Quality Payment
PROGRAM

2022 Field Testing Feedback Summary Report for 5 Episode-Based Cost Measures:

- Emergency Medicine
- Low Back Pain
- Heart Failure
- Major Depressive Disorder
- Psychoses/Related Conditions

May 2022

Table of Contents

1.0 Introduction 4
1.1 Project Title 4
1.2 Project Background 4
1.3 Measure Development and Field Testing Overview 4
2.0 General Feedback Summary 7
2.1 Episode-Based Cost Measures Framework 7
2.1.1 Assignment of Costs 7
2.1.2 Risk Adjustment 7
2.1.3 Alignment of Cost and Quality 7
2.1.4 Measure Reporting and/or Use in MIPS 7
2.2 Measure Development Approach 8
2.2.1 Stakeholder Engagement 8
2.2.2 COVID-19 Implications on Providing Feedback 8
2.3 Field Test Report and Supplemental Documentation 8
2.3.1 Field Test Report Access and Format 8
2.3.2 Mock Field Test Reports 9
2.3.3 Measure Testing Information 10
3.0 Measure-Specific Field Testing Feedback 11
3.1 Emergency Medicine 11
3.1.1 Definition of an Episode Group 11
3.1.2 Attribution of the Episode Group to Clinicians 12
3.1.3 Assignment of Costs to the Episode Group 13
3.1.4 Risk Adjustment 14
3.1.5 Exclusions 15
3.1.6 Alignment of Cost with Quality 15
3.1.7 Measure Reporting and/or Use in MIPS 15
3.2 Heart Failure 15
3.2.1 Definition of an Episode Group 15
3.2.2 Attribution of the Episode Group to Clinicians 16
3.2.3 Assignment of Costs to the Episode Group 17
3.2.4 Risk Adjustment 18
3.2.5 Exclusions 20
3.2.6 Alignment of Cost with Quality 20
3.2.7 Measure Reporting and/or Use in MIPS 20
3.3 Low Back Pain 21
3.3.1 Definition of an Episode Group 21
3.3.2 Attribution of the Episode Group to Clinicians 22
3.3.3 Assignment of Costs to the Episode Group 23
3.3.4 Risk Adjustment 24
3.3.5 Exclusions 25
3.3.6 Alignment of Cost with Quality 25
3.3.7 Measure Reporting and/or Use in MIPS 25
3.4 Major Depressive Disorder 25
3.4.1 Definition of an Episode Group 25
3.4.2 Attribution of the Episode Group to Clinicians 25
3.4.3 Assignment of Costs to the Episode Group 25
3.4.4 Risk Adjustment 26
3.4.5 Exclusions 27
3.4.6 Alignment of Cost with Quality 27

3.4.7 Measure Reporting and/or Use in MIPS 27
3.5 Psychoses/Related Conditions 27
3.5.1 Definition of an Episode Group 27
3.5.2 Attribution of the Episode Group to Clinicians 27
3.5.3 Assignment of Costs to the Episode Group 27
3.5.4 Risk Adjustment 28
3.5.5 Exclusions 28
3.5.6 Alignment of Cost with Quality 28
3.5.7 Measure Reporting and/or Use in MIPS 28
4.0 Next Steps for Measure Specification Refinements 30
4.1 Emergency Medicine 30
4.2 Heart Failure 30
4.3 Low Back Pain 31
4.4 Major Depressive Disorder 31
4.5 Psychoses/Related Conditions 32
Appendix A: List of Commenters 33

1.0 Introduction

1.1 Project Title

MACRA Episode-Based Cost Measures: 2022 Cost Measures Field Testing

1.2 Project Background

The Medicare Access and CHIP Reauthorization Act of 2015 (MACRA) established the Quality Payment Program (QPP). Under the QPP, clinicians are incentivized to provide high quality and high value care through Advanced Alternative Payment Models or the Merit-based Incentive Payment System (MIPS). MIPS eligible clinicians will receive a performance-based adjustment to their Medicare payments. This payment adjustment is based on a MIPS 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.

CMS has contracted with Acumen, LLC, to develop new episode-based cost measures for potential use in the cost performance category of MIPS. This work is under the contract, "Physician Cost Measures and Patient Relationship Codes (PCMP)" (contract number 75FCMC18D0015, Task Order 75FCMC19F0004). Acumen has implemented a measure development process that relies on input from a large number of stakeholders, including multiple groups of clinicians affiliated with a broad range of professional societies and Patient and Family partners, to develop clinically appropriate and transparent measures that provide actionable information to clinicians.

This document summarizes the feedback we received from stakeholders on the 5 episode-based cost measures that were field tested as part of the measure development process from January 10 to March 25, 2022. Section 1.0 provides background on the episode-based cost measure development process and the 5 episode-based cost measures being developed. Section 2.0 summarizes the general feedback Acumen received on the cost measure framework, the measure development process, and the field testing materials. Finally, section 3.0 provides more detailed feedback on each of the 5 episode-based cost measures that underwent field testing.

1.3 Measure Development and Field Testing Overview

The Wave 4 episode-based cost measure development process started in 2020, when Acumen began gathering input from stakeholders to help inform which measures to develop and then continued through 2021 gathering input on measures specifications:

- The public comment period from December 2020 to February 2021 invited stakeholders to provide feedback on prioritizing episode-based cost measures and on preliminary specifications for several measure areas.
- Clinician Expert Workgroups that convened from June to September 2021 provided clinical specifications for Heart Failure, Major Depressive Disorder, Low Back Pain, and Emergency Medicine episode-based cost measures.

The Psychoses/Related Conditions measure was first developed during Wave 2 (2018) and underwent clinician input through a Clinician Subcommittee, instead of the public comment period, and through the Clinician Expert Workgroup from June 2018 to February 2019. Due to concerns from the Measure Applications Partnership (MAP) Committee, the Psychoses/Related Conditions measure wasn't finalized in MIPS. Acumen continued to develop the measure in Wave 4 by reconvening the Clinician Expert Workgroup in October 2021 to review stakeholder

feedback received on the measure and use their input on potential refinements needed to the current measure specifications. For more detailed information on field testing and the episode-based cost measure development process, please refer to the MACRA Feedback Page.

Once the 4 newly developed episode-based cost measures and the revised Psychoses/Related Conditions episode-based cost measure were specified through stakeholder input and empirical analyses, CMS and Acumen conducted field testing on the draft measures from January 10 to March 25, 2022. Field testing is a key part of the measure development process and is an opportunity for clinicians and other stakeholders to learn about episode-based cost measures and provide input on the draft specifications.

During field testing, clinicians and clinician groups had the opportunity to view a Field Test Report on the QPP website with information about their performance. Field Test Reports were available to clinicians and clinician groups who had 20 or more episodes for at least 1 of the Wave 4 episode-based cost measures during the measurement period (1/1/2019 - 12/31/2019). The 5 episode-based cost measures that were tested and the services that clinicians Medicare costs were evaluated on are provided in Table 1 below.

Table 1. Wave 4 Episode-Based Cost Measures and Services

Measure Name	Services
Emergency Medicine	Clinicians who provided care in the emergency department (includes visits leading to both discharges and hospital readmission).
Low Back Pain	Clinicians who managed the ongoing management and treatment of low back pain.
Heart Failure	Clinicians who managed the ongoing management and treatment of heart failure.
Major Depressive Disorder	Clinicians who managed the ongoing management and treatment of major depressive disorder.
Psychoses/Related Conditions	Clinicians who managed the acute inpatient hospitalization for the treatment of psychoses/related conditions.

1,562 clinicians and clinician groups downloaded a Field Test Report from the QPP website during field testing. 1,442 clinician groups (identified by Tax Identification Number or TIN) downloaded a report and 120 clinicians (identified by Tax Identification Number-National Provider Identifier or TIN-NPI) downloaded a report.

For the duration of field testing, all stakeholders were invited to provide feedback on the measures by completing an online survey or submitting a comment letter,[1] regardless of whether they received a report. Acumen and CMS made several materials publicly available for stakeholders' review: (i) draft measure specifications, (ii) mock Field Test Reports, and (iii) supplemental documentation.[2] Acumen and CMS also posted the MACRA 2022 Cost Measures Field Testing Webinar to the QPP Webinar Library at the start of the field testing period to

[1] The survey was previously available online at this link: 2022 Cost Measures Field Testing Feedback Survey (qualtrics.com).
[2] Field testing materials are available for download on the MACRA Feedback Page: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/MACRA-Feedback.html.

provide stakeholders with details regarding the field testing process and draft measure specifications.[3]

Acumen also held 2 specialty society office hours before and during field testing to help provide information about Field Test Reports and to provide an opportunity for stakeholders, including specialty societies who represent clinicians likely to be attributed the measures, to ask questions about field testing and the measure specifications. Acumen had 36 participants attend office hours: 24 attended the first office hours session and 19 attended the second office hours session. Some participants attended both office hour sessions.

