

MACRA Episode-Based Cost Measures: Wave 5 Measure Development

Call for Public Comment

We are seeking feedback from stakeholders regarding which cost measures to develop for the upcoming cycle of episode-based cost measure development (“Wave 5”). This document includes the following:

Table of Contents

MACRA Episode-Based Cost Measures: Wave 5 Measure Development.....	1
1. Introduction	1
2. Approach to Gathering Stakeholder Input in Wave 5	2
3. Selecting Measure Concepts.....	2
4. Questions about Clinical Areas and Measure Concepts	4
5. General Questions across Candidate Measure Concepts and Measure Development.....	11
6. Next Steps	12
Appendix A : Overview of Measure Frameworks	13

1. Introduction

The Centers for Medicare & Medicaid Services (CMS) has contracted with Acumen, LLC to develop episode-based cost measures for potential use in the Merit-based Incentive Payment System (MIPS) to meet the requirements of the Medicare Access and CHIP Reauthorization Act of 2015 (MACRA). MACRA established the Quality Payment Program, which incentivizes clinicians to provide high-quality and high-value care through Advanced Alternative Payment Models or MIPS. Under MIPS, eligible clinicians receive an adjustment based on performance to their Medicare payments. The performance adjustment is derived from a final score that assesses evidence-based and practice-specific data in 4 performance categories: (i) Quality, (ii) Cost, (iii) Improvement Activities, and (iv) Promoting Interoperability. The MIPS cost performance category has 25 measures for 2022:

- 2 population-based or global cost measures
- 23 episode-based cost measures for a range of procedures, acute inpatient medical conditions, and chronic conditions.

The episode-based cost measures in the MIPS Calendar Year 2022 performance period were developed in cycles or “Waves” in 2017 - 2020. We anticipate finishing Wave 4 in mid-2022 and are looking ahead to Wave 5 of development.

For Wave 5, we are gathering input on episode groups to consider for development through a Call for Public Comment. Stakeholders are invited to submit their feedback in response to the information and questions included in this document by **11:59pm ET at end of the public comment period.**¹ Stakeholders may submit a response² to the **Wave 5 Measure Development Survey:**

https://acumen.qualtrics.com/jfe/form/SV_0qB0oApVFxbj8tU

¹ Please refer to the web posting for the specific dates: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/PC-Currently-Accepting-Comments>.

² Stakeholders may also email a comment letter to macra-cost-measures-info@acumenllc.com.

2. Approach to Gathering Stakeholder Input in Wave 5

In Waves 1-3, Acumen obtained input on measure prioritization by convening experts in Clinical Subcommittees (CS). These CS were structured around a clinical area or type of measure. We met with the CS to discuss and vote on preferred episode groups. The start of the COVID-19 public health emergency in 2020 presented unique challenges for the community, particularly front-line clinicians, making it unlike previous Waves. This along with stakeholder input indicating interest in having more flexible participation options for measure development led us to gather input on prioritization through a call for public comment for Wave 4.

For Wave 5, we are also adopting this approach, given the ongoing public health emergency. This will allow for more flexibility to specialty societies and professional associations to review and provide their input on priority clinical areas for Wave 5 (as well as widen the range of stakeholders who can provide feedback). Furthermore, we've received input about many clinical topics and potential measures to consider for this Wave, so we are interested in addressing remaining targeted questions on them. A public comment period will also provide stakeholders with more flexibility to engage over a longer period.

3. Selecting Measure Concepts

To identify a starting point for Wave 5 measure prioritization, we drew on feedback that we've heard over the years through our Technical Expert Panel (TEP), Person and Family Engagement (PFE), Clinical Subcommittees (CS), Clinician Expert Workgroups ("workgroups"), and public comment.³ We considered how to prioritize among measure concepts and what features are necessary to make an effective cost measure. Each are discussed below in turn.

3.1 Criteria for Measure Prioritization

The criteria to consider when assessing candidate episode groups for potential measure development includes:

- Clinical coherence of measure concept to ensure valid comparisons across clinicians
- Impact and importance to MIPS, including cost coverage, clinician coverage, and patient coverage
- Opportunity for performance improvement
- Alignment with quality measures to ensure meaningful assessments of value⁴

3.2 Essential Features of Cost Measures

To ensure that a cost measure is effective in assessing clinician cost performance, we've worked with stakeholders to define and vet standards for essential measure features. This includes discussions over the past years with a TEP and clinician panels convened around areas of care and specific measures (i.e., CS and workgroups).

The standards that guide this development work are as follows:

- Attribution of measures to clinicians is clear and accurately captures their role.
- Episode definitions have clinical face validity and consistency with practice standards.
- Construction of episodes/measures is readily understandable to providers.

³ More information on the TEP, PFE, CS, Clinician Expert Workgroups, and public comment is described in the [2022 Episode-Based Cost Measures Field Testing Wave 4 Measure Development Process document](https://www.cms.gov/files/document/wave-4-measure-development-process-macra.pdf), <https://www.cms.gov/files/document/wave-4-measure-development-process-macra.pdf>.

