MACRA Episode-Based Cost Measures: Wave 4 Measure Development

Call for Public Comment: December 2020 to February 5, 2021

We are seeking feedback from stakeholders regarding which cost measures to develop for the upcoming cycle of episode-based cost measure development ("Wave 4"). This document includes the following:

Table of Contents

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 a performance-based adjustment to their Medicare payments based on 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.[1] The MIPS cost performance category currently has 20 measures:

- 2 population-based or global cost measures
- 18 episode-based cost measures for a range of procedures and acute inpatient medical conditions

The episode-based cost measures in the MIPS Calendar Year 2020 performance period were developed in cycles or "Waves" in 2017 and 2018. We are currently wrapping up Wave 3 with 5 new measures and now looking ahead to Wave 4 of development.

For Wave 4, 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 February 5, 2021. Stakeholders may either: (i) email a comment letter to macra-episode-based-cost-measures-info@acumenllc.com, or (ii) submit a response to the Wave 4 Measure Development Survey:

https://www.surveymonkey.com/r/wave_4_development

[1] CMS, "Episode-Based cost measure development for the quality payment program." (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.

2. Approach to Gathering Stakeholder Input in Wave 4

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 a type of measure. Acumen met with the CS in-person for a one-day meeting to discuss and vote on preferred episode groups.

We understand that 2020 is unlike these previous Waves, and that it has presented new challenges for the community, particularly front-line clinicians. We also heard feedback during the August – September 2020 field test for the Wave 3 measures that stakeholders were interested in ways to engage in measure development with more flexible participation options. As such, for Wave 4, we are seeking input on candidate episode groups through a public comment period that will:

- Allow more time for individuals, specialty societies, and professional associations to review and provide input
- Allow for flexibility around stakeholders' schedules
- Widen the range of stakeholders who can provide feedback on priority clinical areas for Wave 4 and future Waves of measure development

We expect to revisit this approach in Wave 5.

3. Selecting Episode Groups and Clinical Areas

To identify a starting point for Wave 4 measure prioritization, we drew on feedback that we heard over the years through our Technical Expert Panel (TEP), Person and Family Engagement (PFE), Clinical Subcommittees (CS), Clinician Expert Workgroups, and public comment.[2] We consider 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 value

3.2 Essential Features of Cost Measures

To ensure that a cost measure is effective in assessing clinician cost performance, we have 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 Clinician Expert Workgroups). The standards that guide this development work are as follows:

- Attribution of measures to clinicians is clear.
- Episode definitions have clinical face validity and consistency with practice standards.
- Construction of episodes/measures is readily understandable to providers.
- Providers are held accountable for costs of assigned services they can reasonably influence, which accurately captures their role.

---

[2] More information on the TEP, PFE, CS, and Clinician Expert Workgroups is described in the 2020 Episode-Based Cost Measures Field Testing Wave 3 Measure Development Process document, https://www.cms.gov/files/document/macra-cmft-ebcm-process-2020.pdf.

- Measures convey concrete guidance indicating how providers can alter practice to improve measured performance.
- Variation in measures helps distinguish quality of care across individual providers.
- Measure specifications allow for consistent calculation and reproducibility using Medicare data.

3.3 Previous Stakeholder Input

We reviewed the feedback that stakeholders have shared over the years for candidate clinical areas and episode groups. We started by gathering input on the draft list of episode groups and trigger codes through a call for comment posting in 2016 ("December 2016 posting")[3] and have used feedback received at that time as a starting point for selecting episode groups and determining preliminary specifications for Waves 1 through 3. In recent years, we have heard interest from stakeholders in expanding the list of episode groups in the December 2016 posting, as it does not include some types of care (e.g., post-acute care) for which there was no existing framework of measure construction at the time.