Acumen received 64 survey responses, including 19 comment letters. Of the survey responses, 14 were from patients or caregivers. The list of stakeholders who submitted feedback through the online field testing feedback survey is provided in Appendix A. The feedback about each measure was shared with the Clinician Expert Workgroups to help inform measure refinement recommendations after field testing. Acumen and CMS will also evaluate the general feedback on measure specifications, the measure development process, and field testing and consider ways to improve future episode-based cost measure development processes.

[3] MACRA 2022 Cost Measures Field Testing Webinar materials are available on the Quality Payment Program Webinar Library: https://qpp.cms.gov/about/webinars.

2.0 General Feedback Summary

This section summarizes general feedback received on the episode-based cost measures. Section 2.1 summarizes feedback on the episode-based cost measures framework, section 2.2 summarizes feedback on the measure development approach, and section 2.3 provides feedback on the field testing documents.

2.1 Episode-Based Cost Measures Framework

2.1.1 Assignment of Costs

A stakeholder recommended not including any Part D costs in the measures because clinicians aren't able to negotiate formularies, coverage, or price of these medications. The commenter believed that it would be inappropriate to assign Part D costs until more data was available to indicate what impact these costs may have on the measure or clinicians. They believe accessing detailed, aggregate-level data about variation in spending per drug class, plan, and tier with different copayments would be helpful in assessing the impact of including Part D costs in the measures. They also requested that CMS release more information about how payment standardization methodology can distinguish costs in a way that's actionable for clinicians to influence.

2.1.2 Risk Adjustment

A few commenters stated that the risk adjustment model should account for patient preferences because some patients might only feel comfortable with a specific treatment course and this shouldn't affect a clinician's measure score.

Stakeholders also suggested that relying solely on claims data limited the accuracy of the risk adjustment model, urging that CMS consider using alternative sources, such as clinical data registries.

Additionally, one stakeholder requested that CMS release testing on the impact of COVID-19 on the newly developed episode-based cost measure and the cost measures in use in MIPS to determine if the surges are affecting certain geographic regions or specialties. They're concerned whether the current risk adjustment model can mitigate the impact of COVID-19 on service utilization.

2.1.3 Alignment of Cost and Quality

Several stakeholders stressed the importance of linking cost and quality within the construction of the cost measures to avoid disincentivizing appropriate care. One commenter specifically noted they don't believe it was sufficient to assess quality measures alongside cost measures in MIPS, because the measures are developed independently and often don't capture the same scope. Stakeholders recommended exploring alternative approaches to claims data, such as using clinical registry data or incorporating functional outcomes, patient-reported outcomes and preferences, and clinical guidelines, to account for the appropriateness and quality of care within the cost measures.

2.1.4 Measure Reporting and/or Use in MIPS

A stakeholder recommended that for measures to serve as effective motivators for behavioral change, there needs to be a clear relationship between the intended behavioral change and changes in the reported measure. They note that it's been difficult for clinicians to draw this

connection because they find the measures complex, so this may limit the effectiveness of efforts to measure costs as part of the MIPS program.

A stakeholder also requested that measures are implemented in MIPS on an informational only basis their first year in use if Part D costs are included in the measures. They requested that the first year performance feedback include 2 rates, one with Part D costs included in the measure and one without Part D costs.

Finally, a stakeholder highlighted the need for clinicians to have access to claims data, raising concerns that clinicians don't receive timely, actionable feedback on their resource use and attributed costs in Medicare. They noted that this makes it difficult for clinicians to reduce avoidable costs.

2.2 Measure Development Approach

2.2.1 Stakeholder Engagement

Stakeholders appreciated the opportunity to provide feedback on the Wave 4 episode-based cost measures and that CMS and Acumen are developing the episode-based cost measures with significant input from clinicians across many sizes, specialties, and locations. A couple of stakeholders also expressed appreciation for meeting with Acumen during the field testing period to discuss the measures and field test reports, finding this engagement helpful.

A few stakeholders requested additional training on how to interpret the data in reports and provide examples for how to validate the information. Commenters also requested that Acumen hold targeted webinars with stakeholders to walk them through specific reports, and to explain the intent of the data and the rationale for its inclusion.

Some stakeholders requested Acumen continue to find ways to alert clinicians and clinician groups who have a report available. One commenter suggested making a public list of NPIs that have a report, similar to the list that CMS creates each year for certain Qualifying Participants in Advanced Alternative Payment Models (APMs).

2.2.2 COVID-19 Implications on Providing Feedback

Due to the Omicron COVID-19 variant surge coinciding with the field testing period, many commenters expressed concern that stakeholder participation would be limited and recommended extensions to the field testing feedback period. In response to this feedback, CMS and Acumen extended field testing by one month, concluding field testing on March 25.

2.3 Field Test Report and Supplemental Documentation

2.3.1 Field Test Report Access and Format

While some stakeholders found that they could easily access the reports, many stakeholders reported experiencing difficulty accessing their QPP accounts and/or their reports. A stakeholder noted that clinicians in larger health systems or facilities experienced difficulties getting in touch with appropriate administration to access their individual reports or might only be able to view the clinician group report. Some clinicians felt that although there were documents with guidance on accessing reports, it was spread across too many documents and that clinicians didn't have time to navigate the instructions. One commenter requested that Acumen work towards simplifying the process for accessing the field test reports.

While a few stakeholders found the reports to be easy to read and informative, several stakeholders advised that clinicians are confused about how to act on the data presented in the field test reports to improve their performance. They emphasized that providing more guidance in the report or conducting more in-depth clinician education about the measures and how to interpret their reports would be helpful. Some commenters noted that the visualizations, such as flowcharts, were the most helpful resource. They recommended making the flowcharts more prominent and accessible in the future. One commenter also recommended including additional information, potentially using visual illustrations, on how cost measures are scored under MIPS and how benchmarks are applied so that clinicians can better understand how the measure results might be used in the future.

A stakeholder noted that the glossary wasn't enough to give clinicians an understanding of the data included in the report and that Acumen should explore ways to make the data more intuitive and easy to understand. Additionally, a stakeholder recommended including direct links to the supplemental materials (e.g., the methodology documents) within the field test report so that clinicians don't need to download the materials separately. They noted that some clinicians experienced difficulties navigating the field testing materials and found the amount of materials that accompanied the field test reports to be overwhelming.

Some stakeholders requested additional information about data provided in the field test reports. In particular, some clinicians were interested in having additional information about sub-groups and requested sub-group level data to figure more prominently or earlier in the report, as they're most meaningful to the clinicians. Additional recommendations for the specific tables in the report include:

- Provide the number of episodes that are included in the national averages to help compare against a clinician's or clinician group's score.
- Include medians instead of averages to minimize any impact from outliers and include risk bracket medians to reflect whether a clinician or group is above or below the national median.
- Include more information at the beginning of the report about what risk bracket the clinician or clinician group fell into so they can tell initially how complex their patient population is compared to that of other clinicians.
- Make it clearer which Part B drugs are included in the measure.
- Include the volume of services, not just percentages, in the "Service Use and Cost by Setting and Category" table and provide more granular break downs of the procedures or interventions contributing to the episode.
- Break out Part D drugs into drug types or drug categories to provide additional data around prescribing practices, if Part D drugs are included in the measure.
- Clarify the rationale for outlining the top providers within a clinician's TIN and outside of their TIN that are contributing to their costs and how they should interpret this information.

One commenter noted that clinicians experienced confusion with how the field test reports relate to the MIPS performance feedback reports, and the differences between these reports should be clarified. They stated that, ideally, the format and type of information in the field test reports and the MIPS performance feedback reports should align more closely.

2.3.2 Mock Field Test Reports

Stakeholders requested that Acumen make mock field test reports available for all measures, rather than each type of measure (i.e., acute inpatient medical, procedural, and chronic condition) so that stakeholders who don't have access to reports can still review the measure-specific mock report and provide more meaningful feedback.

Some stakeholders also suggested that Acumen make de-identified reports or other data on practices that fall within the top and bottom deciles so that clinicians can better understand what is contributing to their scores.

