MACRA Wave 5 Cost Measure Development Webinar
February 23, 2022

-Thank you for viewing the MACRA Wave 5 Cost Measure Development presentation. The purpose of this presentation is to provide information about the current MACRA Wave 5 cost measure development call for public comment. I will now turn it over to Dr. Ronique Evans from CMS to begin the presentation.

-Thank you. My name is Ronique Evans. I am the cost measures lead from the Division of Quality Measures at CMS. I will go through the next few slides for our introduction before handing it over to our measure development contractor. Next slide, please.

This CMS disclaimer outlines that this presentation is a general summary intended as a reference document for stakeholders. Next slide, please.

Today I will provide an introduction overview on how episode-based cost measures are part of the Merit-based Incentive Payment System, or MIPS, within the QPP program. Then we'll provide an overview of Wave 5 development and provide some information on our approach for prioritizing clinical areas and episode groups. We will then discuss the request for public comment on candidate episode groups. Finally, we will conclude with next steps and relevant resources. Next slide, please.

Here are a few acronyms we use throughout the presentation that may be useful for your reference. Next slide, please.

To begin, I will go over the background for MIPS and the Quality Payment Program. Next slide, please.

The Quality Payment Program has two tracks to choose from for clinicians -- Advanced Alternative Payment Models and the Merit-based Incentive Payment System, or MIPS. Clinicians who participate in MIPS have their performance assessed over four categories --

Quality, Cost, Interoperability and Improvement Activities. A clinician's performance score for each category is rolled up to a MIPS final score which then determines a payment adjustment. CMS has developed a new participation framework called MIPS Value Pathways, or MVPs, as a way to connect the quality, cost and improvement activities categories. The goal of the MVPs is to simplify participation in MIPS and to provide clinicians with a set of aligned measures across performance categories that are meaningful to their care practice. Starting in 2023, seven MVPs will be available. Today's presentation will focus on the cost performance category of MIPS. Next slide, please.

For the 2022 performance period, the cost performance category makes up 30 percent of the MIPS final score. The score for the cost performance category averages all applicable cost measures. The MIPS cost performance category started with two global or population-based cost measures in 2017 and 2018 and has now expanded to also include 23 episode-based cost measures to cover a wide range of conditions and procedures. One key feature of these cost measures is that they're calculated using administrative claims data. There is no additional reporting required for clinicians. Next slide, please.

This slide lists MIPS cost measures we mentioned. Of note, the population-based cost measures were revised for 2020. The episode-based cost measures that are new for calendar year 2022 are denoted with an asterisk. Next slide, please.

An episode-based cost measure represents the Medicare payment for the medical care provided to a patient related to a specific episode of care. These measures only include the cost of relevant services in the episode such as preoperative testing, direct cost of treatment, routine follow-up care or services resulting from the treatment such as complications. The measures are intended to inform clinicians on the cost of care for their patients and identify opportunities for improvement. They are based on three primary types of care, acute in-patient medical condition episode groups which focus on the treatment for an exacerbation of condition requiring hospitalization such as hospitalization for stroke or pneumonia, procedural episode group focused on procedure of a defined purpose or type such as spine fusion surgery or screening

colonoscopy, client and condition episode groups focused on treatment for an ongoing clinical condition at the time of medical visit, such as diabetes, asthma, or COPD. Next slide, please.

The 2017 episode-based cost measure have developed in waves or cycles where a number of measures follow the same development process and timeline. So far, we have completed three waves of development. We developed eight measures in Wave 1, 11 measures in Wave 2 and five measures in Wave 3. Over the years, we've worked with hundreds of clinician experts from various specialty societies and organizations for our convened panel. We anticipate finishing Wave 4 of development this year with four new measures. This includes condition measures for heart failure, low back pain and major depressive disorder, as well as a measure for emergency medicine.

During this time frame, we started working on refinements for an acute in-patient medical condition cost measure of psychoses and related conditions which was originally developed in Wave 2. We also began respecifying chronic condition cost measures for chronic kidney disease, or CKD, and end-stage renal disease, or ESRD, for future potential use in MIPS. These were originally developed for a different model and are now being respecified in MIPS. We are now seeking input via the current public commentary to start Wave 5 cost measure development. I will now turn the presentation over to Acumen, the measure development contractor.

