Low Back Pain Post-Field Test Refinement (PFTR) Meeting Summary

MACRA Episode-Based Cost Measures: Clinician Expert Workgroups
PFTR Webinar, April 11, 2022
June 2022

Contents

Project Overview **1**
Low Back Pain PFTR Webinar, April 11, 2022 ..... **2**
1. Overview 2
2. Summary of Sessions and Discussion 2
2.1 Person and Family Partner (PFP) Findings and Discussion 3
2.2 Defining the Episode Group 3
2.3 Addressing Sub-Populations for Meaningful Clinical Comparison 5
2.4 Assigning Services to the Episode Group 8
2.5 Next Steps 9

Project Overview

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). Acumen's measure development approach involves convening clinician expert panels to provide input in cycles of development ("Waves").[1] In Wave 4, instead of the typical Clinical Subcommittee (CS) process for episode group prioritization and selection, we obtained stakeholder input on candidate clinical areas and episode groups through a public comment period from December 16, 2020, to February 5, 2021.[2] This approach provided flexibility for a wider range of stakeholders to participate around their schedule. The prioritization criteria used to identify strong candidate episode groups and concepts were developed based on input from our technical expert panel (TEP), Person and Family Engagement (PFE), CS, and Clinician Expert Workgroups ("workgroups"). The following Wave 4 episode groups were finalized based on the prioritization criteria, public comments received, and discussions with CMS: (i) Emergency Medicine, (ii) Heart Failure, (iii) Low Back Pain, and (iv) Depression.

We held a nomination period for workgroup members between April 26, 2021, and May 21, 2021. The workgroups are composed of clinicians with expertise directly relevant to the selected episode groups. Workgroups (of about 15-20 members) were finalized in June 2021, and they provided detailed input on the development of the selected episode groups during their first workgroup webinars from June 21 to June 24, 2021. Acumen convened the workgroups again for a Service Assignment and Refinement (SAR) Webinar to revisit the specifications recommended during the workgroup webinar and refine the measures prior to national field testing. After the national field test from January 10, 2022 to March 25, 2022, Acumen convened

[1] For information on measure development in Waves 4, refer to the 2022 Episode-Based Cost Measures Field Testing Wave 4 Measure Development Process document (https://www.cms.gov/files/document/wave-4-measure-development-process-macra.pdf).
[2] For a summary of comments we received during the public comment period, refer to the MACRA Episode-Based Cost Measures: Wave 4 Measure Development Public Comment Summary Report document (https://www.cms.gov/files/document/wave-4-public-comment-summary.pdf).

the workgroups for a third meeting to continue measure specification and refinement discussions in April 2022. For Wave 4, all workgroup meetings were held virtually.

Low Back Pain PFTR Webinar, April 11, 2022

This meeting summary document outlines the purpose, discussion, and recommendations from the Low Back Pain PFTR Webinar. Section 1 provides an overview of the webinar goals and process. Section 2 summarizes the discussion and recommendations from the workgroup.

1. Overview

The goals of the Low Back Pain PFTR Webinar on April 11, 2022, were the following:

(i) Discuss field testing feedback
(ii) Review empirical analyses
(iii) Confirm refinements to finalize the measure prior to submitting for potential consideration in MIPS

The meeting was held online via webinar and attended by 16 of the 22 workgroup members. The webinar was facilitated by an Acumen moderator, Walter Park. The Low Back Pain workgroup chair was Dheeraj Mahajan, who also facilitated meeting discussions. Person and Family Partner (PFP) representatives Lynne Ferguson and Lisa Freeman attended the webinar to discuss and address questions regarding the PFP findings. The MACRA Episode-Based Cost Measure Workgroup Composition List contains the full list of members, including names, professional roles, employers, and clinical specialties.[3]

Stakeholders beyond the workgroup members had access to a public dial-in number to observe the meeting as part of Acumen's continued effort to increase the transparency of the measure development process.