2.3.3 Measure Testing Information

A stakeholder highlighted the importance of having access to reliability, validity, and actionability data for the measures undergoing field testing so that they can assess the appropriateness and scientific rigor of the measures. They also requested that CMS release information about the alignment between the measures undergoing field testing and corresponding quality information.

3.0 Measure-Specific Field Testing Feedback

This section includes the measure-specific feedback received on the 5 episode-based cost measures during the field testing period. The feedback was shared with the Clinician Expert Workgroups prior to the Post-Field Test Refinement (PFTR) webinars in April 2022 for their review as they considered potential refinements to the measures.

Each section provides detailed feedback on the Wave 4 episode-based cost measures. Section 3.1 summarizes feedback on the Emergency Medicine measure, section 3.2 summarizes feedback on the Heart Failure measure, section 3.3 summarizes feedback on the Low Back Pain measure, section 3.4 summarizes feedback on the Major Depressive Disorder measure, and finally, section 3.5 summarizes feedback on the Psychoses/Related Conditions measure.

3.1 Emergency Medicine

3.1.1 Definition of an Episode Group

3.1.1.1 Episode Window

- A few stakeholders suggested shortening the 30-day episode window to 7 days. Another commenter suggested a 7- to 14-day episode window. Stakeholders noted the following reasons for shortening the episode window:
    - Some commenters noted that related Emergency Department (ED) visits tend to happen within 7 days, and ED visits beyond that time are often unrelated. A stakeholder said that nearly 97% of ED visits related to the index visit occur within 7 days.
    - Many other services beyond a 7-day window aren't within the influence of the ED clinician.
    - One commenter noted that the service exclusion rules could be simpler and less voluminous for a shorter episode window. The commenter further noted that with a simpler measure, clinicians would be able to more easily and proactively monitor their measure performance to prepare interventions and assess care effectiveness.

3.1.1.2 Sub-groups and ED Visit Types

- A few stakeholders agreed with the current set of ED visit types, noting that the current visit types are clinically appropriate, understandable, and provide actionable information to clinicians and clinician groups.
- Several commenters also approved of the hierarchy of visit types (the hierarchy used in field testing has been defined empirically to categorize episodes into their most expensive and resource-intensive potential visit type, so that clinicians aren't penalized for furnishing high-cost, medically necessary services).
- A few comments suggested updates to the ED visit types, as follows:
    - Combine the "Stroke" and "Neurologic" visit types with the "Altered Mental State" visit type due to the presentation of these conditions.
    - Combine the "Poisoning" and "Female Disorders" visit types with "Abdominal Pain, Nausea, and Vomiting" due to the presentation of these conditions.
    - Separate a skin visit type from the "Oral, Nasal, And Skin" visit type, and create a separate category for skin conditions such as rashes and abscess.
    - Exclude patients with surgical conditions from the "Medical Complications" visit type.
    - Create a distinct category for all otolaryngologic and eye disorders.

    - Separate non-diabetic endocrine disorders from the “Diabetes” visit type due the different outpatient costs - and the high-cost intervention - of diabetes cases.
    - Re-label “Female Disorders” as “Gynecological Disorders” with exclusions for operative procedures.
    - Create a visit type for hypertension, unless hypertension already falls under “Other Cardiovascular.”
    - Create a separate type for procedures (e.g., laceration repair, chest tubes, treatment of abscesses), given the large variation in cost due to the time and intensity required for these procedures
    - Include the following to the list of diagnosis codes that are used to specify visit types: COVID-19 (ICD-10 diagnosis code U07.1), COVID-19 exposure (Z20.822), headache (R51.9), and fever (R50.9). The commenter noted that these codes represent top diagnoses that the commenter’s ED patients had.
    - Include the “Health Care Maintenance” visit type.
- A stakeholder suggested excluding ED visits that are more appropriate for chronic, outpatient care.
    - These patients are frequently lost to follow-up (i.e., patients who usually just have one or a few visits to treat a particular problem and don’t return for additional services needed to continue treating that condition), so these ED visits may skew costs for clinicians who have a larger share of these episodes.
- Several commenters suggested stratifying episodes with admission and episodes with discharge rather than risk adjusting for this.
    - This would ensure comparability of costs that risk adjustment may not provide and account for differing case mix of patients whose conditions are appropriate for admission vs. discharge.
- Several Patient and Family Engagement (PFE) representatives provided input on their experience with being diagnosed in the ED or upon admission:
    - One PFE representative noted that their diagnosis was initially correct, but less specific prior to when they were admitted.
    - Another PFE representative noted that their diagnosis (pancreatitis) was accurate, and noted good communication among the members of their care team.
    - One PFE noted that their diagnosis changed after admission (the patient was first diagnosed with twisted bowel, but then was diagnosed with simple obstruction with septic UTI); while another PFE noted that they received their diagnosis of osteoporosis / arthritis only after admission.

3.1.2 Attribution of the Episode Group to Clinicians

- A few commenters provided feedback on what specialties they thought should and shouldn’t be attributed the measure:
    - Several commenters suggested that the measure only be attributed to clinicians who provide emergency medical services (e.g., emergency medicine physicians and other non-physician emergency medical workers).
    - Several stakeholder noted the need to have an explicit exclusion of other specialties such as ophthalmologists, dermatologists, or orthopedists, to avoid creating incentives for those non-emergency specialties to stint care.
- The PFE representatives provided input on which clinicians comprised their care teams:
    - Emergency trauma surgeons, an otolaryngologist, a clinical social worker, and clergyperson for an abdominal trauma admission.

- Emergency Medical Technician (EMT), registered nurse, resident physician, attending ED physician, and radiology technician for a pancreatitis admission. The patient consulted with their primary care provider to discuss the admission decision, and consulted with a gastroenterologist during and after their inpatient stay.
- Hospitalist and the ED staff.
- Nurses, doctors, and technicians in the hospital, with a physical therapist and orthopedist managing their post-acute care in a home health facility.
- Inpatient psychiatrists and psychologists, once the patient was admitted to the inpatient psychiatric unit with severe depression.
- Emergency medicine physician who was responsible for the decision to admit the patient to an inpatient setting, with a post-discharge team comprising a primary care physician and behavioral health specialist.

3.1.3 Assignment of Costs to the Episode Group

- One commenter expressed support for broad and uniform exclusions of service categories across the ED visit types, because emergency providers shouldn't be penalized for the cost factors of patients that have continuums of care managed by other disciplines.
- Several commenters provided suggestions on the following services that should be excluded from the measure:
  - Part B drugs due to their irrelevance to the ED visit.
  - Post-discharge home health services, given that they're almost always ordered by the clinician managing the patient's discharge from inpatient care, rather than by the ED clinician.
  - Major outpatient surgical procedures and anesthesia costs after 7 or 14 days from the ED visit, due to the ED physician's limited control over these services.
  - Part B drugs, chemotherapy, and infusions performed in the ED, as these services are subject to defined hospital fees that don't reflect the involvement of the emergency provider.
  - The following categories of services that aren't under the influence of emergency physicians, especially due to their high cost, to avoid having the measure inaccurately reflect the cost of emergency care:
    - Radiation oncology services
    - Chemotherapy
    - Dialysis
    - Ambulance services
    - Post-acute care services in Inpatient Rehabilitation Facilities (IRF), Skilled Nursing Facility (SNF), and Long-Term Care Hospital (LTCH) settings
    - Inpatient rehabilitation
  - Surgical interventions in the abdominal pain and orthopedic cases, and, if the episode window continues to be 30 days, excluding all surgical procedure costs after 14 days from the episode start date.
- Several PFE commenters stated that they or their family member received lab tests during their visit. They also further specified receiving the following services:
  - Fluids and antibiotics
  - Pain medications
  - Imaging, including advanced imaging, CT scan, X-ray, other scans
  - Electrocardiogram (EKG)
  - Blood work

- PFEs shared a range of experiences with follow-up care after discharge, including:
    - No follow-up care
    - Ongoing self-administered antibiotics post-discharge
    - Physical therapy
    - Wound care
    - Home health services - one PFE stakeholder noted that this was a condition of discharge

3.1.4 Risk Adjustment

3.1.4.1 Adjusting for Observation Status

- Stakeholders agreed with the need to account for observation stays.
- A few commenters highlighted particular types of observation stays that should be included in the measure:
    - A patient is placed in observation while still in the ED.
    - Observation stays for admitted patients, similarly to the methodology under the American College of Emergency Physicians' (ACEP's) Acute Unscheduled Care Model (AUCM) model, as it's difficult to differentiate between services provided in an ED observation unit vs. on observation status when admitted. This is also complicated by transfers from observation status to inpatient status that often occur.
- A stakeholder expressed concern about including observation stays where the care is provided by non-emergency clinicians as ED clinicians would mostly no longer be in control of direct care costs.
    - Many observation units are no longer near or in the ED - especially due to COVID - and might be managed by other providers, such as hospital medicine clinicians.