⁴ Regarding this criterion, we considered the shift towards MIPS Value Pathways (MVPs), as these are an approach that combine related cost and quality measures into sets of measures and activities that are more meaningful to a clinician's practice, specialty, or public health priority. More information on MVPs can be found here: <https://qpp.cms.gov/mips/mips-value-pathways>.

- Providers are held accountable for costs of assigned services they can reasonably influence.
- Measures can convey concrete guidance indicating how providers can alter practice to improve measured performance.
- Variation in measures helps distinguish performance across individual providers.
- Measure specifications allow for consistent calculation and reproducibility using Medicare data.

3.3 Approach for Identifying Candidate Clinical Areas

We reviewed feedback that stakeholders have shared over the years regarding candidate clinical areas and episode groups. In 2016, we started by gathering input on the draft list of episode groups and trigger codes through a call for comment posting (“December 2016 posting”)⁵ and have used that feedback as a starting point for selecting episode groups and determining preliminary specifications for Waves 1 – 3. In recent years, we’ve heard interest from stakeholders in expanding beyond the list of episode groups from the December 2016 posting, as it does not include some types of care. We revisited this list of episode groups to consider for development, which includes reviewing our past CS and TEP prioritization discussions to identify strong candidate episode groups. At the February 2020 and the July 2021 TEP meetings, we discussed coverage gaps, how to prioritize filling in gaps in cost measures, and ideas on how to overcome challenges that we know exist with particular gaps.⁶ We also gathered stakeholder input on measure prioritization during the Wave 4 public comment period.⁷

Our approach for identifying candidate clinical areas involved assessing the benefits and drawbacks of various candidate episode groups we’ve gathered from prior stakeholder input. This assessment of measure viability was based on the criteria for measure prioritization described in Section 3.1, as well as each episode group’s ability to meet the essential features of cost measures described in Section 3.2. We gathered measure concepts we’ve heard about from our TEP, CS, PFE, and public comment. For each candidate episode group, we assessed them on clinical coherence, measure importance (impact), opportunity for improvement, and alignment with MIPS quality measures. For example, we explored whether a certain procedure or condition is likely to have sufficient variation across clinicians in cost performance (opportunity for improvement), and we evaluated each candidate episode group’s potential for alignment with the current set of MIPS quality measures (and a cost measure’s potential to be incorporated into a relevant MVP). We also conducted research to identify key challenges that require further investigation and stakeholder input, which informs the questions presented in Section 4. Based on this review, we identified the clinical areas and candidate episode groups described in Section 4 below.

⁵ CMS, “Draft List of MACRA Episode Groups and Trigger Codes”, MACRA Feedback Page (December 2016), <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/Draft-list-of-episode-groups-and-trigger-codes-December-2016.zip>.

⁶ CMS, “Physician Cost Measures and Patient Relationship Codes (PCMP) Technical Expert Panel Summary Report: July 20, 2021.” (September 2021) <https://www.cms.gov/files/document/physician-cost-measures-and-patient-relationship-codes-pcmp-technical-expert-panel-summary-report.pdf>.

⁷ The summary report for the Wave 4 Measure Development Public Comment Period is here: <https://www.cms.gov/files/document/wave-4-public-comment-summary.pdf>.

4. Questions about Clinical Areas and Measure Concepts

Based on the considerations outlined above and in discussion with CMS, we identified 8 candidate measure concepts for potential development in Wave 5. Their clinical areas represent 2 main ideas. Section 4.1 contains topics that would apply to clinicians who currently do not have or only have limited episode-based cost measures, although they may have global cost measures. These topics could increase the range of specialties who have their costs assessed by a measure tailored to the types of care that they provide. Section 4.2 contains high-cost clinical areas that could increase the share of clinicians within a particular specialty who have an episode-based cost measure, which could help build out the depth of coverage, have high impact on saving costs, and lead to a more holistic assessment of clinician cost performance.

We are seeking input on these clinical topics so that we can better inform our recommendation to CMS for which 4 measures to develop in Wave 5. For example, certain clinical topics have specific challenges that would need to be addressed, so we are particularly interested in feedback on these so that we can assess the likely impact, scientific acceptability, and feasibility of a measure in that area. Each clinical area has specific questions, but stakeholders can comment on any aspect of the measure concept. Stakeholders can also consider the cross-cutting questions in Section 5.1, which apply to all the clinical topics in this section.

4.1 Clinical Topics for Clinicians with Limited Episode-Based Cost Measures

The clinical areas in this group represent opportunities to capture types of care that would apply to clinicians who currently have limited or no coverage by episode-based measures. Each of these areas have particular challenges regarding the measure construction that we've heard from stakeholders in previous years. As such, we are especially interested in your feedback and ideas on how to address these challenges to create effective measures.