-Thanks, Ronique. My name is Joyce Lam, and I'm the project manager from Acumen for the cost measure development project. I'll start by providing a brief overview of Wave 5 of cost measure development. Next slide, please.

Our process for measure development involves collecting a wide range of stakeholder feedback so that we can incorporate it into each stage of the process from prioritizing which measure is developed through to the detailed specifications. This feedback helps ensure that measures meet CMS' Meaningful Measures goals. The Meaningful Measures framework is an initiative which identifies the highest priorities for measurement and improvement so that we only assess the

core issues that are the most critical to providing high quality care and improving patient outcomes.

In addition, feedback from stakeholders allows for an iterative development and testing process so that we can combine stakeholder expertise and experience and environmental scans with empirical data. In Waves 1 through 3, we obtained stakeholder input by convening experts in clinical subcommittees structured around a clinical area, like cardiovascular disease management, or a measure framework, such as the chronic condition framework. The clinical subcommittees met for a 1-day in person meeting to review data, discuss different options and ultimately vote on preferred episode groups. We considered this input and worked with CMS who would then confirm which measures to develop in that particular wave.

In Wave 4, we changed this process to provide more flexibility in how stakeholders can share their input on measure prioritization given competing demands on clinician time. As such, we held a public comment period instead of convening clinical subcommittees so that clinicians, specialty societies and patients, families and caregivers could participate around their schedule. We'll continue this approach for Wave 5 as many of the remaining clinical topics require targeted input to assess measure viability to help inform prioritization. In addition, this approach continues to provide stakeholders with more flexibility to share input over a longer period of time. Next slide, please.

This slide provides a bit more detail about the development process and the role that stakeholders play at each stage. We have a standing technical expert panel, or TEP, which provides high level guidance for topics across the project. This does include discussing questions about new measure frameworks and which clinical areas should be priorities for development. This has helped shape the focus areas for Wave 5. As I mentioned, within each wave we gather input on candidate episode groups either through clinical subcommittees or public comment period using the prioritization criteria that the TEP helped to develop. Finally, we have clinician expert work groups that are composed of 15 to 20 members with expertise in the care on which the measure focuses, including clinicians across the patient-care continuum. Work group members provide detailed input on every aspect of the measure specifications.

Additionally, person and family partners, or PFPs, share their input with work group members throughout development. Next slide, please.

Here we provide an overview of the development timeline for the Wave 4 measures as referenced. We began with a public comment period at the end of 2020 and into early 2021. CMS selected the measures to develop based on the input from the comment period. We held an open nomination period for the clinician expert work groups in the spring of 2021. Once selected, the work groups met several times at virtual meetings since mid-2021 and discussed all components of their measures and their specifications. We then held a 7-week national field testing of the measures with their draft specifications which started in January 2022 and ended in February. During this period, we sent out around 270,000 reports to clinicians and clinician groups via the QPP portal so that everyone could review how they would have performed on the measures and share any feedback about the measures with us. After field testing, the clinician expert work group members will convene once more for a refinement webinar to discuss field testing feedback and any refinements to the measures to finalize the specifications. Next slide, please.

Here we have a preliminary timeline for Wave 5 cost measure development, which is anticipated to last through May 2023. The process is similar to the Wave 4 timeline from the previous slide. Next slide, please.

I'll now walk through how we prioritize clinical areas and episode groups for development in preparing for Wave 5. Next slide, please.

Our approach is something that we have developed and refined over the years. This includes incorporating feedback from the TEP, patient perspective, public comment, clinical subcommittees, clinician expert work groups about what are the most important factors to consider when deciding which measures to develop. As mentioned earlier, this prioritization process helps ensure that cost measures are aligning with CMS' Meaningful Measures initiative.

To provide some background, we first posted a draft list of episode groups and trigger codes years ago, which we often call the “December 2016 posting” for short. This was a list of 117 episode groups that clinician stakeholders had identified as contenders for development that year and which we sought public comment on.