We revisited the list of episode groups to consider for future Waves at the February 2020 TEP[4] and through our past CS prioritization discussions to identify strong candidate episode groups. The TEP considered factors like capturing episode groups in high-impact areas, the tradeoffs between expanding coverage by focusing on new clinical areas or building in more in-depth measurement for high-cost areas, and specialty gaps to prioritize for future Waves. The TEP expressed support for developing measures for specialties not covered by the current measure set, particularly for mental/behavioral health, physical therapy, and radiology. The TEP supported developing measures such as Chronic Heart Failure that capture high costs. Panelists also supported measures for broader types of care such as Pain Management, Low Back Pain, and Spinal Injection. In general, TEP panelists agreed that there should be an effort to align with quality measures and focus on areas that CMS has identified as priority areas.

In identifying priority areas to focus on for Wave 4, we also closely examined the MIPS Quality Performance Category, in keeping with our continued effort to align cost with quality to ensure meaningful assessment of value, and also in consideration of potential MIPS Value Pathways (MVPs).

3.4 Empirical Analyses

We conducted a preliminary, exploratory analysis to estimate the potential impact of various episode groups using:

- Preliminary codes used to open an episode (trigger codes), including from previous stakeholder input[5]
- A simplified version of our existing measure frameworks (e.g., procedural and chronic condition) and a draft of the new therapy framework[6]

---

[3] 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.

[4] CMS, "Physician Cost Measures and Patient Relationship Codes (PCMP) Technical Expert Panel: February 6-7, 2020 Summary Report," (May 2020), https://www.cms.gov/files/zip/physician-cost-measures-and-patient-relationship-codes-pcmp.zip.

[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] This is described in the Preliminary Specifications of Wave 4 Candidate Episode Groups workbook.

These results provide a starting point to ascertain the potential impact of various measure concepts. Appendix A provides a table with the priority episode groups and estimates on their potential impact. While there were additional candidate episode groups we explored among the 4 clinical areas,[7] the priority candidate episode groups included in this posting were selected due to their performance on the measure prioritization criteria and other factors related to the overall viability for cost measurement.

The current episode group frameworks are:

- **Procedural episode groups**: These episode groups focus on procedures of a defined purpose or type, such as surgeries.
- **Acute inpatient medical condition episode groups**: These represent treatment for self-limited acute illness or treatment for flares or an exacerbation of a condition that requires a hospital stay.
- **Chronic condition episode groups:** These account for the ongoing management of a disease or condition. Wave 3 was the first Wave to include development of chronic condition episode groups.

[7] Other candidate episode groups that were considered, but were not selected as priority candidate episode groups include the following: (i) Heart Failure: an acute inpatient medical condition episode group for Heart Failure and Shock, (ii) Mental and Behavioral Health: Dementia, Alzheimer's Disease, Parkinson's Disease, Migraine, Post-Traumatic Stress Disorder, and various substance use disorder categories like alcohol and opioids, and (iii) Rheumatology/Arthritis: Lupus and Localized Lupus.

4. Questions about Clinical Areas and Candidate Episode Groups

Considering the factors outlined above, we identified 4 clinical areas with 7 priority episode groups for consideration of development in Wave 4. These clinical areas were selected based on discussions with CMS regarding the criteria for measure prioritization with an emphasis on previous stakeholder input, discussed in turn below.

4.1 Heart Failure

The **Heart Failure** clinical area would focus on a chronic condition measure that would apply to a broad set of clinicians, including general cardiologists, advanced heart failure cardiologists, and primary care clinicians that manage care for patients with the condition. This area has been a high priority for CMS, as it is a common condition among the Medicare population, representing high cost and opportunity for improvement (e.g., optimization of guideline-directed medical therapy and closer follow-up to reduce complications). Heart Failure would be a strong candidate for a MVP, as it is a broad and common condition with applicable MIPS quality measures. Additionally, we have received input suggesting the development of a Heart Failure measure in previous Waves of development and also from our TEP, noting the potential for high patient, clinician, and cost coverage.

