult); 133: Other lower respiratory disease; 237: Complication of device; implant or graft. The measure has several outcome exclusions: Planned admissions; Admissions from a skilled nursing facility

	(SNF) or acute rehab facility; Admissions within 10 days of discharge from a hospital, SNF, or acute rehab; Admissions after patient has entered hospice; Admissions before first visit to provider if no prior year visit; Admissions at time of or following: LVAD implantation, home inotropic therapy, or heart transplant.
Denominator	The measure includes Medicare FFS beneficiaries ≥65 years of age with at least one inpatient principal diagnosis for heart failure/cardiomyopathy, or at least two outpatient or inpatient heart failure/cardiomyopathy diagnoses in any coding position (e.g., primary or secondary position) within the two years prior to the measurement year. Beneficiaries must be enrolled full-time in Medicare Part A and B during the year prior to measurement and during the measurement period. Additionally, the cohort excludes: Patients with internalized left ventricular assist devices (LVADs); Patients with heart transplants; Patients on home inotropic therapy; Patients on hospice for any reason; Patients with end-stage renal disease (ESRD) – defined as chronic kidney disease stage 5 or on dialysis. Provider types included for measurement (vetted by TEP and Clinician Committee): Primary care providers (PCPs): CMS designates PCPs as physicians who practice internal medicine, family medicine, general medicine, or geriatric medicine, and non-physician providers, including nurse practitioners, certified clinical nurse specialists, and physician assistants; Cardiologists: Cardiologists are covered by the measure because they provide overall coordination of care for patients with HF and manage the conditions that put HF patients at risk for admission due to acute cardiovascular-related conditions. Outcome attribution: We begin by assigning each patient to the clinician most responsible for the patient's care, based on the pattern of outpatient visits with PCPs and relevant specialists. The patient can be assigned to a PCP, a cardiologist, or can be left unassigned. A patient who is eligible for attribution is assigned to a cardiologist if they have 2 or more visits with a single cardiologist, regardless of how many visits that patient has with a PCP. There are two scenarios where a patient can be assigned to a PCP. First, if the patient has seen the PCP at least once but has no visits with a cardiologist, the patient is assigned to the PCP. Second, if the patient has seen the PCP more than 2 or more times and has only one visit with a cardiologist, the patient is assigned to the PCP. If the patient has 1 visit each with a cardiologist and a PCP, the patient is assigned to the cardiologist. If the patient has 1 visit with a cardiologist and no visit with a PCP, the patient is assigned to the cardiologist. Finally, the patient will be unassigned if they had no visits with a PCP or cardiologist. Patients are then assigned at the Taxpayer Identification Number (TIN) level, which includes solo clinicians and groups of clinicians who have chosen to report their quality under a common TIN. Patients "follow" their clinician to the TIN designated by the clinician (i.e. they are assigned to their clinician's TIN). Patients unassigned at the individual clinician-level, therefore, continue to be unassigned at the TIN level.
Exclusions	Numerator Exclusions: The measure does not include the following types of admissions in the outcome because they do not reflect the quality of care

	provided by ambulatory care clinicians who are managing the care of HF patients: Planned admissions (utilizes the adapted planned admission algorithm (PAA) to identify and exclude admissions that are planned); Admissions that likely do not reflect the quality of heart failure management provided by ambulatory clinicians including: Admissions that occur within 10 days of discharge from a hospital, skilled nursing facility, or acute rehabilitation facility ("10-day buffer period"); Admissions that occur while patients are enrolled in Medicare's hospice benefit; Admissions that occur prior to the first visit with the assigned clinician. Admissions on the date or after any of the following: LVAD implantation, home inotropic therapy, or heart transplant (censored at the time of transition to advanced care). Denominator Exclusions: The measure excludes: 1. Patients without continuous enrollment in Medicare Parts A and B for the duration of the measurement period. 2. Patients who (or until death), were ever in hospice during the year prior to the measurement year or in hospice at the start of the measurement period. 3. Patients who have had no Evaluation & Management (E&M) visits to a MIPS eligible clinician. 4. Patients who have had a heart transplant, been on home inotropic therapy, or who have had a left ventricular assist device (LVAD) placed.
Measure type	Outcome
What is the NQF status of the measure?	Never submitted
NQF ID number	0000
Year of next anticipated NQF CDP endorsement review	The measure will be submitted to the NQF in the spring of 2021 NQF cycle (January 2021).
Year of most recent NQF Consensus Development Process (CDP) endorsement	None
Is the measure being submitted exactly as endorsed by NQF?	NA
If not exactly as endorsed, describe the nature of the	N/A

differences	
What data sources are used for the measure?	Administrative Claims; Other (enter here): Medicare Enrollment Database; 2017 and 2018 MIPS provider eligibility files; Agency for Healthcare Research Quality (AHRQ) Socioeconomic Status (SES) Index derived from American Community Survey data; Area Health Resources File
If EHR or Administrative Claims or Chart-Abstracted Data, description of parts related to these sources.	N/A
At what level of analysis was the measure tested?	Individual; Group – The measure was tested at the Taxpayer Identification Number (TIN) level; the TIN could be either a group of providers or a solo provider. About half of the TINs with at least one HF patient assigned to them had only one provider.
In which setting was this measure tested?	Ambulatory/office-based care
What NQS priority applies to this measure?	
What one primary meaningful measure area applies to this measure?	Management of chronic conditions
What secondary meaningful measure area applies to this measure?	Admissions and readmissions to hospitals
What one primary healthcare priority applies to this measure?	Promote effective prevention and treatment of chronic disease
What secondary healthcare priority applies to this measure?	Promote effective communication and coordination of care

What area of specialty best fits the measure?	Cardiovascular disease (cardiology)
What is the target population of the measure?	Medicare Fee-for-Service patients, 65 years or older who are enrolled in both Part A and Part B
Is this measure an eCQM?	No
If eCQM, enter Measure Authoring Tool (MAT) number	No
If eCQM, does the measure have a Health Quality Measures Format (HQMF) specification?	No
Comments	We anticipate this measure also being applicable to the following specialties: Primary care, Family practice, General practice, Geriatric medicine, and Internal medicine. Additional data sources used for this measure include the following: Medicare Enrollment Data.
Measure steward	Centers for Medicare & Medicaid Services
Long-Term Measure Steward (if different)	NA
Measure Steward Contact Information	Nicole Hewitt, Ph.D., CCSQ/QMVIG/Division of Quality Measurement; Centers for Medicare and Medicaid Services; (410) 786-7778; nicole.hewitt@cms.hhs.gov
Primary Submitter Contact Information	Altaf, Faseeha; Center for Outcomes Research and Evaluation; 860-752-5471; faseeha.altaf@yale.edu
Long-Term Measure Steward Contact	NA

Information	
Secondary Submitter Contact Information	NA
Was this measure proposed for a previous year's MUC list?	No
In what prior year(s) was this measure proposed?	None
What were the programs that NQF MAP reviewed the measure for in each year?	N/A
Why was the measure not recommended in those year(s)?	N/A
What were the MUC IDs for the measure in each year?	N/A
NQF MAP report page number being referenced for each year	N/A
What was the NQF MAP recommendation in each year?	N/A
List the NQF MAP workgroup(s) in each year	N/A
What is the	New measure never reviewed by MAP Workgroup or used in a CMS

history or background for including this measure on the new MUC list?	program
Range of years(s) this measure has been used by CMS Program(s)	N/A; this is a new measure. However, another version of this measure specified for Accountable Care Organizations (ACOs), “Risk-standardized Acute Admission Rates for Patients with Heart Failure” (ACO-37, NQF ID 2886) was previously used in the CMS Medica
What other federal programs are currently using this measure?	NA
Evidence that the measure can be operationalized	As the measure uses Medicare administrative claims data, the measure presents no additional data collection burden to providers or CMS.
How is the measure expected to be reported to the program?	Claims; Other: Medicare Part B administrative claims
Is this measure similar to and/or competing with measure(s) already in a program?	Yes
Which existing measure(s) is your measure similar to and/or competing with?	There is one related measure: ACO-37/NQF #2887 – Risk-Standardized Acute Admission Rates for Patients with Heart Failure. The measure being submitted is adapted from NQF #2887 to assess the quality of ambulatory care provided by individual clinicians and clinician groups caring for patients with heart failure and cardiomyopathy, and the measure is being adapted for the MIPS setting.
How will this measure be distinguished from other similar and/or competing	The MIPS HF admission measure is adapted from the ACO HF admission measure (ACO-37), which was implemented in the Medicare Shared Savings Program in 2015. There are three main ways that the newly developed MIPS measure differs from its predecessor used in the ACO setting. Cohort: The cohort additionally includes patients with cardiomyopathy. Outcome: CMS narrowed the outcome to focus on admissions whose risk can be reduced by clinicians/groups providing high-

measures?	quality ambulatory care, so that the measure can be used to assess ambulatory (rather than ACO-wide) care quality. As such, the outcome is acute cardiovascular-related admissions, which is different than the ACO HF measure's outcome of all-cause acute unplanned admissions. Risk-adjustment: CMS added a social risk factor (the AHRQ SES Index) to the risk-adjustment model.
Rationale for how this measure will add to the CMS program	N/A
If this measure is being proposed to meet a statutory requirement, please list the corresponding statute.	Medicare Access and Children's Health Insurance Program (CHIP) Reauthorization Act of 2015 (MACRA). Of note, MACRA specifically calls for outcome measures.
Evidence of performance gap	Across the 45,093 TINs who had at least one heart failure patient, RSCAR measure scores, ranged widely from 9.6 to 62.4 per 100 person-years, with a median of 24.8 and an IQR of 24.0 to 25.9 per 100 person-years. Overall, measure results suggest that there is substantial opportunity to reduce the number of admissions for this patient population and decrease the variation in admissions across providers, and that improvement goals are achievable.
Unintended consequences	MIPS eligible clinicians may choose to focus their quality improvement efforts on reducing admissions for a certain limited number of reasons, at the expense of implementing strategies that could affect a broader range of potentially preventable hospitalizations. In addition, there may be the potential for gaming and avoiding coding reasons for admission that are included in the outcome. These consequences may be more likely when effective mechanisms to reduce rates vary substantially in the implementation cost for MIPS providers. The shift to outpatient care may also result in admissions of higher complexity patients and thus may increase readmissions – an outcome captured by several measures, including this one. Moreover, clinicians may be incentivized to avoid clinically complex patients, patients that require elevated levels of extra-clinical support (e.g. coordination services) or patients with a poor history of compliance with prescribed therapies. To mitigate this unintended consequence, thoughtful risk adjustment of patient clinical, frailty and social risk characteristics during measure development has been conducted.
Which clinical guideline(s)?	N/A

Briefly describe the peer reviewed evidence justifying this measure	Hospital admission rates are an effective marker of ambulatory care quality. Hospital admissions from the outpatient setting reflect a deterioration in patients' clinical status and, as such, reflect an outcome that is meaningful to both patients and providers. In addition, hospitalization increases potential exposure to iatrogenic injury and the increasingly recognized toxic effects of hospitalization (e.g., sleep deprivation; poor nourishment; deconditioning from inactivity; confusion from medications; stress from mental exhaustion) leading to "post hospitalization syndrome [1]," which may contribute to the risk of readmission. Patients receiving optimal, coordinated high-quality care should use fewer inpatient services than patients receiving fragmented, low-quality care. Thus, high population rates of hospitalization may, at least to some extent, signal poor quality of care or inefficiency in health system performance. There is evidence that ambulatory care clinicians can influence admission rates by providing high quality of care [2-8]. For example, Brown et al. pointed to four ambulatory care focused Medicare Coordinated Care Demonstration programs that reduced hospitalizations for high-risk patients by 13-30 events per 100 beneficiaries per year (8-33% of hospitalizations). Brown et al. highlighted six program features that were associated with successfully reducing hospitalizations: 1) supplementing patient telephone calls with in-person meetings; 2) occasionally meeting in-person with providers; 3) acting as a communication hub for providers; 4) providing patients with evidence-based education; 5) providing strong medication management; and 6) providing comprehensive and timely transitional care after hospitalizations [2]. In addition, van Loenen et al. found that higher levels of provider continuity decreased the risk of avoidable hospitalizations for ambulatory care-sensitive conditions (ACSCs) and chronic diseases [7]. Hussey et al. [9] found that among Medicare beneficiaries, greater continuity of care was associated with lower hospitalization odds (OR=0.94, CI=0.93-0.95). Favorable results (declines in admissions) were also shown by Dorr et al. (2000), Levine et al. (2012), Littleford et al. (2010), and Zhang et al. (2008) [3-5,8]. Several studies have demonstrated positive impact of early follow-up after hospitalization to reduce readmissions for HF [10-13]. Data from the Centers for Disease Control indicate that "heart failure costs the nation an estimated $30.7 billion in 2012 [18]. This total includes the cost of health care services, medicines to treat heart failure, and missed days of work". Therefore specifically, the reduction of heart failure admissions may be attributed to clinical care that represents interventions that prevent overall admission for these types of patients. Several studies have estimated the cost of HF care. One study found that between 2002-2011, direct expenditures for HF patients were four times as high as for those without HF or $3,446 after adjusting for demographics and comorbidities. Direct costs increased by about 28%, from $21,316 to

$27,152, over 10 years (2002/2003 – 2010/2011), largely driven by increases in inpatient costs [14].
Another study revealed that, on average, 79% of lifetime costs of HF care are accrued during hospitalization [15]. Thus, reducing the number of hospitalizations should significantly reduce healthcare costs related to HF. Second, the measure can promote processes of care that have also been identified as cost-effective therapies. A cost-effectiveness analysis of three medical therapies compared to diuretics alone found that ACE inhibitors and the combination ACE inhibitors and beta-blockers were both cost-saving and more effective therapies. Furthermore, adding aldosterone inhibitors resulted in an additional $501/life-year, compared to ACE inhibitors and beta-blockers alone [16]. Third, effective care can reduce morbidity and, as a result, costly emergency interventions [17]. In particular, provider-led teams that work cooperatively to improve patient outcomes have been shown to reduce admission risk and, therefore, costly care. Johnston et al. found that having a disease-relevant specialist involved in a care episode resulted in a 21.3% lower incidence of ambulatory care sensitive (ACS) hospitalizations ($p<0.05$) but was not associated with lower rates of admissions for heart-failure related ACS or HF [9].

References 1. Krumholz HM. Post-Hospital Syndrome — An Acquired, Transient Condition of Generalized Risk. New England Journal of Medicine. 2013;368(2):100-102. 2. Brown RS, Peikes D, Peterson G, Schore J, Razafindrakoto CM. Six Features of Medicare Coordinated Care Demonstration Programs That Cut Hospital Admissions of High-Risk Patients. Health Affairs. 2012;31(6):1156-1166. 3. Dorr DA, Wilcox AB, Brunker CP, Burdon RE, Donnelly SM. The Effect of Technology-Supported, Multidisease Care Management on the Mortality and Hospitalization of Seniors. Journal of the American Geriatrics Society. 2008;56(12):2195-2202. 4. Levine S, Steinman BA, Attaway K, Jung T, Enguidanos S. Home care program for patients at high risk of hospitalization. The American journal of managed care. 2012;18(8):e269-e276. 5. Littleford A, Kralik D. Making a difference through integrated community care for older people. Journal of Nursing and Healthcare of Chronic Illness. 2010;2(3):178-186. 6. Sommers LS, Marton KI, Barbaccia JC, Randolph J. Physician, Nurse, and Social Worker Collaboration in Primary Care for Chronically Ill Seniors. Archives of Internal Medicine. 2000;160(12):1825-1833. 7. Van Loenen T, Faber MJ, Westert GP, Van den Berg MJ. The impact of primary care organization on avoidable hospital admissions for diabetes in 23 countries. Scandinavian journal of primary health care. 2016;34(1):5-12. 8. Zhang NJ, Wan TTH, Rossiter LF, Murawski MM, Patel UB. Evaluation of chronic disease management on outcomes and cost of care for Medicaid beneficiaries. Health Policy. 2008;86(2):345-354. 9. Hussey PS, Schneider EC, Rudin RS, Fox DS, Lai J, Pollack CE. Continuity and the Costs of Care for Chronic Disease Care Continuity and Costs for Chronic Disease Care Continuity and Costs for Chronic Disease. JAMA Internal Medicine. 2014;174(5):742-748. 10. Donaho EK, Hall AC, Gass JA, et al. Protocol-Driven Allied Health Post-Discharge Transition Clinic to Reduce Hospital Readmissions in Heart Failure. Journal of the American Heart Association. 2015;4(12):e002296. 11. Lee KK, Yang J, Hernandez AF, Steimle AE, Go

	AS. Post-discharge Follow-up Characteristics Associated With 30-Day Readmission After Heart Failure Hospitalization. Medical Care. 2016;54(4):365-372. 12. Murtaugh CM, Deb P, Zhu C, et al. Reducing Readmissions among Heart Failure Patients Discharged to Home Health Care: Effectiveness of Early and Intensive Nursing Services and Early Physician Follow-Up. Health Services Research. 2017;52(4):1445-1472. 13. Ryan J, Kang S, Dolacky S, Ingrassia J, Ganeshan R. Change in Readmissions and Follow-up Visits as Part of a Heart Failure Readmission Quality Improvement Initiative. The American Journal of Medicine. 2013;126(11):989-994.e981. 14. Echouffo-Tcheugui JB, Bishu KG, Fonarow GC, Egede LE. Trends in health care expenditure among US adults with heart failure: The Medical Expenditure Panel Survey 2002-2011. Am Heart J. 2017;186:63-72. 15. Dunlay SM, Shah ND, Shi Q, et al. Lifetime costs of medical care after heart failure diagnosis. Circulation Cardiovascular quality and outcomes. 2011;4(1):68-75. 16. Banka G, Heidenreich PA, Fonarow GC. Incremental Cost-Effectiveness of Guideline-Directed Medical Therapies for Heart Failure. Journal of the American College of Cardiology. 2013;61(13):1440-1446. 17. Bilchick K, Moss T, Welch T, et al. Improving Heart Failure Readmission Costs and Outcomes With a Hospital-to-Home Readmission Intervention Program. American Journal of Medical Quality. 2018:1062860618788436. 18. Center for Medicare & Medicaid Services. National Health Expenditures 2017 Highlights

Preliminary Analysis – MUC ID: MUC20-0034 Risk-Standardized Acute Unplanned Cardiovascular-Related Admission Rates for Patients with Heart Failure for the Merit-based Incentive Payment System Title

Criteria	Yes/No	Justification and Notes
Does the measure address a critical quality objective not currently adequately addressed by the measures in the program set?	Yes	MUC20-0034 addresses three areas that CMS has identified as high-priority measure areas for MIPS: patient outcomes, communication/care coordination and cost reduction. It also meets the Meaningful Measures priority of hospital admissions and management of chronic conditions. Cardiovascular outcomes associated with heart failure (HF) are an important quality focus for Medicare. There are currently three HF measures and two atrial fibrillation measures in MIPS, as well as two readmission measures (all-cause readmissions and readmissions after principal procedure).
Is the measure evidence-based and either strongly linked to outcomes or an outcome measure?	Yes	This measure is an outcomes measure. The measure developer cites research suggesting that HF costs the United States $30.7 billion/year, and 79% of lifetime costs from HF occur during hospitalization (Heidenreich, et al., 2011; Dunlay, et al., 2011). Hospitalization rates for HF patients can be reduced by 20-30% through intensive outpatient support programs that provide high-quality care (ex. multidisciplinary teams, frequent in-person contact, patient education) (Feltner, et al., 2014; Thomas, et al., 2013; McAlister, et al., 2004; Gwawy-Sridhar et al., 2004; Holland, et al., 2005; Jovicic et al., 2006).
Does the measure address a quality challenge?	Yes	MUC20-0034 addresses readmissions for patients with heart failure due to acute cardiovascular events. While readmission rates for HF have fallen since passage of the ACA, readmissions still occurred within a month for a fifth of patients hospitalized with heart failure in 2016 (Desai, et al., 2016; Dharmarajan et al., 2016). The measure developer shared that during testing, risk-standardized acute cardiovascular-related admissions rate (RSCAR) scores ranged widely across the 45,000 TINs in their sample, from 9.6 to 62.4 per 100 person-years, suggesting wide performance variation and opportunity for improvement.

Does the measure contribute to efficient use of measurement resources and/or support alignment of measurement across programs?	Yes	There are currently no outcome measures in MIPS related to heart failure. A version of this measure with slightly different specifications was implemented in MSSP from 2015-2019 (ACO-37/NQF 2886); ACO-37 was specified for ACO-wide use and was retired from MSSP due to overlap with ACO-38 (Risk-Standardized Acute Admission Rates for Patients with Multiple Chronic Conditions) (CMS, 2018). There is an HF readmission measure included in the Hospital Readmissions Reduction Program. HF outcomes measures are present in Inpatient Quality Reporting and Hospital Value Based Purchasing.
Can the measure be feasibly reported?	Yes	The measure uses Medicare claims data which is feasibly reported and a low-burden data source.
Is the measure applicable to and appropriately specified for the program's intended care setting(s), level(s) of analysis, and population(s)?	Yes	The measure is specified for Clinician: Individual and Clinician: Group/Practice levels, which aligns with MIPS reporting categories. The developer noted that cutoff values of ≥32 or ≥21 patients per TIN are needed to achieve provider-level measure score reliabilities of 0.5 and 0.4. At these thresholds, median reliability scores of 0.70 and 0.60 would be achieved. The developer also noted wide variation in the number of patients per TIN with a median of 7 and an interquartile range (IQR) of 2 to 19. This suggests that >75% of clinicians did not meet the developer's threshold for reliability. The measure has not been reviewed for endorsement by an NQF Standing Committee.
If the measure is in current use, have negative unintended issues to the patient been identified? Have implementation challenges outweighing the benefits of the measure have been identified?	No	This measure is new and is not currently being used. The developer identified the following potential unintended consequences: Clinicians may focus quality improvement efforts on specific reasons for readmissions instead of broader prevention, or may "game" the system and avoid coding reasons for admission – especially if implementing programs to reduce HF readmissions are costly. Clinicians may avoid clinically complex patients, patients that require elevated levels of support/coordination, or patients with poor compliance with therapies, but developer has tried to mitigate this through risk adjustment model.
PAC/LTC Core Concept?		N/A
Impact Act Domain		N/A

Hospice High Priority Areas		N/A
Rural Workgroup Input	[redacted]	Relative priority/utility: • Heart failure is a significant problem in rural settings and therefore also relevant. • It was suggested that extending this measure to MIPS from ACOs makes sense for rural providers as well. Data collection issues: • None identified Calculation issues: • None identified Unintended consequences: • None identified Votes: Range is 1 – 5, where higher is more relevant to rural. Average: 3.9 • 1 – 0 votes • 2 – 0 votes • 3 – 1 vote • 4 – 18 votes • 5 – 0 votes
Preliminary Analysis Recommendation	Conditional Support for Rulemaking	Conditional support for rulemaking is contingent on NQF endorsement.
Summary: What is the potential value to the program measure set?	[redacted]	MUC20-0034 addresses MIPS high-priority areas including patient outcomes, care coordination and cost reduction, as well as the Meaningful Measures areas of admissions and readmissions to hospitals and management of chronic conditions. If included, MUC20-0034 would be the only outcome measure in MIPS related to heart failure.

Summary: What is the potential impact of this measure on quality of care for patients?	[redacted]	6.5M Americans are living with heart failure, and a fifth of patients hospitalized with heart failure are readmitted to the hospital within 30 days. Hospitalization is costly and accounts for 79% of lifetime costs associated with heart failure. However, a 20-30% reduction in hospitalization rates can be achieved for heart failure patients through high-quality care with patient support programs. MUC20-0034 encourages clinicians to reduce readmissions through high-quality ambulatory care. Conditional support for rulemaking is contingent on NQF endorsement.