Since then, we've used this December 2016 posting as a starting point for Waves 1 to 3. We've also received feedback expanding that list. At the time of creating that list, some types of care weren't included, such as post-acute care. In February 2020 and July 2021, we talked with our TEP about different ways of approaching measured prioritization. For example, we could focus on developing measures in clinical areas that don't currently have cost measures.

Another way of thinking about prioritization could be to build more cost measures for areas where there might already be some cost measures, but which represent high cost areas and have strong opportunities for improvement. One consideration that the TEP highlighted was identifying the specialties that don't currently have cost measures in MIPS. The TEP expressed support for developing measures for clinicians and specialties who have limited availability of episode-based measures in MIPS. Next slide, please.

In addition to looking at clinical areas which could be measurement gaps, we also have a set of criteria to help determine which of these areas to prioritize. This builds on input we've received from our TEP and the patient and family perspective over the past years. The first criterion is clinical coherence. In other words, is this an aspect of care where we can compare clinician decisions for similar patients? This could include considerations of staging, severity, and what we can distinguish from data. It's important to remember that measure components are for multiple techniques to ensure clinical comparability of patients. These include stratifying the measure into smaller, more similar cohorts and using risk adjustment to account for patient and other factors present at the start of care such as pre-existing conditions.

The second criterion is the impact of the measure and its potential importance to MIPS. This includes looking at the potential for covering a large number of clinicians, patients and costs.

The third criterion is the opportunity for cost performance improvement. Importantly, a measurement gap is not the same as a performance gap, so this criterion ensures that we're considering the ways in which a cost measure can help clinicians to make decisions for cost-effective, high quality care, such as by improving care coordination.

The fourth and final criterion is alignment with quality. These cost measures are intended for use in MIPS alongside the quality measures, so we need to consider how the measures could work holistically to assess value. Next slide, please.

Once we've prioritized measures for development, measures must also have certain essential features to ensure that they can effectively assess clinician cost performance. Similar to the prioritization criteria, we've worked with stakeholders including the TEP to define and vet these standards. These measures must consider the clinician role in care which can include factors like practice standards, what services are clinically related to the condition or procedure being assessed and whether there is sufficient variation in performance to be able to distinguish between providers. Further, it's important that measure construction is both readily understandable and able to convey information to clinicians about ways to improve performance. Finally, the measure should be specified in a way that allows the measure to be consistently calculated and reproduced using Medicare administrative claims data. Next slide, please.

The candidate clinical areas and episode groups that we'll be talking about today were identified by a broad assessment approach focused on assessing potential benefits and drawbacks. This assessment was based on prior stakeholder input and framed around the prioritization criteria and the essential features that we just talked about. As an example, we assess various candidate clinical areas and episode groups for whether there may be sufficient potential for improvement or the degree of alignment with MIPS quality measures. Through our assessment, we found various key challenges that need more investigation and importantly, your input on approaches to overcome these challenges. Next slide, please.

Now that we've covered the measure development process and the criteria for measures to be effective, this section will walk through the request for public comment on candidate clinical areas and episode groups for Wave 5. Next slide, please.

Using the factors that we've discussed, we identified eight clinical areas with priority episode groups centered on two main ideas for this wave. The first one is to capture clinicians with limited episode-based cost measures. For this category, we identified anesthesia care, diagnostic radiology procedures, oncological care, and post-acute care, or PAC, for short.

The second goal is to expand coverage of high-cost clinical areas where there are some episode-based cost measures for which we identified rheumatoid arthritis, ophthalmological conditions, kidney care, and gastrointestinal surgery. For all of these areas, we need your input on the following -- identifying opportunities for improvement, refining draft trigger codes, which can be reviewed in the preliminary specifications work group, and finding ways to align with quality measures. Of course, we also welcome additional comments about the measures such as any concerns or recommendations. Your comments are very important to helping identify the most impactful measures to develop. We're also gathering input in preparation for convening work groups. This includes identifying the specialties and types of experience that should be represented for a given work group if that measure is confirmed to be developed. If you would be interested in participating in a Wave 5 work group should a specific measure concept be selected for development, we are also collecting any expressions of interest. Next slide, please.