3.1.4.2 Risk Adjustment Model and Variables

- Several stakeholders provided feedback on the risk adjustment model that's currently used for the draft Emergency Medicine measure.
    - One commenter expressed general support for the risk adjustment model.
    - A few commenters noted concerns with the risk adjustment model.
        - One stakeholder said that rare but very costly events potentially bias parameter estimates and cost predictions.
- A few commenters provided feedback on the risk adjustment variables:
    - One commenter said that the primary factors for risk adjustment are accounted for in the current model, while several commenters provided suggestions on additional risk adjustors:
        - Use of hospice or palliative care because of the impact on the patient's cost profile.
        - One commenter noted that the costs for post-acute care services were much higher in the field test reports than expected, and that these services should be considered for adjustment.
    - Some stakeholders suggested further testing into whether there are differences in cost for the volume and timing of ED visits, including:
        - The difference between ED visits that occur during holidays, on working days, and weekends, as the ED is the only U.S. system that's open daily. The timing during the day may also be of interest.

- Total ED volume as Medicare only accounts for 20% of ED volume in the United States.
- Hospital size and type (i.e., critical access hospital vs. other hospitals).

3.1.5 Exclusions

- One stakeholder suggested excluding episodes that begin with a transfer to ED from post-acute care settings (i.e., IRF, long-term care homes, or long-term acute care [LTAC] facilities) instead of risk-adjusting for them.
    - These are highly complex patients with infections and other complications that result from both comorbidities and the care provided to the patient at facilities outside of the supervision of the emergency physician.

3.1.6 Alignment of Cost with Quality

- One commenter asked that CMS and Acumen share information comparing scores on quality measures with those on the Emergency Medicine cost measure.
- A few PFE representatives with experience with ED care provided input on the aspects of care that could have been improved:
    - Several PFE commenters emphasized the importance of communication to the patient experience. One commenter expressed satisfaction with the level of communication, while another suggested increased staffing to better accommodate patients.
    - One PFE representative noted that wait times are unpredictable for the discharge/admission decision, while another commenter said that the general discharge and care management processes by the hospital medicine department could be greatly improved.
    - One commenter shared their story about the neglect of a loved one in the inpatient setting, noting that the memory care floor would have been a more appropriate setting for a patient in that condition.
    - One PFE commenter flagged ED boarding as a major issue for time spent in the ED, which they saw as a sign of low quality.
    - One PFE commenter who was readmitted to the ED said that the repeat visit could have been avoided with better pre-discharge communication and coordination with the care team.

3.1.7 Measure Reporting and/or Use in MIPS

- N/A

3.2 Heart Failure

3.2.1 Definition of an Episode Group

3.2.1.1 Episode Group Definition

- A couple of commenters said that the triggering methodology is appropriate. A stakeholder agreed that inpatient encounters shouldn't trigger an episode.
- A few stakeholders provided suggestions about the set of services used to trigger episodes.
    - A commenter suggested removing codes such as Current Procedural Terminology (CPT) codes 99304, 99324, and telephone and internet

assessments (e.g., 99425) as they wouldn't be likely to initiate the patient-provider relationship.

- Some stakeholders suggested changes to the list of trigger diagnosis codes:
    - Focus on episodes where heart failure is the main diagnosis rather than an underlying condition
    - Only use a limited set of the International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) code I50 family. While it may reduce the beneficiary coverage of the measure, it would capture true cases of Heart Failure (HF) and exclude patients such as those with End-stage renal disease (ESRD) who present with signs of HF.

3.2.1.2 Terminating Conditions

- One commenter agreed with the use of left ventricular assist device (LVAD) and transplant as terminating conditions, and the way that the current methodology accounts for them (i.e., including all costs in the episode window until the relevant admission, but excluding the costs of the transplant or LVAD procedure).
- Some stakeholders suggested additional conditions to consider as terminating events as they fundamentally alter a patient's care trajectory:
    - Initiation of continuous inotropic support
    - Initiation of hospice and palliative care
    - Chronic Kidney Disease (CKD)
    - Extracorporeal membrane oxygenation (ECMO)
    - Cardiac shock
    - ESRD

3.2.2 Attribution of the Episode Group to Clinicians

- Some commenters said that the attributed specialties under the draft Heart Failure cost measure were appropriate.
- A few stakeholders noted that the following specialties should be attributed:
    - Nephrologists for the ongoing management of patients with HF and chronic CKD/ESRD
    - Emergency physicians
- PFE commenters identified the following specialties as being part of their care team, beyond a cardiologist and primary care clinician which were common across all respondents to this question:
    - Primary care clinician
    - Infectious disease specialist
    - Orthopedist
    - Hematologist
    - Internist
    - Electrophysiologist
    - Integrative clinician
    - Emergency physician
    - Transplant surgeon
    - Nephrologist
    - Registered nurses who monitored their cardiac resynchronization therapy device (CRT-D)
- Some stakeholders commented on the different roles of particular specialties:

    - HF cardiology is distinct from general, electrophysiology, and interventional cardiology, as there are differing board certifications; they also treat different levels of severity.
    - Electrophysiologists perform high-cost services that general cardiologists don't (e.g., device implantation and ongoing monitoring/adjustment over time). Other specialties provide lower-cost services like medication management.
    - Advanced practice providers (APPs) are engaging in more patient encounters that lead to better patient management due to an increase in team-based care approaches and therefore, APP encounters shouldn't be attributed because there's the potential for individual APPs to be attributed full costs of care for the episode. If APP encounters must be included, then they should be compared to the appropriate cardiology specialty and not default to primary care attribution.
- A few stakeholders commented on adjusting for different costs associated with specialty.
    - One suggested subgroups based on clinician type in order to group physicians with each other.
    - Another asked for more information about how the specialty adjustment accounts for higher costs of care.
- One commenter expressed concern about attributing individual clinicians rather than clinician groups because of the limited role one individual may play in a patient's relationship to a large practice.
- A few PFE stakeholders shared their experience with the ongoing management of HF.
    - One noted they had encounters several times per month, with different intervals for each specialty: cardiology every 3 months with yearly advanced cardiology visits, monthly hematology visits, and pacemaker readings every 3 months.
    - A PFE commenter said that regular visits indicated quality.
- PFE commenters shared a range of experiences with care coordination.
    - A couple of PFE commenters who had diverse care teams reported that the responsibility for care coordination fell to them and their families.
    - One PFE commenter expressed satisfaction with the level of communication and coordination facilitated via Epic.

3.2.3 Assignment of Costs to the Episode Group

- Stakeholders suggested various services to include in the measure:
    - Inpatient stays
    - Part D drugs
        - Sacubitril/valsartan, empagliflozin, dapagliflozin, and other Sodium-glucose Cotransporter-2 (SGLT2) inhibitors
        - Entresto
        - A few stakeholders suggested adding more drugs without specifying which ones.
    - Hospice, end-of-life, and palliative care services
- One commenter agreed with the list of Parts A/B services and Part D drugs that are included in the measure.
- One commenter expressed concern about including Part D drug costs without measuring the impact on quality and outcomes, because for patients with multiple medications, additional follow-ups are required to assess medication tolerance and effectiveness, and to manage titration. The commenter also noted that insurance coverage and adherence issues could influence utilization.