Measure Comments

Author	Submitted Comment
University of Colorado Medicine	Overall we think this is a good measure. One of the main criticisms of the inpatient HRRP is that most avoidable admissions are the index admission, not the readmission. This puts more effort there. The exclusions seem thoughtful.
C-TAC	We support this measure since, with appropriate and ongoing monitoring of symptoms, people with heart failure can be kept out of the hospital. We hope that the implementation of such a MIPS measure will also promote home-based primary and palliative care as ways to reduce such unplanned hospitalizations
American Medical Association	The AMA does not support inclusion of this measure in the Merit-based Incentive Payment System (MIPS). We do not believe that continuing to include measures based on administrative claims meets the intended goals of this program and while this measure may be useful at the community or population level, it is not appropriate to attribute this utilization to an individual physician or practices. This concern is due to several factors. Specifically, the lack of evidence to support applying this measure to individual physicians or practices is particularly concerning since much of what is provided in the preliminary analysis demonstrates that improved care coordination and programs focused on care management can lead to reductions in hospital admissions but required the involvement of multiple partners such as a disease management program, health system, and/or hospital. We do not believe that sufficient evidence was provided to support the theory that physicians or practices, in the absence of some coordinated program or payment offset (e.g., care management fee), can implement structures or processes that can lead to improved outcomes for these patients. We are also concerned with the attribution methodology given the measure compares all physicians, regardless of specialty. A patient who is eligible for attribution is assigned to a cardiologist if they have 2 or more visits with a single cardiologist, regardless of how many visits that patient may have had with a primary care physician. Based on this approach, heart failure specialists will appear to have very poor patient outcomes. Therefore, the measure disincentives physicians from

	providing necessary lifesaving care. Medicare is essentially capping the number of patients a physician may hospitalize by creating a measure that is tied to payment that states all admission related to heart failure are bad. The AMA is also extremely concerned with the continued retrospective approach that prevents this measure from providing timely, meaningful and actionable data at the point of care, the need for this measure to demonstrate a high level of reliability (0.70 at a minimum), the continued lack of robust testing of the validity of the measure including demonstrating that the assignment of the measure to specific physicians, groups, and specialties is clinically appropriate and tied to their ability to meaningfully influence the outcome as well as empiric validity testing. In addition, it remains unclear whether the risk adjustment model for this measure adequately addresses the ongoing concerns around socioeconomic factors since there is no rationale provided on why CMS determined that adjusting for dual eligibility was not warranted. As a result, the AMA recommends that the highest level of MAP recommendation be "Do Not Support."
Association of American Medical Colleges (AAMC)	This measure has not been submitted for NQF endorsement. The AAMC has long held that measures should not be proposed for addition to public reporting programs unless vetted and endorsed by the NQF. The AAMC recommends that the highest level of MAP recommendation be "Conditional Support for Rulemaking" on condition that measure is NQF-endorsed.
AdvaMed	AdvaMed strongly supports this measure as it would provide useful information to support the development of new technologies for early detection of progressive heart failure for patients at home in order to provide early therapy and avoid re-admission.
Heart Failure Society of America	The Heart Failure Society of America (HFSA) appreciates the opportunity to provide comments on the Risk-Standardized Acute Unplanned Cardiovascular-Related Admission Rates for Patients with Heart Failure measure (MUC 20-0034) under consideration by CMS for the Merit-based Incentive Payment System (MIPS). HFSA is a multidisciplinary organization working to improve and expand heart failure care through collaboration, education, research, innovation, and advocacy. Its vision is to significantly reduce the burden of heart failure. The HFSA urges the MAP to recommend that CMS not support this measure for inclusion in MIPS. We have significant concerns about assigning hospitalization rates per capita to a single clinician (or even clinician groups), particularly when our current health care system is increasingly team-based. As such, we do not believe this measure is appropriate for a physician-level accountability program like MIPS. A more appropriate strategy for measurement, particularly in a pay-for-performance program, would be to focus on actions that are in the direct control of the physician or else to use this type of measure for facility or system-level (e.g., ACOs, the VA, etc.) accountability. In addition to the attribution problem, creating an individual MIPS metric of hospitalizations per capita could create perverse incentives for physicians to withhold care. On a hospital level, "success" on the 30-day readmission metric

	(relative to "predicted", the latter based on a weak predictive model) has been found to be associated with an excess mortality over the same time frame. If CMS were to penalize individual providers and essentially tell them that they have a cap on the number of patients "they" may hospitalize, this would create a powerful disincentive to deliver potentially life-saving care. Every major heart failure trial looking at hospitalizations as an adverse event does so accounting for the competing risk of death. If your patient dies, he/she will not be hospitalized. MUC 20-0034 does not seem to account for the competing risk of death and it is unclear if CMS would simultaneously evaluate excess number of deaths per capita. HFSA agrees with the measure steward that hospitalizations put patients at risk of exposure to adverse events, and we recognize the importance of continuity of follow-up post-discharge. However, we believe that clinician-level measurement of heart failure management needs to shift its focus from pure utilization metrics to coupling utilization with quality care delivery and reducing adverse events. For example, clinician-level metrics should focus on providing guideline-directed medical therapy (GDMT) and improving management of hypertension and diabetes, which all have the potential to reduce hospitalizations by making our patients healthy. Outcomes, namely survival, should be measured at the hospital-level. Similarly, it would be much more valuable to count a hospital readmission if the patient did not have a follow-up arranged in 7-10 days or the hospital did not discharge a patient on GDMT. Many patients make appointments and just do not show for follow-up. It is also not uncommon that they do not fill medications— often these patients are underprivileged or underinsured and cannot afford medications (especially in January of each year when copays start over). Thus, if you your patient does not own a car and does not have a smart phone or internet access for e-visits, the clinician is limited in his/her ability to prevent readmissions. HFSA does not believe that MUC 20-0034 adequately adjusts for these social determinants and other risk factors.
The Society for Cardiovascular Angiography and Interventions (SCAI)	January 6, 2021 National Quality Forum Measure Application Partnership VIA NQF WEBSITE On behalf of the Society for Cardiovascular Angiography and Interventions (SCAI), I am writing to recommend two preliminary items contained on the 2020-2021 Measure Under Consideration (MUC) list published and released by the Centers for Medicare and Medicaid Services (CMS) on December 21, 2020. Specifically, we are highly supportive of the following two Measures Under Consideration: • Appropriate Treatment for ST Segment Elevation Myocardial Infarction (STEMI) Patients in the Emergency Department (ED), and

	• Risk-Standardized Acute Unplanned Cardiovascular Related Admission Rates for Patients with Heart Failure for the Merit-based Incentive Payment System The Society for Cardiovascular Angiography and Interventions (SCAI) is a non-profit professional association with over 5,000 members representing the majority of practicing interventional cardiologists and cardiac catheterization teams in the United States, including those providing percutaneous coronary interventions (PCI). SCAI promotes excellence in invasive and interventional cardiovascular medicine through education, representation and the advancement of quality standards to enhance patient care. SCAI diligently participated in the development of these MUCs, working in cooperation with the American College of Cardiology (ACC), the American Heart Association (AHA) and others. SCAI members added the experience, expertise, clinical judgment and especially the value of those physicians that have earned the FSCAI and MSCAI specialty designations to this important work. Only after completing the rigors of medical school, three years of training in internal medicine, 3 more years of training in cardiology and 1 to 2 years of additional cardiology specialization is the value of the SCAI designation is earned. We believe that adding this measure to the MUC list will add value and improve patient outcomes that will likely become a de facto standard of care in this highly complex area. We stand ready to work with you and the Centers for Medicare and Medicaid Services (CMS) to ensure that the benefits of these measures do not outweigh the burden of data collection and reporting now and throughout the challenging process of implementation. We also pledge to continue to provide our experience and expertise related to Quality Improvement, certification and recognition, regulatory and accreditation, public reporting, disease surveillance and adequate payment, to this critical process. As you review these MUC list items and provide input into Medicare programs, including the Merit-based Incentive Payment System (MIPS) and Medicaid Savings Programs (MSSP), please consider SCAI and its members as a critical resource that remains available to you at any time. Please contact Emily Senerth, Senior Manager, Clinical Documents & Quality, should you have questions. Sincerely, Cindy Grines, MD MSCAI President, SCAI
Federation of American Hospitals	The Federation of American Hospitals (FAH) strongly advocates that any measure that is proposed for use in payment programs should be appropriate for accountability purposes at the designated level of attribution and demonstrated to be reliable and valid. As a result, the FAH does not support inclusion of this measure in the Merit-based Incentive Payment System (MIPS).

CMS must ensure that the data produced yields scores that more accurately and consistently represent the quality of care provided by an individual clinician or practice. As such, the FAH recommends that CMS increases the minimum sample size to a higher number to produce a minimum reliability threshold of sufficient magnitude (e.g. 0.7 or higher) in light of the reliability range from 0.50 to nearly 1.00 for clinicians and 0.40 to nearly 1.00 for practices. The FAH also does not believe that face validity is sufficient to demonstrate that the measure as attributed provides appropriate and evidence-based representations of the care provided by these clinicians. We strongly encourage CMS to validate these measures through additional testing,

such as predictive and construct validity, to ensure that the application of the measure to an individual clinician or practice is appropriate and yields scores that are valid and useful. Additional information on how the risk model performs is needed as well as robust testing to assess the impact that social risk factors must be provided, particularly since no reason was given on why CMS chose not to include dual eligibility in the model.

The FAH believes that there is insufficient evidence to support attribution to individuals or groups, particularly with the attribution assigned retrospectively while beneficiaries for accountable care organizations are assigned prospectively, the minimum sample size and reliability threshold remain too low, and additional testing on the validity of the measure when applied at these levels is needed. The National Quality Forum should also endorse the measure prior to finalization. As a result, we request that the highest level of MAP recommendation be “Do Not Support.”

Measure Comments (Post-Workgroup Meeting)

Author	Submitted Comment
American Heart Association/American Stroke Association	The AHA is strongly opposed to attributing this claims-based measure to individual physicians or small group practices and we urge the MAP not to recommend it for this purpose. Use at this level of attribution is simply not appropriate. It is not clear that measures holding individual physicians or practices accountable would be equitable or constructive. We are also concerned about the potential avoidance of patients more likely to be non-compliant and the possibility of exacerbating disparities in care. It is also very questionable whether individual providers or even smaller multi-specialty practices could achieve adequate case minimums to fairly apply a population-level measure such as this. We believe it is would be much more appropriate to attribute such measures at the Accountable Care Organization (ACO) or system level.
Federation of American Hospitals	The Federation of American Hospitals (FAH) recommends that attribution be evidence-based and the minimum reliability threshold should 0.7 or higher both at the individual clinician and practice levels. In addition, empiric validity testing must be completed prior to use in MIPS. The FAH asks that the MAP recommendation also include these conditions.

American Medical Association	The American Medical Association (AMA) does not support inclusion of this measure in the Merit-based Incentive Payment System (MIPS) as it is based on administrative claims and is not appropriate to attribute this utilization to an individual physician or practices. Specifically, the lack of evidence to support applying this measure to individual physicians or practices is particularly concerning since much of what is provided in the preliminary analysis demonstrates that improved care coordination and programs focused on care management can lead to reductions in hospital admissions but required the involvement of multiple partners such as a disease management program, health system, and/or hospital. We do not believe that sufficient evidence was provided to support the theory that physicians or practices, in the absence of some coordinated program or payment offset (e.g., care management fee), can implement structures or processes that can lead to improved outcomes for these patients. The AMA is also extremely concerned with the continued retrospective approach that prevents this measure from providing timely, meaningful and actionable data at the point of care, the need for this measure to demonstrate a high level of reliability (0.70 at the minimum threshold), the continued lack of robust testing of the validity of the measure including demonstrating that the assignment of the measure to specific physicians, groups, and specialties is clinically appropriate and tied to their ability to meaningfully influence the outcome as well as empiric validity testing. In addition, it remains unclear whether the risk adjustment model for this measure adequately addresses the ongoing concerns around social risk factors since there is no rationale provided on why CMS determined that adjusting for dual eligibility was not warranted. As a result, the AMA requests that the highest level of MAP recommendation be "Do Not Support."
AdvaMed	AdvaMed strongly supports this measure as it would provide useful information to support the development of new technologies for early detection of progressive heart failure for patients at home in order to provide early therapy and avoid re-admission.
Heart Failure Society of America	The Heart Failure Society of America (HFSA) appreciates the opportunity to comment on the MAP's draft recommendations on CMS' 2020 MUC List. As noted in our January 6th comment letter to the MAP, the HFSA has significant concerns about MUC 20-0034: Risk-Standardized Acute Unplanned Cardiovascular-Related Admission Rates for Patients with Heart Failure, which is a measure under consideration by CMS for the Merit-based Incentive Payment System (MIPS). We strongly urge the MAP to re-consider its recommendation of "conditional support for rulemaking contingent on NQF endorsement" for this measure as we do not believe it is an appropriate measure of clinician-level quality. As we expressed in our January 6th comments, HFSA shares CMS' goal of improving patients' health. However, we continue to strongly advise against the use of MUC 20-0034 for clinician-level accountability. Metrics that count hospitalizations are misguided in that they focus purely on utilization, without regard to quality, and create perverse incentives by rewarding clinicians who up-code, avoid certain high-risk patients, or whose patients

die without being admitted to the hospital. We are already seeing the impact of these perverse incentives in hospital-level programs that target readmissions. If CMS now shifts this framework to MIPS, it could be disastrous for our patients, particularly the sickest and most vulnerable ones. Utilization metrics such as this measure are simply not appropriate for clinician-level accountability. Individual clinicians instead should be measured on their adherence to the principle of improving health. They should be incentivized to adopt processes and tools, such as guideline-directed medical therapy (GDMT) and systems to arrange follow-up care, that drive quality and favorable outcomes, including reducing both hospitalization rates and mortality. Finally, we remind the MAP that heart failure patients have multiple comorbidities. In fact, more than half of hospitalizations among these patients are unrelated to worsening heart failure. As we expressed earlier, the risk adjustment methodology associated with this measure is inadequate in that it relies exclusively on claims data and on generally rigid variables that do not fully account for severity of illness, medical complexity, and social determinants of health, all of which are critical drivers of heart failure admissions.
Below, we once again share our earlier concerns:
HFSA's January 6th comments

The HFSA appreciates the opportunity to provide comments on MUC 20-0034: Risk-Standardized Acute Unplanned Cardiovascular-Related Admission Rates for Patients with Heart Failure, a measure under consideration by CMS for the Merit-based Incentive Payment System (MIPS). HFSA is a multidisciplinary organization working to improve and expand heart failure care through collaboration, education, research, innovation, and advocacy. Its vision is to significantly reduce the burden of heart failure.

The HFSA urges the MAP to recommend that CMS not support this measure for inclusion in MIPS. We have significant concerns about assigning hospitalization rates per capita to a single clinician (or even clinician groups), particularly when our current health care system is increasingly team-based. As such, we do not believe this measure is appropriate for a physician-level accountability program like MIPS. A more appropriate strategy for measurement, particularly in a pay-for-performance program, would be to focus on actions that are in the direct control of the physician or else to use this type of measure for facility or system-level (e.g., ACOs, the VA, etc.) accountability.

In addition to the attribution problem, creating an individual MIPS metric of hospitalizations per capita could create perverse incentives for physicians to withhold care. On a hospital level, "success" on the 30-day readmission metric (relative to "predicted", the latter based on a weak predictive model) has been found to be associated with an excess mortality over the same time frame. If CMS were to penalize individual providers and essentially tell them that they have a cap on the number of patients "they" may hospitalize, this would create a powerful disincentive to deliver potentially life-saving care.

Every major heart failure trial looking at hospitalizations as an adverse event

	does so accounting for the competing risk of death. If your patient dies, he/she will not be hospitalized. MUC 20-0034 does not seem to account for the competing risk of death and it is unclear if CMS would simultaneously evaluate excess number of deaths per capita. HFSA agrees with the measure steward that hospitalizations put patients at risk of exposure to adverse events, and we recognize the importance of continuity of follow-up post-discharge. However, we believe that clinician-level measurement of heart failure management needs to shift its focus from pure utilization metrics to coupling utilization with quality care delivery and reducing adverse events. For example, clinician-level metrics should focus on providing guideline-directed medical therapy (GDMT) and improving management of hypertension and diabetes, which all have the potential to reduce hospitalizations by making our patients healthy. Outcomes, namely survival, should be measured at the hospital-level. Similarly, it would be much more valuable to count a hospital readmission if the patient did not have a follow-up arranged in 7-10 days or the hospital did not discharge a patient on GDMT. Many patients make appointments and just do not show for follow-up. It is also not uncommon that they do not fill medications— often these patients are underprivileged or underinsured and cannot afford medications (especially in January of each year when copays start over). Thus, if you your patient does not own a car and does not have a smart phone or internet access for e-visits, the clinician is limited in his/her ability to prevent readmissions. HFSA does not believe that MUC 20-0034 adequately adjusts for these social determinants and other risk factors.

MUC20-0040 Intervention for Prediabetes

Measure Information

Characteristic	Submitted Information
MUCID	MUC20-0040
Other Measure Identification Numbers	N/A
Title	Intervention for Prediabetes
Program	Merit-based Incentive Payment System-Quality
Workgroup	Clinician
In what state of development is the measure?	Fully Developed
State of Development Details	This measure is fully developed. The measure was submitted and not endorsed. Data element validity testing, face validity, and feasibility testing was completed on this measure in Q3 and Q4 2019. Using specifications defined by the measure developer, two testing sites were able to access and extract the critical data elements including all components of the numerator, all components of the denominator, and all components of the exclusions, demonstrating that the measure is feasible to collect (see feasibility scorecard for details).Data element validity testing was conducted utilizing Parallel Forms Reliability Testing methodology. The two sites provided a patient-level extract of data from the EHR for the measure. A sample of these data was then compared against a manual review of the medical record. Sample size requirements were calculated based on estimated rates for the measure by site using a calculator based on the calculation defined by Donner-Eliasziw. The data element validity results including Kappa scores are presented below[Table]Site 1, Denominator, % Agree = 93; Kappa = 0.775; N = 112Site 1, Numerator, % Agree = 72; Kappa = 0.448; N = 25Site 2, Denominator, % Agree = 81; Kappa = 0.609; N = 74Site 2, Exclusions, % Agree = 83; Kappa = 0.657; N = 52Site 2, Numerator, % Agree = 100; Kappa = **; N = 20**Kappa scores not calculable with multiple non-responses by raters (i.e., all No/No or all Yes/Yes) For this measure, Kappa scores ranged from .448 to .775 which is considered moderate to substantial. We found instances across the measures where more full and accurate information could be found in the manual abstraction process than through electronic reporting: Numerator – Referrals to diabetes prevention program or dietician are often automated messages. These can be seen in manual abstraction and depending on level of access to the EHR system, not all medical staff can see these messages. With increased use of this measure, you anticipate that the capture of this information in discrete fields and consistently across visits will

	improve. For face validity, an external group of clinical and methodological experts assessed the measure for face validity through an on-line survey. After the measure was fully specified, the expert panel was asked to rate their agreement with the following statement: “The scores obtained from the measure as specified will accurately differentiate quality across providers”. Scale 1-5, where 1= Strongly Disagree; 3=Neither Agree nor Disagree; 5=Strongly Agree, N/A = Not Applicable The panel rating of the validity statement for the measure were as follows: N = 22; Mean rating = 4.05 and 82% of respondents either agree or strongly agree that this measure can accurately distinguish good and poor quality. Frequency Distribution of Ratings5 (Strongly Agree) – 8 4 (Agree) – 103 (Neither Agree nor Disagree) – 2 2 (Disagree) – 1 1 (Strongly Disagree) – 1 X (Not Applicable) – 0.
Measure Description	Percentage of patients aged 18 years and older with identified abnormal lab result in the range of prediabetes during the 12-month measurement period who were provided an intervention
Numerator	Patients who were provided an intervention.* NOTE: *Intervention must include one of the following: referral to a CDC-recognized diabetes prevention program; referral to medical nutrition therapy with a registered dietician; prescription of metformin.
Denominator	All patients aged 18 years and older with identified abnormal lab result in the range of prediabetes during the 12-month measurement period. NOTE: **Abnormal lab result in the range of prediabetes includes a fasting plasma glucose level between 100 mg/dL (5.6 mmol/L) to 125 mg/dL (6.9 mmol/L) OR a 2-hour glucose during a 75g oral glucose tolerance test between 140 mg/dL (7.8 mmol/L) to 199 mg/dL (11.0 mmol/L) OR and A1C between 5.7-6.4% (39-47 mmol/mol).
Exclusions	Denominator exclusions: Patients who are pregnant. Patients who have any existing diagnosis of diabetes (Type 1, Type 2, latent autoimmune diabetes of adults [LADA], monogenic diabetes [MODY]), hospice care in the ambulatory setting.
Measure type	Process
What is the NQF status of the measure?	Not endorsed
NQF ID number	3570e
Year of next anticipated NQF CDP endorsement review	Undetermined
Year of most recent NQF	Spring 2020 NQF PCCI Committee review – not endorsed.

Consensus Development Process (CDP) endorsement	
Is the measure being submitted exactly as endorsed by NQF?	No information given
If not exactly as endorsed, describe the nature of the differences	NA
What data sources are used for the measure?	EHR
If EHR or Administrative Claims or Chart-Abstracted Data, description of parts related to these sources.	NA
At what level of analysis was the measure tested?	Clinician; Group;
In which setting was this measure tested?	Ambulatory/office-based care
What NQS priority applies to this measure?	NA
What one primary meaningful measure area applies to this measure?	Preventive care
What secondary meaningful	Care is personalized and aligned with patient's goals

measure area applies to this measure?	
What one primary healthcare priority applies to this measure?	Promote effective prevention and treatment of chronic disease
What secondary healthcare priority applies to this measure?	Strengthen person and family engagement as partners in their care
What area of specialty best fits the measure?	Internal Medicine/Primary Care
What is the target population of the measure?	All Payer
Is this measure an eCQM?	Yes
If eCQM, enter Measure Authoring Tool (MAT) number	Yes, 0c9106d1-6f9d-4ea0-b2f0-b7cbd4817852
If eCQM, does the measure have a Health Quality Measures Format (HQMF) specification?	Yes
Comments	Disclaimer Notice: The Measures are not clinical guidelines, do not establish a standard of medical care, and have not been tested for all potential applications. The Measures, while copyrighted, can be reproduced and distributed, without modification, for noncommercial purposes, e.g., use by health care providers in connectio with their practices.
Measure steward	American Medical Association
Long-Term Measure	NA

Steward (if different)	
Measure Steward Contact Information	Beth A. Tapper, MA; Senior Program Manager; Improving Health Outcomes; American Medical Association; Beth.Tapper@ama-assn.org
Primary Submitter Contact Information	Beth A. Tapper, MA; Senior Program Manager; Improving Health Outcomes; American Medical Association; 312-933-6636; Beth.Tapper@ama-assn.org
Long-Term Measure Steward Contact Information	NA
Secondary Submitter Contact Information	NA
Was this measure proposed for a previous year's MUC list?	No
In what prior year(s) was this measure proposed?	NA
What were the programs that NQF MAP reviewed the measure for in each year?	NA
Why was the measure not recommended in those year(s)?	NA
What were the MUC IDs for the measure in each year?	NA
NQF MAP report page number	NA

being referenced for each year	
What was the NQF MAP recommendation in each year?	NA
List the NQF MAP workgroup(s) in each year	NA
What is the history or background for including this measure on the new MUC list?	NA
Range of years(s) this measure has been used by CMS Program(s)	NA
What other federal programs are currently using this measure?	NA
Evidence that the measure can be operationalized	This measure is an eCQM and all data that are collected in this measure will be collected from a practice or health system EHR or data warehouse. This measure was tested for feasibility in two sites using 2 different EHR systems (Epic, Cerner) and they were able to operationalize the measure. The feasibility scorecard is attached and demonstrates that most data elements can be captured electronically, some available in structured fields and notes. The two above referenced testing sites are now making plans to fully implement and utilize the measures health system-wide beginning in October 2020.
How is the measure expected to be reported to the program?	eCQM; CQM (Registry)

Is this measure similar to and/or competing with measure(s) already in a program?	No
Which existing measure(s) is your measure similar to and/or competing with?	None
How will this measure be distinguished from other similar and/or competing measures?	This measure is the first measure to address the treatment of prediabetes, and is part of a set that will produce the first measurement set in the U.S. intended to prevent type 2 diabetes
Rationale for how this measure will add to the CMS program	This measure is critical to proving interventions to patients with prediabetes who are not screened, thus missing potential cases that progress to diabetes. This measure is part of a set that will produce the first measurement set in the U.S. intended to prevent type 2 diabetes. Currently, eighty-four million Americans have prediabetes and 9 out 10 patients are unaware that they have this condition. CDC-recognized lifestyle change programs are included in the health benefit plans and the Medicare Diabetes Prevention Program, including Medicare beneficiaries.
If this measure is being proposed to meet a statutory requirement, please list the corresponding statute.	NA
Evidence of performance gap	Significant gaps exist in the treatment of patients that have prediabetes. Patients who are diagnosed with prediabetes benefit from referral to intervention programs, however research shows that most patients with prediabetes are not referred for intervention. Moreover, data from the 2012 National Ambulatory Medical Care Survey shows that only 23% of visits that were associated with prediabetes showed that any type of referral or intervention was made. In another study, survey data show that while providers report following patients with prediabetes closely, only 11% reported referring to a behavioral weight loss program. Data support that there is room for improvement in providing patients with prediabetes an intervention. Specifically related to opportunities to improve referrals to

	DPPs, a study of primary care physicians (PCPs) reported they provide referrals to DPPs on average to 45% of their newly diagnosed patients with pre-diabetes. Another study in the American Journal of Preventive Medicine showed that only 23% of physicians report referring any patients to the DPP. A recent survey determined that primary care physicians have significant knowledge gaps regarding prediabetes screening, diagnosis, and management, with less than 20% of physicians correctly answering questions in those domains around the evidence and appropriate treatment and management A study assessing the rates of prediabetes recognition and treatment documented in the EHR, found that in the 6 months after identification of prediabetes, 18% of patients had their blood glucose levels retested; 13% received a physician diagnosis of prediabetes/hyperglycemia; 31.0% had prediabetes, diabetes, or lifestyle documented in the clinical notes; and <0.1% initiated metformin, demonstrating a significant gap in treatment and management. Additionally, there is good evidence that individualized Medical Nutrition Therapy (MNT) provided by a registered dietitian nutritionist (RDN) is successful in deterring the progression of prediabetes to type 2 diabetes. Dietitians who provide individualized MNT demonstrate the use of the extended care team in partnering with patients to prevent type 2 diabetes. When we tested this measure, performance for this measure was 25%, so a significant area for improvement.
Unintended consequences	No unintended consequences were identified during testing of this measure. AMA will continue to monitor whether any are identified during implementation of the measure.
Which clinical guideline(s)?	This measure is based on evidence-based guidelines from the United States Preventive Services Task Force (USPSTF) and from the American Diabetes Association (ADA). Clinicians should offer or refer patients with abnormal blood glucose to intensive behavioral counseling interventions to promote a healthful diet and physical activity. (USPSTF, 2015) (B recommendation)Patients with prediabetes should be referred to an intensive behavioral lifestyle intervention program modeled on the Diabetes Prevention Program to achieve and maintain 7% loss of initial body weight and increase moderate-intensity physical activity (such as brisk walking) to at least 150 min/week. (ADA,2020) (A).Metformin therapy for prevention of type 2 diabetes should be considered in those with prediabetes, especially for those with BMI ≥35 kg/m2, those aged <60 years, women with prior gestational diabetes mellitus. (ADA,2020) (A)As is the case for those with diabetes, individualized medical nutrition therapy is effective in lowering A1C in individuals diagnosed with prediabetes. Medical Nutrition Therapy Recommendation: An individualized medical nutrition therapy program as needed to achieve treatment goals, provided by a registered dietitian nutritionist (RD/RDN), preferably one who has comprehensive knowledge and experience in diabetes care, is recommended for all people with type 1 or type 2 diabetes, prediabetes, and gestational diabetes mellitus. (ADA, 2020) (A)This measure will support the USPSTF and ADA guidelines for treatment of patients with prediabetes.

Briefly describe the peer reviewed evidence justifying this measure	This measure is based on evidence-based guidelines from the United States Preventive Services Task Force (USPSTF) and from the American Diabetes Association (ADA). Clinicians should offer or refer patients with abnormal blood glucose to intensive behavioral counseling interventions to promote a healthful diet and physical activity. (USPSTF, 2015) (B recommendation). Numerous peer reviewed evidence-based publications exist on the treatment and management of patients with prediabetes. One of the largest publications to date comes from the January 2020 ADA Journal, Diabetes Care: Standards of Medical Care in Diabetes—2020. This the largest compilation of evidence and recommendations for the diagnosis and treatment of patients with prediabetes and diabetes. Patients with prediabetes should be referred to an intensive behavioral lifestyle intervention program modeled on the Diabetes Prevention Program to achieve and maintain 7% loss of initial body weight and increase moderate-intensity physical activity (such as brisk walking) to at least 150 min/week. Metformin therapy for prevention of type 2 diabetes should be considered in those with prediabetes, especially for those with BMI ≥35 kg/m2, those aged <60 years, women with prior gestational diabetes mellitus. As is the case for those with diabetes, individualized medical nutrition therapy is effective in lowering A1C in individuals diagnosed with prediabetes. Lifestyle Management: An individualized MNT program, preferably provided by a registered dietitian, is recommended for all people with type 1 or type 2 diabetes, prediabetes, or gestational diabetes mellitus.

Preliminary Analysis – MUC ID: MUC20-0040 Intervention for Prediabetes

Criteria	Yes/No	Justification and Notes
Does the measure address a critical quality objective not currently adequately addressed by the measures in the program set?	Yes	MUC20-0040 addresses the Meaningful Measure area of Preventive Care. Clinicians who identify patients with prediabetes can reduce risk of diabetes onset through clinical and lifestyle interventions. There are currently eight MIPS quality measures related to diabetes, but no measures in MIPS related to prediabetes.