- **Question 1**: What are ways to categorize different subtypes of Heart Failure (e.g., reduced or preserved ejection fraction) using administrative claims data? We can use techniques like risk adjusting or sub-grouping[8] based on services that are indicative of a given Heart Failure subtype. Are there certain types of services or diagnoses available via claims that may be useful in identifying subtypes?
- **Question 2**: What are ways to categorize different levels of severity of Heart Failure using administrative claims data? Are there certain types of services or diagnoses available via claims that may be useful in determining disease severity?
- **Question 3**: We aim to capture care provided by clinicians for the chronic management of Heart Failure. Are there different roles in providing care that we should consider (e.g., do internal medicine and cardiology specialists play distinct roles)?[9]
- **Question 4**: Are there any other concerns that may be present with assessing the chronic care for Heart Failure patients? If so, what are some potential approaches to address these concerns for a cost measure?

4.2 Mental and Behavioral Health

The **Mental and Behavioral Health** clinical area would focus on a chronic condition measure for either Major Depressive Disorder, Schizoaffective Disorder, or Bipolar Disorder, applying to a broad set of psychiatrists and other related specialties. Mental health is also a CMS priority area, and these are common conditions that are impactful for patients. A measure for one of these conditions would address the variation in outpatient treatment and how proper care may avoid costly complications (e.g., hospitalizations). A broad mental health topic such as Major Depressive Disorder would be a strong candidate for a MVP, as it is a broad and common condition with applicable MIPS quality measures. Additionally, we have received input suggesting the development of a Major Depressive Disorder measure from our TEP, noting the need for cost measurement for mental and behavioral health topics.

---

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

[9] More detailed information on attribution for chronic condition measures, as well as other elements of the framework, are included in the Draft Cost Measure Methodology document for Diabetes within the Episode-Based Cost Measure Specifications ZIP file: https://www.cms.gov/files/zip/macra-2020-wave-3-ft-specs.zip.

- **Question 1**: Among the priority candidate episode groups for this clinical area of Major Depressive Disorder, Bipolar Disorder, or Schizoaffective Disorder, which area should we focus on for Wave 4, and why? Are there additional concepts besides these 3 that would be valuable to explore within this clinical area? For a cost measure focused on the ongoing treatment and care for these conditions, what are some areas of opportunity for improvement a measure may be able to capture regarding care and potential mitigation of complications?
- **Question 2**: What are ways to account for different severity levels for the candidate episode groups (i.e., Major Depressive Disorder, Bipolar Disorder, or Schizoaffective Disorder)? We may use techniques like risk adjusting or sub-grouping for services that are indicative of various levels of severity. Are there certain types of services or diagnoses available via claims that may be useful in identifying various levels of severity?
- **Question 3**: Are there any other concerns that may be present with assessing the chronic care for patients with these conditions? If so, what are some potential approaches to address these concerns for a cost measure?

4.3 Therapy and Rehabilitation

The **Therapy and Rehabilitation** clinical area would represent the new measure framework we are exploring for Wave 4, focusing on the care and treatment provided by physical therapists and related specialties for broad conditions such as Low Back Pain or both Low Back and Neck Pain. This is a framework that is ripe for development, as it may build from the chronic condition framework to capture the care provided by therapists. Low back and neck pain are also common conditions that are impactful for patients, and a measure for these conditions would address the variation in treatment techniques (e.g., duration/frequency, use of higher cost interventions like imaging). A broad topic such as Low Back Pain would also be a strong candidate for a MVP, as it is a broad and common condition with applicable MIPS quality measures. Additionally, we have received input suggesting the development of a Low Back Pain measure from our TEP and other stakeholders, noting the need for cost measurement centered on the care provided by therapists participating in MIPS.