- A few stakeholders suggested removing services from the measure because they're concerned that their inclusion could be punitive for clinicians who are providing good quality care for the patient:
    - Medicines for pulmonary hypertension and immunosuppression for inflammatory disorders and cardiomyopathies, such as Cellcept
    - Inpatient procedures, including:
        - Heart valve procedures (surgical and percutaneous procedures)
        - Extracorporeal circulation auxiliary to open heart procedures
        - Cardiac Defibrillator Implant With Cardiac Catheterization With Acute Myocardial Infarction (AMI), HF, or Shock
        - Cardiac Defibrillator Implant With Cardiac Catheterization Without AMI, HF, or Shock
        - Cardiac Defibrillator Implant Without Cardiac Catheterization
        - Ecmo Or Tracheostomy With Mv >96 Hours Or Principal Diagnosis Except Face, Mouth And Neck With Major Operating Room (OR) Procedures
        - Tracheostomy For Face, Mouth And Neck Diagnoses Or Laryngectomy
        - Other OR heart procedures (LVAD, cardiac massage, great artery repair, congenital surgery, etc.)
        - Other Cardiothoracic Procedures
        - Other Major Cardiovascular Procedures
        - Other Endovascular Cardiac Valve Procedures
        - Permanent Cardiac Pacemaker Implant
        - Cardiac Pacemaker Device Replacement
        - Cardiac Pacemaker Revision Except Device Replacement
        - Automatic Implantable Cardioverter Defibrillator (AICD) Generator Procedures
        - AICD Lead Procedures
    - Devices, including:
        - Mitraclip
        - Cardiac Contractility Modulation
        - BaroStim
        - CardioMEMS
    - Outpatient inotropes therapy
- Several PFE stakeholders provided input on the services that were the most helpful for managing heart failure for them. These include:
    - Education and guidance on management of fluids, weight, sodium, or diet from an integrative clinician
    - Pacemaker and ablation helped to address the arrhythmia; medications for arterial fibrillation and cardioversion to reinstate the patient's rhythm didn't help with symptoms
    - Mental health support for HF-related depression; accessing this care was challenging due to the reluctance of their cardiologist and primary care doctor to refer the individual for these services
    - Anti-rejection medications after their heart transplant were a high-value service. Access to these medications are very important and could be improved.
    - CRT-D and daily exercise were the most valuable interventions for health
    - Telemedicine was a high-value source of care that helped to address lack of access to care due to distance

3.2.4 Risk Adjustment

3.2.4.1 Chronic Kidney Disease and Use of Dialysis

- Several stakeholders suggested accounting for CKD and ESRD. Some were in favor of risk adjusting, while a few suggested excluding this patient cohort.
    - One commenter specifically requested adjusting for CKD stage 3.
- A commenter noted that risk adjustor or exclusion could be based on the family of exclusively diastolic HF codes; volume overload from inadequate dialysis is often miscoded as diastolic heart failure, because many ESRD patients have diastolic dysfunction. These codes are:
    - I5030 Unspecified Diastolic (Congestive) HF
    - I5031 Acute Diastolic (Congestive) HF
    - I5032 Chronic Diastolic (Congestive) HF
    - I5033 Acute On Chronic Diastolic (Congestive) HF
- A few stakeholders suggested focusing on the use of dialysis:
    - One commenter suggested excluding patients with acute HF while on hemodialysis due to the challenges this intersection introduces to standard HF therapy, and to the difficulty of determining between renal and HF-induced volume overload.
    - Another stakeholder recommended testing the dialysis risk adjustor for reliability, and removing it from the model if it's not reliable.
- A commenter noted that CKD treatment may vary due to glomerular filtration rate (GFR) and tendency to hyperkalemia and other medications.

3.2.4.2 HFrEF and HFpEF

- Stakeholders generally agreed that the measure should differentiate between HFrEF versus HFpEF.
    - Most were in favor of risk adjusting for these even if there aren't meaningful differences in observed cost between the two groups based on Acumen's draft classification. One commenter noted that future therapies might have meaningful impacts that would further support the need to risk adjust.
    - One commenter favored stratifying episodes into HFrEF and HFpEF sub-groups rather than risk adjustment.
- Stakeholders suggested other ways of distinguishing between HFrEF and HFpEF:
    - Use modeling to act as a proxy for evaluating ejection fraction (EF) class, and review the updated American College of Cardiology (ACC) Universal Definition and Classification of HF, which provides a revised definition of HF and a classification revision of HF by left ventricular ejection fraction (LVEF). This could be beneficial in classifying patients based on risk with potential for enhanced treatment indications and assessment of guideline-directed medical therapy.
    - Blood tests
    - LVEF < 55% and LVEF > 55%

3.2.4.3 Other Variables

- A few commenters noted the need to account for severity of HF.
    - One stakeholder also acknowledged that it may be challenging to locate a patient on their trajectory of the syndrome based on claims alone (e.g., lack of information on left ventricular ejection fraction or HF type).
- Stakeholders suggested other factors to account for:
    - Obstructive sleep apnea

    - Non-adherence to medications and/or to medical recommendations as these cases can sometimes contribute substantial cost despite best efforts of medical team
    - High-output failure identified through ICD-10 diagnosis code I50.8 and its sub-codes (except for the sub-codes I50.814 and I50.82), as there's no proven treatment for this small and complex subset of patients. This cohort could be accounted for through risk adjustment or exclusions. High-output HF is currently excluded from the measure.
    - Social determinants of health, socioeconomic status, dual eligibility status in Medicare/Medicaid, other health equity factors
- Stakeholders also flagged patient populations with complex case mix, as follows:
    - One commenter expressed concerns with holding one clinician or clinician group accountable for health outcomes that aren't directly attributable only to the care provided by one clinician or clinician group over time, potentially leading to perverse incentives not to treat patients with underlying HF.
    - One commenter noted that valvular heart disease and CKD might not be easily categorized.
    - One PFE commenter noted that fatigue, edema, and an understanding of the cause of symptoms (i.e., HF versus arterial fibrillation) drove their frequency of care and choice of clinician.

3.2.5 Exclusions

- One commenter requested that the Heart Failure measure exclude patients with cardiac sarcoidosis and arrhythmogenic right ventricular cardiomyopathy (ARVC).
    - Neither of these conditions are true infiltrative cardiomyopathies,
    - Sarcoidosis often requires multiple expensive services, particularly within the first year of diagnosis (magnetic resonance imaging [MRI] and 2 or more fluorodeoxyglucose-positron emission tomography (FDG-PET) scans), and ARVC often requires ablation if ventricular tachycardia (VT) worsens which is a consequence of the specific pathology and isn't indicative of general care provided.
- One commenter suggested considering excluding an episode where the patient is diagnosed with acute heart failure with ESRD while on hemodialysis, and standard HF therapies are incapable of volume management.

3.2.6 Alignment of Cost with Quality

- A commenter suggested evaluating the following quality measures alongside the cost measure:
    - Qualified Clinical Data Registry (QCDR) measure ACCPIN3: HF: Patient Self Care Education
    - CMS #008: Beta Blocker use for LV dysfunction
    - CMS #005: ACEi/ARB/ARNI use for LV dysfunction
    - CMS #236: Controlling high blood pressure
- One stakeholder noted that many indicators of quality are important for value-based care: quality of life, care coordination, disparity reduction, palliative and end-of-life care, patient education, disease management, and safety.
- One commenter suggested accounting for appropriate use criteria for (Implantable Cardioverter Defibrillators) ICD implants.

3.2.7 Measure Reporting and/or Use in MIPS

- N/A

3.3 Low Back Pain

3.3.1 Definition of an Episode Group

3.3.1.1 Attribution Window

- One commenter advised that a 60-day confirming diagnosis window is too long because physical therapy-related visits typically occur within 24 hours and up to one month after the initial trigger event. They note that the 60-day confirming diagnosis window could capture claims not related to physical therapy care.
- A couple of stakeholders expressed concerns about weighting the number of days assigned to an episode.
    - One commenter suggested that an episode with a longer period of time between the initial evaluation to surgery, where a surgeon treats a patient with non-operative services first or has a busy schedule, could be weighed more heavily and that this could penalize surgeons for providing patient relief while awaiting surgery.
    - Another commenter stated that the scaling methodology for the measure is unclear and potentially penalizes clinicians who provide high-value services over shorter periods. They were concerned that this may incentivize clinicians to extend follow-up periods to reduce the concentration of costs.

3.3.1.2 Episode Triggers

- A stakeholder recommended including additional codes related to dry needling and remote therapeutic monitoring to the trigger and confirming diagnoses code list because they're effective and low-cost options for treating low back pain. The commenter recommended including the following codes: 20560, 20561, 98975, 98977, 98980, & 98981.
- One commenter recommended removing diagnoses regarding the thoracic spine or cervicothoracic junction and trigger codes for x-rays, CT scans, and MRIs of the "middle spine" because they believe these aren't related to low back pain.
- One commenter raised concerns that the current trigger code logic only captures chiropractors' Chiropractic Manipulative Treatment (CMT) codes, which doesn't align with how they provide care. They noted that patients receiving services like therapeutic exercises, modalities, and other non-CMT manual therapies when seen in a chiropractic office won't be captured because they're statutorily non-covered by Medicare when performed by a chiropractor. The commenter further noted that the CMS limits Chiropractic diagnoses codes to segmental dysfunction and subluxations. They also mentioned that secondary codes are limited to one per region, which in total will prohibit coding for the post-laminectomy subpopulation they often encounter.