Is the measure evidence-based and either strongly linked to outcomes or an outcome measure?	Yes	The developer presents evidence that individualized Medical Nutrition Therapy (MNT) provided by a registered dietitian nutritionist (RDN) is successful in deterring the progression of prediabetes to type 2 diabetes (Early, et al., 2018; Parker, et al., 2017). Current evidence supports a role for metformin in diabetes prevention when coupled with lifestyle interventions in people with prediabetes (Hostelek, et al., 2015).
Does the measure address a quality challenge?	Yes	Developer notes a median performance for this measure of 25%, implying opportunities for improvement. Developer cites data from the 2012 National Ambulatory Medical Care Survey where 23% of visits associated with prediabetes had a referral or intervention made (Mainous, et al., 2016). In another study, survey data show that while providers report following patients with prediabetes closely, only 11% reported referring to a behavioral weight loss program (Tseng, et al., 2017). A study of primary care physicians (PCPs) reported they provide referrals to diabetes prevention programs (DPP) on average to 45% of their newly diagnosed patients with prediabetes (Kiefer, et al., 2015).
Does the measure contribute to efficient use of measurement resources and/or support alignment of measurement across programs?	Yes	This measure is not duplicative of other measures currently within the MIPS program. No prediabetes measures were identified within other CMS quality programs, though diabetes measures are common.
Can the measure can be feasibly reported?	Yes	The measure is an electronic clinical quality measure (eCQM) that draws upon EHR data documented as part of the routine delivery of care.
Is the measure applicable to and appropriately specified for the program's intended care setting(s), level(s) of analysis, and population(s)?	No	MUC20-0040 was evaluated by the NQF Primary Care and Chronic Illness (PCCI) Standing Committee during the Spring 2020 measure evaluation period as NQF 3570e. The measure was not endorsed. The Committee was concerned that this measure has limited interventions available to meet the numerator, requiring clinicians to either prescribe metformin or refer the patient out to another service. The PCCI Committee consider this is be not representative of the range of interventions available to primary care clinicians to address prediabetes and to be burdensome to providers and patients.

If the measure is in current use, have negative unintended issues to the patient been identified? Have implementation challenges outweighing the benefits of the measure have been identified?	N/A	MUC20-0040 has not been implemented. The PCCI Standing Committee identified the unintended consequence associated with the limitations of interventions for the measure where providers are required to either prescribe metformin or to refer the patient out.
PAC/LTC Core Concept?		N/A
Impact Act Domain		N/A
Hospice High Priority Areas		N/A

Rural Workgroup Input	[redacted]	Relative priority/utility: • The measure was noted to have very specific interventions that will introduce limitations in the rural setting with limited access to referral services to meet the numerator. Data collection issues: • None identified Calculation issues: • None identified Unintended consequences: • None identified Votes: Range is 1 – 5, where higher is more relevant to rural. Average: 3.6 • 1 – 1 vote • 2 – 3 votes • 3 – 0 votes • 4 – 12 votes • 5 – 2 votes
Preliminary Analysis Recommendation	Do Not Support With Potential for Mitigation	Mitigation points include re-specifying the measure to include an adequate range of interventions for prediabetes available to the clinician beyond prescription of metformin or referring the patient to an external service. The measure should also receive NQF endorsement.
Summary: What is the potential value to the program measure set?	[redacted]	MUC20-0040 addresses the Meaningful Measure area of Preventive Care. Clinicians who identify patients with prediabetes can reduce risk of diabetes onset through clinical and lifestyle interventions. Prevention measures are of high value to MIPS and there are currently no prediabetes measures in MIPS.

Summary: What is the potential impact of this measure on quality of care for patients?	[redacted]	Prediabetes and diabetes are important conditions within the Medicare population resulting in high mortality, morbidity and cost of care. Diabetes has preventable risk factors and can be addressed through intervention. Medical Nutrition Therapy has been shown to be successful in deterring the progression of prediabetes to type 2 diabetes. Current evidence supports a role for metformin in diabetes prevention when coupled with lifestyle interventions in people with prediabetes. However, the measure was noted by the NQF Primary Care and Chronic Illness Committee to offer too few options for intervention. The measure is recommended as do not support for rulemaking with potential for mitigation. The mitigation points are that the measure should be re-specified to address the PCCI Committee concerns and then receive NQF endorsement.

Measure Comments

Author	Submitted Comment
American Physical Therapy Association	We encourage physical therapists to be recognized as eligible clinicians who can report this measure. Physical therapy interventions focused on lifestyle intervention (diet, exercise, behavior modification) can reduce the risk of developing diabetes.
Novo Nordisk	We support the adoption of MUC20-0040 Intervention for Prediabetes into the Merit-Based Incentive Payment System. Novo Nordisk is committed to driving change to improve outcomes in diabetes and obesity, both diseases impacted by the actions of health care providers and patients, and that can benefit from early detection and intervention. Performance measures are essential tools for driving change in the U.S. healthcare system and we urge CMS to consider this prediabetes measure as an important starting point for process change and the need to address conditions earlier in their lifecycle to prevent long-term negative health outcomes.
University of Colorado Medicine	It is unclear from the information provided if it is only a referral that is needed, or if it is the delivery of the intervention, or both. It would also be nice to know what data is being used for this measure. Medicare will not receive information that a referral was made, as that is not claims data. If the purpose is to measure what a clinician is recommending, then a referral-alone would be sufficient. If the purpose is to measure how many patients received the intervention, then evidence via a claim or EHR evidence of an encounter with RD/nutrionists is needed. For measures based on medications it is important to clarify how exclusions for persons with prior side effects or allergies, renal function decline that prohibits metformin prescribing will be defined. Lastly, if claims data is being used as an indicator of RD services, this might be a problem for CPC+ or PCF practices who are covering these resources via other funds and cannot therefore bill for their services.
Diabetes Advocacy	January 6, 2021

Alliance

TO: Measure Application Partnership
MAP MUC 2020 Comment Period

FROM: Hannah Martin, Academy of Nutrition and Dietetics
Diabetes Advocacy Alliance Co-Chair
hmartin@eatright.org

Kate Thomas, Association of Diabetes Care & Education Specialists
Diabetes Advocacy Alliance Co-Chair
kthomas@adces.org

RE: Comments to MUC20-0040, Intervention for Prediabetes

The Diabetes Advocacy Alliance (DAA) appreciates the opportunity to submit comments related to Measure Under Consideration: MUC20-0040, Intervention for Prediabetes.

The DAA is a coalition of 26 diverse member organizations, representing patient, professional and trade associations, other non-profit organizations, and corporations, all united in the desire to change the way diabetes is viewed and treated in America. Since 2010, the DAA has worked with legislators and policymakers to increase awareness of, and action on, the diabetes epidemic. The organizations that comprise the DAA share a common goal of elevating diabetes on the national agenda so we may ultimately defeat this potentially devastating chronic disease.

The DAA recognizes that health inequities have had, and continue to have, a tremendous negative impact on our society's ability to identify those at risk, prevent new cases of diabetes, and effectively treat people with diabetes and obesity to help prevent the many serious complications of the disease. The DAA believes that quality measures are important, but we urge CMS to recognize that inequities have an effect on the data that are collected and thus some quality measure data may not be representative of populations most affected by the proposed measures.

MUC20-0040, Intervention for Prediabetes

Description: Percentage of patients aged 18 years and older with identified abnormal lab result in the range of prediabetes during the 12-month measurement period who were provided an intervention.

Measure Type: Process

Measure Steward: American Medical Association

CMS Program: Merit-Based Incentive Payment System (MIPS)

The American Medical Association (AMA) – the Measure Steward for the proposed

Intervention for Prediabetes measure – is a member organization of the DAA. The DAA's comments reflect input and review from multiple members of the DAA, including the AMA, that support this measure.

Measure MUC20-0040, Intervention for Prediabetes, would add value and help improve patient outcomes in these ways:

• It would encourage dialogue between health care providers and their patients with prediabetes, so that patients could learn what it means to have prediabetes and why it is important to take action to prevent or delay the onset of type 2 diabetes.

• It would provide an incentive for health care providers to identify and refer their Medicare and Medicaid patients with prediabetes to evidence-based treatments including the diabetes prevention lifestyle change programs that have received recognition from the Centers for Disease Control and Prevention. These programs are available in health care and community-based settings for in-person instruction and online through virtual programs. (The DAA recognizes that CMS does not currently cover virtual-only programs for Medicare beneficiaries; however, the DAA continues to advocate for such coverage and is currently seeking statutory support. Further, a growing number of state Medicaid agencies do cover diabetes prevention program (DPP) services provided by CDC fully-recognized DPP suppliers, for example California and Maryland.)

• It would further align the process of care with preventive care services recommended as grade B or higher by USPSTF, and therefore, offered without having copays and deductibles payable by a beneficiary, consistent with Affordable Care Act standards for preventive care. (See the USPSTF final recommendation statement dated October 26, 2015, for Abnormal Blood Glucose and Type 2 Diabetes Mellitus: Screening. This recommendation is currently being updated.)

• Of the 88 million Americans with prediabetes, more than 8 out of 10 don't know they have it. Adoption of this quality measure would increase the numbers of people aware of their condition and increase the numbers being treated. (https://www.cdc.gov/diabetes/pdfs/data/statistics/national-diabetes-statistics-report.pdf)

• It could contribute to reduced future costs of care for people with prediabetes who complete the program and lose sufficient weight to have a positive impact on blood glucose and A1c levels, so as to prevent or delay the onset of type 2 diabetes. People with prediabetes who go on to develop type 2 diabetes incur greater heath care costs than those who do not develop type 2 diabetes. (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5649409/)

The DAA believes that the data collection and reporting burdens for this measure are low and that the benefits of this measure outweigh any small burdens that may exist. Physicians and other health care providers, during routine care, are screening for type 2 diabetes risk factors which mirror those for prediabetes, such as family history, obesity/BMI and hypertension. These data are being included in reports and

	registries generated by the electronic health record (EHR) according to clinical guidelines and can generate the relevant referral through the EHR or e-referral system. All that remains, then, is documenting the referral in a patient's record and ensuring it can be tabulated to generate the measured result. As the tabulation mechanism should occur through EHR software development, the burden on a clinical practice is merely documenting that the referral has been made. The DAA would use data captured through this measure to establish a baseline of the percentage of health care provider referrals, and then monitor the percentage of referrals in successive years. Some DAA-member organizations or their respective members offer evidence-based diabetes prevention programs that are CDC-recognized, and some also work to inform and persuade health care providers of the importance of counseling their patients with prediabetes and telling them about the potential benefits of participation in a diabetes prevention program. This measure would help DAA member organizations assess the progress of their efforts. The DAA does not foresee any significant implementation challenges with this measure.
American Medical Association	The AMA developed three measures on prediabetes to reduce data collection burden, facilitate widespread implementation, and drive improvements around screening and interventions for individuals at risk of progressing to Type 2 diabetes. We welcome the MAP's consideration of this measure as we strive to increase access to CDC-recognized lifestyle change programs and the Medicare Diabetes Prevention Program. The AMA is working with several groups including the CDC and Diabetes Advocacy Alliance and is in discussions with CMMI and Medicaid regarding their implementation.
Vista Consulting, LLC	I support this measure, it aligns with the diabetes prevention strategy we are encouraging across South Carolina as we work with Patient Centered Medical Homes, and this measure is valid and feasible for our docomentation/tracking efforts.
Tannaz Moin, MD	I fully support the proposed prediabetes measures, which align nicely with ongoing national diabetes prevention efforts. CMS has endorsed the importance of diabetes prevention and providing beneficiaries with DPP coverage. These proposed measures will help enhance health system and provider engagement in this critical area. Implementing these measures in health systems is feasible and the approach is valid.
Prisma Health System	I am writing in support of MUC20-0040. The proposed measure provides a mechanism to track performance in an important population health related measure - intervention for pre-diabetes. This measure is important in many regards: (1)- The measure provides a definition for determining who is in the pre-diabetes population and with this a framework for identifying the scope of the opportunity within a practice, health system and/or a community. (2) The exclusions are minimal, easily tracked and make good clinical sense. (3) The potential measures that would meet the measure are clinically relevant and documented to have effective outcomes in reducing the risk of progression to diabetes. (4) The interventions that would satisfy

	the measure are significantly underutilized in the communities in which they exist and the inclusion of this measure in our ACO and health system's quality toolbox could assist in overcoming the therapeutic inertia necessary to increase the utilization of these effective programs. (5)The data collection utilizing an EMR based methodology seems minimal at most. (6)As a clinician and as a leader in a large health system and ACO, we would use the measure to asses and improve quality of care with proven patient outcomes from the interventions meeting the measure. Our health system and ACO would use the measure in a multitude of formats including internal value based performance measures for our practitioners, public reporting regarding our performance on the measure and disease surveillance in the community to insure we are meeting our goals of improving the state of health in South Carolina.
Federation of American Hospitals	The Federation of American Hospitals (FAH) notes that this measure (#3570e) did not receive NQF endorsement during the most recent endorsement review cycle and revisions may need to be considered to address the NQF Standing Committee's concerns. As a result, the FAH recommends that the highest level of MAP recommendation be "Do Not Support with Potential for Mitigation."

Measure Comments (Post-Workgroup Meeting)

Author	Submitted Comment

American Academy of Family Physicians	The American Academy of Family Physicians (AAFP) agrees with the MAP recommendation of "Do Not Support with Potential for Mitigation." The AAFP is extremely concerned with the lack of evidence for improvement of outcomes and the large potential for harms of treatment with medications for prediabetes. We are also concerned with the proposition of taking a risk factor for a disease (diabetes mellitus) and attempting to make it a disease that is treated pharmacologically. This opens the gate for use of other drugs outside of metformin and the movement to combat therapeutic inertia by subspecialists and pharma with little regard to the complex patient or social determinants of health, concerns that family physicians respect and consider as a core component of their relationship with patients. Evidence: Supporting documentation for these measures rest solely on expert opinion. • “At least annual monitoring for the development of diabetes in those with prediabetes is suggested. (ADA, 2018) (E Recommendation)”. This is expert opinion according to ADA. • The developer provides evidence of disease prevalence and systematic misses of opportunities to intervene by clinicians, but does not provide studies that offer evidence that annual monitoring is associated with positive outcomes. While NQF staff recommend an exception to the evidence, the AAFP does not agree with the NQF recommendation. Risk Factor Versus Disease: We offer insight from Steven R. Brown, MD, FAAFP, University of Arizona College of Medicine and former chair of the AAFP Commission on Health of the Public and Science, in his editorial comments and supporting references regarding treating adults with prediabetes by prescribing metformin. (1) “Prediabetes is not a disease. It is a risk factor for a disease (diabetes mellitus). Most people with so-called prediabetes will not develop diabetes within five years.”(2) The threshold values set by the ADA and reflected in these measures for defining prediabetes dramatically expand the number of people who meet the definition of prediabetes, with one-third of Americans meeting the criteria.(3) Many people will be mislabeled as having prediabetes using a single A1C. Cutoff values for diabetes and prediabetes are arbitrary.(4) Although some randomized controlled studies have shown a lower incidence of diabetes with metformin treatment versus lifestyle or placebo changes, treating borderline glucose values does not improve quality of life, mortality or other patient-oriented outcomes.(5)(6) Treating the Whole Patient: Dr. Brown points out that, “It is unlikely that initiating metformin before diabetes is diagnosed improves outcomes compared with waiting for a formal diagnosis. When the clinically meaningful benefit of a treatment is marginal, harms become more important to consider. Although metformin is inexpensive and simple to use, many patients experience gastrointestinal symptoms such as diarrhea, flatulence, nausea, and vomiting.(5) Long-term use is associated with vitamin B12 deficiency.”(7) Labeling patients with a diagnosis of “prediabetes” is an example of selling a sickness to grow the markets for those that sell and deliver treatments. The medical community should not cave-in to such practices. The AAFP opposes these three measures and favors discussing with patients the pros and cons of medications for borderline glucose values along with

lifestyle changes, smoking cessation, blood pressure control, and cardiovascular prevention.

Measure 3569e: Prediabetes: Screening for Abnormal Blood Glucose:The AAFP opposes Measure 3569e for the following reasons. First, the measure creates a “disease” out of a risk factor. The term “prediabetes” should be deleted from the title and remaining specifications/discussion. The measure should be titled “Abnormal blood glucose.”
In addition, the measure has no upper age limit, which is not consistent with the AAFP or the USPSTF recommendations, both of which recommend screening for abnormal blood glucose as part of cardiovascular risk assessment in adults aged 40 to 70 years who are overweight or obese.”
Finally, the AAFP recommends abnormal results be confirmed prior to intervention, and this proposed measure does not require confirmation of results, which may lead to unnecessary, excessive, and harmful treatments.
Although the AAFP concludes there is currently inadequate evidence whether early detection of abnormal blood glucose or diabetes leads to improvements in mortality or cardiovascular morbidity, screening is consistent with AAFPs' recommendations in adults who are obese or are overweight with additional cardiovascular risk factors.
The AAFP also has concern with data capture. Fasting glucose status may not be captured in EHR in distinct field; and the exclusions (comfort care, hospice, palliative care) may not be captured in distinct field. Only Site 2 was assessed for reliability of exclusions at a rate of 76%. We have concern that the measure was only tested in EPIC and Cerner EHRs. Independent practices frequently use other less expensive EHRs which may not be as robust.

Measure 3570e: Intervention for Prediabetes:The USPSTF and the AAFP recommend physicians "Offer or refer patients with abnormal blood glucose to intensive behavioral counseling interventions to promote a healthful diet and physical activity." The proposed quality measure requires physicians to either prescribe metformin or refer the patient out, which is not consistent with AAFP or USPSTF recommendations, and does not reflect scope of practice for primary care physicians. Primary care physicians are well-qualified to offer intense behavioral counseling and numerous other interventions to manage abnormal blood glucose without the need for a referral. Referrals can unnecessarily drive up costs and may lead to unnecessary treatment. In addition MDDPs or CDC-recognized programs are not accessible to many areas in the country, so are not a viable alternative, particularly in rural areas. (https://innovation.cms.gov/innovation-models/medicare-diabetes-prevention-program/mdpp-map; https://www.cdc.gov/diabetes/programs/national-dpp-maps/index.html).
Of equal concern are the issues discussed in our general comments.
• The evidence rests solely on expert opinion;
• Prediabetes is not a disease and should not be treated as such. Most patients with the risk factors will not develop diabetes within five years. Cut-off points are arbitrary. A large number of patients (1/3 of American adults) would be labeled as having “prediabetes” according to these thresholds. Labeling a patient with “prediabetes” can lead to emotional stress and treating borderline glucose values does not improve quality of life, mortality

or other patient-oriented outcomes. The options for intervention presented are too limited.
• The measure promotes treatment with pharmaceutical and ignores the potential harms associated with such treatment.
• The measure does not take a whole person view of managing risk factors, which must consider co-morbid conditions, socioeconomic factors, and patient goals, values, and readiness for change. The AAFP favors discussing with patients the pros and cons of medications for borderline glucose values along with lifestyle changes, smoking cessation, blood pressure control, and cardiovascular prevention and determining a patient-centered plan of care.
• The measure promotes use of pharmaceuticals and other treatments by labeling patients with a "diagnosis" and then recommending a treatment for that diagnosis.

Measure 3571e: Retesting of Abnormal Blood Glucose in Patients with Prediabetes: AAFP policy states, "There is limited evidence on the best rescreening intervals for adults with normal results; however screening every 3 years is a reasonable option." In contrast, this measure requires re-testing at least annually. In addition, the exclusions for this measure are different from the others.

(1) Brown, Steven R., MD, FAAFP, University of Arizona College of Medicine–Phoenix, Phoenix, Arizona. Am Fam Physician. 2019 Aug 1;100(3):136-137. https://www.aafp.org/afp/2019/0801/p136.html
(2) American Diabetes Association. Classification and diagnosis of diabetes. Standards of medical care in diabetes–2018. Diabetes Care. 2018;41(suppl 1):S13–S27.
(3) Centers for Disease Control and Prevention. National Diabetes Statistic Report, 2017. Accessed February 5, 2019. https://www.cdc.gov/features/diabetes-statistic-report/index.html
(4) Yudkin JS, Montori VM. The epidemic of pre-diabetes: the medicine and the politics [published correction appears in BMJ. 2014;349:g4683]. BMJ. 2014;349:g4485.
(5) Knowler WC, Barrett-Connor E, Fowler SE, et al.; Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med. 2002;346(6):393–403.
(6) Marrero D, Pan Q, Barrett-Connor E, et al.; DPPOS Research Group. Impact of diagnosis of diabetes on health-related quality of life among high risk individuals: the Diabetes Prevention Program Outcomes Study. Qual Life Res. 2014;23(1):75–88.
(7) Aroda VR, Edelstein SL, Goldberg RB. Long-term metformin use and vitamin B12 deficiency in the Diabetes Prevention Program Outcomes Study. J Clin Endocrinol Metab. 2016;101(4):1754–1761.

Prisma Health Department of Medicine	The Intervention for Prediabetes measure is an important one if we are to improve the health of our nation. The increasing numbers of patients at risk for prediabetes is large and continues to rise. Giving our patients, physicians and health systems the right tools to measure and use effective proven techniques to change the course of the patients with prediabetes. The three mechanisms to meet the measure are realistic in a team based care model today. We have implemented all of these programs and have seen the significant impact the programs can make. The physicians look at these proven methods as an extension of the office team that allows them to be most efficient in delivering care while using experts and proven programs to truly effect change in the at risk population. The time invested would never be possible in the primary care setting given the pace of care today and the number of patients needing primary care and prediabetes care. Interestingly, in the presence of Covid, we have shifted to a virtual delivery model for DPP and MNT and have noted similar outcomes which mitigates concern that these programs cannot be scaled to large populations. The measures and interventions are easily measured in today's EMR's and thus makes this an easily implemented high impact measure. Due to all of the above, I would urge NQF to move forward with approval and would encourage CMS to move ahead with or without NQF approval given the importance of this measure to improve the health of our patients at risk.
Johns Hopkins University School of Medicine	I am a general internist and research expert on diabetes prevention. Clinicians play a huge role in preventing diabetes by diagnosing prediabetes, providing brief counseling on lifestyle changes and referring patients to evidence-based yearlong Diabetes Prevention Programs. Unfortunately, there is no good evidence to support that clinicians have the training or time to adequately provide lifestyle counseling, but studies do confirm that when clinicians begin the discussion, patients are more likely to attempt lifestyle changes. There is an abundance of evidence supporting the effectiveness of Diabetes Prevention Programs. As accessibility to these programs grows nationally, this measure will make clinicians accountable for providing evidence-based interventions to prevent diabetes. Endorsing this measure does not remove the vital role that clinicians play in engaging their patients in diabetes prevention, but rather acknowledges that evidence-based interventions exist and more patients should be referred to them. Finally, there is growing number of Diabetes Prevention Programs that are delivered through community-based organizations. At Hopkins, the Brancati Center works with local organizations such as churches to develop Diabetes Prevention Programs. These programs have had a great success in reaching communities where the burden of prediabetes is high. Therefore, this measure supports evidence-based interventions that can strengthen the relationship between the community and healthcare organizations, which is important for tackling the diabetes epidemic in our country.

American Heart Association/American Stroke Association	The AHA strongly supports intervention for patients in the range of pre-diabetes, however, we question whether prescription of metformin by itself should be among the first-line interventions, even if diet and exercise have not been tried. We recognize that the American Diabetes Association1 has recommended that metformin should be considered in those with prediabetes, however, they also acknowledge that no pharmacologic agent has been approved by the U.S. Food and Drug Administration specifically for diabetes prevention. Since this use for metformin is an off-label indication—and some controversy remains on whether it prevents serious complications of diabetes2--we question whether it should be included in a quality measure at this point. 1. Prevention or Delay of Type 2 Diabetes: Standards of Medical Care in Diabetes 2021 Diabetes Care 2021;44(Suppl. 1):S34–S39. Available at: https://care.diabetesjournals.org/content/44/Supplement_1/S34 2. Davidson MB. Metformin Should Not Be Used to Treat Prediabetes. Diabetes Care. 2020 Sep;43(9):1983-1987. Available at: https://care.diabetesjournals.org/content/43/9/1983
Federation of American Hospitals	Support
American Medical Association	The AMA is aware of recommendations to expand the denominator. However, the recommendations to expand the denominator are not supported by evidence.
Northwestern University Feinberg School of Medicine	I am a physician and diabetes prevention researcher, who has focused on the Diabetes Prevention Program as a member of the DPP study group and with my own independent NIH funding in this area. The DPP study has generated all of the evidence on which this prediabetes treatment is based. There is a large body of evidence supporting the use of DPP (and related MNT) and metformin for preventing diabetes or delaying its onset. This is based on our large randomized trial, and hundreds of translational studies of these same interventions in many different populations and settings. There is no evidence to support brief interventions for preventing or delaying diabetes. Therefore, this measure is well aligned with the existing evidence base, and should be adopted. There are many examples of effective health interventions offered by non-physicians, including MNT and also physical therapy. The DPP is yet another example of a highly evidence-based intervention that physicians can refer their patients to. Uptake of the Diabetes Prevention Program has been low, which limits the potential to prevent diabetes and reduce its significant population burden. This measure will help to promote uptake of evidence-based treatment for diabetes prevention, which is essential for improving our nation's health.
Cerner Corporation	We support tracking of this preventative care measure and do not foresee any data implementation challenges.

Prisma Health	Diabetes is currently the 7th leading cause of death in South Carolina; and as the number of patients at risk for diabetes continues to grow, the Intervention for Prediabetes measure is critical to improving overall health of our patients, particularly in the Medicare population which results in high mortality, morbidity and total cost of care. The tool would allow for tracking of proven methods to help those specifically at risk; all of which can be easily documented in EMRs. Through emphasis on DPP and MNT, both of which are able to be shifted easily to a virtual care environment, the measure reinforces the need for integration across the continuum and encouragement of community links. Through my career and current role as Chief Clinical Officer at Prisma Health, a large integrated health system, and as CEO of the Care Coordination Institute (CCI), a population health and data enablement company, and from experience and participation on multiple national including America's Physician Groups (APG) Board of Directors and on the Guiding Committee for the CMS Health Care Payment Learning and Action Network (LAN); I have witnessed the need for a measure such as this one. I strongly advocate for the reconsideration and approval by NQF and CMS of this measure.