4.1.1 Anesthesia Care

We identified **Anesthesia Care** as a clinical topic that could provide anesthesiologists and Certified Registered Nurse Anesthetists (CRNAs) with an episode-based cost measure. A candidate measure within this general area could focus on the provision of anesthesia services. An episode could be triggered by Current Procedural Terminology/Healthcare Common Procedure Coding System (CPT/HCPCS) codes for all or specific types of anesthesia; these codes would also identify the attributed clinician who bills these services (more on procedural episode groups in the [Appendix](#)). This measure concept could have high impact as there is large potential beneficiary coverage. Also, this could potentially align with the clinical focus of the MVP for Patient Safety and Support of Positive Experiences with Anesthesia (finalized in the CY 2022 PFS final rule to be available starting 2023). However, we seek input on the types of care that would be within the scope of anesthesia care provided by an anesthesiologist or CRNA to ensure that there is sufficient variation and opportunities for improvement.

- **Question 1:** Previous stakeholder feedback has identified some anesthesia-related complications such as airway injury from intubation, untreated hypothermia, and nerve injury for a peripheral block. Since these may be infrequent, are there other services for complications or other follow-up care that could differentiate good care from poor care? That is, if a cost measure is centered on anesthesia services for a type of surgery, what sort of complications and other follow-up services may be reasonably influenced by the clinician providing the anesthesia services rather than the surgeon alone?
- **Question 2:** Should we develop a broad anesthesia measure for all types of procedures, or would it be better to develop something narrower (e.g., anesthesia for joint replacement)? If a narrower measure is preferred, what scope of services would help capture anesthesia care services provided by anesthesiologists and CRNA broadly? If a

broad measure is preferred, would sub-grouping⁸ by procedure type be useful? What categorization of procedure type would be clinically coherent for a broad anesthesia measure?

Another candidate measure within this area could focus on interventional pain management. A key area for improvement with this measure concept would be to reduce any unnecessary use of costly injections; however, this would have much smaller beneficiary coverage and apply to a smaller subset of clinicians. This could mean that a substantial share of anesthesiologists and CRNAs would still not have an episode-based cost measure. In addition, the measure would need to have sufficient variation in costs beyond injections.

- **Question 3:** What other related services, besides injections, could be included in an interventional pain management measure? For example, if injections are not successful at managing pain, what would a clinician focusing on pain management care provide as the next line of treatment? What sorts of services would a patient with poorly managed pain receive that would be different in frequency or intensity than a patient with well-managed pain?
- **Question 4:** Should a measure on interventional pain management focus on acute pain management (e.g., local anesthetics such as facet injections), chronic pain management (e.g., local pain intervention such as treatments for tendonitis or carpal tunnel), or both? Using claims data, what approaches could we consider to help identify chronic versus acute interventional pain management?

4.1.2 Diagnostic Radiology Procedures: Screening Mammography

Diagnostic Radiology Procedures is a clinical area that could create an episode-based cost measure for diagnostic radiologists. Within this general topic, we have heard from stakeholders through the Wave 4 public comment period that screening mammography would be the strongest candidate measure, citing the degree of influence radiologists may have over this area as well as well-established quality metrics. Our TEP agreed, noting that it would be a compelling cost measure due to its high frequency. An episode could be triggered by CPT/HCPCS codes for mammography procedures and attributed to the clinician billing these services. However, we are seeking input on how to identify and account for differences in the patient care trajectory depending on the findings of the scan. We are also interested in your feedback on what types of services should be included so that the measure captures opportunities for improvement and differentiates between clinician performances.

- **Question 1:** What should be the scope of a mammography measure, given the expected differences in cost depending on the result of the scan? For example, stakeholders have suggested that the measure should focus on undifferentiated cases. Stakeholders have suggested a measure that includes screening mammography through cancer diagnosis or return to annual screening, as that is under the diagnostic radiologist's control. Since cost measures use administrative claims data, would newly occurring cancer diagnoses after a scan be an appropriate proxy to identify these cases?

⁸ Sub-grouping is stratifying the episode group into mutually exclusive and exhaustive sub-groups to define more homogeneous patient cohorts.

The share of ambiguous results from screening mammography is expected to be quite low (e.g., recall rates are estimated at around 12%⁹). Are there clinically sound ways to define a broader patient cohort to increase the potential coverage and impact while still accounting for expected differences in cost? Is focusing on undifferentiated cases appropriate to capture the type of care with the greatest opportunity for cost improvement?

- **Question 2:** Currently, the MIPS cost measures span episode windows from 14 days to 1 year or more. What are some suitable timeframes for capturing radiologists' overall effects of their work on mammography? What short- and long-term outcomes are influenced by diagnostic radiologists (e.g., achievement of diagnosis, appropriate use of additional studies to achieve diagnosis, etc.) that could be included as services in the cost measure? Is there sufficient variation in the occurrence, frequency, or intensity of these services to be able to distinguish good care from poor care by the radiologist?