3.3.1.3 Episode Sub-groups

- A few commenters highlighted that they agree with the sub-groups as presented in the draft measure specifications.
- Several stakeholders suggested considerations for additional sub-groups:
    - Chronic low back pain and new or acute low back pain to better account for the patient heterogeneity in the cost measure.
    - Previous spinal surgery as failed or recurrent spinal surgery help determine if spinal surgery treatment is successful.

    - Patients with post-acute care after spinal surgery, because this population has significantly different and unique courses of care.
    - Episodes with and without MRI imaging as these are costly services that can be avoided.
    - Patient weight and overall diet, as poor nutrition contributes to slower improvement rates and high systematic inflammation.
- A couple of commenters noted that sub-grouping for surgical and non-surgical episodes removes an opportunity to capture cost improvement, even though it's a way of accounting for patient heterogeneity.
    - This could encourage cost savings and better patient outcomes, as some spinal surgeries may be avoidable if clinicians provide appropriate earlier in an episode.
    - The presence of a spine surgery during a low back pain episode could be an outcome of low-value care of the attributed provider.

3.3.2 Attribution of the Episode Group to Clinicians

- A few commenters raised concerns about the role of surgeons being captured by this measure. For example, one stakeholder questioned why non-operative patients are being attributed to surgeons as that seems to assess surgeons on care outside of their role.
    - One stakeholder recommends that neurosurgeons and orthopedic surgeons be removed from the list of eligible specialties for attribution. They were specifically concerned with scenarios where a surgeon who advises a patient that they can't perform a surgery would be considered less costly than a surgeon who also advises that a patient can't perform a surgery, but refers the patient to alternative care (e.g., physical therapy, injections).
    - The stakeholder noted that the Lumbar Spine Fusion, Levels 1-3 measure in MIPS is intended to evaluate surgical cost and outcomes, which would be more appropriate to assess surgeons than under this measure as a rare surgical treatment of a chronic non-surgical disease.
    - They requested an analysis of clinicians attributed both the Lumbar Spine Fusion, Levels 1-3 measure and the Low Back Pain measure to assess potential overlap.
- A few stakeholders expressed concern about how physical therapists are being attributed and what care they're assessed on.
    - One stakeholder recommended that a physical therapist should be attributed by the first physical therapy service they provide and not via a referral from a medical provider, as a referral relationship can complicate whether a physical therapist can practice autonomously on the patient, or must coordinate care with the medical care team.
    - Commenters identified services over which a physical therapist has little control: behavioral health (e.g., outpatient psychotherapy and depression screening), imaging, injection, neuro-stimulators hospitalizations, post-acute care, spinal surgery, and medications.
- One commenter was concerned that the measure could be distorted and produce less meaningful results due to referrals and the availability of specialists within a practice.
    - The specific example was to compare the scenario where a practice could be assigned a trigger claim and confirming claim if a neurosurgeon refers a patient internally for an injection where the practice has pain management specialists, whereas if the patient receives the injection elsewhere and the neurosurgeon doesn't see the patient again, this breaks the trigger-confirming claim link.

- One commenter noted that grouping the costs of physicians with physical therapists could abnormally skew the episode costs, as physician costs are much higher than physical therapist costs.
- PFE stakeholders had different comments about care coordination. A few highlighted that their specialty care and primary care provider or care team coordinated well, while others noted that there was minimal coordination of care or it was dependent on which provider was ordering services.

3.3.3 Assignment of Costs to the Episode Group

- Most stakeholders responding to the question about nutritional and weight management services were in favor of including them as a cost effective way to reduce low back pain that can play a pivotal role in low back pain recovery.
    - A few commenters opposed the inclusion of these services over concerns it could disincentivize clinicians from using them in an effort to limit cost of care, weight management isn't under the control of attributed clinicians, it could be challenging to determine meaningful results from measuring diet or nutrition counseling, or that Medicare provides only a limited set of services.
- Almost all stakeholders responding to the question about including post-acute care services supported these for service assignment because it captured an appropriate scope of care for low back pain.
    - A few commenters expressed hesitation because it could skew individual clinician's performance and may be more directly related for certain episodes, such as those with a spinal surgery or hospitalization.
- A few commenters agreed that the Medicare Part D drugs included in the measure accurately captures clinically related medications and don't recommend any adjustments.
- Several commenters recommended adding services to the measure, some of which are already included in the measure:
    - Home exercise regimens and self-care (e.g., stretching)
    - Back braces
    - Additional physical and occupational therapy codes to teach about fear of movement and its impact on pain and function.
    - Inpatient and outpatient physical therapy; these can obviate the need for expensive imaging
    - Chiropractic adjustments
- A few commenters suggested specific services to exclude:
    - Certain operative Clinical Physical Therapist codes, specifically procedures on the thoracic spine, because they're out of scope for treatment of low back pain
- A few stakeholders noted limitations in what clinicians can influence.
    - A commenter noted that physical therapists may have limited care management authority over the patient as many of these services are dependent on the willingness of a primary care provider or other clinicians to collaborate.
    - A couple of stakeholders suggested that clinicians' measure scores should only include their individual costs, not the total cost for the episode, which includes services that might have been provided by other clinicians.
- One commenter expressed that the timing or sequence of services should be factored into the measure so that conservative care can be distinguished as good clinician performance.
    - The stakeholder gave the example that physical therapy prior to imaging or surgical intervention would reduce costs and recovery time for patients.

- Several PFE commenters highlighted services that they received to address low back pain, including pain management, medical cannabis, physical therapy, MRIs, muscle relaxers, acupuncture, injections, post-acute care, chiropractic adjustments, home self-care, and nutrition counseling.
- PFE stakeholders noted various services that they found effective or low-value. Some noted that post-acute care, MRIs, muscle relaxers, acupuncture, and injections were low-value, while some did find that muscle relaxers were valuable. Individuals most commonly listed physical therapy and chiropractic adjustments as the most helpful services.
- A few PFE commenters noted some barriers to their care, including transportation costs.

3.3.4 Risk Adjustment

- Several commenters suggested that the cost measure should include additional variables to the risk adjustment model to better account for patient heterogeneity, some of which are already in the measure:
    - Facet syndrome
    - Cognitive status
    - Patient age
    - Use of durable medical equipment (DME) (e.g., walkers, wheelchairs)
    - Diabetes mellitus
    - History of opioid use
    - Obesity
    - Anxiety or depression
    - Social determinants of health (e.g., income level, education attained, and ZIP codes)
    - Patient "care seeking" (i.e., plays a significant factor in physical therapy care that's difficult to control and can require educating patients)
- Many commenters supported risk-adjusting for the history of spine surgery, as failed spine surgery can result in long-term, continued low back pain. A few stakeholders recommended extending the lookback period for history of spine surgery.
    - One commenter suggested including the following codes to indicate a history of spinal surgery: M43.26 lumbar fusion, M43.27 lumbosacral fusion, and M43.28 sacroiliac combined with Z98.1: surgically caused fusion. They also recommended adding a modifier or additional codes to Z.98.1, (e.g., Z98.12 or Z98.13) to indicate past history of surgical fusion for a certain number of years prior. Another commenter recommended diagnosis codes for an encounter for other orthopedic aftercare (Z47.89), pain due to internal orthopedic prosthetic devices (T84.84XA, T84.84XD, T84.84XS), and other idiopathic scoliosis (M41.27). M41.27 is included in the risk-adjustment variable for scoliosis and other spinal deformities.
    - A few stakeholders preferred extending the lookback period (e.g., to 5 years) as patients very commonly have long-term low back pain, lower extremity (LE) symptoms, and gait difficulty following a low back surgery and patients may initiate additional care (e.g., physical therapy) between 1 and 4 years post-surgery.
- One commenter recommended risk-adjusting for intensity of coding.
- A few stakeholders suggested exploring other sources for risk adjustment.
    - One commenter noted that the quantity of diseases doesn't predict severity, and suggested exploring how to use the Patient-Reported Outcomes Measurement

Information System (PROMIS) developed by the National Institutes of Health (NIH) in the risk adjustment model to measure low back pain severity.
    - One commenter noted that obesity is a common comorbidity and suggested exploring additional ways to measure exercise habits, such as medical chart data, Choosing Wisely standards, or through self-audit results.
- One commenter was supportive of adjusting for variation across specialties, but was concerned that the adjustments would be applied to overly broad clinician categories (e.g., nurse practitioner, physician assistant) and so they may not be accurate. They also noted that the process for applying these specialty adjustments was unclear.