MUC20-0042 Person-Centered Primary Care Measure Patient Reported Outcome Performance Measure

Measure Information

Characteristic	Submitted Information
MUCID	MUC20-0042
Other Measure Identification Numbers	
Title	Person-Centered Primary Care Measure Patient Reported Outcome Performance Measure (PCPCM PRO-PM)
Program	Merit-based Incentive Payment System-Quality
Workgroup	Clinician
In what state of development is the measure?	Fully Developed
State of Development Details	This measure – the PCPCM PRO-PM – is fully developed and currently entering its third round of pilot testing. Identifying a gap. The measure has been submitted to NQF. We first conducted a thorough environmental scan of the literature, publishing our findings. As mentioned below in response to Row 21:"A number of measures have been developed to assess different aspects of primary care.16,17,20-25 Unfortunately, they tend to be long and seldom used outside of the research setting. Clinical primary care settings often turn to patient experience surveys, such as the Clinician and Group

Consumer Assessment of Healthcare Providers and Systems (CG CAHPS), and researchers have recently sought to shorten the CG CAHPS in order to increase its use.26 Patient experience measures focus important attention to the consumer experience of care delivery and receipt of services but fall short of focused attention to the broad scope of primary care functions and care.13,15Our team has conducted an extensive survey of measures used to assess primary care.1 No patient reported measure is previously existing that offers a patient reported assessment of full scope primary care. "Research to determine what is meaningful to stakeholders. Understanding the lack of a primary care measure fully aligned with primary care's purpose, function, and expectation of stakeholders, we began development by surveying hundreds of patients, hundreds of clinicians, and close to one hundred employers directly responsible for purchasing health care plans for their organizations. We asked open ended questions, received over 10,000 comments, and analyzed these using a grounded-theory approach to content and thematic analysis, enlisting a multidisciplinary team, several of whom have received lifetime achievement awards for their work in primary care teaching, primary care research, and psychometrics. Through intensive analysis, we identified overlapping interests among clinicians and patients that fell into 18 quality indicator areas. Aligning areas significant to stakeholders with 'fit for purpose.' Having used this very grounded approach, we then understood that a measure created based on these findings must also be fit for purpose – it must align with the needs of practice, education, payment, patients, the health system, and other payors. Our next step was then to put these 18 quality indicator areas in front of a group of 70 national and international primary care experts. These were people from all primary care disciplines and stakeholder groups. We had them work on these 18 areas for 2.5 days in a conference held in DC. The group included patients, insurers, employers, all primary care disciplines, psychometricians, actuaries, CEOs of Boards, and allied professionals co-located in primary care settings. With the IOM Vital Signs report as our standard for creating a meaningful set of parsimonious measures, and having provided all participants with a primer on the current state of measures, the group engaged in a collaborative process to identify a subset of 11 quality indicator areas. Designing the initial instrument using "native language". Our team then used over 80 hours of digitally recorded conversation from the conference, along with dozens of interviews and thousands of open text comments in surveys, to develop a single item per quality indicator area using the language of stakeholders. Using stakeholder terms and the phrases they most often identified as important would help to ensure the reliability, reading level, and fit for purpose that is often not well addressed when asked about directly. Cognitive testing, validation testing, concurrent validity testing. We conducted testing well beyond the sample size required. After interviewing patients to ensure the questions assessed the information intended and collected the information intended, we went beyond the 110 participants required to meet statistical standards and tested the 11-item patient reported PCPCM PROM on over 1000 patients. All items loaded strongly onto a single factor. We repeated the exercise and found the same results. Both times, this instrument was administered electronically to people who identified as having seen primary care in the past year. Our patient participants were a paid for sample designed to reflect the US census in terms of age, gender, education, employment, income, minority status, and regional distribution. We next tested the

	PCPCM PROM in point-of-care patient settings: a community health center, a private practice, a hospital owned clinic, and a pediatric practice. In each case, we used a population of consecutive patients, excluding only those who opted not to participate, and fielding until the sample was sufficient to validate the instrument within that setting. It was again validated with similarly strong statistical findings. In person findings were more positively skewed than online samples but the curve created by the 11 items in each circumstance was identical. A review of the 4-point scale used demonstrated use of the entire scale and a wide distribution. While fielding, an additional item, reversed coded, was used to test whether patients were attentive to the purpose of the question. Patients were also asked if the questions were easily understood and if they would be meaningful to their care if shared with their clinicians. In both cases, the answers were positive. Testing of the PCPCM PROM along-side two other patient reported and validated instruments – one that assessed patient self-management and one that assessed cost and utilization of services – offered evidence that the PCPCM PROM was also strong correlated to desirable system and clinical population health outcomes. Since validation, we have faced the first pandemic in our living history. Concerned that use of the PCPCM PROM during the pandemic would not alter its reliability, we fielded the PCPCM PROM to an online audience of ~1400, selected in the same manner as previously selected. That publication is being prepared – we found the measure still valid. While typical quality measures are likely to not be collected or to suggest a dramatic dip in performance during the pandemic, the PCPCM PROM was fielded without the motivation of connection to a payment model, without adding any burden to practices, and able to measure quality even during the very difficult time. The PCPCM PROM is now being tested by Anthem in Colorado, by VCU Health System in Richmond VA, by the University of Missouri, and by the PRIME quality registry of the American Board of Family Medicine. The measure has been validated in 28 languages and 35 country settings using a similar online process and subject matter experts for the translation service. Teams in Japan, Hong Kong, the Netherlands, Australia, and Canada currently have on the ground research projects to validate the PCPCM PROM in those cultural settings. All but Australia have publications in process demonstrating measure validity. In the VCU health system, results are already in and were successfully used to determine differences in performance among clinicians and among practices. We fielded the PCPCM PROM among 6 practices and 50 clinicians, using data from 16 clinicians for our validation effort. In this case, the measure was fielded using a free-standing electronic platform designed by our team to ease implementation of the measure in practices. It required less than 2 hrs of time from a practice administrator for onboarding. Results have been shared with the practices.
Measure Description	The Person-Centered Primary Care Measure Patient Reported Outcome Performance Measure (PCPCM PRO-PM) uses the PCPCM PROM (a comprehensive and parsimonious set of 11 patient-reported items) to assess the broad scope of primary care. Unlike other primary care measures, the PCPCM PRO-PM measures the high value aspects of primary care based on a patient's relationship with the provider or practice. Patients identify the PCPCM PROM as meaningful and able to communicate the quality of their care to their clinicians and/or care team.

	The items within the PCPCM PROM are based on extensive stakeholder engagement and comprehensive reviews of the literature.
Numerator	The target population is all active patients in a practice during the performance reporting period. A patient is defined as active if the patient has had a documented interaction with the practice within 12 months of their birth month within the measurement period. The PCPCM PROM is the same for all patients, regardless of age. Because the PCPCM PROM applies to all patients and is not particular to a clinical encounter, it is administered once a year to each patient during their birth month. The target population is defined the same, regardless of unit of analysis (clinician, practice, or system).The numerator is the sum of all PCPCM PROM scores for active patients. To use the numerator for calculating the PCPCM Performance Score, please refer to the Calculation Algorithm/Measure Logic section of the attached Measure Information Form. Current national benchmark for the PCPCM Performance Measure was established by the national sample used in the published validation of the measure. Individual Benchmarks from National Pilot (n=2229) by Item Score1. My practice makes it easy for me to get care. Mean 3.1, 78%2. My practice is able to provide most of my care. Mean 3.1, 78% 3. In caring for me, my doctor considers all the factors that affect my health. Mean 3.2. 80%4. My practice coordinates the care I get from multiple places. Mean 2.8, 70%5. My doctor or practice knows me as a person. Mean 2.9, 73%6. My doctor and I have been through a lot together. Mean 2.2, 55%7. My doctor or practice stands up for me. Mean 2.7, 68%8. The care I get takes into account knowledge of my family. Mean 2.7, 68%9. The care I get in this practice is informed by knowledge of my community. Mean 2.4, 60%10. Over time, my practice helps me to stay healthy. Mean 2.8, 70%11. Over time, my practice helps me to meet my goals. Mean 3.0, 75%National PCPCM Performance Score Benchmark: 2.8, 70%.
Denominator	The denominator is the total number of complete PCPCM PROM instruments received in the reporting period. A completed PROM instrument is defined as a PROM instrument for which the patient has responded to at least 8 of 11 items. The target population is all active patients in a practice during the performance reporting period. A patient is defined as active if the patient has had a documented interaction with the practice within 12 months of their birth month during the measurement period. The PCPCM PROM is the same for all patients, regardless of age. Because the PCPCM PROM applies to all patients and is not particular to a clinical encounter, it is administered once a year to each patient during their birth month. The target population is defined the same, regardless of unit of analysis (clinician, practice, or system).
Exclusions	None
Measure type	Patient Reported Outcome
What is the NQF status of the	Submitted

measure?	
NQF ID number	3568
Year of next anticipated NQF CDP endorsement review	2020
Year of most recent NQF Consensus Development Process (CDP) endorsement	NA
Is the measure being submitted exactly as endorsed by NQF?	NA
If not exactly as endorsed, describe the nature of the differences	NA
What data sources are used for the measure?	Administrative clinical data; Other (enter here): Patient Reported Outcome Measure (PROM)
If EHR or Administrative Claims or Chart-Abstracted Data, description of parts related to these sources.	NA
At what level of analysis was the measure tested?	Clinician; Group
In which setting was this measure tested?	Ambulatory/office-based care
What NQS priority applies	

to this measure?	
What one primary meaningful measure area applies to this measure?	Care is personalized and aligned with patient's goals
What secondary meaningful measure area applies to this measure?	NA
What one primary healthcare priority applies to this measure?	Strengthen person and family engagement as partners in their care
What secondary healthcare priority applies to this measure?	
What area of specialty best fits the measure?	Primary care
What is the target population of the measure?	All primary care populations regardless of payor type.
Is this measure an eCQM?	No
If eCQM, enter Measure Authoring Tool (MAT) number	No
If eCQM, does the measure have a Health Quality Measures Format (HQMF) specification?	No

Comments	One test bed used to field the PCPCM PROM-PRO was within The American Board of Family Medicine's PRIME Registry where it was endorsed as a QCDR measure for measurement year 2020.
Measure steward	The American Board of Family Medicine
Long-Term Measure Steward (if different)	The American Board of Family Medicine; The American Board of Family Medicine Foundation
Measure Steward Contact Information	Shuemaker, Jill and Pavletic, Denise; 202-600-9447; DPavletic@theabfm.org; JShuemaker@theabfm.org
Primary Submitter Contact Information	Shuemaker, Jill; The American Board of Family Medicine; 202-600-9447; JShuemaker@theabfm.org
Long-Term Measure Steward Contact Information	Shuemaker, Jill and Pavletic, Denise; 202-600-9447; DPavletic@theabfm.org; JShuemaker@theabfm.org
Secondary Submitter Contact Information	Etz, Rebecca; The Larry A. Green Center; 804-827-4995; Rebecca.etz@vcuhealth.org
Was this measure proposed for a previous year's MUC list?	No
In what prior year(s) was this measure proposed?	None
What were the programs that NQF MAP reviewed the measure for in each year?	N/A
Why was the measure not recommended in	N/A

those year(s)?	
What were the MUC IDs for the measure in each year?	N/A
NQF MAP report page number being referenced for each year	N/A
What was the NQF MAP recommendation in each year?	N/A
List the NQF MAP workgroup(s) in each year	N/A
What is the history or background for including this measure on the new MUC list?	New measure never reviewed by MAP Workgroup or used in a CMS program
Range of years(s) this measure has been used by CMS Program(s)	NA
What other federal programs are currently using this measure?	NA
Evidence that the measure can be operationalized	NA
How is the measure expected to be	CQM (Registry); Web interface

reported to the program?	
Is this measure similar to and/or competing with measure(s) already in a program?	No
Which existing measure(s) is your measure similar to and/or competing with?	N/A
How will this measure be distinguished from other similar and/or competing measures?	N/A
Rationale for how this measure will add to the CMS program	Current measure sets presume that quality primary care is the sum of quality measures for individual diseases and health screening. Value-based payments to primary care physicians frequently employ measures that are not aligned or recognize the higher-level integrating, personalizing, and prioritizing functions of primary care and the needs of patients, communities or health care systems.1,2 These measures are then tied to financial incentives which drive behavior to maximize these rudimentary measures. Driving clinicians' behavior toward low-value measures produces burnout and diminishes the value of primary care for people and populations. Clinical quality measurements should drive improved patient-centered care, align physician assessment and payment to produce high-value care, reduce physician burden, reduce high-cost behaviors, prevent low-cost physicians from changing their behaviors, and enable assessment and comparison of health systems that employ primary care physicians. Moreover, quality measurement should support the quadruple aim: improve health outcomes, improve patient experience, decrease clinician burnout, and lower health care costs.1. Etz RS, Gonzalez MM, Brooks EM, Stange KC. Less AND More Are Needed to Assess Primary Care. J Am Board Fam Med. 2017;30(1):13-15. doi:10.3122/jabfm.2017.01.1602092. Stange KC, Etz RS, Gullett H, et al. Metrics for assessing improvements in primary health care. Annul Rev Public Health. 2014; 35: 423–442.
If this measure is being proposed to meet a statutory	N/A

requirement, please list the corresponding statute.	
Evidence of performance gap	Using the data collected electronically from 6 practices and 16 clinicians, we assessed PCPCM performance measure score reliability and differences in performance among participants. The variation of PCPCM performance measure scores as found during our validation process, illustrates a difference of moderate (0.5) to large effect size (0.8) among clinicians in our validation tests, is evidence of a performance gap and opportunities for improvement. In a recently submitted manuscript regarding the score validity and reliability of the PCPCM PROM, among 6 practices, there were significant differences (p=0.004) in PCPCM PROM scores with a moderate effect size (at least .5 standard deviation).[See attachment titled "Tables for Template Row 51."]
Unintended consequences	There are no known or observed unintended consequences to date related to the fielding and use of the PCPCM performance measure. On implementation, some clinicians worry that certain measure items, such as “this doctor and I have been through a lot together” will disadvantage them if they have new patients or patients who have not had big health issues. Since the purpose of the measure is cross-comparison and comparison to a national benchmark, we have not found this to be an issue among the over 5,000 completed PCPCM PROs thus far. Others worry if, without risk stratification, they will be disadvantaged either because they have more pediatric patients, more elderly patients, sicker patients, more Medicaid patients, or more minority patients. Again, through all of our testing, we have not found this to be the case. We have published on the absence of score association to a particular minority or gender status. As expected, state of health and age group can be rank ordered but again does not require risk stratification.
Which clinical guideline(s)?	The IOM Report on Primary Care calls for care to be personalized at the patient level, with care integrated for whole people to overcome the many problems of fragmented and depersonalized care. The PCPCM PRO-PM complements more narrow disease-specific quality measures, and can be used to integrate care for whole people.51 The PCPCM PRO-PM also addresses a critical quality measure gap as identified by the MACRA-MDP Technical Expert Panel, of which Dr. Etz – the developer of the PCPCM PROM – was a part.41 Primary care reports on 94 Centers for Medicare and Medicaid Services (CMS) measures, more than any other specialty.23 In a national survey of primary care physicians, fewer than 25% expressed a positive opinion of quality measures used.24 Other studies found family medicine physicians among the most dissatisfied with levels of clerical burden and with some of the highest rates of burnout.25-29 The National Academy of Medicine, previously the IOM, has stated that there is no national consensus regarding how best to measure primary care delivery and performance.42 The Vital Signs report of the IOM defined the need for

	stakeholder created and informed measures, as opposed to measures that begin with subject matter experts and only later, if at all, rely on stakeholder input.4 The PCPCM PRO-PM was designed with this as its basis. 4. In: Blumenthal D, Malphrus E, McGinnis JM, eds. Vital Signs: Core Metrics for Health and Health Care Progress. Washington (DC)2015.23. Etz RS, Zyzanski SJ, Gonzalez MM, Reves SR, O'Neal JP, Stange KC. A New Comprehensive Measure of High-Value Aspects of Primary Care. Ann Fam Med. 2019 May;17(3):221-230.24. Howie JG, Heaney DJ, Maxwell M, Walker JJ. A comparison of a Patient Enablement Instrument (PEI) against two established satisfaction scales as an outcome measure of primary care consultations. Fam Pract. 1998; 15(2): 165–171.25. Wasson JH, Ho L, Soloway L, Moore LG. Validation of the What Matters Index: A brief, patient-reported index that guides care for chronic conditions and can substitute for computer-generated risk models. PLoS One. 2018; 13(2): e0192475.26. Wasson JH, Soloway L, Moore LG, Labrec P, Ho L. Development of a care guidance index based on what matters to patients. Qual Life Res. 2018; 27(1): 51–58.27. 2003. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care 41, 582-592.28. Meaningful Measures Framework of CMS. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/CMS-Quality-Strategy Accessed June 27, 2020.29. Jerant A, Fenton JJ, Franks P. Primary care attributes and mortality: a national person-level study. Ann Fam Med. 2012;10(1):34-41.42. Vital Directions for Health and Healthcare. Accessed September 2017. https://nam.edu/initiatives/vital-directions-for-health-and-health-care/.51. Institute of Medicine. Donaldson MS, Yordy KD, Lohr KN, and Vanselow NA, editors. Committee on the Future of Primary Care, Division of Health Care Services. National Academy Press. Washington, D.C. 1996.
Briefly describe the peer reviewed evidence justifying this measure	The PCPCM PRO-PM is based on data collected using the PCPCM PRO instrument. Validation of this instrument has been published in Annals of Family Medicine, globally the top ranked primary care peer-reviewed journal.10 It was created after conducting a thorough review of the literature of primary care measures,11 and after surveying over 1000 stakeholders (patients, clinicians, and payors) to identify what overlap currently exists between what is measured in primary care and what is most valued by those who seek care, those who provide it, and those who purchase it.12Primary care's effects are known to be better healthcare, better health, contained expenditures, and reduced disparities.1,2 Sustaining the platform and its focus is a high priority and requires measures able to promote continual improvements and investment in primary care. Most attempts to create measures for primary care focus on disease pathways, work pathways, or decisional pathways, and fail to address key elements through which primary care provides value.3 Leaders in primary care, including CMS Administrator Seema Verma, the Institute of Medicine, have noted an absence of meaningful measures and have called for measures appropriate to the task of assessing primary care, public health, stakeholder identified

needs, and the certainty of health equity.3-9 The measure we describe here – the Person-Centered Primary Care Measure (PCPCM) – fulfills the call from the Institute of Medicine and from CMS to create a stakeholder informed, meaningful measure that is an assessment of quality, low burden for implementation and collection, and provides adequate ability to compare performance across clinicians and practices while providing great face validity, transparency and actionable information. The PCPCM does that. It is unusual in its combination of robust internal consistency together with breadth and brevity. Its combination of parsimony - with a single item for each of 11 diverse primary care components - and conceptual coherence - exemplified by the fact that all 11 items load onto a single factor - is the result of an unusually broad and deep amount of preparatory work grounded in diverse stakeholder engagement. This stakeholder engagement enabled the development of meaningful measure items and is the reason why the PCPCM covers 4 of the 8 “cross cutting connections” in the Meaningful Measures Framework (identified as patient-centered and meaningful to patients; fulfill requirements in programs’ statues; minimize level of burden for providers; significant opportunity for improvement).28 In addition to being a useful new measure, the PCPCM adds to the field by empirically demonstrating that the broad focus of primary care is conceptually coherent, as seen and reported by the key stakeholder - patients. “A number of measures have been developed to assess different aspects of primary care.16,17,20-25 Unfortunately, they tend to be long and seldom used outside of the research setting. Clinical primary care settings often turn to patient experience surveys, such as the Clinician and Group Consumer Assessment of Healthcare Providers and Systems (CG CAHPS), and researchers have recently sought to shorten the CG CAHPS in order to increase its use.26 Patient experience measures focus important attention to the consumer experience of care delivery and receipt of services but fall short of focused attention to the broad scope of primary care functions and care.13,15Our team has conducted an extensive survey of measures used to assess primary care.1 No patient reported measure is previously existing that offers a patient reported assessment of full scope primary care. "In its combination of breadth, internal consistency, and parsimony, the PCPCM complements other existing measures of primary care. The measure’s detailed exposition of specific attributes of primary care, grounded in extensive advance work and member-checking with patients, clinicians and policymakers, allows evaluation of the specific mechanisms by which primary care adds value, and thus complements more global assessments of primary care, such as having a usual source of care.29-31 In its parsimony, the PCPCM complements other patient-report measures of primary care that measure fewer domains, but with multiple items per domain,32-36 or that measure aspects of primary care for specific purposes.37-40 The PCPCM-PM is the first measure developed to meet these nationally identified needs. The PCPCM also addresses a critical quality measure gap as identified by the MACRA-MDP Technical Expert Panel, of which Dr. Etz – the developer of the PCPCM – was a part.41 Primary care reports on 94 Centers for Medicare and Medicaid Services

(CMS) measures, more than any other specialty.23 In a national survey of primary care physicians, fewer than 25% expressed a positive opinion of quality measures used.24 Other studies found family medicine physicians among the most dissatisfied with levels of clerical burden and with some of the highest rates of burnout.25-29 The National Academy of Medicine, previously the IOM, has stated that there is no national consensus regarding how best to measure primary care delivery and performance.42 The Vital Signs report of the IOM defined the need for stakeholder created and informed measures, as opposed to measures that begin with subject matter experts and only later, if at all, rely on stakeholder input.14 The PCPCM was designed with this as its basis. References: 1. Phillips RL and Bazemore AW. Primary Care and Why It Matters for US Health System Reform. Hlth Aff. 2010;29(5):806-810. 2. Starfield B, Shi LY, Macinko J. Contribution of primary care to health systems and health. Milbank Q, 2005;83:457-502. 3. Stange KC, Etz RS, Gullet H, et. al. Metrics for Assessing Improvements in Primary Health Care. ARPH. 2014:423-42. 4. The Center for Medicare and Medicaid Services. Remarks by Administrator Seema Verma at the Health Care Payment Learning and Action Network (LAN) Fall Summit press@cms.hhs.gov; 2017. 5. Berwick DM. Era 3 for Medicine and Health Care. JAMA. 2016 Apr 5;13:1329-30. 6. Blumenthal D, Malphrus E, McGinnis JM, eds. Vital Signs: Core Metrics for Health and Health Care Progress. 2015. 7. Berenson RA, Rich EC. US approaches to physician payment: the deconstruction of primary care. J Gen Intern Med. 2010;25:613-618. 8. Conway, PH and the Core Quality Measures Collaborative Workgroup. The Core Quality Measures Collaborative: A Rationale And Framework For Public-Private Quality Measure Alignment. June 23, 2015 Health Aff Blog http://healthaffairs.org/blog/2015/06/23/the-core-quality-measures-collaborative-a-rationale-and-framework-for-public-private-quality-measure-alignment/ Accessed July 17, 2016. 9. Rich EC, O'Malley AS. Measuring what matters in primary care. October 6, 2015. Health Aff Blog http://healthaffairs.org/blog/2015/10/06/measuring-what-matters-in-primary-care/ Accessed December 3, 2015. 10. Etz RS, Zyzanski SJ, Gonzalez MM, Reves SR, O'Neal JP, Stange KC. A New Comprehensive Measure of High-Value Aspects of Primary Care. Ann Fam Med. 2019 May;17(3):221-230. 11. Stange KC, Etz RS, Gullett H, et al. Metrics For Assessing Improvements In Primary Health Care. Annual review of public health. 2014;35:423-442. 12. Etz RS, Gonzalez MM, Brooks EM, Stange KC. Less AND more are needed to assess primary care. J Am Board Fam Med. 2017; 30(1): 13–15. 13. Stange KC, Etz RS, Gullett H, et al. Metrics For Assessing Improvements In Primary Health Care. Annual review of public health. 2014;35:423-442. 14. In: Blumenthal D, Malphrus E, McGinnis JM, eds. Vital Signs: Core Metrics for Health and Health Care Progress. Washington (DC)2015. 15. Etz RS, Gonzalez MM, Brooks EM, Stange KC. Less AND More Are Needed to Assess Primary Care. J Am Board Fam Med. 2017;30(1):13-15. 16. Howie JG, Heaney DJ, Maxwell M, Walker JJ. A comparison of a Patient Enablement Instrument (PEI) against two established satisfaction scales as an outcome measure of primary care consultations. Fam Pract. 1998;15(2):165-171. 17. Wasson JH, Ho L,

Soloway L, Moore LG. Validation of the What Matters Index: A brief, patient-reported index that guides care for chronic conditions and can substitute for computer-generated risk models. PLoS One. 2018;13(2):e0192475. 20. Jabbarpour Y. Measures in Primary Care. An annotated bibliography. Starfield Summit; April 2016, 2016; Washington, DC. 21. 2017. Etz, RS and the Starfield Writing Team. Conference Brief: Framework of PC Measure Domains and Key Elements. Starfield Summit III: Washington, DC. Accessed January 6, 2020, http://www.starfieldsummit.com/resources3 22. Etz RS, Gonzalez MM, Brooks EM, Stange KC. Less AND more are needed to assess primary care. J Am Board Fam Med. 2017; 30(1): 13–15. 23. Etz RS, Zyzanski SJ, Gonzalez MM, Reves SR, O'Neal JP, Stange KC. A New Comprehensive Measure of High-Value Aspects of Primary Care. Ann Fam Med. 2019 May;17(3):221-230. 24. Howie JG, Heaney DJ, Maxwell M, Walker JJ. A comparison of a Patient Enablement Instrument (PEI) against two established satisfaction scales as an outcome measure of primary care consultations. Fam Pract. 1998; 15(2): 165–171. 25. Wasson JH, Ho L, Soloway L, Moore LG. Validation of the What Matters Index: A brief, patient-reported index that guides care for chronic conditions and can substitute for computer-generated risk models. PLoS One. 2018; 13(2): e0192475. 26. Wasson JH, Soloway L, Moore LG, Labrec P, Ho L. Development of a care guidance index based on what matters to patients. Qual Life Res. 2018; 27(1): 51–58. 27. 2003. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care 41, 582-592. 28. Meaningful Measures Framework of CMS. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/CMS-Quality-Strategy Accessed June 27, 2020. 29. Jerant A, Fenton JJ, Franks P. Primary care attributes and mortality: a national person-level study. Ann Fam Med. 2012;10(1):34-41. 30. DeVoe JE, Tillotson CJ, Wallace LS, Angier H, Carlson MJ, Gold R. Parent and child usual source of care and children's receipt of health care services. Ann Fam Med. 2011;9(6):504-13. 31. Ettner SL. The timing of preventive services for women and children: the effect of having a usual source of care. Am J Public Health. 1996;86(12):1748-54.32. Jabbarpour Y. Measures in Primary Care. An annotated bibliography. Starfield Summit; April 2016; Washington, DC: Robert Graham Center / Eugene S. Farley, Jr. Health Policy Center; 2016. p. 15. 33. Shi L, Starfield B, Xu J. Validating the adult primary care assessment tool. J Fam Pract. 2001;50(2):161W-75W. 34. Safran DG, Kosinski M, Tarlov AR, Rogers WH, Taira DH, Lieberman N, et al. The Primary Care Assessment Survey: tests of data quality and measurement performance. Med Care. 1998;36(5):728-39. 35. Flocke SA, Stange KC, Zyzanski SJ. The association of attributes of primary care with the delivery of clinical preventive services. Med Care. 1998;36(8 Suppl):AS21-30. 36. Flocke SA. Measuring attributes of primary care: development of a new instrument. J Fam Pract. 1997;45(1):64-74. 37. Mercer SW, Howie JG. CQI-2--a new measure of holistic interpersonal care in primary care consultations. Br J Gen Pract. 2006;56(525):262-8. 38. Mercer SW, McConnachie A, Maxwell M, Heaney D, Watt GC. Relevance and practical

use of the Consultation and Relational Empathy (CARE) Measure in general practice. Fam Pract. 2005;22(3):328-34. 39. Given CW, Branson M, Zemach R. Evaluation and application of continuity measures in primary care settings. J Community Health. 1985;10(1):22-41. 40. Solomon LS, Hays RD, Zaslavsky AM, Ding L, Cleary PD. Psychometric properties of a group-level Consumer Assessment of Health Plans Study (CAHPS) instrument. Med Care. 2005;43(1):53-60. 41. CMS Quality Measure Development Plan Technical Expert Panel Meeting Summary, November 17, 2016. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/MDP_TEP_Nov17_MtgSummary.pdf Accessed July 24, 2020. 42. Vital Directions for Health and Healthcare. Accessed September 2017. https://nam.edu/initiatives/vital-directions-for-health-and-health-care/.