We also considered other measure concepts within this clinical area. Lung cancer screening and pulmonary nodule screening are much less frequent than screening mammography. Also, almost all radiologists interpret chest x-rays or chest computerized tomography (CT) scans, whereas screening mammography is generally performed by a subset of radiologists. A broader measure concept could be for outpatient chest scans; this would likely have high patient and clinician coverage but face challenges with defining a clinically coherent patient cohort.

- **Question 3:** Should these other types of screening (e.g., outpatient chest scans) be considered as candidates for development within this clinical area? If so, what features make them more compelling or viable than mammography?

4.1.3 *Oncological Care: Cancer*

The **Oncological Care** clinical area would focus on creating a measure that would capture the care provided by oncologic specialties. **Prostate cancer** is one of the most common cancer diagnoses and has multiple treatment options, which means it may have more cost variation than other types of cancer. An episode could be triggered by a pair of services billed by the same clinician group that indicate the start of a care relationship to treat prostate cancer, such as outpatient evaluation and management (E&M) services or chemotherapy when paired with prostate cancer diagnosis codes (more on chronic condition episode groups in the [Appendix](#)). We've heard that a major limitation to developing any cost measure to assess the costs of cancer care is the lack of coding specificity for cancer staging.

- **Question 1:** Other cost measures have used algorithms as proxies to identify conditions of interest to account for differences in expected costs. For example, the Diabetes cost measure that was added to MIPS in 2022 stratifies patients into sub-groups for Type 1 and Type 2 diabetes based on 4 independent indicators (e.g., share of Type 1 or Type 2 diagnosis codes over a year-long period), and the degree of agreement across these tests.¹⁰

⁹ Brian L. Sprague, et al. "Assessment Of Radiologist Performance In Breast Cancer Screening Using Digital Breast Tomosynthesis Vs Digital Mammography". JAMA Network Open 3, no. 3 (2020): e201759. doi:10.1001/jamanetworkopen.2020.1759.

¹⁰ For more details on the sub-grouping methodology for the Diabetes cost measure, see Appendix B of the Diabetes Measure Information Form (MIF) Methodology Appendices document ("2021-12-13-mif-ebcm-diabetes-appendices.pdf") within the [2022 MIPS Measure Information Forms](https://qpp-cm-prod-content.s3.amazonaws.com/uploads/1729/2022-cost-measure-information-forms.zip) ZIP file (<https://qpp-cm-prod-content.s3.amazonaws.com/uploads/1729/2022-cost-measure-information-forms.zip>).

How should this measure account for differences in costs due to cancer severity using administrative claims data?

Some examples of claims-based identifiers for severity include:

- International Classification of Diseases, 10th Revision (ICD-10) diagnosis codes with or without secondary malignancy site codes, cancer as a primary diagnosis, first use of cancer diagnosis following a designated lookback period, the number of medical claims on separate dates in combination with identified ICD-10 diagnosis codes¹¹
- A combination of healthcare claims data, such as ICD-10 diagnosis codes and hospitalizations or multiple outpatient office visits¹²

As part of testing the algorithm, we would compare any classification system against other indicators in claims data. To do this, we could check whether patients identified through the indicators above are receiving services that would be expected for that stage or level of severity. What are services that are typically distinct for different stages of prostate cancer that could be used for this kind of validation check?

Since this clinical topic was identified to fill in episode-based cost measure gaps for oncologic specialties, one consideration for prostate cancer is that urologists frequently treat this, so the coverage of oncologists may be limited. As such, we considered other types of cancer (e.g., **breast cancer, lung cancer**), or a **broader set of cancer care** that could be part of one measure, stratified by cancer type. The advantage of a broader measure is that a larger share of oncologic specialties could be covered by the episode-based measure.

- **Question 2:** Are other types of cancer preferable for measure development, such as breast or lung cancer? Should we consider a broad cancer measure that stratifies patients by type of cancer and stage, and if so, what would that measure need to account for to ensure clinically meaningful comparisons?
- **Question 3:** Given that drug costs dominate costs of care across types of cancer, what are other opportunities for cost improvement? That is, what types of services are clinically related to the treatment and management of cancer that could distinguish variation in care?

4.1.4 Post-Acute Care (PAC)

PAC is a high-cost area of care that includes Skilled Nursing Facilities (SNF), Inpatient Rehabilitation Facilities (IRF), Home Health (HH), and Long Term Care Hospitals (LTCH). An episode-based cost measure focusing on clinicians providing care in these settings would have the potential to have large impact on cost improvement. A PAC measure could also help ensure that all providers and clinicians have similar incentives to coordinate for cost-effective care: for example, a PAC measure for clinicians could align with the Medicare Spending Per Beneficiary – Post-Acute Care (MSPB-PAC) measures that are attributed to PAC providers.¹³ Opportunities to improve cost effectiveness include supporting care transitions, reducing transfers to emergency departments and hospitals, and reducing pressure ulcers and falls. A measure for clinicians providing PAC care could be constructed similarly to inpatient measures where an

¹¹ Joanna L. Whyte et al. "An evaluation of algorithms for identifying metastatic breast, lung, or colorectal cancer in administrative claims data." *Medical care* 53, no. 7 (2015): e49-e57.