- **Question 1**: We identified 2 concepts for this clinical area, which includes Low Back Pain or an alternative approach for both Low Back and Neck Pain. Given the criteria for measure prioritization and the essential features of cost measures described above, which of the options would be preferable for the first therapy episode-based cost measure, and why? Are there additional concepts that would be valuable to explore within this clinical area? For a cost measure focused on the ongoing treatment and care for these conditions, what are some areas of opportunity for improvement a measure may be able to capture regarding care and potential mitigation of complications?
- **Question 2**: Based on the draft approach described in the “Appendix_Framework” tab of the Preliminary Specifications of Wave 4 Candidate Episode Groups workbook, which refinements would you recommend? What are types of services to use as indication of ongoing therapy management and care?
- **Question 3**: Based on the draft triggering approach, how should a therapy cost measure address the variation across patients for the need of therapy regarding Low Back Pain or for Low Back and Neck Pain (e.g., patients with chronic pain versus patients with a recent spinal surgery)? Some options include sub-grouping, risk adjusting, or excluding. Similarly, what recommendations do you have for how a measure may address patients with radicular syndrome/pain or arthritis?
- **Question 4**: Are there any other concerns that may be present with assessing the care for patients with these conditions? If so, what are some potential approaches to address these concerns for a cost measure?

4.4 Rheumatology/Arthritis

The **Rheumatoid Arthritis** clinical area would focus on a chronic condition measure that would apply to rheumatologists and primary care clinicians that manage care for patients with the condition. This area is also a priority, as it is a common condition among the Medicare population, representing opportunity for improvement (e.g., variation in treatment/drug options and efficient monitoring/imaging/therapy, including for adverse effects to treatments). Rheumatoid Arthritis would be a good candidate for a MVP, as it is a broad and common condition with applicable MIPS quality measures. Additionally, we have received input suggesting the development of measure concepts that address arthritis and related conditions.

- **Question 1**: What are ways to account for different severity levels for Rheumatoid Arthritis? Are there considerations like the specialty of the attributed clinician (e.g., internal medicine versus rheumatology) that may help inform different severity levels? We may use techniques like risk adjusting or sub-grouping for services that are indicative of various levels of severity. Are there certain types of services or diagnoses available via claims that may be useful in identifying various levels of severity?
- **Question 2**: Are there any concerns regarding the attribution of Rheumatoid Arthritis episodes to clinicians from certain specialties (e.g., internal medicine versus rheumatology)? For reference, chronic condition measure attribution for clinicians includes the requirement that the clinician within the attributed clinician group must bill at least 30% of "primary care" evaluation and management (E&M) codes with a relevant chronic condition diagnosis and/or chronic condition-related Current Procedural Terminology/Healthcare Common Procedure Coding System (CPT/HCPCS) codes for related services with a relevant chronic condition diagnosis on Part B Physician/Supplier claim lines during the episode (along with other requirements).[10]
- **Question 3**: For a cost measure focused on the ongoing treatment and care for Rheumatoid Arthritis, what are some areas for opportunity for improvement a measure may be able to capture regarding care and potential mitigation of complications?
- **Question 4**: Are there any other concerns that may be present with assessing the chronic care for patients with Rheumatoid Arthritis? If so, what are some potential approaches to address these concerns for a cost measure?

[10] The other requirements for chronic condition measure attribution of episodes to clinicians is that the clinician within the attributed clinician group must also bill at least: (i) 1 "primary care" E&M code with a relevant chronic condition diagnosis or chronic condition-related CPT/HCPCS code for related services with a relevant chronic diagnosis within 1 year prior to or on the episode start date, and (ii) 2 condition-related prescriptions at different time points to 2 different patients.

5. General Questions across Candidate Episode Groups and Measure Development

Beyond the 4 clinical areas and their corresponding episode groups that we considered, we are also interested in thoughts on other potential clinical areas and episode groups that we may consider for future Waves of development.