The first clinical area is anesthesia care. We identified this as a clinical topic where a general candidate cost measure could focus on the provision of anesthesia services. This would cover anesthesiologists and certified registered nurse anesthetists, or CRNAs. We've heard in the past from many stakeholders that this clinical area could use its own episode-based cost measure. This measure concept could have high impact and potentially align with the clinical focus of the MVP for patient safety and support of positive experiences with anesthesia. This MVP was finalized in the 2022 PFS final rule and will be available starting 2023.

There are several potential measures within this general area of anesthesia care. One is to focus on anesthesia services. This could be all types of anesthesia or just specific types of anesthesia services, such as for specific surgeries. Another option could be to focus on interventional pain management. So there are several areas that we're seeking your input on.

For an anesthesia measure, we want to hear about what is the range of complications and other follow-up services that may be reasonably influenced by the clinician providing the anesthesia services rather than the surgeon alone. We've previously heard feedback from stakeholders about some complications like airway injury from intubation, but these might be quite infrequent, so we're interested in hearing about other related services that could help distinguish good from poor care.

Second, we want to hear about whether the measure should take a narrow approach, focusing on something like anesthesia for joint replacement as an example, or a broader approach, such as all anesthesia services that could then be stratified or risk adjusted to define smaller patient cohorts. In general, starting with a broader measure is preferred because it has greater potential to be impactful, but it must also represent a type of care that can allow for meaningful clinical comparisons.

For an interventional pain management measure, we want to know about what additional services besides injections could be included in this measure. Injections can be costly, so could represent a strong opportunity for improvement, but a measure would need to include other types of services to be able to distinguish between clinician performance. We're also interested in your input on whether an interventional pain management measure ought to focus on acute pain management, chronic pain management or both. Relatedly, what are approaches we can consider to either limit to one of these or accurately distinguish these types of care management for a broader measure? Next slide, please.

The next clinical area is diagnostic radiology. For this area, we are thinking of creating an episode-based cost measure that would focus on the care provided by diagnostic radiologists. We've heard from stakeholders including in the Wave 4 public comment period that screening

mammography would be a strong candidate cost measure, noting the degree of influence that radiologists may have over this area and the availability of well-established quality metrics. Our TEP agreed noting that it would be a compelling cost measure due to its high frequency. Here is what we're seeking your input on.

First, we want to know what should be the scope of a mammography measure. For example, should it focus on undifferentiated or ambiguous cases? One of the key challenges is that since we use claims data, we aren't able to directly identify the results of a scan. However, there could be proxies in claims data that could be used to identify an appropriate scope of the measure, such as by looking at new cancer diagnoses after a scan.

We're also interested in your input on what time frame and services would reflect the reasonable influence of the radiologist. For example, what are some short and long-term types of care that a diagnostic radiologist can influence? What types of services should be included in the measure, and is there enough variation in these services that would allow the measure to distinguish between clinician cost performance?

Finally, while we've heard mostly about mammography as the best place to start when thinking about measures for diagnostic radiologists, we want to hear your thoughts on whether there are other measure concepts in this general area that should be considered. For example, would something like outpatient chest scans be something that could provide higher patient clinician coverage but still be defined to be clinically coherent? Next slide, please.

The third clinical area is oncological care. A candidate episode group in this area is prostate cancer as it's one of the most common cancer diagnoses and has multiple treatment options which means that it may have more cost variation than treatment for other types of cancer. However, this may capture the care more by a urologist than oncologist, and one of the reasons we identified this clinical area is wanting to develop a measure that could apply to oncologists. A major concern for a cancer care measure is that claims data doesn't have the coding specificity for cancer staging. Being able to identify the stage or severity of cancer would be important in a cost measure as the treatment would differ depending on this. So, we're

interested in hearing your thoughts on how to account for cancer staging and severity using Medicare claims data, using proxy approaches or algorithms. Using proxies from claims data is something that we've done for other cost measures.