¹² Alyson L. Mahar et al., "Validating an algorithm to identify metastatic gastric cancer in the absence of routinely collected TNM staging data." *BMC health services research* 18, no. 1 (2018): 1-7.

¹³ CMS, "Measure Specifications: Medicare Spending Per Beneficiary – Post-Acute Care Resource Use Measures (July 2016)," https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/LTCH-Quality-Reporting/Downloads/2016_07_20_mspb_pac_ltch_irf_snf_measure_specs.pdf.

episode is triggered by a clinician billing certain E&M services on Part B Physician/Supplier claims during an inpatient hospitalization.

- **Question 1:** Patients use PAC for many different reasons and with varying levels of care needs. How should the measure account for the heterogeneity of patients in PAC? We can use techniques like risk adjusting or sub-grouping¹⁴ by PAC setting, or based on services that are indicative of a given group of diagnoses or conditions (e.g., distinguishing patients with complex medical conditions, dementia, or those who receive rehabilitation). Are there certain types of services, diagnoses, or other data available via claims that may be useful in defining more homogeneous patient cohorts (e.g., separating the short term post-acute residents from the longer term ones)?
- **Question 2:** Given the similarities in facility-based care, would an attribution methodology similar to that used by the MSPB Clinician measure and acute inpatient medical condition episode-based cost measures (more on acute inpatient medical condition episode groups in the [Appendix](#)) appropriately capture the role of clinicians providing PAC services? Would any modifications be needed for different PAC settings, such as home health? If so, what approach would be clinically sound in terms of identifying responsible clinicians for those PAC settings?

4.2 High-Cost Clinical Areas for Clinicians with Some Episode-Based Cost Measure Coverage

This section contains clinical topics that could have particularly strong opportunities for cost improvement due to the high costs involved. These measure concepts would likely apply to specialties that already have some coverage from episode-based cost measures. That is, these measures would help capture more clinicians within given specialties for certain types of care.

4.2.1 Rheumatoid Arthritis

The **Rheumatoid Arthritis** measure would apply to rheumatologists and primary care clinicians that manage the ongoing care for patients with rheumatoid arthritis, with the former currently having limited or no episode-based cost measures. This measure concept was included as part of the Wave 4 public comment posting and stakeholders provided valuable feedback on measure construction. Ultimately, CMS prioritized other areas for Wave 4 after weighing all the criteria. We are considering this measure concept again for Wave 5, as it meets many of the prioritization criteria. It is a common condition among the Medicare population and represents opportunity for improvement (e.g., variation in treatment/drug options and efficient monitoring/imaging/therapy, including for adverse effects to treatments). An episode-based cost measure for rheumatoid arthritis could also align with the MVP for Advancing Rheumatology Patient Care, finalized for use starting in 2023. However, there are challenges around identifying the patient cohort and accounting for certain costs.

- **Question 1:** Stakeholders have suggested focusing on newly diagnosed rheumatoid arthritis patients. Since this would result in lower beneficiary and cost coverage, is there a way to define a broader (yet still clinically coherent) patient cohort that could represent a viable measure? For example, are there opportunities for improvement in later stages of the disease?
- **Question 2:** Using claims data, how should the measure account for differences in costs due to rheumatoid arthritis severity or patients' responses to medication? Some example

¹⁴ Sub-grouping is stratifying the episode group into mutually exclusive and exhaustive sub-groups to define more homogeneous patient cohorts.

approaches include: linking severity to prescription/dialogic use,¹⁵ using the claims-based index of rheumatoid arthritis severity (CIRAS),^{16,17} using the presence of extra-articular manifestations (e.g., pulmonary, ocular), and looking for the presence of other comorbidities or services (e.g., coronary artery disease, lymphoma, lung disease, vasculitis, and side effects from medications).

4.2.2 *Ophthalmologic Conditions*

We identified Ophthalmologic Conditions as a clinical area that could increase the share of ophthalmologists who have an episode-based cost measure, as the Routine Cataract Removal with Intraocular Lens (IOL) Implantation measure only applies to clinicians performing this specific procedure. We've heard feedback about the diversity of practice across ophthalmology sub-specialties. As such, we identified 2 potential measure concepts that may improve specialty coverage in this clinical area: the chronic management of **Age-related Macular Degeneration (AMD)** and the **Retinal Detachment** procedure. AMD could have substantial variation in cost (e.g., injections), representing opportunities for improvement in cost performance.