5.1 Cross-Cutting Questions for All Wave 4 Candidate Episode Groups

We considered share of Medicare expenditures, clinician coverage, 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:

1. **Opportunity for improvement.** What kinds of services can reflect the opportunities for improvement? For example, cost measures generally include services reflecting variation in treatment options, intensity/duration, follow-up care, complications, and others.
2. **Trigger codes**: Trigger codes define the patient cohort for the measure. The preliminary set of codes that we used are in the accompanying Preliminary Specifications of Wave 4 Candidate Episode Groups workbook. We solicit comment on this list of trigger codes.
3. **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.

5.2 Participating in Wave 4 Development

1. **Are you interested in participating in Wave 4?** Please fill out this expression of interest form[11] if you or your organization has experience in one of the clinical areas in Section 4 and are interested in finding out more in early 2021 when we confirm the measures that we will be developing. The form also has space for any feedback about the types of specialties or expertise that should be considered for the composition of the Clinician Expert Workgroups.

5.3 Other Measure Concepts and Measure Development for Future Waves

In addition, we are interested in your feedback on a range of topics that are not being considered for Wave 4 but may be considered in future Waves.

1. **Cost measures for non-patient facing clinicians (e.g., anesthesiologists, radiologists).** We have heard interest in measures for care provided by radiologists, anesthesiologists/anesthetists, pathologists, and other similar specialties. We agree that these are important areas to cover and seek comment on the following:
    a. Overall, what is the opportunity for improvement via cost measurement for these types of specialties? How could a cost measure influence decision-making for these types of clinicians?
    b. How could we distinguish the responsibility of non-patient facing clinicians versus patient-facing clinicians (e.g., the anesthesiologist versus the surgeon)? Which types of services would be reasonably influenced by non-patient facing clinicians? For example, if there is a cost measure for anesthesiologists centered

[11] This is the MACRA Episode-Based Cost Measures Mailing List (https://www.surveymonkey.com/r/macra_cost_measures_mailing_list), which contains questions regarding interest in particular Wave 4 clinical areas or candidate episode groups.

on a type of surgery, what sort of complications and other follow-up services may be reasonably influenced by the anesthesiologist rather than the surgeon alone?
    c. Currently, the MIPS cost measures span episode windows from 14 days to 1 year or more. What are some suitable timeframes for concepts like mammography or anesthesiology for orthopedic surgery?
2. **Cost measures for head and neck disease clinicians**. We also heard interest in measures for care provided by otolaryngologists, or ear, nose, and throat (ENT) clinicians. We agree that this is an important area to cover and seek comment on the following:
    a. What clinically coherent candidate episode groups for head and neck disease care would you recommend as cost measure concepts? Some options we looked into include Sinusitis, Sinus Surgery, and Hearing Loss. What episode-based cost measures may help capture a broader set of ENT clinicians?
    b. Based on the types of conditions treated or procedures performed by otolaryngologists, would episode-based cost measures around procedures or around chronic condition care be more impactful to capture performance gaps among ENT clinicians, and why?
    c. Are there any challenges with assessing the care for head and neck disease conditions or procedures? If so, what are some potential approaches to address these concerns for a cost measure?
3. **Cost measures for emergency room clinicians.** We also have heard interest in measures for care provided by emergency room clinicians. We agree that this is an important area to cover and seek comment on the following:
    a. How can a measure on emergency room care appropriately address the result of the visit (e.g., release to community versus transfer to hospital)? How should cases where the emergency room visit leads to an inpatient admission (or transfer to another facility for follow-up care) be handled by an emergency room episode group?
    b. What would be clinically coherent scopes for candidate episode groups for emergency room care (e.g., visits for shortness of breath, chest pain, wounds)?
    c. Based on the type of care provided in the emergency room setting, how should individual clinicians and clinician groups be attributed episodes?
    d. In terms of episode window, what are suitable timeframes that can assess care, treatment, and subsequent outcomes that may be reasonably influenced by attributed emergency room clinicians?
4. **Oncology coding**. In Wave 2, we developed a lumpectomy and mastectomy measure, and in Wave 3, we are developing a melanoma resection measure. These tend to cover the care provided by surgeons who deal with oncologic care and dermatologists, respectively. For oncologists, we have heard that a major limitation is the lack of coding specificity for staging. What are some ways to account for the lack of cancer staging information in codes? Are there any other benefits or drawbacks to consider for the potential development of further oncologic episode groups?
5. **Information gathering process**. For Wave 4, we are seeking input through public comment to provide a more flexible way for stakeholders to participate during these times. In future Waves, would you prefer exploring a CS option again (i.e., convening a group of clinicians from various specialties and societies to provide targeted input during a meeting)?