Prior to the webinar, workgroup members were provided with information and materials to inform their meeting discussions. After the webinar, workgroup members were sent a recording of the webinar, supplemental analyses and documentation, and were polled on their preferences to ensure the measures are developed based on well-documented stakeholder input. Based on National Quality Forum practices, the threshold for support was greater than 60% consensus among poll responses. This document summarizes the workgroup members' input from both the discussion as well as the polls.

This meeting was convened by Acumen as part of the measure development process to gather expert clinical input; as such, these are preliminary discussions and materials, which don't represent any final decisions about the measure specifications or MIPS.

2. Summary of Sessions and Discussion

This section describes workgroup member discussions and recommendations. The first sub-section summarizes the PFP findings discussed in the webinar (Section 2.1). The remaining sub-sections describe workgroup member discussions and recommendations on defining the episode group (Section 2.2), addressing sub-populations of interest for meaningful clinical

[3] CMS, "MACRA Episode-Based Cost Measures: Wave 4 Clinician Expert Workgroup Composition (Membership) List" (https://www.cms.gov/files/document/wave-4-measure-specific-workgroup-composition-list.pdf).

comparison (Section 2.3), and assigning services to the episode group (Section 2.4). Section 2.5 describes the next steps.

2.1 Person and Family Partner (PFP) Findings and Discussion

The attending PFPs summarized PFP responses to the field testing survey. PFPs and other person and family stakeholders provided their thoughts on the helpfulness of the services they received for their low back pain. All 10 respondents agreed physical therapy was beneficial for managing low back pain. Some respondents identified additional services like chiropractic care, Pilates, stretching, strengthening exercises, and nutrition as helpful for managing low back pain. Other respondents noted that acupuncture and muscle relaxants were less helpful. Two PFPs reported undergoing surgery to alleviate low back pain; one reported being in good health while the other reported chronic pain persisting 10 years later.

PFPs also identified barriers impacting the management of their low back pain. They noted limitations on the number of physical therapy sessions covered by Medicare and how it made it difficult to access physical therapy as a primary source of low back pain management. One PFP respondent explained that if physical therapy sessions were more accessible and had greater availability, they wouldn't need to engage in costlier treatments like surgery or post-acute care (PAC) services. PFP respondents said better care coordination and communication between primary care teams, physical therapists, pain specialists, and orthopedists could improve low back pain management.

2.2 Defining the Episode Group

Workgroup members discussed the extent to which the measure captures the clinician role in managing and treating low back pain (Section 2.2.1) and additional refinements to the trigger logic (Section 2.2.2).

2.2.1 Capturing the Clinician Role in Managing and Treating Low Back Pain

The intent of the measure is to assess the treatment and management of low back pain, defined as lower spinal conditions excluding malignancies, infections, and fractures, amongst others. When development work on this measure began, the workgroup discussed and voted on the services and diagnoses that identify the start of a care relationship for low back pain, referred to as "trigger logic," "trigger services," and "confirming services." The trigger logic was designed to capture the range of clinicians who have this role in treating and managing low back pain, including physicians, chiropractors, physical therapists, and others.

One concern raised during field testing was whether the trigger logic is inadvertently picking up a solely pre-operative or consultative type of care, typically provided by surgeons, rather than treatment or management of low back pain. Some commenters questioned why neurosurgeons were being attributed non-surgical episodes in their field test reports and questioned what care was being assessed. Separately, a stakeholder noted that some neurosurgeons may refer patients to other types of care to manage and treat low back pain, such as physical therapy or spinal injections.

Acumen presented analyses on the length of time between the trigger service and spine surgery. The findings were intended to support discussions about potential trigger logic modifications to minimize the risk of including episodes that are purely pre-operative or consultative, rather than treatment and management of low back pain. One way of interpreting this analysis is to consider whether the shorter the gap, the more likely that it's a pre-operative role. Another potential signal is to compare the length of time across the specialties that would be likely to have a pre-operative or consultative role surrounding a spine surgery (i.e., those

who perform or assist surgeries). When examining episodes that have a spinal surgery within the year following the trigger service, the mean number of days from trigger to surgery tended to be lowest when the trigger and confirming codes were outpatient evaluation and management (E&M) services paired with imaging services.