- **Question 1:** For AMD, what is the most appropriate patient population (e.g., all AMD versus only wet AMD) and episode window for this condition?
- **Question 2:** Would anti-vascular endothelial growth factor (VEGF) intervention be appropriate for identifying whether a clinician is managing a patient's AMD chronic condition? What other interventions may indicate that a clinician is managing a patient's AMD?
- **Question 3:** Can we infer AMD clinical outcomes from claims data? How can a measure avoid penalizing clinicians who treat patients for whom the more expensive injection is the only clinical option or more frequent injections are needed (e.g., requiring a higher case volume per clinician)?
- **Question 4:** For retinal detachment, how should the measure account for differences across patients based on pre-existing conditions that may impact the likelihood of treatment success? How can the measure be constructed to be as broad as possible for measure impact and viability (e.g., risk adjusting or sub-grouping certain patient cohorts rather than excluding)?
- **Question 5:** Besides AMD and retinal detachment, are there other concepts in this clinical area (e.g., glaucoma care) that would be strong candidates for development considering the prioritization criteria and essential features of cost measures? The goal would be to capture the care provided by different types of ophthalmologists for which there is sufficient opportunity for improvement.

4.2.3 *Kidney Care: Kidney Transplant Management*

The **Kidney Transplant Management** measure would provide a cohesive set of kidney care measures alongside the Chronic Kidney Disease (CKD) and End-Stage Renal Disease (ESRD) measures. These 2 measures are currently being re-specified for use in MIPS.¹⁸ Stakeholder input on the CKD/ESRD measures emphasized the importance of including kidney transplant

¹⁵ Urmila Chandran et al., "Inferring Disease Severity in Rheumatoid Arthritis Using Predictive Modeling in Administrative Claims Databases", PLOS ONE 14, no. 12 (2019): e0226255, doi:10.1371/journal.pone.0226255.

¹⁶ CIRAS includes tests for inflammatory markers, number of chemistry panels/platelet counts ordered, rheumatoid factor test, number of rehabilitation and rheumatology visits, and Felty's syndrome. CIRAS correlated moderately well with a previously validated RA medical records-based index of severity.

¹⁷ Gladys Ting et al., "Development Of A Health Care Utilization Data-Based Index For Rheumatoid Arthritis Severity: A Preliminary Study", Arthritis Research & Therapy 10, no. 4 (2008): R95, doi: 10.1186/ar2482.

¹⁸ CMS, "CKD/ESRD Workgroup Meeting Summary," <https://www.cms.gov/files/document/summary-ckdesrd-workgroup-webinar.pdf>.

recipients as part of a cost measure to capture the full spectrum of kidney care. Furthermore, kidney care is a high-cost area with strong opportunities for improvement; this potential measure would capture costs such as return to maintenance dialysis if the transplant fails. An episode would be triggered by a pair of services billed by the same clinician group that indicate the start of a care relationship for kidney transplant management, such as outpatient E&M services (more on chronic condition episode groups in the [Appendix](#) and the draft trigger codes are included in the Preliminary Specifications of Wave 5 Candidate Episode Groups workbook).

- **Question 1:** The CKD/ESRD measures currently under development target the ongoing outpatient management of these conditions,¹⁹ with an emphasis on comprehensive assessment. The CKD/ESRD measures do not include kidney transplant recipients. How could a transplant measure best be developed to align with the CKD/ESRD measures and jointly assess the high costs of kidney care?
- **Question 2:** Stakeholder input received to date has emphasized the importance of assessing the costs for kidney transplant recipients. Other kidney payment models, such as the ESRD Treatment Choices Alternative Payment Model, have emphasized the importance of assessing transplant-related care and reducing disparities among Medicare beneficiaries with kidney disease.²⁰ Are there potential unintended consequences of including or excluding the transplant recipient population in kidney cost measures? How can the kidney care measures promote high-value care and health equity?

4.2.4 Gastrointestinal Surgery: Cholecystectomy

There are currently 15 episode-based cost measures in MIPS that focus 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 surgeons. An episode would be triggered by applicable CPT/HCPCS codes (more on procedural episode groups in the [Appendix](#)). Cost improvement opportunities include reducing length of stay and hospital readmissions or emergency room visits, mitigating complications (e.g., bile leaks, bleeding, infection, injury to nearby structures), and improving post-surgical instructions.

- **Question 1:** Is this procedure an area with sufficient opportunity for improvement? How broadly could the patient cohort be defined for cholecystectomy? For example, should trigger codes include bile duct surgery or laparoscopic and open surgery? Are there any interventional radiology procedures on the bile duct that should be included? Are these interventions comparable to one another? What types of features (that can be defined from claims data) influence the method of intervention?
- **Question 2:** Should we create different populations of cholecystectomy based on whether the procedure is emergent or planned? If so, what are ways we can identify these populations from claims data? Are there other important distinctions to make in the larger population of gallbladder disease?
- **Question 3:** What types of services and costs can be influenced by proceduralists performing cholecystectomies in the short and long-term?