3.3.5 Exclusions

- A few stakeholders suggested further exclusions:
    - Spinal and metastatic tumors (i.e., HCC8: Metastatic Cancer and Acute Leukemia) because the patient population requires more aggressive treatment and is too complex to add as a risk adjustor
    - History of rheumatoid arthritis
    - Osteoporosis

3.3.6 Alignment of Cost with Quality

- N/A

3.3.7 Measure Reporting and/or Use in MIPS

- N/A

3.4 Major Depressive Disorder

3.4.1 Definition of an Episode Group

- One commenter noted that there's a lack of clarity on defining episodes and the relationship of a defined episode with reporting year for MIPS.
- One commenter suggested not including trigger codes for nursing facility visits because the codes could pick up patients who are not receiving ongoing management, including: 99304, 99305, 99306, 99307, 99308, 99310, 99309, and 99315.

3.4.2 Attribution of the Episode Group to Clinicians

- One commenter raised concerns that the chronic condition measures aren't capturing clinical psychologists appropriately because the methodology doesn't align with how psychologists work within healthcare (e.g., span multiple specialties, work within numerous settings, and treat patients with more than one diagnosis), or how mental and behavioral health conditions are best treated.
    - They note only 8% of clinical psychologists would be attributed to the cost measure for Major Depressive Disorder, which would leave them out of the cost performance category of the MIPS program.
- One PFE commenter identified the following as part of their care team: psychiatrist, therapist, psychopharmacology advanced practice registered nurse (APRN), internist, home health provider, neurologist, nursing practitioner, and social workers.

3.4.3 Assignment of Costs to the Episode Group

- One stakeholder raised concerns that certain medications outlined in the measure may not be used just for Major Depressive Disorder (MDD), are used in patients with MDD

but prescribed by other clinicians (e.g., for pain management), or aren't typically prescribed. These include:
    - Glucocorticosteroids
    - Antihistamines – Piperidines (aren't best practice to use for depression or anxiety)
    - Alpha-2 Receptor Antagonists (Tetracyclics)
    - GABA Receptor Modulator - Neuroactive Steroid
    - N-Methyl-D-aspartic acid (NMDA) Receptor Antagonists (can be used for dementia)
    - Premenstrual Dysphoric Disorder (PMDD) Agents
    - Vasomotor Symptom Agents
    - Movement Disorder Drug Therapy
    - Combination Psychotherapeutics
- Stakeholders had mixed feedback on the inclusion of Part D drugs in the measure.
    - Two commenters agreed with the Part D medications included in the measure.
    - One commenter recommended that the measure only include costs for Part D drugs that the attributed clinician provides.
- A couple of stakeholders suggested considering how to include additional services:
    - Emerging medications and treatments, such as new drugs for tardive dyskinesia that are costly but effective
    - Medication management, which should be considered as a downstream consequence of care
    - Care provided in extended stay assisted living facilities
- Stakeholders raised concerns that including certain services could disincentivize their use to keep costs low, especially if they're already underutilized. These include:
    - Patient transportation services
    - Psychotherapy
    - Pre-operative/lab work prior to Electroconvulsive Therapy (ECT)
    - Medication monitoring
    - Collaborative care services
- One commenter noted that it's difficult for clinicians to control the amount of psychotherapy (i.e., counseling) services delivered if they're not providing those services themselves.
- PFE commenters identified the following services as being the most helpful: nutrition counseling, physical therapy, in-home care, yoga, and talk therapy.

3.4.4 Risk Adjustment

- Some commenters would like to account for social determinants of health (e.g., homelessness, domestic violence), suicidal ideations or attempts, and potential treatment interruptions (e.g., physical hospitalizations) because they may be unrelated to treatment but could negatively affect the cost of care.
    - The commenter recognizes that capturing some of these variables may be difficult through claims data.
- Stakeholders stated that there are significant geographic variations in costs and ability to access mental health services, as some communities are severely lacking mental health care resources.
    - Specifically, a commenter highlighted its role in depression treatment costs. The commenter expressed that if the cost differences related to ECT and Transcranial Magnetic Stimulation (TMS), medications, psychotherapy, and more aren't accounted for, then psychologists will be attributed high cost interventions

like ECT, when they predominantly deliver much lower costs treatments such as psychotherapy.

- Stakeholders suggested the following indicators of treatment-resistant depression (TRD):
    - Two or more failed trials of antidepressant medication (or 3 to 4 failed trials for a higher threshold)
    - Emergency department visits or crisis admissions for depression, often classified as observation stays and not as inpatient admissions
    - A history of living permanently in assisted living facilities
- One stakeholder believes that the measure would pick up TRD as it's currently specified, but that it would only capture the most severe cases, missing a large number of patients with TRD.

3.4.5 Exclusions

- One commenter suggested excluding episodes for patients above 75 or 80 years of age.

3.4.6 Alignment of Cost with Quality

- N/A

3.4.7 Measure Reporting and/or Use in MIPS

- Stakeholders recommended that if the measure is used in MIPS, it should be informational-only for the first 1 to 2 years so that attributed clinicians can familiarize themselves with the measures' scoring methodology before it affects their performance scores.

3.5 Psychoses/Related Conditions

3.5.1 Definition of an Episode Group

- One commenter asked for clarification on how episodes are defined and the relationship of a defined episode within a MIPS reporting year.

3.5.2 Attribution of the Episode Group to Clinicians

- N/A

3.5.3 Assignment of Costs to the Episode Group

- One stakeholder recommended that the measure only include costs for Part D drugs that the attributed clinician provides.
- Some commenters expressed concern that post-discharge services are included in the measure because they believe the services are outside the control of inpatient psychiatrists.
    - Commenters expressed concerns that by including these services, the measure could exacerbate problems with accessing care and potentially have unintended negative effects on the vulnerable patient population if clinicians take cost-cutting actions.
- One commenter recommended excluding patient transportation costs because it might disincentivize their use to keep costs low.
- A couple of commenters advised that increasing care coordination efforts by inpatient psychiatrists or their treatment team would have little or no effect when there's a fragmented mental health system and insufficient community resources available.

    - Commenters noted that most readmissions occur at other facilities and the initial psychiatrist will have no feedback on the costs or readmission status of individual patients.
- One stakeholder noted that it's difficult for clinicians to control the amount of psychotherapy (i.e., counseling) services delivered if they're not providing those services themselves.
- One PFE commenter highlighted experiencing a well-coordinated inpatient psychiatric team and a smooth transition from an inpatient to outpatient program. They noted that the outpatient program didn't coordinate well when a higher level of treatment was required.
    - They noted that mental health diagnoses and treatment seem behind other disciplines because treatments were based on the clinicians' personal field experience. They highlighted the need for additional diagnostic tests, and better collaborative efforts for diagnosing and treating mental health issues.

3.5.4 Risk Adjustment

- Some commenters raised concerns about accounting for social determinants of health (e.g., homelessness) and potential treatment interruptions (e.g., physical hospitalizations) because they may be unrelated to treatment but could negatively affect the cost of care.
    - One commenter acknowledged that the performance metrics demonstrated that the social determinants didn't consistently affect costs after correcting for multiple other comorbidities. They continued to raised concerns that many of these individuals are dual-eligible and the number who are receiving Medicare due to significant disability is reflected in the low age of the psychosis patients relative to patients in all other episode-based measures.
- One commenter suggested examining other psychiatric comorbidities as factors outside the influence of the clinician that could affect the expected spending.
- Stakeholders stated that there are significant geographic variations in costs and ability to access mental health services, as some communities are severely lacking mental healthcare resources.
- Specifically, a commenter highlighted its role in depression treatment costs. For example, they note that there are cost differences related to Electroconvulsive Therapy (ECT) and Transcranial Magnetic Stimulation (TMS), medications, and psychotherapy.

3.5.5 Exclusions

- N/A

3.5.6 Alignment of Cost with Quality

- N/A

3.5.7 Measure Reporting and/or Use in MIPS

- Stakeholders recommended that if the measure is used in MIPS, it should be informational-only for the first 1 to 2 years so that attributed clinicians can familiarize themselves with the measures' scoring methodology before it affects their performance scores.
- One commenter stated that there are significant costs associated with compliance with the MIPS program, and that the costs for initiating programs to be responsive to measures such as the Psychoses/Related Conditions measure are likely to be greater

than the cost incentives of improved performance. For example, a practice adopted the Patient Relationship Category (PRC) codes to assist with cost measure attribution, but it required costly interventions to comply with data collection, coding, and submission requirements.