For example, in MIPS, there's a cost measure which focuses on outpatient diabetes care, and it distinguishes between type 1 and type 2 diabetes by looking at four independent indicators and the degree of agreement across these indicators. Other questions we want input on are, should a cancer cost measure focus on a specific type of cancer like prostate, breast or lung, or should it focus on all cancers, stratifying by types and staging or severity? As mentioned, the benefit of a broad measure is that it can have greater impact and importance to the program, but it must still be narrow enough to be defined in a clinically coherent way. Another question is about what types of non-drug services may capture opportunities for improvement. Drugs can be very costly, so we want to see if there are other services so that the measure isn't dominated just by drug costs. Next slide, please.

The next clinical area is PAC. A cost measure in PAC represents potential impact because it includes a wide range of settings where care coordination may be improved including skilled nursing facilities, inpatient rehabilitation facilities, long-term care hospitals and home health. A measure in this area could help ensure that all clinicians providing PAC care have similar incentives to coordinate for cost effective care. This could include supporting care transitions, reducing transfers to emergency departments or hospitals, reducing pressure ulcers and preventing falls. A potential cost measure in this area might also align with cost measures for post-acute care facilities. This includes the Medicare spending for, per beneficiary, post-acute care measures used in various quality reporting programs. A measure for PAC could be constructed similarly to inpatient measures where an episode is triggered by a clinician billing certain E&M services, or evaluation and management, on Part B physician/supplier claims during the event. We're seeking input on a couple of areas.

One is how to account for the heterogeneity of patients receiving PAC. For instance, should we stratify by diagnosis groups, medical complexity, short-term versus long-term cases or a

combination of these? We're also interested in input on attribution methodologies including any differences in approach across PAC settings that may be needed. Next slide, please.

The next clinical area is rheumatoid arthritis. A measure in this area would apply to rheumatologists and primary care clinicians that manage the ongoing care for patients with this common condition. As with some of the other areas, there is a coverage gap in rheumatologists as rheumatologists in MIPS currently don't have a targeted episode-based cost measure. We considered this area during Wave 4 in the public comment period and received valuable feedback from stakeholders. We're considering this measure concept again for Wave 5 as it meets many of the prioritization criteria.

It's a common condition amongst the Medicare beneficiary population and represents variation in treatment such as drug options and therapy, efficient monitoring and imaging including for adverse effects to treatments. A measure for rheumatoid arthritis could also align with the MVP for advancing rheumatology patient care which is finalized for use starting in 2023. We're interested in hearing your thoughts on how to determine what would be an appropriate clinically coherent patient cohort that could capture opportunities for improvement and differentiate between clinician cost performance. We also want to know any suggestions for proxy approaches or algorithms such as those described just now for the cancer care measure for how we could account for severity and patient's responses to medication. Next slide, please.

Next, we have ophthalmological care. We identified a few candidate episode groups within this general area including age-related macular degeneration, or AMD, and retinal detachment. These potential measure concepts could improve the coverage of ophthalmologists as a specialty and capture some high-cost services such as injections. Ophthalmology is a large specialty with a diversity of practice across sub-specialties, and this is something that we've heard from stakeholders. MIPS does have a cataract episode-based cost measure, and we've heard that there are various subspecialties that aren't directly involved with cataract removal procedures. So we're interested in your input on both of these potential episode groups.

For AMD, we want to know which services, such as anti-VEGF, would be appropriate for identifying whether a clinician is managing a patient's AMD condition, and we also want to find approaches to avoid penalizing clinicians who treat patients who require more expensive treatment and how we could identify this from administrative claims data. For retinal detachment, we seek your input on accounting for differences across patients based on pre-existing conditions that may impact the likelihood of treatment success as this will impact the expected costs. Finally, we're interested in whether there are other strong candidates for measure development within ophthalmological care that could increase the coverage of this specialty for clinicians who don't provide cataract removal procedures. Next slide, please.

The next clinical area is kidney care. Currently, measures for chronic kidney disease, or CKD, and end-stage renal disease, ESRD, are in the process of being respecified for potential use in MIPS. Throughout this respecification process, work group members on CKD/ESRD emphasize the importance of including kidney transplant recipients as part of cost measurement to represent the full spectrum of kidney care. Kidney care is a high-cost area with strong opportunities for improvement. Specifically for kidney transplant management, a potential measure could capture costs such as the return to maintenance dialysis if the transplant fails. We're seeking input on what the best way to align a kidney transplant measure with the existing CKD/ESRG measures to make sure that they are all working together to assess the high costs of kidney care. In addition, we want to know what are potential unintended consequences of including or excluding the transplant recipient population in kidney care cost management. Next slide, please.