¹⁹ More detailed information on attribution for chronic condition measures, as well as other elements of the framework, are included in the Measure Information Form for Diabetes within the [2022 Cost Measure Information Forms](#) ZIP file: <https://www.cms.gov/files/zip/2022-cost-measure-information-forms-macra.zip>.

²⁰ Center for Medicare & Medicaid Innovation, “ESRD Treatment Choices (ETC) Model,” <https://innovation.cms.gov/innovation-models/esrd-treatment-choices-model>.

5. General Questions across Candidate Measure Concepts and Measure Development

This section contains questions that apply across all the potential measure concepts in Section 4. Section 5.1 provides cross-cutting questions about characteristics of measures that would be important to consider in determining which measures to develop. Section 5.2 includes questions about gathering input for the development process, should a given measure concept be selected for development.

5.1 Cross-Cutting Questions for All Wave 5 Candidate Episode Groups

We considered measure importance and the opportunity for improvement in various care settings. We also considered the essential features of cost measures described above, focusing on clinical coherence. We welcome comment on the candidate episode groups regarding the following topics:

- **Opportunity for improvement.** What kinds of services can reflect that the candidate episode group has sufficient opportunities for improvement? For example, cost measures generally include services reflecting variation in treatment options, intensity/duration, follow-up care, complications, and more.
- **Trigger codes:** Trigger codes define the patient cohort for the measure. The preliminary set of draft trigger codes we propose is in the accompanying Preliminary Specifications of Wave 5 Candidate Episode Groups workbook. We solicit comment on this list of draft trigger codes to help inform the patient cohort. What modifications can we apply to these draft trigger codes to ensure a measure represents a clinically coherent patient cohort and also sufficient impact and coverage?
- **Quality alignment for assessing value:** We solicit comments regarding alignment of quality of care with cost measures as well as comments on any indicators of quality that would be valuable to assess alongside the cost performance for the candidate episode groups.
- **Additional concerns:** Are there any other concerns that may be present with assessing the care of patients in any of the clinical areas listed above? If so, what are some potential approaches to address these concerns for a cost measure?

5.2 Participating in Wave 5 Development

We are interested in gathering input on the development of the candidate episode groups, should they be selected for development. While we will only develop 4 measures during Wave 5, feedback on these items may be considered for a future Wave.

1. **Wave 5 Workgroup Composition.** What types of specialties or particular types of experience ought to be represented in the Clinician Expert Workgroup for the candidate episode groups? Some examples for types of experience include: “direct treatment of condition in outpatient setting,” “performing the procedure,” “post-procedure rehabilitation care,” or “medical coding knowledge relevant to the episode group.”
2. **Are you interested in participating in Wave 5?** Please indicate in the online survey if you or your organization has experience in one of the clinical areas listed in Section 4 and are interested in finding out more about participation in a potential Wave 5 workgroup when we confirm the measures we’ll be developing.

6. Next Steps

Please share your feedback by submitting a response to the online survey by **11:59 p.m. Eastern Time** on the last day of the public comment period:²¹

https://acumen.qualtrics.com/jfe/form/SV_0qB0oApVFxbj8tU

We will consider stakeholder feedback in our recommendations to CMS to finalize the set of episode groups to develop in Wave 5. The decision on the Wave 5 cost measures will be made by CMS, which may be informed based on the input provided by stakeholders from this public comment posting. In addition, we will prepare a public comment summary report that will be posted on the MACRA Feedback Page²² later this year.

Once the final set of measures are approved, we will hold a public Call for Nominations to convene a workgroup for each measure. The workgroups will be composed of approximately 15 members with experience and expertise in the condition, procedure, or type of care being assessed by the measure. The workgroups will meet multiple times over an 18-month period to provide input on measure specification. We will contact stakeholders who expressed interest in participating in development for a specific measure to encourage them to submit a nomination form. When the workgroups convene later this year, they'll be able to review data informed by public comments from this posting.

²¹ Please refer to the web posting for the specific dates: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/PC-Currently-Accepting-Comments>.

²² CMS, MACRA Feedback Page, <https://www.cms.gov/Medicare/Quality-Payment-Program/Quality-Payment-Program/Give-Feedback>.

Appendix A: Overview of Measure Frameworks

This appendix provides a brief overview of the 3 main types of episode groups: (i) procedural episode groups are of a defined purpose or type and can be performed in different settings depending on the specific measure's intended focus, (ii) acute inpatient medical condition episode groups require a hospital stay and can represent treatment for a self-limited acute illness or treatment for a flare-up or an exacerbation of a condition, and (iii) chronic condition episode groups require ongoing management of a long-term health condition. A more detailed description of these can be found in the following examples from measures we've developed within the 2022 Cost Measure Information Forms ZIP file:²³ (i) the "2021-12-13-mif-colrec-rsct.pdf" file for procedural episode groups, (ii) the "2021-12-13-mif-sepsis.pdf" file for acute inpatient medical condition episode groups, and (iii) the "2021-12-13-mif-diabetes.pdf" file for chronic condition episode groups.