Several workgroup members questioned what type of care being provided by surgeons is being assessed and posited that surgeons should be excluded from attribution. Their comments included the following:

- Surgeons should only be assessed on their performance for surgeries; this is the core nature of their role and potential payment adjustments should be determined based on this.
- The existing Lumbar Spine Fusion procedural episode-based cost measure already measures some of the care provided by surgeons. Spinal fusions and other lumbar spine surgical interventions are currently included in the Low Back Pain measure.
- The majority of patients seen aren't candidates for surgery. In these cases, the surgeon's role can include referring the patient to physical therapy or pain management services. There were comments during field testing and the webinar about surgeons seeing this reflected in their field test reports, where their attributed episodes were largely non-surgical.
- The distinction between this and other cost measures is that low back pain involves a range of diagnoses, rather than a specific condition. A workgroup member noted this may pose challenges for attribution.

Following the webinar, the Acumen team provided workgroup members with additional analyses about the trigger logic and overlap with the Lumbar Spine Fusion measure to inform their poll responses about potential modifications to the measure specifications (e.g., removing imaging as a confirming service, excluding surgical episodes when a surgery occurs soon after the initial trigger service).

2.2.2 Additional Refinements to Trigger Logic

The workgroup also reviewed other potential changes to trigger and confirming codes suggested during field testing. Dry needling and remote therapeutic monitoring services were suggested as confirming services during field testing. During the meeting, workgroup members inquired about Medicare coverage for the services. The services are covered by Medicare, though coverage may be limited to certain circumstances.

During field testing, imaging for the middle spine was suggested to be removed, as it isn't related to low back pain. However, removing these codes might reduce the discriminatory power of the measure due to variation in coding practices. Note that these will be removed if consensus is reached to remove imaging services entirely from the trigger logic. They will still be included as assigned services, regardless of any updates to trigger logic.

Thoracic spine, cervicothoracic junction, and non-spine specific diagnoses were suggested during field testing to be removed, as they aren't related to low back pain.

<u>Key Takeaways from Poll Results for Defining the Episode Group</u>:

- Workgroup members agreed the role described by neurosurgeons is appropriate to include in the measure and aligns with the type of care intended to be measured.

- A majority of workgroup members recommended updating the measure specifications to minimize the risk of identifying relationships that are solely pre-operative or consultative:
    - Workgroup members recommended excluding surgical episodes when the surgery occurs soon after the initial trigger service, with the majority recommending excluding episodes in which surgery occurs within 60 days.
    - Workgroup members didn't recommend removing imaging services from the trigger logic.
- Workgroup members recommended adding codes for dry needling and remote therapeutic monitoring as confirming services.
- Workgroup members didn't recommend removing codes for X-ray, computed tomography (CT), and magnetic resonance imaging (MRI) of the middle spine as confirming services.
- Workgroup members recommended removing thoracic, cervicothoracic, and non-spine specific codes from the trigger logic.

2.3 Addressing Sub-Populations for Meaningful Clinical Comparison

Members also engaged in a detailed discussion about how to account for patient cohort heterogeneity among various sub-populations within the low back pain episode group. Sub-populations refer to patient cohorts as defined by their pre-existing conditions and characteristics. Workgroup members discussed:

(i) Stratifying the patient population into mutually exclusive and exhaustive sub-groups to define more homogenous patient cohorts[4]
(ii) Defining covariates in the risk adjustment model[5]
(iii) Identifying measure exclusions[6]

Workgroup members discussed the patient sub-populations and their preferences for how to address them. They discussed their preferences for episode sub-groups (Section 2.3.1), risk adjustment (Section 2.3.2), and measure exclusions (Section 2.3.3).