The final clinical area is gastrointestinal surgery. Currently, MIPS includes 15 episode-based cost measures focused on particularly high frequency and/or costly procedures, and cholecystectomy is one of the remaining procedures that could build out further coverage of specialties such as general surgery. Potential opportunities for improvement in this area include reducing lengthier stays and hospital or ED visits, mitigating complications such as bile leaks, bleeding, infection, injuries in nearby structures and risks of general anesthesia like blood clots and pneumonia and improving post-surgical instructions.

We're seeking your input on whether a measure in cholecystectomy should include bile duct surgery, laparoscopic surgery, open surgery, annual interventional radiology procedures on the bile duct. We're also interested in how emergent and non-emergent cases could be distinguished and accounted for using administrative claims data, and finally, we would like to hear your input on the types of services that proceduralists can reasonably influence in the short and long-term. Next slide, please.

The public comment posting materials are currently available on a CMS Measures Management System, or MMS, web page. The page is called CMS Currently Accepting Comments, and you can see the link here. The first document is the call for public comment. This provides more detail for what we've talked about today including the background of cost measure development, an overview of cost measure frameworks, an essential features of cost measures and the list of questions for clinical areas and candidate episode groups. The other document available for download on this page is a workbook containing the preliminary specifications for the Wave 5 candidate episode groups. This work group lists the draft and very much preliminary trigger codes that could be used to define the patient cohorts. Next slide, please.

The public comment period for Wave 5 measure prioritization opened on February 18, 2022 and is open until the 1st of April. To provide feedback, stakeholders may submit their response or upload a comment letter as a PDF or Word document in the Wave 5 measure development survey by 11:59 p.m. Eastern Time on April 1st.

Stakeholders may also express interest in participating in Wave 5 should a particular measure be selected for development in the appropriate sections of the survey or e-mail our team for any questions. The e-mail address is macra-cost-measures-info@acumenllc.com. At the end of the public comment period, we will review all of your input and work with CMS to confirm measures to develop in Wave 5. We will then get started with convening work groups to build out specifications for each of the selected measures. Next slide, please. Finally, I will wrap up this presentation by going over some next steps and resources. Next slide, please.

In terms of next steps, our team will review and compile the stakeholder feedback to make recommendations to CMS about which episode groups to develop in Wave 5. Then CMS will make the final decision on which episode groups to develop in Wave 5. We'll then work on identifying clinician expert work groups and person and family partners by reviewing the expressions of interest that we received through this public comment period and posting a call for nominations later this year.

After we compose the work groups, the work groups will meet multiple times throughout 2022 and 2023 to provide input which will inform each aspect of the measure specifications.

Stakeholders at large will have many opportunities to participate throughout this measure development process. Webinars for the work groups will have a public dial-in option so that folks can join and observe the meetings, and meeting summaries will be publicly posted on the MACRA feedback page. In addition, the measures will go through national field testing, likely in late 2022 or early 2023. During this period, stakeholders will be able to review the draft measures and the specifications and provide feedback on any aspect of these. Next slide, please.

Here we have a list of cost measure resources along with the URLs. This includes the CMS currently accepting comments page which is where you can download the public comment posting materials. You can also see the link to the Wave 5 measure development survey. Next, we have the MACRA feedback page, which contains many useful resources about episode-based cost measures in general including the development process, specifications for other cost measures, and more. You can also submit a response to the MACRA episode-based cost measures mailing list to stay apprised of any cost measure development updates in the future.

The slide deck for today's presentation is available on the Quality Payment Program webinar library. If you have any questions, please feel free to contact the Acumen MACRA cost measures support team at macra-cost-measures-info@acumenllc.com. Thank you for joining today, and we look forward to hearing your feedback during this public comment period. Next slide, please.

This concludes the recording.