6. Next Steps

Please share your feedback by emailing macra-episode-based-cost-measures-info@acumenllc.com or submitting a response to the online survey[12] by **11:59 p.m. Eastern Time** on **February 5, 2021**.

We will consider stakeholder feedback in our recommendations to CMS to finalize the set of episode groups to develop in Wave 4. The decision on the Wave 4 cost measures will be made by CMS, which may be informed based on the input provided by stakeholders from this public comment posting. Once approved, we will begin efforts to identify members of the Clinician Expert Workgroups that will be tasked with convening throughout measure development to provide input that will inform measure specifications. For each Wave 4 measure, there will be a Clinician Expert Workgroup of about 15 members. We will review comments expressing potential interest in workgroup membership for the various candidate episode groups and also have a public call for nominations in 2021. When the Clinician Expert Workgroups convene, we will also prepare analyses that incorporate issues and concerns raised during this public comment period (e.g., specific patient cohorts to evaluate); members will be able to review data informed by public comments and discuss approaches regarding measure specification based on the data and their clinical expertise.

[12] Wave 4 Measure Development Survey, https://www.surveymonkey.com/r/wave_4_development.

Appendix A: Descriptive Statistics for Preliminary Episode Groups

The table below provides estimate data for the priority candidate episode groups based on a preliminary, exploratory analysis. The Preliminary Specifications of Wave 4 Candidate Episode Groups workbook includes the draft list of preliminary trigger codes as well as an appendix describing the construction of episodes used to produce these results.

Table 1: Preliminary Results on Priority Candidate Episode Groups

Initial Sort Order	Clinical Area	Episode Group	Coverage			Top 3 Most Frequently Attributed Specialties[13]
			Beneficiary[14]	Clinician[15]	All-Cost % of Parts A+B[16]	
1	Heart Failure	Heart Failure (Chronic)	1,345,435	47,916 TINs; 199,263 TIN-NPIs	9.87 – 23.07%	Cardiology, Internal Medicine, Family Practice
2	Mental and Behavioral Health	Major Depressive Disorder (Chronic)	1,586,050	53,088 TINs; 202,652 TIN-NPIs	5.51 – 13.43%	Internal Medicine, Family Practice, Psychiatry
3	Mental and Behavioral Health	Schizoaffective Disorder (Chronic)	304,508	25,834 TINs; 74,913 TIN-NPIs	1.39 – 3.45%	Psychiatry, Nurse Practitioner, Internal Medicine
4	Mental and Behavioral Health	Bipolar Disorder (Chronic)	315,167	32,372 TINs; 96,829 TIN-NPIs	1.17 – 2.87%	Psychiatry, Nurse Practitioner, Family Practice
5	Therapy and Rehabilitation	Low Back and Neck Pain (Therapy)	822,358	14,762 TINs; 59,212 TIN-NPIs	2.18%	Physical Therapist, Orthopedic Surgery, Occupational Therapist
6	Therapy and Rehabilitation	Low Back Pain (Therapy)	653,328	14,380 TINs; 56,572 TIN-NPIs	1.76%	Physical Therapist, Orthopedic Surgery, Physical Medicine And Rehabilitation
7	Rheumatology/ Arthritis	Rheumatoid Arthritis (Chronic)	445,157	41,626 TINs; 123,534 TIN-NPIs	2.52 – 6.32%	Rheumatology, Internal Medicine, Family Practice