Preliminary Analysis – MUC ID: MUC20-0042 Person-Centered Primary Care Measure

Criteria	Yes/No	Justification and Notes
Does the measure address a critical quality objective not currently adequately addressed by the measures in the program set?	Yes	MUC20-0042 addresses the Meaningful Measurement area of Care is Personalized and Aligned with Patient's Goals, and the MIPS high-priority measurement area of Person and Caregiver-centered Experience and Outcomes. Capturing the voice of the patient is an important component of delivering high-value primary care. MIPS currently has two patient experience measure, the CAHPS survey and a cataract surgery satisfaction measure.
Is the measure evidence-based and either strongly linked to outcomes or an outcome measure?	Yes	MUC20-0042 is an outcome measure realted to patient experience of care. The measure has been submitted to the NQF Primary Care and Chronic Illness (PCCI) Standing Committee as NQF 3568 for review during the Fall 2020 cycle. The developer's evidence submission includes actions that clinicians can perform to improve performance on each of the 11 items on the instrument. The developer cites patient feedback throughout the measure development process, aiming to ensure the instrument reflects patient priorities in primary care delivery. Developer references a body of evidence that demonstrates a strong connection between patient experience of care and traditional health care outcomes, such as improved intermediate outcomes, greater adherence to recommended treatment, and reduced use of health care services (Anhang, et al., 2014; Doyle, et al., 2013; Strange, et al., 2014)

Does the measure address a quality challenge?	Yes	The assessment of patient experience of care is a critical element in care quality. Developer suggests that patient experience measures focus important attention to the consumer experience of care delivery and receipt of services but fall short of focused attention to the broad scope of primary care functions and care (Etz, et al., 2017). In the NQF endorsement testing submission, the developer provides performance data for six clinician groups and 16 individual clinicians. Among the six practices, there were significant differences (p=0.004) in PCPCM PRO-PM scores with a moderate effect size (at least 0.5 standard deviation). Performance scores for the six sites ranged from 0.84-0.90. Among the 16 individual clinicians, there were significant differences in scores as well (p=0.0001) with moderate effect sizes (standard deviations range from 0.39-0.91, with most above 0.5). ICCs ranged from 0.90-0.91 for the practices and 0.76-0.94 for the individual clinicians, with most ICCs above 0.8.
Does the measure contribute to efficient use of measurement resources and/or support alignment of measurement across programs?	Yes	MIPS currently has an experience of care measure, the MIPS CAHPS survey. There are no other primary care experience measures in other CMS programs.
Can the measure can be feasibly reported?	Yes	MUC20-0042 has multiple formats, including the option of electronic administration. The developer reported in their submission for NQF endorsement that the measure requires an average of 90 seconds for patients to complete.
Is the measure applicable to and appropriately specified for the program's intended care setting(s), level(s) of analysis, and population(s)?	Yes	The measure is specified for Clinician: Individual and Clinician: Group/Practice levels, which aligns with MIPS reporting categories. The measure has been submitted for endorsement, has passed scientific methodological review by the NQF Scientific Methods Panel, and will be reviewed by the PCCI Committee in February of 2021.

If the measure is in current use, have negative unintended issues to the patient been identified? Have implementation challenges outweighing the benefits of the measure have been identified?	N/A	The measure is not in current use.
PAC/LTC Core Concept?		N/A
Impact Act Domain		N/A
Hospice High Priority Areas		N/A

Rural Workgroup Input	[redacted]	Relative priority/utility: • Noted to be similar to CAHPS instruments • Noted to be relevant and meaningful to rural patients and providers Data collection issues: • None identified Calculation issues: • None identified Unintended consequences: • None identified Votes: Range is 1 – 5, where higher is more relevant to rural. Average: 4.2 • 1 – 0 votes • 2 – 1 vote • 3 – 2 votes • 4 – 8 votes • 5 – 7 votes
Preliminary Analysis Recommendation	Conditional Support for Rulemaking	Conditional support for rulemaking is contingent on NQF endorsement.
Summary: What is the potential value to the program measure set?	[redacted]	MUC20-0042 addresses the Meaningful Measurement area of Care is Personalized and Aligned with Patient's Goals, and the MIPS high-priority measurement area of Person and Caregiver-centered Experience and Outcomes. Capturing the voice of the patient is an important component of delivering high-value primary care. There are a limited number of patient experience measures within the MIPS program measure set.

Summary: What is the potential impact of this measure on quality of care for patients?	[redacted]	Developer references a body of evidence that demonstrates a strong connection between patient experience of care and traditional health care outcomes, such as improved intermediate outcomes, greater adherence to recommended treatment, and reduced use of health care services. The assessment of patient experience of care is a critical element in care quality. Patient experience measures focus important attention to the consumer experience of care delivery and receipt of services, but fall short of focused attention to the broad scope of primary care. Conditional support for rulemaking is contingent on NQF endorsement.

Measure Comments

Author	Submitted Comment
University of Colorado Medicine	Yes, would recommend including the MUC in the program under certain circumstances
Blue Cross Blue Shield of Massachusetts	MUC20-0042 Person-Centered Primary Care Measure Patient Reported Outcome Performance Measure (PCPCM PRO-PM) should be removed from the MUC list. This measure requires further development before it can be considered ready for high-stakes uses such as performance-based payment and public reporting. For the PCPCM PRO-PM, the most critical area in need of development is case-mix adjustment. To our knowledge, the PCPCM has not undergone empirical analysis to assess the need for case-mix adjustment and to develop case-mix adjustment methods. It is plausible that PCPCM scores, which include items that implicitly assume a need for care "from multiple places" and a long enough relationship to "have been through a lot together," vary substantially according to patient age, health status, and tenure with the index practice. The clinician-level ICCs reported for the PCPCM (as with interunit reliability calculations generally) are likely to be misleading when the underlying measure is not valid for comparisons—for example, because case-mix adjustment is needed but has not been developed. In the absence of case-mix adjustment, high ICCs can result from differences in case-mix rather than differences in providers' true performance. Before this measure is used as a basis for payment, we suggest that the measure developers analyze, based on a large PCPCM fielding that reflects the wide array of practices eligible for MIPS, the relationships between standard CAHPS case-mix adjustment variables (at a minimum) and PCPCM scores—and then develop case-mix adjustment methods and re-estimate the interunit reliabilities of PCPCM PRO-PM scores based on valid (i.e., case-mix adjusted) comparisons.
American Medical Association	The AMA does not have any concerns with this measure but believes that CMS should assess the long-term impact and implications of implementing multiple patient-reported outcome performance measures (PRO-PMs) within MIPS. We ask that the MAP recommendation include the following condition: "CMS should carefully consider the potential burden on clinicians and practices when requiring

	reporting of patient reported outcome performance measures (PRO-PMs) in MIPS and seek to provide incentives and prioritize PRO-PMs that minimize reporting burden."
Federation of American Hospitals	The Federation of American Hospitals (FAH) supports the development and implementation of patient-reported outcomes performance measures (PRO-PMs) but we also believe that additional questions and work remain before their widespread use such as the degree to which multiple PRO-PMs could lead to survey fatigue for patients, the potential impact additional PRO-PMs may have on the reporting of well-established measures such as CG-CAHPs, and what level of data collection burden for an individual PRO-PM is acceptable for a hospital or other healthcare provider. Specifically, it is critical to understand whether there is a potential for individuals to prioritize the completion of one survey over another and therefore lead to negative unintended consequences on response rates for other PRO-PMs such as HCAHPS. Analysis of response rates for HCAHPS from 2008 (33%) to 2017 (26%) revealed a percentage change of -22% overall and an average 0.8 percentage point drop per year (FAH, 2019). This erosion of participation from patients will likely only increase as PRO-PMs become more prevalent. The FAH believes that CMS must develop solutions to these concerns prior to implementation of this measure in the Merit-based Incentive Payment System. As a result, the FAH requests that the highest level of MAP recommendation be "Conditional Support." Reference: Federation of American Hospitals. Modernizing the HCAHPS Survey. Released June 2019. Available at: https://www.fah.org/fah-ee2-uploads/website/documents/Modernizing_HCAHPS_-Recommendations_from_PELs.pdf.

Measure Comments (Post-Workgroup Meeting)

Author	Submitted Comment
Federation of American Hospitals	The Federation of American Hospitals (FAH) continues to have concerns with the potential implementation of this measure. Specifically, the FAH believes that additional questions and work remain before its widespread use such as the degree to which multiple PRO-PMs could lead to survey fatigue for patients, the potential impact additional PRO-PMs may have on the reporting of well-established measures such as HCAHPs, and what level of data collection burden for an individual PRO-PM is acceptable for a hospital or other healthcare provider. As a result, the FAH requests that the highest level of MAP recommendation be "Do Not Support with Potential for Mitigation."

MUC20-0043: Preventive Care and Wellness (composite)

Measure Information

Characteristic	Submitted Information
MUCID	MUC20-0043
Other Measure Identification Numbers	
Title	Preventive Care and Wellness (composite)
Program	Merit-based Incentive Payment System-Quality
Workgroup	Clinician
In what state of development is the measure?	Early Development
State of Development Details	We used registry-reported CMS program data from 2018 and 2019 calendar years for the seven component measures from the Merit-Based Incentive Payment System (MIPS) to test the reliability and validity of the PCW composite measure. Note that CMS is currently testing this measure with patient-level data (data forthcoming). Reliability: Using registry-reported CMS program data from 2018 and 2019 calendar years for the seven component measures from the Merit-Based Incentive Payment System (MIPS), we computed reliability of the PCW composite using the following methods: 1. Signal-to-noise approach, which captures precision of the measure scores at the clinician level; 2. Test-retest approach, which captures stability of the measure scores across two samples of patients or two consecutive years of data; 3. Internal-consistency (Cronbach's alpha) approach, which captures how closely the individual measure components are related as a group. Signal-to-noise reliability: The results of the signal-to-noise reliability analysis show that the mean reliability score of the composite measure was high (≥0.99) for both measurement years, regardless of the number of component measures reported by clinicians. For context, reliability above 0.70 is considered sufficient to draw conclusions about groups, and values above 0.9 are considered sufficient to draw conclusions about individuals (Adams 2009). We attribute a high reliability of the composite measure to a sufficiently large sample size of patients for each individual component measure, which, in turn, results in very small noise variance (within-clinician variance specific to each individual component measure). Also, for clinicians who do not report all seven measures, high reliability could be driven by imputing missing measure scores

with the national means, which have no variation. Citation: Adams, J. L. The Reliability of Provider Profiling: A Tutorial. Santa Monica, CA: RAND Corporation, 2009. Available at Santa Monica, CA: RAND Corporation, 2009. Available at https://www.rand.org/pubs/technical_reports/TR653.html. Table 2. Signal-to-noise reliability of Preventive Care and Wellness composite scores by year and number of component measures reported. [See supporting file.] Source: Mathematica analysis of 2018 and 2019 MIPS Eligibility File (MIPSEF) data set. The data are restricted to unique clinicians, identified by their individual National Provider Identifier (NPI), reporting to MIPS as individuals. Note: For the reliability and validity analysis, we excluded one clinician whose denominator sample size for QIDs 110 and 111 were considered outliers. Max = maximum; Min = minimum; MIPS = Merit-based Incentive Payment System; N = number of clinicians; pctl. = percentile. Test-retest reliability: For clinicians who reported scores for the same individual measure in both 2018 and 2019, we additionally computed test-retest reliability (also referred to as temporal stability reliability) by correlating clinicians' measure scores in the two consecutive periods. We computed composite scores using the component-level linear combination method, which was driven by the lack of available patient-level data at the time of testing. We are currently pursuing patient-level data to facilitate testing using several composite approaches. The Spearman correlation between composite scores for clinicians who reported at least one, at least four measures, and all seven measures was high indicating adequate temporal stability of the composite measure scores for the same clinicians over time. Table 3. Test-retest reliability for the Preventive Care and Wellness composite scores (for clinicians in both 2018 and 2019 samples). [See supporting file.] Source: Mathematica analysis of 2018 and 2019 MIPS Eligibility File (MIPSEF) data set. The data are restricted to unique clinicians, identified by their individual National Provider Identifier (NPI), reporting to MIPS as individuals. Note: For the reliability and validity analysis we excluded one clinician whose denominator sample size for QIDs 110 and 111 were considered outliers. Composite internal consistency: We assessed the internal consistency of the PCW composite via Cronbach's alpha statistic. The alpha statistic assesses the degree to which the individual measure components measure the same underlying construct. Internal consistency was high (>0.8) in both 2018 and 2019 data, regardless of the approach we chose to handle missing component measures (that is, listwise or pairwise deletion). We tested the correlation of each individual component measure with the PCW composite computed based on the remaining six measures, and the Cronbach's alpha if an individual component was removed from the composite. For example, in 2018, QID 110: Preventive Care and Screening: Influenza Immunization, correlates with the composite score at 0.636 (listwise deletion). If the QID 110component measure was excluded from the composite, internal consistency of the composite would change to 0.783. This indicates that excluding the QID 110 measure would decrease internal consistency of the composite by 0.036 points from the baseline alpha of 0.819. By contrast, QID 226: Tobacco Use: Screening and Cessation, correlates with the composite score at 0.305 (listwise deletion column). If this component measure was excluded from the composite, its internal consistency would increase from 0.819 to 0.843. Results suggest that of all seven measures included in a composite score, six demonstrated moderate-to-strong correlations

(Spearman ρ~0.5 to 0.8) with the composite. One measure (QID 226: Tobacco Use: Screening and Cessation) showed weaker correlation with the composite (Spearman ρ~0.03 to 0.4). This means the internal consistency of the composite would increase if this component measure was excluded from the composite. Table 4. Internal-consistency reliability for the Preventive Care and Wellness composite scores (for clinicians in both 2018 and 2019 samples). [See supporting file.] Source: Mathematica analysis of 2018 and 2019 MIPS Eligibility File (MIPSEF) data set. The data are restricted to unique clinicians, identified by their individual National Provider Identifier (NPI), reporting to MIPS as individuals. Note: For the reliability and validity analysis we excluded one clinician whose denominator sample size for QIDs 110 and 111 were considered outliers. * The sample size varies for each pair of measures (Ns for the individual component measures ranged from 2,737 to 12,494 in the 2018 data and from 707 to 5,985 for the 2019 data). In the listwise deletion (complete case analysis), clinicians were excluded from the analysis if they were missing data on any of the individual component measures. In the pairwise deletion (available case analysis), Cronbach's alpha was calculated on all available data including missing values. Table 5. Item-total correlations and Cronbach's alpha if component measures were removed from a composite (for clinicians in both 2018 and 2019 samples). [See supporting file.] Source: Mathematica analysis of 2018 and 2019 MIPS Eligibility File (MIPSEF) data set. The data are restricted to unique clinicians, identified by their individual National Provider Identifier (NPI), reporting to MIPS as individuals. Note: For the reliability and validity analysis, we excluded one clinician whose denominator sample size for QIDs 110 and 111 were considered outliers. Alpha indicates internal-consistency reliability coefficient if a component measure was excluded from the composite score. Validity: Concurrent validity: Using registry-reported CMS program data from 2018 and 2019 calendar years for the seven component measures from the Merit-Based Incentive Payment System (MIPS), we examined the validity of the PCW composite measure by measuring correlations between the individual component measures (concurrent validity). If the individual component measures tap into the same underlying construct (for example, the quality of the preventive care provided by clinicians), we would expect the individual component measures to show positive correlation with each other. Note that this approach is conceptually different from the internal-consistency reliability analysis. To compute internal-consistency reliability of the PCW composite, we correlated individual component measure scores with the composite scores. In contrast, in the validity analysis we computed correlations between the individual measures. We found low to moderate positive correlations between the individual measures included in this composite in both 2018 and 2019. Correlation coefficients range from 0.045 to 0.602. As expected, individual measures with similar constructs had the strongest correlations. In both years, the highest correlation coefficients were between cancer screening measures (QID 112: Breast Cancer Screening and QID 113: Colorectal Cancer Screening) and between vaccination/immunization measures (QID 110: Influenza Immunization and QID 111: Pneumococcal Vaccination Status for Older Adults). The weakest correlation coefficients in both years were for the tobacco use and cessation intervention measure (QID 226: Tobacco Use: Screening and Cessation) and each of the cancer screening measures (QID 112: Breast Cancer

	Screening and QID 113: Colorectal Cancer Screening). Overall, these positive correlations among component measures indicate alignment with a similar underlying construct. Table 6. Bivariate correlations for the individual component measures (for clinicians in both 2018 and 2019 samples). [See supporting file.] Source: Mathematica analysis of 2018 and 2019 MIPS Eligibility File (MIPSEF) data set. The data are restricted to unique clinicians, identified by their individual National Provider Identifier (NPI), reporting to MIPS as individuals. Note: For the reliability and validity analysis we excluded one clinician whose denominator sample size for QIDs 110 and 111 were considered outliers. [a] represents correlations for 2019 data. [b] represents correlations for 2018 data.
Measure Description	Percentage of patients who received age- and sex-appropriate preventive screenings and wellness services. This measure is a composite of seven component measures that are based on recommendations for preventive care by the U.S. Preventive Services Task Force (USPSTF), Advisory Committee on Immunization Practices (ACIP), and American Association of Clinical Endocrinologists (AACE) and American College of Endocrinology (ACE).
Numerator	Numerator 1: Patients who received an influenza immunization OR who reported previous receipt of an influenza immunization (Previous Receipt – Receipt of the current season's influenza immunization from another provider OR from same provider prior to the visit to which the measure is applied [typically, prior vaccination would include influenza vaccine given since August 1st]). Numerator 2: Patients who have ever received a pneumococcal vaccination before the end of the measurement period Numerator 3: Women with one or more mammograms during the 27 months prior to the end of the measurement period Numerator 4: Patients with one or more screenings for colorectal cancer Numerator 5: Patients with a documented BMI during the encounter or during the previous twelve months, AND when the BMI is outside of normal parameters, a follow-up plan is documented during the encounter or during the previous twelve months of the current encounter Numerator 6: - Patients who were screened for tobacco use at least once within 24 months - Patients who received tobacco cessation intervention

	- Patients who were screened for tobacco use at least once within 24 months AND who received tobacco cessation intervention if identified as a tobacco user Numerator 7: Patients who were screened for high blood pressure AND have a recommended follow-up plan documented, as indicated, if the blood pressure is pre-hypertensive or hypertensive Composite method: To create the composite score for this draft CQM as currently specified, we used component-level linear combination. First, we computed the measure score for each individual component measure (measure numerator divided by the measure denominator), and then computed the average of the seven individual scores.
Denominator	Denominator 1: All patients aged 6 months and older seen for a visit during the measurement period Denominator 2: Patients 65 years of age and older with a visit during the measurement period Denominator 3: Women 51 - 74 years of age with a visit during the measurement period Denominator 4: Patients 50-75 years of age with a visit during the measurement period Denominator 5: All patients aged 18 and older on the date of the encounter with at least one eligible encounter during the measurement period Denominator 6: • All patients aged 18 years and older seen for at least two visits or at least one preventive visit during the measurement period • All patients aged 18 years and older seen for at least two visits or at least one preventive visit during the measurement period who were screened for tobacco use and identified as a tobacco user • All patients aged 18 years and older seen for at least two visits or at least one preventive visit during the measurement period Denominator 7: All patients aged 18 years and older at the beginning of the measurement period with at least one eligible encounter during the measurement period
Exclusions	Denominator Exclusion Population 1: None Denominator Exception Population 1: Influenza immunization was not administered for reasons documented by clinician (e.g., patient allergy or other medical reasons, patient declined or other patient reasons, vaccine not available or other system reasons)

Denominator Exclusion Population 2: Patient received hospice services any time during the measurement period

Denominator Exception Population 2: Not applicable

Denominator Exclusion Population 3:

- Women who had a bilateral mastectomy or who have a history of a bilateral mastectomy or for whom there is evidence of a right and a left unilateral mastectomy

- Hospice services used by patient any time during the measurement period

- Patients age 66 or older in Institutional Special Needs Plans (SNP) or residing in long term care

- Patients 66 years of age and older with at least one claim/encounter for frailty during the measurement period AND a dispensed medication for dementia during the measurement period or the year prior to the measurement period

- Patients 66 years of age and older with at least one claim/encounter for frailty during the measurement period AND either one acute inpatient encounter with a diagnosis of advanced illness or two outpatient, observation, ED or nonacute inpatient encounters on different dates of service with an advanced illness diagnosis during the measurement period or the year prior to the measurement period

Denominator Exception Population 3: Not applicable

Denominator Exclusion Population 4:

- Patients with a diagnosis or past history of total colectomy or colorectal cancer

- Patient was provided hospice services any time during the measurement period

- Patient age 66 or older in Institutional Special Needs Plans (SNP) or residing in long-term care

- Patients 66 years of age and older with at least one claim/encounter for frailty during the measurement period AND a dispensed medication for dementia during the measurement period or the year prior to the measurement period

- Patients 66 years of age and older with at least one claim/encounter for

	frailty during the measurement period AND either one acute inpatient encounter with a diagnosis of advanced illness or two outpatient, observation, ED or nonacute inpatient encounters on different dates of service with an advanced illness diagnosis during the measurement period or the year prior to the measurement period Denominator Exception Population 4: Not applicable Denominator Exclusion Population 5: - BMI not documented, documentation the patient is not eligible for BMI calculation - BMI is documented as being outside of normal limits, follow-up plan is not documented, documentation the patient is not eligible Denominator Exception Population 5: BMI is documented as being outside of normal limits, follow-up plan is not completed for documented reason Denominator Exclusion Population 6: None Denominator Exception Population 6: - Documentation of medical reason(s) for not screening for tobacco use (e.g., limited life expectancy, other medical reason) - Documentation of medical reason(s) for not providing tobacco cessation intervention (e.g., limited life expectancy, other medical reason) - Documentation of medical reason(s) for not screening for tobacco use (e.g., limited life expectancy, other medical reason) -Documentation of medical reason(s) for not providing tobacco cessation intervention if identified as a tobacco user (e.g., limited life expectancy, other medical reason) Denominator Exclusion Population 7: Patient not eligible due to active diagnosis of hypertension Denominator Exception Population 7: Documented reason for not screening or recommending a follow-up for high blood pressure
Measure type	Composite
What is the NQF status of	Never submitted

the measure?	
NQF ID number	0000
Year of next anticipated NQF CDP endorsement review	NA
Year of most recent NQF Consensus Development Process (CDP) endorsement	NA
Is the measure being submitted exactly as endorsed by NQF?	NA
If not exactly as endorsed, describe the nature of the differences	NA
What data sources are used for the measure?	Other (enter here): CMS clinician-level Quality Payment Program CQM data
If EHR or Administra	Information not provided

tive Claims or Chart-Abstracted Data, description of parts related to these sources.	
At what level of analysis was the measure tested?	Clinician
In which setting was this measure tested?	Ambulatory/office-based care
What NQS priority applies to this measure?	Information not provided
What one primary meaningful measure area applies to this measure?	Preventive care
What secondary meaningful measure area applies to this measure?	Admissions and readmissions to hospitals
What one primary healthcare	Promote effective prevention and treatment of chronic disease

priority applies to this measure?	
What secondary healthcare priority applies to this measure?	Promote effective communication and coordination of care
What area of specialty best fits the measure?	Family practice (see Comments)
What is the target population of the measure?	All Payer
Is this measure an eCQM?	No
If eCQM, enter Measure Authoring Tool (MAT) number	No
If eCQM, does the measure have a Health Quality Measures Format (HQMF) specification?	No

Comments	For Row #23 (area of specialty) Alternative choices: Preventive medicine; Primary care. Developed under contract with CMS (5FCMC18D0032/75FCMC19F0004)
Measure steward	Centers for Medicare & Medicaid Services
Long-Term Measure Steward (if different)	NA
Measure Steward Contact Information	Andress, Joel; CMS CCSQ QMVIG; 410-786-5237; Joel.Andress@cms.hhs.gov
Primary Submitter Contact Information	Bandyopadhyay, Jayanti; Mathematica Policy Research (contractor); 609-297-4546; jbandyopadhyay@mathematica-mpr.com
Long-Term Measure Steward Contact Information	NA
Secondary Submitter Contact Information	Holland, Christine; Mathematica Policy Research (contractor); 202-484-5271; cholland@mathematica-mpr.com
Was this measure proposed for a previous year's MUC list?	No
In what prior year(s) was this measure	None

proposed?	
What were the programs that NQF MAP reviewed the measure for in each year?	Not applicable.
Why was the measure not recommended in those year(s)?	Not applicable.
What were the MUC IDs for the measure in each year?	Not applicable.
NQF MAP report page number being referenced for each year	Not applicable.
What was the NQF MAP recommendation in each year?	Not applicable.
List the NQF MAP workgroup(s) in each	Not applicable.

year	
What is the history or background for including this measure on the new MUC list?	New measure never reviewed by MAP Workgroup or used in a CMS program
Range of years(s) this measure has been used by CMS Program(s)	NA
What other federal programs are currently using this measure?	NA
Evidence that the measure can be operationalized	All of the component measures are based on measures currently implemented in MIPS, which supports the feasibility of the composite measure.
How is the measure expected to be reported to the program?	CQM (Registry)
Is this measure similar to and/or	Yes

competing with measure(s) already in a program?	
Which existing measure(s) is your measure similar to and/or competing with?	The composite uses existing measures in the MIPS program: - Quality ID 110: Preventive Care and Screening: Influenza Immunization - Quality ID 111: Pneumococcal Vaccination Status for Older Adults - Quality ID 112: Breast Cancer Screening - Quality ID 113: Colorectal Cancer Screening - Quality ID 128: Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow-Up Plan - Quality ID 226: Preventive Care and Screening: Tobacco Use: Screening and Cessation Intervention - Quality ID 317: Preventive Care and Screening: Screening for High Blood Pressure and Follow-Up Documented
How will this measure be distinguished from other similar and/or competing measures?	We did not identify any competing composite measures. This composite measure is constructed using existing component measures that are supported by recommendations from the USPSTF, the ACIP, or the AACE/ACE, therefore the composite's components exactly align with the parent measures currently used for PY2020 MIPS reporting.
Rationale for how this measure will add to the CMS program	With rising rates of certain chronic conditions in the general population, wellness and preventive care have become increasingly important to improve outcomes and reduce costs. Research shows that performing the preventive services identified in the measure leads to identification of disease earlier in the care process (screenings) or prevention of disease (immunizations), which enables treatment to begin earlier, potentially improving patient outcomes. The composite measure can provide an opportunity for providers and patients to identify and manage a patient's health risks for many preventable conditions. This measure assigns a single performance score reflecting overall eligible clinician delivery of age- and sex-appropriate preventive screenings and wellness services to their patients. The seven services in this measure are (1) influenza vaccination, (2) pneumococcal vaccination, (3) breast cancer screening, (4) colorectal-cancer screening, (5) body mass index screening and follow-up, (6) tobacco use

screening and intervention, and (7) screening for high blood pressure and follow-up. The services contained in the measure are recommended by USPSTF, ACIP, and AACE/ACE and apply to the general population (rather than a specific age group with specific risks, for example, older adults with cardiovascular risk). Although increased use of preventive care services may cause a short-term increase in health care costs, it may result in better quality of life and care. A study of preventive services covered under the Affordable Care Act examined the extent to which lives could be saved if adults over 18 received them, including some addressed by this measure. The article states that preventive services ameliorate 9 of the 10 leading causes of death in America and could save at least 100,000 lives (Fox and Shaw 2015). Among the services referenced are screening for breast cancer, colon cancer, blood pressure, diabetes, and tobacco cessation, as well as influenza and pneumococcal vaccination. Higher rates of patient compliance with the appropriate and recommended preventive services could save additional lives and ensure better health outcomes.