There are 2 processes in calculating cost measure scores: (i) episode construction, and (ii) measure calculation. First, episode construction involves the following:

- **Trigger and define an episode:**
 - Procedural: Episodes are triggered or opened by CPT/HCPSC codes indicating that a procedure has been performed. The episode window is defined as a period after the trigger date and may include a period before the trigger to capture pre-procedure care.
 - Acute Inpatient Medical Condition: Episodes are triggered or opened by an inpatient hospital stay with a specified Medicare Severity-Diagnosis Related Group (MS-DRG). The episode window is defined by the trigger date and includes a period after the trigger date (and may include a period before the trigger to capture pre-admission care such as pre-operative testing, if applicable).
 - Chronic Condition: A trigger event is identified by the occurrence of 2 Part B Physician/Supplier (Carrier) claims billed by the same clinician group within a given time of one another. The pair of services must include a trigger claim and a confirming claim, which indicates that a clinician-patient relationship has begun. A trigger claim is an initial evaluation and management (E&M) code for outpatient services along with a relevant chronic condition diagnosis. The confirming claim can be either another outpatient services E&M code with a relevant chronic condition diagnosis, or a condition-related CPT/HCPSC code with a relevant chronic condition diagnosis.
- **Attribute the episode to a clinician:**
 - Procedural: An episode is attributed to a clinician that bills a trigger code for the episode group, and episodes are attributed to clinician groups by aggregating all episodes attributed to the clinicians that bill to the clinician group.
 - Acute Inpatient Medical Condition: An episode is attributed to a clinician group that bills at least 30% of the inpatient E&M codes during the trigger inpatient stay. An attributed clinician is any clinician within the attributed clinician group that bills any inpatient E&M codes on identified Part B Physician/Supplier claim lines during the inpatient stay.
 - Chronic Condition: An episode is attributed to a clinician group that bills the trigger and confirming claims for the total attribution window. To attribute the

²³ CMS, "2022 Cost Measure Information Forms" ZIP file, <https://www.cms.gov/files/zip/2022-cost-measure-information-forms.zip>.

episode to an individual clinician, we identify any clinician within the attributed clinician group who plays a substantial role in the care for the patient (e.g., billing at least 30% of the outpatient E&M codes with a relevant chronic condition diagnosis and/or condition-related CPT/HCPCS codes with a relevant chronic condition diagnosis on Part B Physician/Supplier claim lines during the episode).

- **Assign costs to the episode and calculate the episode observed cost:**
 - Procedural and Acute Inpatient Medical Condition: Clinically related services occurring during the episode window are assigned to the episode. The cost of these services is summed to determine each episode's standardized²⁴ observed cost.
 - Chronic Condition: It functions the same as procedural and acute inpatient medical condition measures; however, the standardized cost of the assigned services is summed and averaged across the number of days in an episode. This average daily cost is then multiplied by the length of the total attribution window to determine each episode's scaled standardized observed cost.

Lastly, measure calculation involves the following:

- **Exclude episodes:**
 - Procedural, Acute Inpatient Medical Condition, and Chronic Condition: Exclusions remove unique groups of patients from cost measure calculation in cases where it may be impractical and unfair to compare the costs of caring for these patients to the costs of caring for the cohort at large.
- **Calculate expected costs through risk adjustment:**
 - Procedural, Acute Inpatient Medical Condition, and Chronic Condition: A regression analysis is run using the risk adjustment variables as covariates to estimate the expected cost of each episode. Then, statistical techniques are applied to reduce the effect of extreme outliers on measure scores (e.g., winsorization).
- **Calculate the measure score:**
 - Procedural and Acute Inpatient Medical Condition: For each episode, the ratio of standardized total observed cost to expected cost is calculated and averaged across all of a clinician's or clinician group's attributed episodes to obtain the average episode cost ratio. The average episode cost ratio is multiplied by the national average observed episode cost to generate a dollar figure for the cost measure score.
 - Chronic Condition: It functions similarly as procedural and acute inpatient medical condition measures. However, we use the ratio of winsorized scaled observed cost to scaled expected cost. The measure score is also calculated as a weighted average of these ratios based on each episode's number of assigned days.

²⁴ Claim payments are standardized to account for differences in Medicare payments for the same service(s) across Medicare providers. Payment standardized costs remove the effect of differences in Medicare payment among health care providers that are the result of differences in regional health care provider expenses measured by hospital wage indexes and geographic price cost indexes (GPCIs) or other payment adjustments such as those for teaching hospitals. For more information, please refer to the "CMS Price (Payment) Standardization - Basics" and "CMS Price (Payment) Standardization - Detailed Methods" documents posted on the [CMS Price \(Payment\) Standardization Overview page](https://resdac.org/articles/cms-price-payment-standardization-overview) (<https://resdac.org/articles/cms-price-payment-standardization-overview>).