[13] The name of the top 3 most attributed Medicare specialties for the episode group, based on the full set of attributed clinicians across all episodes.
[14] The number of Medicare beneficiaries with an episode from this episode group that meet a few basic inclusion criteria (e.g., beneficiaries that were enrolled in Medicare Parts A and B during the episode window and were not enrolled in Part C).
[15] The number of clinician groups (TINs) and clinicians (TIN-NPIs) that bill to 1 or more episodes for the episode group within the study period.
[16] The share of all costs from this episode group out of all Medicare Part A and Part B expenditures during the study period. This number provides an upper bound estimate of the share of Medicare Part A and B costs captured by this episode group, as there is no service assignment applied for these estimate values. For chronic condition episode groups, there is a range where the approach described above is the upper limit. To approximate a tapered down approach, the lower bound estimate included for chronic condition episode groups was calculated with the following approach: The numerator differs from the All-Cost % of Part A + B numerator as it considers 100% of costs through 60 days, 50% of costs from 61 through 90 days, and 25% of costs thereafter to the end of the episode.

Appendix B: Overview of Cost Measure Frameworks

This appendix provides a brief overview of the 2 existing measure frameworks for the candidate episode groups (i.e., procedural and chronic condition). A more detailed description of the current frameworks is provided in the following files within the Episode-Based Cost Measure Specifications ZIP file:[17] (i) the "2020-08-methods-colrec-rsct.pdf" file for the procedural framework, (ii) the "2020-08-methods-diabetes.pdf" file for the chronic condition framework, and (iii) the "2020-08-methods-sepsis.pdf" file for the acute inpatient medical condition framework. Note that the new therapy framework is not yet developed, and the draft approach (for which we solicit stakeholder input) is described in the "Appendix_Framework" tab of the Preliminary Specifications of Wave 4 Candidate Episode Groups workbook.

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/HCPCS codes indicating that a procedure has been performed. The episode window is defined around the trigger and may include a period before the trigger to capture pre-procedure care.
  - 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 180 days 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.[18] An episode has a minimum one-year segment of a total attribution window.
- **Attribute the episode to a clinician**:
  - Procedural: An attributed clinician is any clinician who 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.
  - Chronic Condition: An attributed clinician group is the clinician group that bills the trigger and confirming claims for the total attribution window. An attributed clinician is any clinician within the attributed clinician group that bills at least 30% of "primary care" E&M codes with a relevant chronic condition diagnosis and/or chronic condition-related CPT/HCPCS codes for related services 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: 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 standardized observed cost.
  - Chronic Condition: It functions the same as procedural measures; however, the cost of the assigned services is summed and averaged across the number of

[17] "Episode-Based Cost Measure Specifications" ZIP file, https://www.cms.gov/files/zip/macra-2020-wave-3-ft-specs.zip.

[18] The trigger claim is an initial "primary care" evaluation and management (E&M) code with a relevant chronic condition diagnosis. The confirming claim can be either another "primary care" E&M code with a chronic condition diagnosis, or a chronic condition-related CPT/HCPCS code for related services with a relevant chronic condition diagnosis.

assigned days in an episode, and this average daily cost is then multiplied by 365 to determine each episode's annualized observed cost.

Lastly, measure calculation involves the following:

- **Exclude episodes**:
    - Procedural 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 for risk adjustment**:
    - Procedural 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: For each episode, the ratio of standardized total observed cost to risk-adjusted 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 measures. However, for this framework, we use the ratio of winsorized annualized standardized observed cost to annualized expected cost. The measure score is calculated as a weighted average of these ratios across all of a clinician's or clinician group's attributed episodes, where the weighting is each episode's number of assigned days. The weighted average episode cost ratio is then multiplied by the national average winsorized annualized observed episode cost to generate a dollar figure for the cost measure score.