2.3.1 Sub-groups

During the previous webinar, the workgroup discussed pros and cons of different approaches of accounting for spine surgeries. The majority preferred to include spinal surgeries in the measure through sub-grouping, though the workgroup didn't meet the consensus threshold. Based on this and other feedback from the workgroup, we field tested the measure using 4 sub-groups: (1) surgical episode with history of complex low back pain, (2) surgical episode without history of complex low back pain, (3) non-surgical episode with history of complex low back pain, and (4) non-surgical episode with no history of complex low back pain.

---

[4] Sub-grouping is a method that's intended for when we would want to compare episodes only with other similar episodes within the same sub-group. This approach is used when sub-groups are very different from one another, and each sub-group requires its own risk adjustment model. Since each sub-group will have its own risk adjustment model, the size of each sub-group should be sufficiently large.

[5] Risk adjusting is a method to account for the case-mix of patients and other non-clinical characteristics that influence complexity. It's meant to be used for sub-populations that make a large share of patients who have a characteristic that's outside of the attributed clinician's reasonable influence. Risk-adjusted cost measures adjust observed episode spending to an expected episode spending (predicted by a risk adjustment model).

[6] Excluding is a method in which we exclude certain patients or episodes to address issues with patient heterogeneity. This approach should be used when the sub-population affects a small, unique set of patients in which risk adjustment wouldn't be sufficient to account for their differences in expected cost.

Acumen reviewed field testing comments, which were generally in favor of accounting for the cost differences of surgery through sub-grouping. A small number of stakeholders noted that while this was a way of accounting for patient heterogeneity, it neutralized what could be a substantial area for cost improvement. There were also comments more generally about the measure intent and what type of care provided by surgeons was being included. Acumen also presented analyses showing the large difference in observed cost between surgical and non-surgical episodes effectively being neutralized once the risk adjustment model is applied.

Workgroup members then discussed the inclusion of surgical episodes in the measure. Among workgroup members who wanted to continue to include episodes with spine surgery and sub-group for surgical and non-surgical episodes, they noted the following:

- Including surgical episodes captures downstream costs, including services provided by other clinicians.
- While surgical episodes are a small percent of all episodes, the observed cost is much greater than non-surgical episodes, and therefore, represent a significant portion of costs.
- The sub-grouping approach assesses surgical judgment and decision-making, while also not penalizing clinicians for having surgical episodes.
- Spine surgeries do take place for patients with low back pain, so these costs should be included.
- The measure is interdisciplinary and should include care provided by a range of clinicians.

Other workgroup members expressed a preference to exclude episodes with spine surgery and provided the following comments:

- Spine surgery isn't part of low back pain care; rather, it's for spinal conditions that can't be determined through claims data.
- It would be easier to exclude surgical episodes given that they make up a small percent of total episodes.
- Including surgeries and neutralizing their cost through sub-grouping makes the intent of the measure unclear.
- The costs of spinal surgery may impact the decision-making and performance of other clinicians who don't provide surgeries (e.g., primary care clinicians).
- Spine fusions are included in the procedural Lumbar Spine Fusion measure.

2.3.2 Risk Adjustment

The base risk adjustment model already includes the standard set of risk adjustors from the CMS-Hierarchical Condition Categories (HCC) version 22 in 2016, disability status, End-Stage Renal Disease (ESRD) status, comorbidity interaction variables, recent long-term care use, HCC count, and measure-specific clinical risk adjustors (spondylolysis, scoliosis and other spinal deformities, recent hospitalization for medical back problems, osteoarthritis, osteoporosis, depression, smoking, and history of prior spine surgery in previous year).

Field testing commenters suggested additional risk factor variables relevant to the low back pain measure (i.e., history of opioid use, cognitive status/dementia, and frailty or frailty proxies). Workgroup members were generally supportive of including the suggested risk factor variables; one workgroup member also suggested risk adjustment for fibromyalgia.