4.0 Next Steps for Measure Specification Refinements

This section outlines the discussion topics and subsequent questions that Acumen brought to the Clinician Expert Workgroups during the Post Field Test Refinement webinars. Acumen identified these topics for discussion largely based on stakeholders' feedback gathered during field testing and subsequent empirical analyses. The Clinician Expert Workgroups' discussions about these questions directly help to inform refinements to the measures' specifications.

4.1 Emergency Medicine

The following discussion topics were brought to the Emergency Medicine workgroup regarding ED visit types:

- Categorization of ED visit types
    - Re-organize existing categories
    - Rename category
    - Account for other symptoms and conditions that are commonly seen in the ED
- Use of information from the inpatient stay if the patient is admitted (i.e. Medicare Severity Diagnosis Related Groups [MS-DRG]) in the ED visit type algorithm
- ED visits leading to admission versus discharge: should these continue to be risk-adjusted or sub-grouped?

Topics also covered the episode window length and service assignment:

- Episode window length: should this be shortened from the current 30 days?
- Service exclusions either across all ED visit types or for specific ED visit types

The workgroup also considered particular care scenarios and patient risk factors, and how the measure should account for them:

- Observation stays: currently risk-adjusted, are any other approaches needed to account for this care pathway?
- Risk adjustor for transfers who enter ED having been transferred from PAC

4.2 Heart Failure

The following discussion topics were brought to the Heart Failure workgroup regarding trigger codes, exclusions, and attribution:

- Inclusion of Cardiomyopathy diagnoses in trigger codes
- Confirmation of existing exclusions and the addition of new exclusions
    - ARVC and sarcoid (ARVC is included)
    - Diastolic HF (included)
    - ESRD (included, but risk-adjusted)
- Use of billing threshold and prescribing patterns in attribution

The workgroup considered the following refinements to the service assignment rules:

- Inclusion of sacubitril and/or valsartan, which are not currently included
- Exclusion of several services, which are currently included:
    - Cardiac devices
    - New technology (e.g., Barostim, cardiomems) and cardiac infusions
    - Respiratory Failure Admission
    - Medication for Pulmonary Hypertension (PHTN)

    - Heart valve, ECMO, tracheostomy, OR heart procedures, other cardiothoracic procedures

The workgroup also discussed risk adjusting for patients with the following:

- CKD and ESRD
- Obstructive sleep apnea
- Non-adherence to medications and/or medical recommendations

4.3 Low Back Pain

The following discussion topics were brought to the Low Back Pain workgroup regarding triggering and attribution methodology:

- Confirming claims: should imaging services be removed and/or dry needling and remote therapeutic monitoring services be added?
- Removing the following from trigger logic:
    - Imaging for the middle spine
    - Diagnoses related to thoracic spine, cervicothroacic junction, pain in the leg or non-spine specific diagnoses

The workgroup considered the following adjustments to the measure's sub-groups, risk-adjustment, and exclusions:

- Sub-groups: should laminectomy codes for surgical episodes be added to the existing list of sub-groups?
- Risk adjustment:
    - Spinal surgery: should this continue to be risk adjusted with a one year look-back period?
    - Any additional risk adjustment variables
- Exclusions:
    - Adding spinal neoplasms
    - Expanding the definition of spinal infection

The workgroup also discussed capturing variation in care through service assignment, specifically the inclusion of PAC services and nutritional services.

4.4 Major Depressive Disorder

The following discussion topics were brought to the Major Depressive Disorder workgroup regarding how to best define the care relationship through the triggering methodology.

- Evaluation & management (E&M) codes: should nursing facility E&M codes be removed?
- Including additional trigger codes to capture more clinical psychologists

Topics also covered reorganizing risk adjustors to account for patient heterogeneity:

- Including indicators of TRD
- Excluding patients aged 75 or older
- Accounting for social risk factor or differences in care by specialty

Additionally, the workgroup considered how to capture variation in care through service assignment:

- Part D costs: are there any reasons to reconsider their inclusion?
- Including costs of common comorbidities (e.g., substance use)

- Including additional diagnosis codes related to MDD
- Considering further action for the following categories of services suggested during field testing:
    - Care in extended stay assisted living facilities
    - Medication Management
    - Emergent medications and treatment
    - Transportation costs

4.5 Psychoses/Related Conditions

The following discussion topics were brought to the Psychoses/Related Conditions workgroup regarding how to account for Inpatient Prospective Payment System (IPPS):

- Risk adjusting for IPPS
- Sub-grouping for IPPS and Inpatient Psychiatric Facility (IPF)
- Not adjusting or sub-grouping

The workgroup also discussed how to account for the measure's potential to exacerbate access to care issues for vulnerable populations.

Finally, the workgroup considered whether to include post discharge codes, such as readmissions, in the post-trigger window.

Appendix A: List of Commenters

This appendix provides an index of stakeholders who submitted a comment during field testing. Though commenters who provided feedback and didn't include their name or organization aren't included in this table, their input has been included in the report.

Table A1. Stakeholders Providing Feedback on the 2022 Field Testing

Name	Individual or Representative	Organization
Rachel Groman	Representative	American Association of Neurological Surgeons
Madison Switalia	Representative	American Academy of Ophthalmology
Matthew Grierson	Representative	American Academy of Physical Medicine and Rehabilitation
Fareen Pourhamidi	Representative	American College of Cardiology
Jeffrey Davis, Debra Clark	Representative	American College of Emergency Physicians
Jennifer McLaughlin	Representative	American Medical Association
Rachel Groman	Representative	Congress of Neurological Surgeons
Jessica Peterson	Representative	Marsden Advisors
Rebecca Yowell, Andrew Lyzenga	Representative	American Psychiatric Association
C. Vaile Wright	Representative	American Psychological Association
Robyn Hursley, Timothy William Flynn, Lauren Masi, Debra Clark	Representative	American Physical Therapy Association Private Practice Section
Kate Gilliard	Representative	American Physical Therapy Association
Paul Kim	Representative	Atrium Health Musculoskeletal Institute and SpineFirst
Lauren Heeter	Representative	Atrium Health
Mark Drazner	Representative	Heart Failure Society of America
Fred Kusumoto	Representative	Heart Rhythm Society
Andrew Waligora	Individual	
Angela Pennisi	Individual	
Bethany Austin	Individual	
Christine Norton	Individual	
Cindy Kosh	Individual	Yale Medicine
Connie M Lewis	Individual	
Craig O'Neil	Individual	Upstream Rehabilitation
Dario Grisales	Individual	Pan America Pain Institute PL
Fred Dery	Individual	
Garth Barbee	Individual	Northwest Acute Specialists
Gaye Hyre	Individual	
Gerri Lynn Baumblatt	Individual	
Glenn Kopelson	Individual	
Harold Kraft	Individual	Laser MD Pain Relief
Jack Janning	Individual	
Janine Sladewski	Individual	
Jason M Shawbel	Individual	US Acute Care Solutions
Julie Rose Clayton	Individual	Lake View Family Chiropractic Clinic
Karen Case	Individual	SCP Health
Karen M Fernandes	Individual	
Kaylin Hansen	Individual	Superior Physical Therapy
Kelley Morrison	Individual	Adult Primary Care, LLC
Leigh Langerwerf	Individual	Butte Premier Physical Therapy

Name	Individual or Representative	Organization
Linnea Singleton Comstock	Individual	
Lynn Ferguson	Individual	
La Wanda Fleisher	Individual	
Marc Gruner	Individual	
Matthew J. Smith	Individual	
Melody Cromer	Individual	
Nicole A Ziehl	Individual	
Patricio Emilio	Individual	Tidewater Emergency Medical Care
Roger A. Lacoy	Individual	
Rosie Bartel	Individual	
Roxane Deakter	Individual	Office of Neil Schultz MD
Sherry LaBonte	Individual	Alexander Orthopaedic Associates
Sri Sundaram	Individual	
Steve Cassabaum	Individual	
Susan Nedza	Individual	
Troy Sturgill	Individual	American Chiropractic Association
Vicky Oldfield	Individual	ATV Health Solutions
Angelina Razo	Individual	
Judith Cheryl Olsen	Individual	
Sharon Roberts	Individual	
Luis Contreras	Individual	El Paso County Hospital District
Leo J Bronston	Individual	American Chiropractic Association
Craig Johnson	Individual	Therapy Partners, Inc.
Sara Smith	Individual	TeamHealth
Mbutambe Miranda	Individual	Cityworld Family Practice
Marcus Nynas	Individual	