CMS prioritized development of this measure because, as a composite, it had several advantages for CMS and stakeholders when compared to using individual measures in a program. Composites can overcome statistical challenges such as small sample sizes while reducing data burden for interpretability (Peterson et al., 2010; Samuel, 2014; van Doorn-Klomberg et al., 2012). Due to the condensed nature of the composite's information, it is more feasible to track a broader, more comprehensive range of metrics than otherwise possible, making composites well suited for pay-for-performance incentives or consumer decisions about clinicians (Peterson et al., 2010). Composite measures are an important strategy to maintain data fidelity as they are more likely to be stable over time, making incentives less sensitive to individual measure performance (Martsolf, 2012; Prentice et al., 2016). Potential implementation of this composite measure not only provides a more comprehensive assessment of a clinician's performance of preventive care than any single measure, but also provides CMS an opportunity to replace the individual measures in the program with a more robust measure, which aligns with the meaningful measure framework's goal to include fewer, more robust measures in the program overall.

Citations:

Fox, J.B., and F.E. Shaw. "Clinical Preventive Services Coverage and the Affordable Care Act." American Journal of Public Health, vol. 105, no. 1, 2015, pp. e7–e10.

Martsolf, G. (2012). Creation and Evaluation of Composite Measures of Physician Practice Quality Using Aggregated Health Insurance Claims. The Pennsylvania State University, 194.

Peterson, E. D., DeLong Elizabeth R., Masoudi Frederick A., O'Brien Sean M., Peterson Pamela N., Rumsfeld John S., Shahian David M., & Shaw Richard E. (2010). ACCF/AHA 2010 Position Statement on Composite Measures for Healthcare Performance Assessment. Circulation, 121(15), 1780–1791. https://doi.org/10.1161/CIR.0b013e3181d2ab98.

Prentice, J. C., Frakt, A. B., & Pizer, S. D. (2016). Metrics That Matter. Journal of General Internal Medicine, 31(1), 70–73. https://doi.org/10.1007/s11606-015-3559-0.

Samuel, C. A. (2014, March). Essays on Health Care Quality and Access: Cancer Care Disparities, Composite Measure Development, and Geographic Variations in Electronic Health Record Adoption. https://dash.harvard.edu/bitstream/handle/1/12274571/Samuel_gsas.harvard_0084L_11583.pdf?sequence=4&isAllowed=y.

van Doorn-Klomberg, A.L., J.C. Braspenning, R.C. Feskens, M. Bouma, S.M. Campbell, and D. Reeves. "Precision of Individual and Composite Performance Scores: The Ideal Number of Indicators in an Indicator Set." Medical Care. doi: 10.1097/MLR.0b013e3182726bf1. Epub 2012.

If this measure is being proposed to meet a statutory requirement, please list the corresponding statute.	Not applicable.
Evidence of performance gap	Using registry-reported CMS program data from 2018 and 2019 calendar years for the seven component measures from the Merit-Based Incentive Payment System (MIPS), we calculated distributions of the 2018 and 2019 composite rates using the component-level linear combination method to determine if the PCW composite measure was "topped out." For the composite measure to be topped out, two conditions had to be met (Analysis of Topped-Out Measures 2014). First, the truncated coefficient of variation (TCV) (calculated by first removing the lower and upper 5th percentiles and then dividing the standard deviation by the mean of this truncated distribution) must be less than or equal to 0.10. Second, the 75th performance percentile must be statistically indistinguishable (within two standard

	errors) from the 90th percentile. Results indicated that PCW composite did not meet the topped-out criteria. Citation: “Analysis of Topped-Out Measures Finalized for the PY 2016 ESRD QIP.” Updated June 19, 2014. Available at https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/ESRDQIP/Downloads/AnalysisofTopped-OutMeasuresFinalizedforthePY2016ESRDQIP.pdf. Table 1. Topped-out Analysis of the Preventive Care and Wellness composite. [See supporting file.] Source: Mathematica analysis of 2018 and 2019 MIPS Eligibility File (MIPSEF) data set. The data are restricted to unique clinicians, identified by their individual National Provider Identifier (NPI), reporting to MIPS as individuals. Notes: MIPS = Merit-based Incentive Payment System; N = number of clinicians; TCV = truncated coefficient of variation; pctl. = percentile; SE = standard error; Δ = difference. Bootstrap estimates for standard errors for the 90th percentile based on 1,000 replications.
Unintended consequences	This is a new measure and has not been submitted for NQF endorsement or MAP review.
Which clinical guideline(s)?	The measures in the Preventive Care and Wellness composite are based on seven preventive services recommended by the USPSTF, the ACIP, or the AACE/ACE. Component 1: Routine annual influenza vaccination is recommended for all persons aged >=6 months who do not have contraindications. Optimally, vaccination should occur before onset of influenza activity in the community. Although vaccination by the end of October is recommended, vaccine administered in December or later, even if influenza activity has already begun, is likely to be beneficial in the majority of influenza seasons (CDC/Advisory Committee on Immunization Practices [ACIP], 2018). Component 2: In 2014, the Advisory Committee on Immunization Practices (ACIP) began recommending a dose of 13-valent pneumococcal conjugate vaccine (PCV13) be followed by a dose of 23-valent pneumococcal polysaccharide vaccine (PPSV23) 6-12 months later in adults aged 65 and older who have not previously received a pneumococcal vaccination, and in persons over the age of two years who are considered to be at higher risk for pneumococcal disease due to an underlying condition. The two vaccines should not be coadministered and intervals for administration of the two vaccines vary slightly depending on the age, risk group, and history of vaccination (Kobayashi, 2015). In 2015, ACIP updated its recommendation and changed the interval between PCV13 and PPSV23, from 6-12 months to at least one year for immunocompetent adults aged >=65 years who have not previously received pneumococcal vaccine. For immunocompromised vaccine-naïve adults, the minimum acceptable interval between PCV13 and PPSV23 is 8 weeks. Both

immunocompetent and immunocompromised adults aged >=65 years who have previously received a dose of PPSV23 when over the age of 65 should receive a dose of PCV13 at least one year after PPSV23 (>=1 year). Immunocompetent and immunocompromised adults aged >=65 who have previously received a dose of PPSV23 when under the age of 65, should also receive a dose of PCV13 at least one year after PPSV23 (>=1 year) and then another dose of PPSV23 at least one year after PCV13. It is recommended that for those that have this alternative three-dose schedule (2 PPSV23 and 1 PCV13), the three doses should be spread over a time period of five or more years (Kobayashi, 2015).

Component 3: The U.S. Preventive Services Task Force (USPSTF) recommends biennial screening mammography for women aged 50-74 years (B recommendation).

Component 4: The U. S. Preventive Services Task Force (2016) recommends screening for colorectal cancer starting at age 50 years and continuing until age 75 years. This is a Grade A recommendation (U.S. Preventive Services Task Force 2016).

Component 5: All adults should be screened annually using a BMI measurement. BMI measurements >25kg/m2 should be used to initiate further evaluation of overweight or obesity after taking into account age, gender, ethnicity, fluid status, and muscularity; therefore, clinical evaluation and judgment must be used when BMI is employed as the anthropometric indicator of excess adiposity, particularly in athletes and those with sarcopenia (Garvey, et al., 2016 AACE/ACE Guidelines, 2016. pp. 12-13) (Grade A).

Component 6: The USPSTF recommends that clinicians ask all adults about tobacco use, advise them to stop using tobacco, and provide behavioral interventions and U.S. Food and Drug Administration (FDA)-approved pharmacotherapy for cessation to adults who use tobacco (Grade A Recommendation) (U.S. Preventive Services Task Force, 2015).

Component 7: The U.S. Preventive Services Task Force (USPSTF) recommends screening for high blood pressure in adults age 18 years and older. This is a grade A recommendation.

Briefly describe the peer	Each component measure corresponds to a measure currently used for PY2020 reporting in MIPS, some of which are also NQF-endorsed measures, meaning the evidence for each measure has been evaluated by CMS, and in some cases

reviewed evidence justifying this measure	by an NQF committee, and determined to have enough evidence to support the measure intent and inclusion in MIPS. Component measures endorsed by NQF: • Quality ID 110: Preventive Care and Screening: Influenza Immunization (NQF #0041) • Quality ID 112: Breast Cancer Screening (NQF #2372) • Quality ID 113: Colorectal Cancer Screening (NQF #0034) • Quality ID 226: Preventive Care and Screening: Tobacco Use: Screening and Cessation Intervention (NQF #0028) Component measures not endorsed by NQF: • Quality ID 111: Pneumococcal Vaccination Status for Older Adults • Qualtiy ID 128: Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow-Up Plan • Quality ID 317: Preventive Care and Screening: Screening for High Blood Pressure and Follow-Up Documented

Preliminary Analysis – MUC20-0043: Preventive Care and Wellness (composite)

Criteria	Yes/No	Justification and Notes
Does the measure address a critical quality objective not currently adequately addressed by the measures in the program set?	Yes	The measure aligns with CMS Meaningful Measure area of Promoting Effective Prevention & Treatment of Chronic Disease; Preventive Care. The composite measure will help identify and manage preventable chronic conditions by using seven identical preventive care PY 2020 MIPS measures: • Quality ID 110: Preventive Care and Screening: Influenza Immunization • Quality ID 111: Pneumococcal Vaccination Status for Older Adults • Quality ID 112: Breast Cancer Screening • Quality ID 113: Colorectal Cancer Screening • Quality ID 128: Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow-Up Plan • Quality ID 226: Preventive Care and Screening: Tobacco Use: Screening and Cessation Intervention • Quality ID 317: Preventive Care and Screening: Screening for High Blood Pressure and Follow-Up Documented The composite consists of measures currently included in MIPS. Submission states the possibility of replacing the individual measures for one composite measure. MIPS currently has no composite measures.
Is the measure evidence-based and either strongly linked to outcomes or an outcome measure?	Yes	This composite is made up of singular process measure which meets clinical guidelines for the various preventative components it aims measures. Research shows that preventive health care is vital aspect of medical practice and can lead to signifcant imrpovements in a patient's overall health (CDC, 1999). Overall, there is strong evidence that the individual components of screening and immunization result in better health outcomes. For example, one study found that delivery of preventive services of cancer screening and vaccines could avert 100,000 deaths per year (Woolf, 2009).

Does the measure address a quality challenge?	Importance subcriteria of CDP worksheet and report; UptoDate	This measure addresses a clinical component of health care for all recipients of care. Immunization rates have opportunities for improvement (CDC, 2018). Data shows beneficiries have high rates of hypertension (58.8%) and hyperlipidemia (49.1%) which can have servere health effects (CDC, 2020). Addionally, beneficiaries over the age of 65 account for 54% of all new cancer cases (American Cancer Society, 2013).
Does the measure contribute to efficient use of measurement resources and/or support alignment of measurement across programs?	No	The seven components of this composite measures are all currently used in MIPS and Part C and D Star Ratings program and are duplicative if component measures are kept. The submission states that replacement of the singular measures for the composite measure may improve data interpretability burden and make tracking of preventive care easier and comprehensive (Prentice et al., 2016, Peterson, et al, 2010).
Can the measure can be feasibly reported?	Yes	This measure uses clinician-level quality payment program clinical quality measure (CQM) registry data that can be feasibly reported and is a low-burden data source.
Is the measure applicable to and appropriately specified for the program's intended care setting(s), level(s) of analysis, and population(s)?	Yes	The measure is specified for the clinician level of analysis which is appropriate for this program.
If the measure is in current use, have negative unintended issues to the patient been identified? Have implementation challenges outweighing the benefits of the measure been identified?	N/A	The measure is not in current use and the developers did not indicate any potential unintended consequences.

PAC/LTC Core Concept?		N/A
Impact Act Domain		N/A
Hospice High Priority Areas		N/A
Rural Workgroup Input	[redacted]	Relative priority/utility: • The measure was noted to be low burden. • The measure was noted to be a good "report card" especially for rural providers and patients. Data collection issues: • None identified Calculation issues: • A concern was expressed for the complexity of the composite score and interpretation of aggregate data. Unintended consequences: • None identified Program gap areas: • None identified Votes: Range is 1 – 5, where higher is more relevant to rural. Average: 3.9 • 1 – 0 votes • 2 – 2 votes • 3 – 1 vote • 4 – 13 votes • 5 – 3 votes
Preliminary Analysis Recommendation	Conditional Support for Rulemaking	Conditional support contingent upon receipt of NQF endorsement and addressing redundancy issues from duplicative component measures in MIPS.

Summary: What is the potential value to the program measure set?	[redacted]	The seven components of this composite measures are all currently used in MIPS and Part C and D program. The resolution of potential redundancy with the singular measures for the composite measure already in MIPS may improve data interpretability burden for reporting entities and would make tracking of preventive care easier and comprehensive.
Summary: What is the potential impact of this measure on quality of care for patients?	[redacted]	This measure may impact the 37 million Medicare beneficiaries who receive one or more preventive services, and the 1 in 6 Medicare beneficiaries who are younger than 65 years old who would seek preventive services (Fox, et al; 2015). Conditional support contingent upon receipt of NQF endorsement and addressing redundancy issues from duplicative component measures in MIPS.

Measure Comments

Author	Submitted Comment
University of Colorado Medicine	Composite measures are not very useful for informing quality improvement programs. While clinicians and systems strive to have their patients be completely up to date with all preventive screenings, there are so many reasons that a patient may fall behind on any given topic that the combination of all topics and all reasons renders it almost impossible for a person to be up to date on all. This same fact renders the information from compositve measues to be virtually meaningless for those monitoring and running QI programs. The reasons for falling into the 'not up to date' categorty- include just aging in, near to aging out, completed external to system and not in the EHR system in manner that allows it be captured, pt desires to postpone, pt disagrees with the recommendation, pt is new to system and we're just getting started with their care. Many of us have worked on early diabetes composite measuers, and our experience with those is that the results were not informative, we needed to tease apart the individual measures to find value, and the results were so discouraging to providers that we have to stop focusing on it/presenting the composite measure. As a side note, what we found was that scores were low due to patients receipt of diabetic eye exams external to our system. What we did was spend a lot of time and resources tracking down external eye results. What we didn't do was spend time developing processes to improve patient care and outcomes.
Pfizer	Finding ways to prevent or detect illness through Value-based care should be prioritized. Preventive care can help to lower health care costs by preventing, treating diseases, and preventing additional comorbid conditions. Preventive care has been shown to save lives by either preventing or detecting an illness in an earlier stage, thereby improving outcomes. Implementing a Preventive care & wellness measure will bring accountability and improvements in provider-patient care plans. However, in respect to colorectal and breast cancer screening, the exclusion for patients over 66 who have at least one claim for frailty during the measurement period and an inpatient encounter or outpatient encounters with a diagnosis of

	advanced illness is discriminatory for those who are disabled and will increase disparities given the higher rates of poor health among Black and Hispanic older adults (JAMA 2020, 11/11) This measure outweighs the burden of data collection or reporting. Evidence has shown that the use of prevention saves lives. Further, the majority of the preventive items can be done in an annual physical, allowing the provider to collect the data through their EMR.
American Medical Association	While the AMA supports the individual components included in this composite, we are concerned that the complexity of the measure with seven numerators, denominators and exclusions/exceptions will directly impact the feasibility of the measure for use in MIPS. We anticipate that implementation will be challenging given that several of the components require longitudinal data. Based on the information available during this comment period, it appears that only existing MIPS data was used to assess the reliability and validity of the measure and feasibility of reporting this composite is assumed because the individual measures are currently available for reporting in MIPS. Additional evaluations of the feasibility of collecting and reporting the data required for this measure is needed. As a result, we recommend that the highest level of MAP recommendation be "Do Not Support with Potential for Mitigation."
Association of American Medical Colleges (AAMC)	This measure is still in early development stage and has not been submitted for NQF endorsement. The AAMC has long held that measures should not be proposed for addition to public reporting programs unless vetted and endorsed by the NQF. This is especially true with composite measures, to ensure that not only the composite is valid and reliable, but that so too are the underlying component measures. It is simply premature for consideration in a public reporting and performance program. The AAMC recommends that the highest level of MAP recommendation be "Do Not Support With Potential For Mitigation."
American Academy of Neurology	The American Academy of Neurology (AAN), an association of more than 36,000 neurologists and neuroscience professionals, appreciates the opportunity to comment on the proposed measure. The AAN understands that neurology providers are not the intended users of the measure but notes concerns about feasibility and burden of collecting this data for practices, particularly solo and small providers. These concerns should be studied and balanced prior to implementation in a payment program to prevent potential disparities.
Federation of American Hospitals	The Federation of American Hospitals (FAH) supports the individual measures included within this composite but questions whether the feasibility of reporting this measure has been fully answered. The developers state that because each of the seven measures are currently reported within the Merit-based Incentive Payment System, individual clinicians and practices should be able to collect and report on this composite. The FAH does not believe that this assumption is practical given the fact that each of the measures has different timeframes and patient populations, and to our knowledge all seven are not typically selected and reported by one practice in a reporting year. Additional work must be completed to understand how

	feasible this composite would be to collect and report prior to its implementation. As a result, the FAH requests that the highest level of MAP recommendation be "Do Not Support with Potential for Mitigation."

Measure Comments (Post-Workgroup Meeting)

Author	Submitted Comment
American Heart Association/American Stroke Association	The AHA supports this important measure for use in the MIPS program. Although not endorsed as a composite, most of the components are—or were previously—NQF-endorsed.
Federation of American Hospitals	Support
American Medical Association	The American Medical Association (AMA) remains concerned that the complexity of the measure with seven numerators, denominators and exclusions/exceptions will directly impact the feasibility of the measure for use in MIPS. This concern was not adequately addressed during the Clinician Workgroup deliberations. In fact, we were alarmed to learn that CMS intends to remove the seven individual measures if this composite is implemented in MIPS as were many of the workgroup members. The AMA does not support the removal of these measures that address important preventive care activities and as a result, cannot support this measure. The AMA asks that the recommendation be changed to "Do not Support."

MUC20-0045 SARS-CoV-2 Vaccination by Clinicians

Measure Information

Characteristic	Submitted Information
MUCID	MUC20-0045
Other Measure Identification Numbers	N/A
Title	SARS-CoV-2 Vaccination in by Clinicians
Program	MIPS
Workgroup	Clinician
In what state of development is the measure?	Under development
State of Development Details	N/A
Measure Description	Percentage of patients aged 18 years and older seen for a visit during the measurement period who have ever received or reported having ever received a SARS-CoV-2 vaccination dose OR who have ever received or reported having ever received a full SARS-CoV-2 vaccination course.
Numerator	Patients who have ever received or reported having ever received a SARS-CoV-2 vaccination dose OR who have ever received or reported having ever received a full SARS-CoV-2 vaccination course.
Denominator	All patients aged 18 years and older seen for a visit during the measurement period.
Exclusions	Exclusion: Patient received hospice services any time during the measurement period. Exceptions: 1. SARS-CoV-2 vaccine dose was not administered, as documented by clinician, due to patient contraindication. 2. SARS-CoV-2 vaccine dose was not administered, as documented by clinician, due to patient refusal. 3. SARS-CoV-2 vaccine dose was not administered, as documented by clinician, due to vaccine being unavailable.
Measure type	Process
What is the NQF status of the measure?	N/A

NQF ID number	N/A
Year of next anticipated NQF CDP endorsement review	N/A
Year of most recent NQF Consensus Development Process (CDP) endorsement	N/A
Is the measure being submitted exactly as endorsed by NQF?	N/A
If not exactly as endorsed, describe the nature of the differences	N/A
What data sources are used for the measure?	Registry
If EHR or Administrative Claims or Chart-Abstracted Data, description of parts related to these sources.	N/A
At what level of analysis was the measure tested?	N/A
In which setting was this measure tested?	N/A
What NQF priority applies to this measure?	N/A

What one primary meaningful measure area applies to this measure?	Preventative Care
What secondary meaningful measure area applies to this measure?	N/A
What one primary healthcare priority applies to this measure?	N/A
What secondary healthcare priority applies to this measure?	N/A
What area of specialty best fits the measure?	Not Listed
What is the target population of the measure?	Not Listed
Is this measure an eCQM?	Not Listed
If eCQM, enter Measure Authoring Tool (MAT) number	N/A
If eCQM, does the measure have a Health Quality Measures Format (HQMF) specification?	N/A
Comments	Not Listed
Measure steward	Centers for Medicare & Medicaid Services

Long-Term Measure Steward (if different)	N/A
Measure Steward Contact Information	Not Listed
Primary Submitter Contact Information	Not Listed
Long-Term Measure Steward Contact Information	Not Listed
Secondary Submitter Contact Information	N/A
Was this measure proposed for a previous year's MUC list?	No
In what prior year(s) was this measure proposed?	N/A
What were the programs that NQF MAP reviewed the measure for in each year?	N/A
Why was the measure not recommended in those year(s)?	N/A
What were the MUC IDs for the measure in each year?	N/A
NQF MAP report page number	N/A

being referenced for each year	
What was the NQF MAP recommendation in each year?	N/A
List the NQF MAP workgroup(s) in each year	N/A
What is the history or background for including this measure on the new MUC list?	N/A
Range of years(s) this measure has been used by CMS Program(s)	N/A
What other federal programs are currently using this measure?	N/A
Evidence that the measure can be operationalized	Not Listed
How is the measure expected to be reported to the program?	Not Listed
Is this measure similar to and/or competing with measure(s) already in a program?	Not Listed
Which existing measure(s) is your measure similar to and/or	Not Listed

competing with?	
How will this measure be distinguished from other similar and/or competing measures?	Not Listed
Rationale for how this measure will add to the CMS program	As of November 15, 2020, the Centers for Disease Control and Prevention (CDC) reported 10,846,373 cases of Coronavirus Disease 2019 (COVID-19) and 244,810 deaths. A vaccine for SARS-CoV-2, the virus that causes COVID-19, will be critically important to stemming the morbidity and mortality caused by this disease. While a SARS-CoV-2 vaccine has not yet been approved by the U.S. Food and Drug Administration, there are a large number of trials underway seeking to find viable vaccines. The Centers for Medicare & Medicaid Services would like to have a measure ready for implementation assessing SARS-CoV-2 vaccination at the earliest time possible, which would be for reporting period (measurement year) 2022. This measure builds off other vaccination measures in the MIPS program as much as possible. Given constraints, the measure likely will not be tested until after it has been proposed for implementation. This measure has not yet been tested and as there is not yet an approved vaccine, there is no available data to describe performance gaps. Other vaccination measures that are already in MIPS indicate performance gaps. According to 2020 benchmarks, the average performance for the CQM version of Preventive Care and Screening: Influenza Immunization was 58.5 percent. The average performance for the CQM version of Pneumococcal Vaccination Status for Older Adults was 61 percent.
If this measure is being proposed to meet a statutory requirement, please list the corresponding statute.	N/A
Evidence of performance gap	Not Listed
Unintended consequences	Not Listed
Which clinical guideline(s)?	Not Listed
Briefly describe the peer reviewed	N/A

evidence justifying this measure

Preliminary Analysis – MUC ID: MUC20-0045 SARS-CoV-2 Vaccination by Clinicians

Criteria	Yes/No	Justification and Notes
Does the measure address a critical quality objective not currently adequately addressed by the measures in the program set?	Yes	This is a new measure that has not been review by a MAP Workgroup or used in a CMS program. SARS-CoV-2 vaccination is a national healthcare priority. There are currently four measures in MIPS related to immunizations including influenza, pneumococcal, childhood immunization status, adolescent immunization status.
Is the measure evidence-based and either strongly linked to outcomes or an outcome measure?	No	This is a process measure. Vaccines to prevent SARS-CoV-2 infection are considered the most promising approach to addressing the current pandemic (Jeyanathan et al., 2020). The Centers for Disease Control (CDC) notes that 8 out of 10 COVID-19 deaths reported in the US have been in adults 65 years old and older. Clinical trials of COVID-19 vaccines must first show they are safe and effective before any vaccine can be FDA approved for emergency use (CDC, 2020). Early evidence for the effectiveness of the vaccines suggest they may be more than 90% effective in the prevention of transmission of the SARS-CoV-2 (Mahase, 2020). Early evidence submitted to FDA for emergency use authorization is promising, but the full range of evidence necessary is still emerging.

Does the measure address a quality challenge?	Yes	This measure covers a topic not currently addressed in MIPS. It will be among a set of the first quality measures to address prevention of COVID-19. In late November 2020, the Johns Hopkins Coronovirus Resource Center reported almost 12.6 million COVID-19 cases with almost 260,000 deaths in the United States. Both numbers were increasing rapidly. At the time of drafting this preliminary analysis (November 2020), no SARS-CoV-2 vaccines have been approved by the Federal Drug Administration (FDA). Performance on the measure is therefore essentially zero, maximizing the performance gap. Developer notes that other immunization measures in MIPS have significant performance gaps with influenza and pneumococcal vaccination rates at 58.5% and 61% respectively.
Does the measure contribute to efficient use of measurement resources and/or support alignment of measurement across programs?	Yes	This measure provides important information not currently available for this setting or level of analysis.
Can the measure can be feasibly reported?	Unclear	This measure has not been specified sufficiently to clearly indicate feasibility.
Is the measure applicable to and appropriately specified for the program's intended care setting(s), level(s) of analysis, and population(s)?	Unclear	Specifications are incomplete pending approved vaccines and vaccination protocols, but what is available is applicable and appropriately specified.

If the measure is in current use, have negative unintended issues to the patient been identified? Have implementation challenges outweighing the benefits of the measure have been identified?	N/A	This is a new measure that is not currently in use.
PAC/LTC Core Concept?		N/A
Impact Act Domain		N/A
Hospice High Priority Areas		N/A

Rural Workgroup Input	[redacted]	Relative priority/utility: • The context of the measure was suggested to be appropriate for rural providers and fitting in accordance to existing vaccine measures. Data collection issues: • None identified Calculation issues: • None identified Unintended consequences: • Potentially higher rates of vaccine hesitancy in rural settings were noted, although patient refusal is included as an exclusion. • Concerns expressed over vaccine availability. • Noted that provider groups can't require vaccination and that FDA's EUA has limitations; members expressed that they would feel more comfortable with full FDA approval. Votes: Range is 1 – 5, where higher is more relevant to rural. Average: 4.0 • 1 – 0 votes • 2 – 1 vote • 3 – 1 vote • 4 – 12 votes • 5 – 3 votes
Preliminary Analysis Recommendation	Do not support with potential for mitigation	The mitigation points for this measure prior to implementation are that the evidence should be well documented, the measure specifications should be finalized, followed by testing and NQF endorsement. The proposed measure represents a promising effort to advance measurement for an evolving national pandemic. The incomplete specifications require immediate mitigation and further development should continue.
Summary: What is the potential value to the program measure set?	[redacted]	This measure would add value to the program measure set by providing visibility into an important intervention to limit COVID-19 infections.

Summary: What is the potential impact of this measure on quality of care for patients?	[redacted]	Collecting information on SARS-CoV-2 vaccination coverage and providing feedback to clinicians will facilitate benchmarking and quality improvement. Vaccination overage will reduce transmission and associated mortality and morbidity. Prior to use in MIPS, this important measure should have the supporting evidence well-documented, be fully developed, followed by testing and receipt of NQF endorsement.

Measure Comments

Author	Submitted Comment
University of Colorado Medicine	Important measure, initially the results for this measure will depend on the vaccine distribution processes at the federal and state level.
Pfizer	This population-based measure adds value as vaccination will reduce the spread of Covid-19, lowering the positivity rate and protecting the most vulnerable individuals. Historically, minority populations are less likely to be vaccinated, and risk adjustment of this measure may be considered for comparability across clinicians. However, as the CDC has recognized, long-standing systemic health and social inequities have put people from racial and ethnic minorities at greater risk of getting sick and dying from COVID-19. Thus, ensuring minority populations have access to COVID-19 vaccines is critical -- not only for these patient populations, but also to successfully combatting the pandemic in the U.S. generally. To further improve outcomes, CMS should consider adding a separate measure of vaccination among minority populations.
American Medical Association	The AMA supports the inclusion of this measure but strongly encourages CMS to harmonize the numerator of this measure with MUC20-0044 and MUC20-0048. We request CMS clarify why the measure fails to capture information on the second dose of the vaccination. Specifically, the numerator does not currently capture whether a patient received a full course; rather, it could include individuals who only received one dose of a two-dose regimen. As a result, rates could be very misleading and misrepresent the true vaccination rate for this virus. We also encourage CMS to expand upon the list of available exclusions to capture information, such as vaccine refusal and allergic reactions. While it is most important to ensure that as many people get two doses as possible, it is also important to know how many don't and the reasons for patients not receiving the second dose. We also encourage CMS to revise and/or update the measure as new evidence comes forward and based on feedback received from the field. Therefore, if CMS moves forward with the measure in MIPS physician should only receive credit for reporting on the measure, not held to a benchmark given the vaccine is new and still learning about implementation and challenges with administering the vaccine.
Association of American	The AAMC is concerned as to why this measure differs from that of the other CoV-2 vaccination measure in that it measures receipt of a single dose of the vaccination,

Medical Colleges (AAMC)	rather than the full course. Is there a benefit to measuring a single dose rather than the full course of the vaccination? Additionally, the AAMC has broader concerns that match with that of the other CoV-2 vaccination measure on the MUC list regarding the premature nature of such a quality measure, which is in early stages of development. (1) The vaccination is still undergoing Phase 3 clinical trials and thus such a quality measure is premature without the full scope of clinical understanding of the vaccination and potential appropriate exclusions (2) Similarly, since the vaccination is currently authorized only under FDA Emergency Use Authorization (EUA) and not full approval, it is concerning to use it as a measure of quality versus CMS using other data reporting mechanisms to gain information on vaccination rates of clinicians. (3) Measure has not been submitted for NQF endorsement. The AAMC recommends that the highest level of MAP recommendation be "Do Not Support With Potential For Mitigation."
American College of Surgeons	On behalf of the over 80,000 members of the American College of Surgeons (ACS), we appreciate the opportunity to submit comments to the Measure Applications Partnership (MAP). The ACS is a scientific and education association of surgeons founded in 1913 to improve the quality of care for the surgical patient by setting high standards for surgical education and practice. ACS has a vested interest in CMS' MAP and the CMS Measures Under Consideration (MUC) list because of our dedication to improving the value of care for surgical patients. With our 100-year history in developing quality programs to optimize the delivery of surgical services, we believe that we can offer valuable insight to the MAPs deliberations. This comment refers to measures MUC20-0044 and MUC20-0045. The COVID-19 Pandemic has had a major impact on many aspects of healthcare and healthcare delivery, and many issues remain unresolved, such as the delays in surgery and other critical care. We suggest that the MAP also consider this as they review the measures for the upcoming years of CMS programs as our communities begin to recover from the adverse effects caused by the pandemic. In addition to the COVID-19 related measures on the MUC list, the healthcare system must establish a means for assessing priorities in restoring care. We should be measuring how we are evaluating restoration of prevention and early detection of critical conditions (such as cancer), and fully resuming surgical procedures. The ACS believes that we need to give physicians and health systems mechanisms to help them assess the priorities of their community and determine the gaps in patient care that are directly or indirectly tied to the Pandemic.
Federation of American Hospitals	The Federation of American Hospitals (FAH) supports the inclusion of this measure in the Merit-based Incentive Payment System but believes that it must be harmonized with MUC20-0044 and MUC20-0048. Specifically, this measure as currently written would not necessarily capture the rate of individuals who received a full course of vaccination. Rather, it only asks if a dose was received. Specifically, in instances where a patient received only one dose of a vaccine that requires two doses to be most effective, it would meet the numerator requirements. As a result, there is the potential for performance scores to provide misleading information on SARS-CoV-2

	vaccination rates. In addition, CMS must continuously revise and update this measure in coordination with other measure developers with similar measures. These revisions should be based on emerging evidence, newly approved vaccinations, and feedback from the field to ensure that each measure reflects the most current knowledge and evidence and can be easily collected and reported. Because we anticipate that this measure will undergo substantial changes within and across reporting years, the FAH does not believe that it should be used for payment decisions nor should it be publicly reported until the underlying evidence is stable and reporting of the measure has occurred for several years.

Measure Comments (Post-Workgroup Meeting)

Author	Submitted Comment
American Heart Association/American Stroke Association	The AHA does not support this measure at this time, given the vaccine supply issues that are likely to persist for some time. It seems especially inappropriate to attribute this measure at the individual physician level since, at present, most clinicians will not have the ability to influence this if they do not have vaccine to administer. Even in large practices at academic medical centers this is handled at the system level with no physician input.
Federation of American Hospitals	Support
America's Health Insurance Plans (AHIP)	America's Health Insurance Plans (AHIP) supports the Clinician Workgroup's recommendation. While AHIP is strongly supportive of efforts to vaccinate against SARS-CoV-2, we agree this measure should complete development and testing before implementation in an accountability program as we are concerned about the potential for inaccuracies in patient-reported data on receipt of the vaccine and completion of the course given the likelihood of different dosing for vaccines by different manufacturers. AHIP recommends the measure developers consider using claims data for this measure. AHIP supports the exclusions chosen by the measure developers to address vaccine refusal and the current limited availability of vaccines.
American Medical Association	Support, under certain circumstances
Association of American Medical Colleges (AAMC)	The Clinician MAP did not support the COVID-19 vaccination process measure (MUC2020-0045), with potential for mitigation. The three areas for mitigation are that prior to implementation the evidence should be well documented, and that the measure specifications should be finalized, followed by testing and NQF endorsement. The AAMC supports the efforts to advance measurement in response to the national pandemic but does not support inclusion of a measure that has not been fully specified and is currently under development. Furthermore, the AAMC is concerned that this measure is premature when no vaccine is fully approved (beyond an emergency use authorization) by the FDA nor is widely available. The AAMC agrees with the MAP's recommendation.

CyncHealth

January 20, 2021

Re: Draft Recommendations National Quality Forum's Measure Applications Partnership (MAP). Measures under consideration: MUC20-0044 SARS-CoV-2 Vaccination Coverage among Healthcare Personnel, SARS-CoV-2 MUC20-0045 SARS-CoV-2 Vaccination by Clinicians and MUC20-0048 SARS-CoV-2 Vaccination coverage for Patients in End-Stage Renal (ESRD) Facilities.

To whom it may concern:

CyncHealth appreciates the opportunity to provide comments on the Measures Under Consideration to the National Quality Forum (NQF) for the Centers for Medicare & Medicaid Services (CMS).

The Nebraska Health Information Initiative, Inc. (NEHII) is now doing business as CyncHealth. CyncHealth is the Nebraska statewide Health Information Exchange (HIE) as well as having a regional presence in Iowa. CyncHealth has transformed into a public health utility informing and aligning clinicians, health economists, and policy makers to respond to population-level needs. As a neutral convener for health systems and clinicians, CyncHealth pioneers efforts in population health, interoperability, and exchange of health information. These efforts include the collection, aggregation, and operationalization of people, tools, and technology to facilitate the best possible health care for communities using economies of scale. As a public health utility, CyncHealth provides services beyond the health data exchange, including a Prescription Drug Monitoring Program (PDMP), Social Determinant of Health (SDOH) platform, and CMS certified Qualified Registry (QR) and Qualified Clinical Data Registry (QCDR) and Qualified Entity.

Given our public utility model, CyncHealth supports measures MUC20-0044, MUC20-0045 and MUC20-0048. We appreciate that NQF recognizes the value of coronavirus vaccinations as high priority for measure development. The tracking of COVID-19 vaccination status for healthcare personnel and patients through these measures will help slow the spread of the SARS-CoV-2 virus. Additionally, monitoring compliance with vaccination efforts will increase visibility to potential community needs and vulnerability. Tracking vaccinations for COVID-19 ensures the opportunity to make data informed decisions and policies supporting population and public health Further, as an advocate for advanced interoperability, CyncHealth recommends further promoting the use of federally funded infrastructure in HIEs to reduce healthcare costs over acquiring new profit-oriented solutions while reducing provider burden for manual entry. HIEs have long leveraged federal funding as a vehicle to promote interoperability in communities and partnerships with state and county public health departments. As such, we agree with the background and rationale that this committee has provided in the proposed measures and measure specifications.

Conclusion
CyncHealth appreciates this opportunity to comment on the measures under consideration and your consideration of our recommendations. We value

the ongoing collaboration between CMS, NQF and CyncHealth on improving healthcare quality.

If you have questions or need clarification on any of our comments, please feel free to contact CyncHealth CEO Jamie Bland at 402-506-9900.

Sincerely,

Ann Polich

Dr. Ann Polich, MD
Chief Medical Officer
CyncHealth

MUC20-0033 ACO-Level Days at Home for Patients with Complex, Chronic Conditions

Measure Information

Characteristic	Submitted Information
MUCID	MUC20-0033
Other Measure Identification Numbers	n/a
Title	ACO-Level Days at Home for Patients with Complex, Chronic Conditions
Program	Medicare Shared Savings Program
Workgroup	Clinician
In what state of development is the measure?	Fully Developed
State of Development Details	The preliminary version of the measure is complete. Testing of this version of the Days at Home measure was completed using Medicare FFS claims from calendar years 2017 and 2018 for patients of SSP ACOs, comprising 610 ACOs with 1,154,779 patients meeting the measure inclusion criteria. CORE completed descriptive analyses of the cohort and ACOs, unadjusted analyses of the outcome, bivariate analyses of candidate clinical risk variables, and calculation of the preliminary risk-adjusted measure results. CORE is currently finalizing the risk model, including consideration of potential social risk factors. CORE will finalize testing using the final outcome definition (including the nursing home adjustment and any social risk factors) by January 2021. Using a preliminary outcome definition (as described here, except without the nursing home adjustment and with a preliminary set of clinical risk factors), CORE assessed the reliability of the measure using a split half methodology, by splitting the cohort randomly in half, calculating the measure on both halves, and comparing the ACO scores between them. CORE found high agreement between the split samples (with an intra-class correlation coefficient of 0.828 in mortality-adjusted days at home), indicating the measure is highly consistent and not greatly sensitive to chance variations in the underlying data. CORE has not completed empirical testing of validity. CORE engaged in a detailed conversation with the Technical Expert Panel (TEP) to discuss the measure concept, key benefits, and

	potential unintended consequences. After presenting the TEP with final specifications and testing results, CORE will systematically solicit the TEP's input on the face validity of Days at Home as a measure of ACO quality.
Measure Description	This is a measure of days at home or in community settings (that is, not in unplanned acute or emergent care settings) for patients with complex, chronic conditions in Shared Savings Program (SSP) Accountable Care Organizations (ACOs). The measure includes risk adjustment for differences in patient mix across ACOs, with an adjustment based patients' risk of death. A policy-based nursing home adjustment that accounts for patients' risk of transitioning to a long-term nursing home is also applied to incentivize community-based care.
Numerator	The measure outcome is days at home for a patient in the measure period, defined as the total number of eligible patient days minus the number of days spent in specified acute care settings (that is, a "day at home" is any day alive and not in care). The specified care settings are: inpatient acute and post-acute facilities, comprising short-term acute care hospitals, critical access hospitals (CAHs), inpatient rehabilitation facilities (IRFs), inpatient psychiatric facilities (IPFs), long-term care hospitals (LTCHs), and skilled nursing facilities; emergency department (ED) visits; and observation stays. Any day on which a patient is admitted to one of these settings is a "day in care", except for obstetric admissions or if the patient is enrolled in hospice (during which a patient will be considered "at home" regardless of care use). Other types of care settings (including outpatient visits and procedures, hospice, residential psychiatric and substance abuse facilities, assisted living facilities and group homes, and home health and telehealth services) are not considered "days in care" for the purpose of this measure; rather they are treated as "days at home." To ensure ACOs are not incentivized to withhold medically necessary care, the Days at Home measure accounts for higher-than-expected mortality rates, by adjusting days at home by the standardized mortality ratio. The numerator does not count days spent in long term (residential) nursing homes, as dates of these services are not reliably captured in Medicare claims. In response to CMS's policy-based recommendation that days in a nursing home should not be considered "days at home," the measure scores are adjusted based on how frequently patients transition from living at home to a residential nursing home during the performance year, such that ACOs with fewer transitions than expected receive better scores. Notably, the measure only considers transitions to nursing homes during the performance

year, which may have been affected by an ACO's performance; patients already living in a nursing home at the start of the performance year are considered to be at home.

The numerator will be calculated based on three risk-adjusted statistical models. First, "excess days in care" for each patient are modeled using a hierarchical negative binomial regression with an offset for days alive. "Excess days in care" is defined as predicted minus expected days in care, where "predicted" includes clinical risk adjustment, survival offset, and an ACO-specific effect, and "expected" includes only clinical risk adjustment and survival offset.

Second, mortality is modeled using a hierarchical logistic regression model with adjustment for the patient case-mix, to calculate a standardized mortality rate (SMR) at the patient level. A high SMR indicates a patient at greater-than-expected risk of death due to their ACO's performance.

Third, a patient's risk of transitioning to a residential nursing home is modeled using a hierarchical logistic regression model with adjustment for patient case-mix and Medicaid dual-eligibility status, to calculate a standardized "nursing home ratio" (NHR) and then scaled to have the same mean and standard deviation as the SMR. A higher NHR indicates a patient at greater-than-expected risk of transitioning to a nursing home due to their ACO's performance.

For the mortality adjustment for each patient, "excess days in care" is multiplied by SMR (if excess days >= 0) or divided by SMR (if excess days < 0), such that a greater SMR results in an absolute increase of "excess days in care" (that is, ACOs are rewarded for lower mortality than expected and penalized for greater mortality than expected). Similarly, for the policy-based nursing home adjustment for each patient, "excess days in care" is multiplied by [0.5*NHR] (if excess days < 0) or divided by [0.5*NHR] (if excess days >= 0) so that ACOs are rewarded for lower rates of transition to the nursing home than expected.

The SMR and NHR adjustments are combined additively to give a "mortality- and nursing home transition risk-adjusted excess days in care," which is subtracted from the patient-level national average of days alive, resulting in a risk-, mortality, and nursing home transition-adjusted measure of "days at home."

Finally, the adjusted days at home are averaged over all patients of each ACO to summarize the ACO's measure

	performance as the “ACO-level adjusted days at home.”
Denominator	The denominator includes patients meeting all of the following criteria: Adult (age 18 or older); Medicare Fee-for-Service beneficiary continuously enrolled in Medicare parts A and B during the full performance year (up to date of death among patients who died) and one full year prior; With an average Hierarchical Condition Category (HCC) composite risk score >= 2.0 in the pre-performance year; and Attributed to (that is, a patient of) a participating ACO as determined by SSP. The measure includes patients alive as of the first day of the performance year. Patients who die during the performance period are included up to date of death.
Exclusions	There are currently no denominator exclusions or exceptions for the measure. All patients meeting the denominator inclusion criteria are included. There are two numerator exclusions from the measure outcome. As noted, all admissions to select care settings are considered “days in care” unless: The patient is enrolled in hospice at the time of service (rationale: to promote effective and appropriate care for terminally ill patients), or The patient is admitted for childbirth, miscarriage, or termination of pregnancy (rationale: these obstetric admissions do not indicate care quality and counting them may create perverse incentives for care of pregnant patients).
Measure type	Outcome
What is the NQF status of the measure?	Never submitted
NQF ID number	0000
Year of next anticipated NQF CDP endorsement review	NA
Year of most recent NQF Consensus Development Process (CDP) endorsement	NA

Is the measure being submitted exactly as endorsed by NQF?	NA
If not exactly as endorsed, describe the nature of the differences	n/a
What data sources are used for the measure?	Chronic condition data warehouse (CCW); Claims
If EHR or Administrative Claims or Chart-Abstracted Data, description of parts related to these sources.	All measure components are taken from Medicare FFS claims-based sources
At what level of analysis was the measure tested?	Group Other: Accountable Care Organization
In which setting was this measure tested?	Ambulatory surgery center; Ambulatory/office-based care; Behavioral health clinic or inpatient psychiatric facility; Community hospitals; Dialysis facility; Emergency department; Federally qualified health center (FQHC); Hospital outpatient department (HOD); Home health; Hospice; Hospital inpatient acute care facility; Inpatient rehabilitation facility; Long-term care hospital; Nursing home; PPS-exempt cancer hospital; Skilled nursing facility
What NQS priority applies to this measure?	NA
What one primary meaningful measure area applies to this measure?	Management of chronic conditions
What secondary meaningful	Preventive care

measure area applies to this measure?	
What one primary healthcare priority applies to this measure?	Promote effective prevention and treatment of chronic disease
What secondary healthcare priority applies to this measure?	Promote effective prevention and treatment of chronic disease
What area of specialty best fits the measure?	Primary Care
What is the target population of the measure?	Medicare FFS patients enrolled in Parts A and B who are attributable to participating Accountable Care Organizations (ACOs)
Is this measure an eCQM?	No
If eCQM, enter Measure Authoring Tool (MAT) number	No
If eCQM, does the measure have a Health Quality Measures Format (HQMF) specification?	No
Comments	Row 41 (Level of Analysis): Other – Accountable Care Organization
Measure steward	Centers for Medicare & Medicaid Services
Long-Term Measure Steward (if different)	NA

Measure Steward Contact Information	Winder-Wells, Teresa; CMS Contracting Officer's Representative; 410.786.4102; Teresa.winder-wells@cms.hhs.gov
Primary Submitter Contact Information	Bernheim, Susannah; Yale New Haven Health Services Corp. / Center for Outcome Research & Evaluation (Yale/CORE); 203.764.5700; susannah.bernheim@yale.edu
Long-Term Measure Steward Contact Information	n/a
Secondary Submitter Contact Information	n/a
Was this measure proposed for a previous year's MUC list?	No
In what prior year(s) was this measure proposed?	NA
What were the programs that NQF MAP reviewed the measure for in each year?	n/a
Why was the measure not recommended in those year(s)?	n/a
What were the MUC IDs for the measure in each year?	n/a
NQF MAP report page number being referenced for	n/a

each year	
What was the NQF MAP recommendation in each year?	n/a
List the NQF MAP workgroup(s) in each year	n/a
What is the history or background for including this measure on the new MUC list?	New measure never reviewed by MAP Workgroup or used in a CMS program
Range of years(s) this measure has been used by CMS Program(s)	n/a
What other federal programs are currently using this measure?	
Evidence that the measure can be operationalized	The measure uses Medicare FFS parts A and B billing claims submitted by hospitals, physicians, and other providers. Attribution of patients to ACOs will follow the existing SSP attribution methodology. No additional data submission or extraction is necessary.
How is the measure expected to be reported to the program?	Claims
Is this measure similar to and/or competing with measure(s) already in a program?	Yes

Which existing measure(s) is your measure similar to and/or competing with?	Excess days in acute care (EDAC) after hospitalization for heart failure (NQF #2880) EDAC after hospitalization for acute myocardial infarction (NQF #2881) EDAC after hospitalization for pneumonia (NQF #2882) ACO Risk-Standardized Acute Admission Rates for Patients With Multiple Chronic Conditions (MCC) (NQF #2888)
How will this measure be distinguished from other similar and/or competing measures?	The EDAC and ACO MCC measures are similar, but not competing, to the Days at Home measure. The EDAC measures apply to a broader cohort of patients aged 65 and older who are enrolled in Medicare FFS. The EDAC outcome limits measurement of excess days in care within 30 days of an admission for short-term acute care hospital settings only (ED visits, observation stays, and unplanned readmissions), and quality of care is attributed to hospitals. The ACO MCC measure applies to a cohort of patients aged 65 or older with multiple chronic conditions (defined as having at two or more of a possible nine chronic conditions) and attributed to ACOs. The measure outcome only captures unplanned admissions to short-term acute care hospitals, not the full range of care utilization as in the Days at Home measure. Moreover, while the MCC measure only counts a dichotomous admitted/not admitted outcome, the Days at Home measure captures additional information about the length of stay.
Rationale for how this measure will add to the CMS program	The primary goal of the Days at Home measure is to promote high-quality coordinated care to keep adults with complex, chronic conditions in home or community settings and out of acute care or long-term care settings. The measure expands the EDAC and ACO MCC measure priorities of discouraging the use of preventable unplanned hospital visits. The Days at Home measure assesses patients' care use across a wide range of settings to incentivize ACOs to improve care coordination, reduce use of acute and post-acute care, and assist older and/or sicker patients in remaining at home. The Days at Home measure expands on the MCC measure by considering the total days spent in care (rather than just total admissions) and additionally accounts for ACO mortality to mitigate potential unintended consequences.
If this measure is being proposed to meet a statutory requirement, please list the corresponding statute.	n/a

Evidence of performance gap	Several studies demonstrate that time spent at home differs substantially among older patients [1, 2], which suggests that there is potential for improving the quality of care and resulting days at home for the elderly population. While the majority of patients spend all or most days at home, one study noted that patients aged 65 and older with multiple chronic conditions spend fewer days at home, with patients having three or more chronic conditions spending an average of 12.3 fewer days at home (mean 336.6 days, SD 3.0) in a one-year period than do all patients ages 65 and older (mean 348.9 days, SD 1.7) [1]. There is some evidence that patients' days at home are related to quality of care at the ACO level. In our preliminary adjusted outcome analysis, (that is, with the mortality adjustment and preliminary risk variables), CORE estimated the mean risk- and mortality-adjusted days at home for patients of each SSP ACO in 2018. CORE found that for the average ACO, the mean adjusted days at home was 330.78 days, with an interquartile range from 328.98 to 332.16 and a total range from 292.24 to 339.59. Notably, because of the use of risk adjustment in both the Days in Care and Mortality component models, this distribution does not simply reflect differences in case mix between ACOs but instead suggests that some variation is due to different ACO quality. References: 1. Burke LG, Orav EJ, Zheng J, Jha AK. Healthy Days at home: A novel population-based outcome measure. Healthcare (Amsterdam, Netherlands). 2019:100378. 2. Wallace, L., et al. (2019). 2019 Condition-Specific Excess Days in Acute Care Measures Updates and Specifications Report. Yale New Haven Health Services Corporation – Center for Outcomes Research & Evaluation (YNHHSC/CORE).
Unintended consequences	CORE is currently considering possible unintended consequences in discussion with CMMI experts, the TEP, and other experts. A potential unintended consequenceof the measure is providers withholding medically appropriate care, particularly from patients near the end of life, in order to reduce attributable healthcare use. To mitigate this potential unintended consequence, we consider all days in hospice care as days at home (even if there is some use of select care settings), and adjust the measure score for patient mortality at each ACO.
Which clinical guideline(s)?	n/a
Briefly describe the peer reviewed evidence	Generally, patients prefer to remain at home and avoid unnecessary hospitalizations and time in institutional settings [1]. Days at home are associated with other important outcomes, including social activity and depression [1]. Timely

justifying this measure	and appropriate primary care and end-of-life care services can increase the number of days patients spend at home [2]. Improved care coordination and care transitions prevent unplanned hospital visits, leading to more days at home and high-quality timely care [3]. High-risk patients not receiving care coordination from a nurse transition guide were more likely to be readmitted than high-risk patient who did, after adjusting for age, gender, and number of comorbidities (odds ratio 1.27, 95% CI 1.12-1.44) [4]. Dewilde et al. found an inpatient and outpatient cost savings of $99 for each additional day at home during the first three months after stroke [5]. A literature review conducted by McCaffrey et al. (2018) of end-of-life care patient studies concluded that home-based palliative care reduced costs per patient and acute care utilization, while increasing patient quality of life and time at home [6]. References: 1. Lee H, Shi SM, Kim DH. Home Time as a Patient-Centered Outcome in Administrative Claims Data. Journal of the American Geriatrics Society. 2019;67(2):347-351. 2. Totten AM, White-Chu EF, Wasson N, et al. Home-Based Primary Care Interventions. Rockville MD2016. 3. Harrison JD, Auerbach AD, Quinn K, Kynoch E, Mourad M. Assessing the impact of nurse post-discharge telephone calls on 30-day hospital readmission rates. Journal of general internal medicine. 2014;29(11):1519-1525. 4. Hoyer EH, Brotman DJ, Apfel A, et al. Improving Outcomes After Hospitalization: A Prospective Observational Multicenter Evaluation of Care Coordination Strategies for Reducing 30-Day Readmissions to Maryland Hospitals. Journal of general internal medicine. 2018;33(5):621-627. 5. Dewilde S, Annemans L, Peeters A, et al. The relationship between Home-time, quality of life and costs after ischemic stroke: the impact of the need for mobility aids, home and car modifications on Home-time. Disability and rehabilitation. 2018:1-7. 6. McCaffrey N, Agar M, Harlum J, Karnon J, Currow D, Eckermann S. Is home-based palliative care cost-effective? An economic evaluation of the Palliative Care Extended Packages at Home (PEACH) pilot. BMJ supportive & palliative care. 2013;3(4):431-435.

Preliminary Analysis – MUC ID: MUC20-0033 ACO-Level Days at Home for Patients with Complex, Chronic Conditions

Criteria	Yes/No	Justification and Notes
Does the measure address a critical quality objective not currently adequately addressed by the measures in the program set?	Yes	MUC20-0033 address the Meaningful Measures areas of Management of Chronic Conditions and Preventive care, and the healthcare priority of Promote Effective Prevention and Treatment of Chronic Disease. The measure aims to promote high-quality coordinated care to keep adults with complex, chronic conditions in home or community settings and out of acute care or long-term care settings. SSP currently has outcomes measures for admissions and readmissions, but none specifically addressing number of days not in acute care settings.
Is the measure evidence-based and either strongly linked to outcomes or an outcome measure?	Yes	MUC20-0033 is an outcome measure. Remaining in the home is generally perferred by patients and associated with other important outcomes, including increased social activity and reduced depression (Lee, et al., 2019). Timely and appropriate primary care and end-of-life care services can increase the number of days patients spend at home (Totten, et al., 2016). Improved care coordination and care transitions prevent unplanned hospital visits, leading to more days at home and high-quality timely care (Harrison, et al., 2014; Hoyer, et al., 2018). Other evidence presented by the developer demonstrated cost savings associated with days at home.
Does the measure address a quality challenge?	Yes	Developer indicates that time spent at home differs substantially between older patients, which suggests that there is potential for improving the quality of care and resulting days at home for the elderly population. Patients having three or more chronic conditions spend an average of 12.3 fewer days at home (mean 336.6 days, SD 3.0) in a one-year period than do all patients ages 65 and older (mean 348.9 days, SD 1.7) (Burke, et al., 2019). Developer found that for the average ACO, the mean adjusted days at home was 330.3 for MUC20-0033 (standard deviation 3.6 days, range 47.4 days).

Does the measure contribute to efficient use of measurement resources and/or support alignment of measurement across programs?	Yes	This measure is not duplicative of other measures currently within the SSP program.
Can the measure can be feasibly reported?	Yes	The measure uses Medicare claims data which can be feasibly be reported.
Is the measure applicable to and appropriately specified for the program's intended care setting(s), level(s) of analysis, and population(s)?	Yes	The measure is specified for the ACO level, which aligns with SSP reporting categories. The measure has not been reviewed for endorsement by an NQF Standing Committee.
If the measure is in current use, have negative unintended issues to the patient been identified? Have implementation challenges outweighing the benefits of the measure have been identified?	N/A	The measure is not in current use. The developer-identified potential unintended consequences include providers withholding medically appropriate care, particularly from patients near the end of life. To mitigate this potential unintended consequence, the measure considers all days in hospice care as days home (even if there is some use of select care settings) and is adjusted for higher ACO patient mortality rates.
PAC/LTC Core Concept?		N/A
Impact Act Domain		N/A
Hospice High Priority Areas		N/A

Rural Workgroup Input		Relative priority/utility: • There was a comment that rural providers could perform well on this measure. • Another comment was shared that rural providers may not have the necessary means in place to provide sufficient home health services, which would render this measure as challenging in the rural environment. Data collection issues: • None identified Calculation issues: • None identified Unintended consequences: • None identified Program gap areas: • None identified Votes: Range is 1 – 5, where higher is more relevant to rural. Average: 3.4 • 1 – 1 vote • 2 – 1 vote • 3 – 5 votes • 4 – 9 votes • 5 – 0 votes
Preliminary Analysis Recommendation	Conditional Support for Rulemaking	Conditional support for rulemaking is contingent on NQF endorsement.
Summary: What is the potential value to the program measure set?		MUC20-0033 address the Meaningful Measures areas of Management of chronic conditions and Preventive care, and the healthcare priority of Promote Effective Prevention and Treatment of Chronic Disease. The measure aims to promote high-quality coordinated care to keep adults with complex, chronic conditions in home or community settings and out of acute care or long-term care settings.

Summary: What is the potential impact of this measure on quality of care for patients?	[redacted]	Remaining in the home is generally perferred by patients and associated with other important outcomes, including social activity and reduced depression. One study indicated that patients having three or more chronic conditions spend an average of 12.3 fewer days at home (mean 336.6 days, SD 3.0) in a one-year period than do all patients ages 65 and older. Timely and appropriate primary care and end-of-life care services can increase the number of days patients spend at home. Improved care coordination and care transitions can prevent unplanned hospital visits, leading to more days at home and high-quality timely care. Conditional support for rulemaking is contingent on NQF endorsement following testing completion and further refinement of measure specifications.

Measure Comments

Author	Submitted Comment
University of Colorado Medicine	Concerned that patients who end up in nursing homes after hospitalization, with or without a SNF stay in between, and usually from a medical condition in which there is a loss of activities of daily living (such as a new stroke) would count against us in this measure.
C-TAC	We appreciate the concept of this measure and agree that most people want to be at home as much as possible. One concern we have is that days at home may not be possible for all patients with chronic conditions, some of whom may benefit from the added care of nursing facilities. Perhaps it would be better, and methodologically easier, to also track “days out of the hospital”, since wherever they are, most patients don’t want to be hospitalized. Additionally, we would encourage NQF and CMS to consider accounting not solely for raw “days at home” but “healthy” days at home, so as to help ensure that any days spent out of institutional care settings were indeed what the patient wanted and contributed to their self-assessed quality of life and quality of health.
American Medical Association	The American Medical Association (AMA) believes that understanding the degree to which individuals spend their time at home is a useful indicator to determine if the health care system is achieving one of its primary goals–to have an individual healthy at home. While this indicator provides a broader viewpoint on the health of an individual, rather than measures such as admissions or readmissions, the many different factors that can affect a patient’s “healthy days at home” raises serious concerns about whether differences in performance on this measure can be reliably attributed to the services delivered by ACOs and whether this measure could be used to truly distinguish the quality of care ACO participants receive. While the initial information provided on reliability and performance variation of the measure is useful, the state of development details indicates that the risk model development is still underway and validity testing has not been completed. Additional information is needed prior to recommending this measure for inclusion

	in MSSP. For example, we are concerned that the preliminary performance scores confirm our concerns that the measure will not produce sufficient variation to enable anyone to distinguish high versus low performers. The recent work by the Harvard School of Public Health and the Medicare Payment Advisory Commission (MedPAC) highlights statistical issues that will likely be encountered when testing is completed on this measure (Burke, 2019). Specifically, the analysis found that the difference between the minimum and maximum days at home was less than 11 days for Medicare beneficiaries over the age of 65, and for beneficiaries with 3 or more chronic conditions, the differences were only between 12 and 14 days across 306 markets. When the range of geographic markets were compared to the national mean, it was a difference of 5.8 days in the worst performing markets, and 5.0 days in the best performing markets across all Medicare beneficiaries aged 65 and older. Those with three or more chronic conditions showed more variation; the range of days was only 9.1 below the national mean and 7.9 above. Based on the sample used, good reliability of at least 0.7 required at least 2,000 beneficiaries and an analysis of market socioeconomic status (SES) characteristics identified several factors that were significantly associated with this measure including income, poverty, and physician and primary care physician density. While the MedPAC concluded that the study yielded results that could provide meaningful information to compare performance across populations and guide care planning, this does not mean that it would be appropriate to use the measure to penalize ACOs. We caution CMS that this measure must be attributed at a level where the outcome can be meaningfully influenced, is closely linked to structures and processes that are actionable by ACOs, feasible to implement without unnecessary burden, and demonstrably reliable and valid with appropriate risk adjustment, including social risk factors. In addition, simply adding this measure to the existing set would duplicate what is already measured through admissions and readmissions, and what is already encouraged by the overall financial incentives in MSSP. As a result, we recommend that the highest level of MAP recommendation be “Do Not Support.” Reference: Burke, Laura & Orav, E. & Zheng, Jie & Jha, Ashish. (2019). Healthy Days at home: A novel population-based outcome measure. Healthcare. 8. 100378. 10.1016/j.hjdsi.2019.100378.
Premier	Premier does not support adoption of this measure as it is not a true outcomes-based measure and there are other means that already address the measure's intended outcomes. Since ACOs are held accountable for both cost and quality of care, the program is inherently designed to ensure patients receive care in the most appropriate setting. Differences in this measure could be driven by a variety of factors that are unrelated to quality and cost of care, including the ACO's structure and care redesign plans and/or available resources in the community. As a result, an ACO should not be held accountable for patient days at home when it is providing high quality care and meeting its ACO benchmark. It is also unclear why the measure adjusts risk based on mortality rather than by HCCs.

Association of American Medical Colleges (AAMC)	CMS recently finalized a change to MSSP ACO quality reporting policies to align it with other MIPS APMs under the new APM Performance Pathway (APP). Part of the rational for this change was comparability of quality across APMs, with a benefit of reducing the number of measures ACOs must report on under the new APP. The AAMC is unclear on the intent of the APP if CMS is also contemplating adding further MSSP-ACO specific measures on top of the APP that are not more broadly aligned with MIPS APMs. Furthermore, two of the six measures under the APP are admission-related measures, and this measure is similarly based around inpatient utilization. The financial structure of the MSSP generally incentivizes reducing unnecessary acute and emergent care utilization. Thus, this measure duplicates the incentive/penalty structures in the MSSP payment model with a complex, imperfectly risk-adjusted quality measure. Finally, the measure has not been submitted for NQF endorsement, rendering it premature for consideration for inclusion in the MSSP. The AAMC recommends that the highest level of MAP recommendation be "Do Not Support for Rulemaking."
National Association for Home Care & Hospice (NAHC)	Since 1982, the National Association for Home Care & Hospice (NAHC) has been the leading association representing the interests of hospice, home health, and home care providers across the nation, including the home caregiving staff and the patients and families they serve. Our members are providers of all sizes and types -- from small rural agencies to large national companies -- and including government-based providers, nonprofit voluntary hospices, privately-owned companies and public corporations. As such, we welcome the opportunity to comment on the CMS List of Measures Under Consideration for December 21, 2020. We are commenting on the following two measures: MUC20-0030 Hospice Index MUC20-033 ACO-Level Days at Home for Patients with Complex, Chronic Conditions MUC20-0033 ACO-Level Days at Home for Patients with Complex, Chronic Conditions This proposed measure drives ACOs toward increased home health utilization during a performance year. NAHC supports this measure as it drives accountability for care in the home setting. Numerous studies have shown that care in the home results in better outcomes most of the time. We thank you, as always, for the opportunity to submit comments on these pending measures. If you have any questions or if I can be of assistance in any way, please do not hesitate to contact me. Sincerely,

	Katie Wehri Director of Home Health and Hospice Regulatory Affairs National Association for Home Care & Hospice Katie@nahc.org
Federation of American Hospitals	The Federation of American Hospitals (FAH) cautions CMS on the potential implementation of this measure as it could be considered the inverse of many of the measures currently included within Medicare Shared Savings Program (MSSP), such as the Hospital-Wide, 30-day, All- Cause Unplanned Readmission (HWR) Rate for MIPS Eligible Clinician Groups and the newly finalized Risk Standardized, All-Cause Unplanned Admissions for Multiple Chronic Conditions for ACOs. Because this new measure would present the opposite viewpoint of the time spent in health care facilities, it could be viewed as a form of double counting. As such, the FAH encourages CMS to reassess the set of measures used for MSSP if and when this measure is ever proposed and MAP to consider this issue when they begin to make recommendations regarding the removal of measures from programs. The FAH believes that the recent work by the Medicare Payment Advisory Commission (MedPAC) and the Harvard School of Public health to explore the usefulness of a Healthy Days at Home measure underscores several of the challenges associated with this measure (Burke, 2019). For example, the time that individuals aged 65 years of age and older spent at home ranged from 343.1 to 353.9 days during a 12-month period. While the days at home were slightly lower for those with two or more chronic conditions (minimum: 334.0 and maximum: 348.7) and those with three or more chronic conditions (minimum: 327.5 and maximum: 344.5), these ranges demonstrate minimal variation across 306 markets. The analysis of the number of markets that performed better or worse than the national mean is also useful in understanding the degree to which differences in performance across ACOs could be meaningfully distinguished. Across all Medicare beneficiaries in the sample, their time at home was just under six days, and the best was 5 days greater than the national mean. Those beneficiaries with more complex health needs (3 or more chronic conditions) receiving care in the worst performing market spent 9 days less at home and just under 8 days more in the best performing markets. The researchers also found that there were several socioeconomic factors that would be significantly associated with healthy days at home including but not limited to, median income, percentage below the poverty line, physician and primary care physician density, and acute care hospital beds per 1,000 residents. These findings indicate that a similar measure applied to ACOs may not provide sufficient variation to enable assessments of which are better or worse performers and will likely require inclusion of social risk factors within any risk adjustment. The FAH notes that the risk model is still under development and the developer must provide sufficient information on the social risk factors tested and incorporated into the risk model. In addition, the submission states that validity testing is not yet completed. The FAH believes strongly that face validity alone should not be considered sufficient for this measure and the developer must provide results from empirical validity testing. This information and the question of “double counting” must be addressed along with NQF endorsement prior to implementation in MSSP.

	As a result, the FAH recommends that the highest level of MAP recommendation be "Do Not Support with Potential for Mitigation." Reference: Burke, Laura & Orav, E. & Zheng, Jie & Jha, Ashish. (2019). Healthy Days at home: A novel population-based outcome measure. Healthcare. 8. 100378. 10.1016/j.hjdsi.2019.100378.

Measure Comments (Post-Workgroup Meeting)

Author	Submitted Comment
National Association for Home Care & Hospice (NAHC)	January 20, 2021 Public Comment 2020 Measures Under Consideration Submitted via: https://share.qualityforum.org/portfolio/MeasureApplicationsPartnership/Lists/MAP%20MUC%202020%20Comment%20Period/NewForm.aspx Since 1982, the National Association for Home Care & Hospice (NAHC) has been the leading association representing the interests of hospice, home health, and home care providers across the nation, including the home caregiving staff and the patients and families they serve. Our members are providers of all sizes and types -- from small rural agencies to large national companies -- and including government-based providers, nonprofit voluntary hospices, privately-owned companies and public corporations. As such, we welcome the opportunity to comment on the CMS List of Measures Under Consideration. We are commenting on the following two measures: MUC20-0030 Hospice Index MUC20-033 ACO-Level Days at Home for Patients with Complex, Chronic Conditions MUC20-0033 ACO-Level Days at Home for Patients with Complex, Chronic Conditions This proposed measure drives ACOs toward increased home health utilization during a performance year. NAHC supports this measure as it drives accountability for care in the home setting. Numerous studies have shown that care in the home results in better outcomes most of the time.

	We thank you, as always, for the opportunity to submit comments on these pending measures. If you have any questions or if I can be of assistance in any way, please do not hesitate to contact me. Sincerely, Katie Wehri Katie Wehri Director of Home Health and Hospice Regulatory Affairs National Association for Home Care & Hospice Katie@nahc.org
American Heart Association/ American Stroke Association	The AHA supports this measure if reviewed and endorsed by NQF and if carefully monitored by CMS to detect any unintended consequences. We believe that days at home is preferable to—and more patient-centric than--other outcome measures such as 30-day readmission, which we remain concerned may have the unintended consequence of increasing mortality.
Federation of American Hospitals	The Federation of American Hospitals (FAH) continues to believe that there is the potential double counting as this measure could be considered the inverse of many of the measures currently included within Medicare Shared Savings Program (MSSP), such as the Hospital-Wide, 30-day, All- Cause Unplanned Readmission (HWR) Rate for MIPS Eligible Clinician Groups and the newly finalized Risk Standardized, All-Cause Unplanned Admissions for Multiple Chronic Conditions for ACOs. Testing must be completed as well as an evaluation of whether the measure produces sufficient variation to enable assessments of which ACOs are better or worse performers. This information along with NQF endorsement must be completed prior to implementation in MSSP.
American Medical Association	The American Medical Association (AMA) continues to have significant concerns with the potential implementation of this measure in MSSP. Testing has not been yet completed on this measure and preliminary performance scores confirm our concerns that the measure will likely not produce sufficient variation to enable anyone to distinguish high versus low performers. This measure must be attributed at a level where the outcome can be meaningfully influenced, is closely linked to structures and processes that are actionable by ACOs, feasible to implement without unnecessary burden, and demonstrably reliable and valid with appropriate risk adjustment, including social risk factors. In addition, simply adding this measure to the existing set would duplicate what is already measured through admissions and readmissions, and what is already encouraged by the overall financial incentives in MSSP. As a result, the AMA requests that the MAP recommendation be “Do not Support.”
NAACOS	For the Days at Home Measure, NAACOS supports the concept of the measure but

	has concerns with the exceptions for the measure and risk adjustment issues with the measure. We also have outstanding questions regarding how the measure will be implemented, who will be assessed on the measure and what data will be shared with ACOs (such as quarterly data) in conjunction with the measure. Given our remaining concerns and questions, we do not support addition of the measure at this time.
Association of American Medical Colleges (AAMC)	"The Clinician MAP conditionally supported the days at home measure (MUC2020-0033) for future rulemaking for the Medicare Shared Savings Program (MSSP) pending NQF endorsement. CMS recently finalized a change to MSSP ACO quality reporting policies to align it with other MIPS APMs under the new APM Performance Pathway (APP). Part of the rational for this change was comparability of quality across APMs, with a benefit of reducing the number of measures ACOs must report on under the new APP. The AAMC is unclear on the intent of the APP if CMS is also contemplating adding further MSSP-ACO specific measures on top of the APP measure set that are not more broadly aligned with MIPS APMs. Furthermore, two of the six measures under the APP are admission-related measures, and this measure is similarly based around inpatient utilization. The financial structure of the MSSP generally incentivizes reducing unnecessary acute and emergent care utilization. Thus, this measure duplicates the incentive/penalty structures in the MSSP payment model.

General Public Comments

Author	Submitted Comment
American College of Surgeons	On behalf of the over 80,000 members of the American College of Surgeons (ACS), we appreciate the opportunity to submit comments to the Measure Applications Partnership (MAP). The ACS is a scientific and education association of surgeons founded in 1913 to improve the quality of care for the surgical patient by setting high standards for surgical education and practice. ACS has a vested interest in CMS' MAP and the CMS Measures Under Consideration (MUC) list because of our dedication to improving the value of care for surgical patients. With our 100-year history in developing quality programs to optimize the delivery of surgical services, we believe that we can offer valuable insight to the MAPs deliberations. In our comments below we first recommend fundamental changes to the national quality measurement framework which we believe are critical for the transition toward patient-centered value-based care. Following our comments on the strategic direction for assessing quality, we provide specific feedback to measures on the MUC list. Introduction: A Pathway to Value To provide context to our comments, it is helpful to describe the pathway to value from the perspective of the surgical patient. Surgical Quality is a program within a care model that is delivered by a medical team and lead by a surgeon. Quality metrics should inform the patient and the team as to the outcomes of care and the opportunities for team-based improvement of care. It is difficult to manifest quality in a single set of metrics locked into payment programs which divide and separate the surgical team from the patient's goals. Thus, using payment to establish quality in the U.S healthcare system, which is largely a Fee-For-Service (FFS) Model, fails to connect the team's care and incentivize coordination of the care team around the surgical patient. The longstanding payer-based quality metrics have failed to recognize patient-centered quality programs while giving little attention to patient goals and utilizing measures that are not generally reflective of the care delivered. We believe a necessary step toward value-based care is by first building a well-structured, verifiable surgical quality program that defines value based on patient perspectives, gives the team the infrastructure needed to provide optimal care, and aligns clinicians with facilities. We believe this strategy will help optimize the need to reward excellence in care by turning insurers' attention to the major elements of creating and sustaining a functioning and effective quality program, and simultaneously move beyond the use of a few non-systematically organized performance measures that often are of comparatively limited value to patients. The ACS believes that incentivizing the quality as program in its entirety is the way to implement a quality program. With the right framework, surgical care teams will define surgical care fit for quality measurement and improvement, while patients and insurers will be able to better assess surgical value. In our comments on the quality model below, we discuss the following:

1. The nations response to the COVID-19 pandemic illustrates the need for a programmatic model for quality
2. Guiding principles for implementing quality as a program as part of a national incentive program based on ACS quality program experience, and
3. Long-term goals for the implementation of programmatic quality as a pathway to achieve patient-centric value-based care

National Response to COVID-19 Demonstrates Necessity for A New Model for Quality

As terrible a pandemic that COVID-19 is, it is an important case study to consider when considering the pathway to value. Winston Churchill is often quoted when he was working to form the United Nations after WWII, "Never let a good crisis go to waste." We should make note of several lessons learned in the pandemic when it comes to understanding quality and payment. As a planet, we knew almost nothing about the virus, how it spreads, the impact it has on humans, acute treatment and the consequences, or long-term sequela. The first order of care was to understand the medical condition and begin to formulate a care model. Resources played a major role in supporting care team needs, patients' needs, as well as clinical protection for caregivers. Data systems sprung up, and shared knowledge became the goal across the entire globe. The world turned into a massive observational data registry with every expert and scientific filter applied. Using data to gain an advantage on COVID-19 led to redesign of the care model and how best to implement the resources for optimal care. Patients and their condition were the centerpiece. To incent quality, the revenue models and payment systems have to be secondary to the care, its resources, and the data models which inform improvement.

Within all of these efforts in response to COVID, we find the model for a quality program. It begins with the patient, their condition, and their care team. The right structures and processes must be in place in order to effectively and efficiently deliver the intended outcomes with alignment at the facility and clinician levels. We learned of the role of ICU care, oxygenation protocols, prone ventilation, steroids, antiviral agents and more. All of the lessons learned developed from details in structured care, in care processes and in measured and observed outcomes. Knowledge sharing from all sources continues to inform the care team and drives its improvement cycles. Payment models were secondary, with a clinical focus on optimally meeting the patient's goals and outcomes, while minimizing avoidable harms. These are the quality program lessons learned from a pandemic. The entire world shared knowledge as to how to structure care, deliver it and learn from each outcome. COVID-19 is a lesson in quality as a program and how shared knowledge can spread across all of healthcare. We should be careful to absorb them and not brush aside these lessons as a passing fad.

These lessons are especially important because they contrast with the framework of the current CMS quality incentive programs whose measures are being considered by the National Quality Forum (NQF) MAP. For some time, CMS, NQF, and the MAP have divided quality into a few metrics scattered across payment programs with the hope that chasing individual measures would result in better care. This hypothesis is true for singular moments in a doctor patient interface but it is ill-suited for complex

surgical care. Instead of a surgical quality program, the current CMS programs measure erratic components of care discretely, measuring the individual surgeon separately from the hospital, separately from the anesthesiologist, separately from the pathologist, and so on. This is very well illustrated in the columns and rows of sporadic measures on the MUC list. The result is an overly burdensome measurement system and a fragmented picture of "quality." The current CMS framework has failed to drive improvements in surgical care because it is disconnected from the health care delivery process and sporadic in nature. The result is that clinicians end up chasing measures for payment rather than building teams for gainful improvement in outcomes of care.

The College has advocated for the approach used in ACS quality programs that is fundamentally similar to the model for a quality program we saw in the response to COIVD-19. This approach focuses overarchingly on the care of the patient, including the goals and outcomes important to the patient, while also valuing the infrastructure, resources, and processes needed to deliver optimal care and improvement.

"Quality" is a Program; Not a Measure: Guiding Principles

The model for "quality is a program" has been developed and supported by ACS for more than half a century, and below we offer guiding principles to translate a programmatic quality approach to a national incentive program.

Well known examples of ACS quality programs include the Trauma Center Verification Program, the Commission on Cancer Accreditation, and the Metabolic and Bariatric Surgery Verification program. The latest addition to the ACS library of verified quality programs now includes Geriatric Surgery Verification program, tailored to improve care for some of the most vulnerable and frail patients. Each of the ACS quality programs is built on a four-part model, known as the ACS Quality Model, that includes: 1.) program-specific standards, 2.) infrastructure needed for delivering high–quality care, 3.) data collection and its use for care delivery and improvement, and 4.) verification site visits to ensure implementation of the critical elements for optimal care. The evidence supporting this model demonstrates the concept that quality is a multi-component program involving a team of clinicians and surgeons operating in a culture of excellence, with systems engineering for efficiency, appropriateness, proper resources applied within structure and processes, as well as measures for conformance and outcomes. In order to assure quality, the ACS' experience shows that setting standards for care (both at the facility and individual clinician levels) and assuring, with rigor, that those standards are implemented is indispensable.

With the concept that quality is a program, we ask the MAP to consider rethinking some of its longstanding approaches to promoting quality. First, it is important to define surgical quality as a program and then consider how it fits within the guardrails of payment and within the statutory limits for CMS. We believe it is possible to reset the CMS and MAP compass to meet the clinical concepts of quality, but the result is much different than the historical approach of simply counting measures and tabulating for payment.

Below are guiding principles for what a Quality Incentive Program would look like for surgery:

1. Should span across most or all of surgical care to verify structure, processes, and outcomes are applied and tracked for the comprehensive patient journey and patient goals
ü The prioritization of surgical patient goals across the five phases of surgical care by the patient's team is paramount.

2. Should link clinicians, facilities to achieve patient goals (create shared accountability)
ü It is necessary that the individual provider
(e.g., surgeon/anesthesiologist/radiology/ pathology, etc.) and the facilities are aligned in their accountability. These measures which demonstrate alignment are interrelated. That is, for example, the facility has a clinical chair-led, quality committee which is data driven based on outcomes where the clinicians participate in active improvement cycles based on data work within the committee.

3. Reflect proper programmatic alignment, structure, processes, and outcomes
ü We should have alignment of these measures so that surgical team members are sufficiently equipped by hospital structure to care for patients.

4. Structure and processes for performing QI (across the surgical team)
ü The surgical team and hospital measures should not just include process and outcomes but also structure.

5. Incentives for physicians and hospitals/facilities should rely on interrelated quality measures
ü The CMS payment systems should reward measures which assess hospital structure and resources needed to care for the surgical patient AND the clinicians delivering care in the quality improvement program.

6. Incentive programs should reward those willing to make a special effort toward programmatic alignment
ü The greater the level of alignment (and performance) of hospital structures and the care provided by the surgical team members, the greater the reward.

Long-term Goal for CMS Implementation of a Quality Framework

The undergirding framework illustrates ACS' long-term goals for implementing a quality program and is based on the Donabedian Quality Model for Evaluating Care. Donabedian's structure, process, and outcomes quality model is a proven way to conceptualize quality of care. CMS, the payer community, and NQF has downplayed the significance of structure and process as too much of a check-the-box measure. ACS does not disagree when using these types of measures in isolation. When each of these are individually measures, they do become overly simplistic and lose their significance. However, when measured together as part of a verification program, the assessment relates to how structure and process yield outcomes, how failure points are noted and addressed. Decades of ACS experience in trauma, cancer and bariatrics have defined the importance of how to measure the effectiveness of the

Donabedian elements in assessing the level of quality as a program in a care delivery system. In this way, the sum of all the verified standards working together to create a quality program become a meaningful measure.

The ACS' dogma supports surgical quality should be delivered (and measured) as a full program that fundamentally operationalizes the entire Donabedian quality model. This program involves expert reviews (now virtually) by trained assessors in evaluating a care team for meeting the standards for structure and process components by defining the resources, infrastructure, and processes needed to achieve optimal quality improvement (QI). The ACS Clinical Programs set the standards for clinical care—these programs are where condition or specialty-specific standards are added (e.g. Bariatric, Trauma, Geriatrics). Layering on top of clinical accreditation are appropriate and adequate processes which further help to implement the care model. Patients should know that their care team meets or excels in the care they expect to receive.

Verified, standards-based programs drive improvement cycles which are based on reliable, valid, risk-adjusted outcomes. ACS experience shows that the best programs have structured meetings and processes which monitor the clinical outcomes with accurate, clinical, risk-adjusted data (e.g. National Surgical Quality Improvement Program (NSQIP)) measured at the hospital level, followed by outcomes reporting by the patient, or PROs, measured at the individual level. The framework presented takes concrete steps to facilitate the value transformation by incorporating Patient-reported Outcomes (PROs) for the patient perspective, it measures the team around the patient, not the individual, and it creates alignment between clinicians and facilities to keep the focus on the patient. Each component of the quality model builds on and is interrelated to the others, pulling the information to assess the essential components for a patient, allowing for patients, clinicians, and payers to assess (more completely) the quality of care. The ideal for the systematically organized set of measures is to represent the spectrum of an effective quality program by focusing and crediting each layer of this pyramid.

CMS and the NQF have limited approaches such as we have outlined due to a concern of overly burdensome and costly measurement. ACS experience demonstrates the opposite effect. By building quality as a program, the entire team engages in the effort enthusiastically because of its impact on the overall outcome. The ACS approach appeals to clinical professionalism and pride for the entire staff. Also, quality programs, once verified, are maintained by the team in each individual contributor's daily workflow and allows quality payment programs to credit the services delivered for meeting a standard for a set period.

It is critical that CMS and the MAP appreciate that this concept cannot be taken apart into individual components for implementation because it is the four-part model that has demonstrated improvements in care and fits the delivery system. Through the ACS experience in creating quality programs, we know that the optimal and most advanced clinical patient care is given by providers who routinely perform both optimal clinical processes and optimal quality evaluation/improvement processes ALL THE TIME—not just in an incentive program. This type of program culture is what should be incentivized in CMS incentive programs.

We appreciate the opportunity to comment on the important work of the MAP. The ACS looks forward to continuing dialogue on these important issues. If you have any questions about our comments, please contact Jill Sage, Quality Affairs Manager, at jsage@facs.org.

Sincerely,

David B. Hoyt, MD, FACS
Executive Director