Additionally, workgroup members revisited previous discussions about the importance of including patient history of spinal surgery in the risk adjustment model. Field testing comments noted that longer lookback periods can be useful in capturing history of spinal surgery that may impact ongoing care for low back pain. While deciding to use a longer lookback period results in some trade-offs, namely measuring fewer episodes due to data completeness requirements, testing indicates that a 365-day lookback period captures twice the number of patients than 120 days. The workgroup seemed to generally be in agreement during the meeting that the 365-day lookback period was preferable to 120 days.

Workgroup members also reviewed additional service and diagnosis codes that could be used to identify and risk adjust for history of prior spinal surgery. There was one set of service codes representing laminectomy procedures (currently not part of the risk adjustment model) that seemed related to the other types of spinal surgery that are included in the definition. There were also suggestions from field testing comments for diagnosis codes that are non-specific to spine surgery; as such, they could inadvertently capture unrelated care that doesn't predict higher expected costs.

2.3.3 Measure Exclusions

Field testing commenters identified potential exclusions to apply for the measure. Several suggestions (cancer, rheumatoid arthritis, and osteoporosis) are already part of the measure through the standard risk adjustment model and measure-specific risk adjustment variables. During the webinar, the workgroup reviewed analyses that showed differences in observed costs among these episodes were neutralized after risk adjustment, suggesting the measure already adequately accounts for these variables.

The workgroup also discussed that the trigger logic doesn't include diagnoses for spinal neoplasms or spinal infections, as these conditions weren't intended to be captured through the measure. Additionally, certain spinal infections are already expressly excluded from the measure. As a further step to ensure that episodes for patients with spinal neoplasms and spinal infections aren't included in the measure, additional diagnosis codes can be added to the list of expressly excluded conditions.

<u>Key Takeaways from Poll Results for Addressing Sub-Populations for Meaningful Clinical Comparison</u>:

- Workgroup members agreed with continuing to include surgical and non-surgical episodes using the sub-grouping approach implemented during field testing.
- Workgroup members agreed laminectomy procedures should be part of the surgical sub-group definition.
- Workgroup members agreed with continuing to use a 365-day lookback period to risk adjust for patient history of prior spinal surgery.
- Workgroup members voted on including additional codes in the definition of the risk adjustor for patient history of prior spinal surgery; they recommended to:
    - Add laminectomy codes
    - Add codes for pain due to internal orthopedic prosthetic devices, implants and grafts
    - Not add the code for encounter for other orthopedic aftercare
- Workgroup members recommended adding the following variables to the risk adjustment model:
    - History of opioid use
    - Cognitive status/ dementia
    - Fibromyalgia
    - Frailty proxies (e.g., use of wheelchair, walkers)

- Workgroup members recommended excluding episodes for patients with spinal neoplasms.
- Workgroup members recommended using additional spinal infection codes to exclude episodes from the measure.

2.4 Assigning Services to the Episode Group

Acumen described the purpose of service assignment so that members could continue discussing which services associated with the attributed clinician's role in managing the patient's care should be included in the cost measure. These assigned services should be inclusive enough to identify a measurable performance difference between clinicians but also not introduce excessive noise. The following sub-sections summarize discussions of the categories of assigned services.

Workgroup members engaged in discussions on whether to assign PAC (Section 2.4.1) and nutritional services (Section 2.4.2) to the episode group.

2.4.1 Post-Acute Care Services

After the previous webinar, the majority of workgroup members supported the inclusion of home health, inpatient rehabilitation facility (IRF), and long-term care hospital (LTCH) services, though the consensus threshold wasn't met for LTCH. Field testing commenters were also generally supportive of including PAC services related to low back pain, and emphasized the importance of only assigning PAC services when provided in relation to low back pain. This aligns with the current service assignment approach. One PFP commenter also noted PAC services were of low value to them, suggesting this may be an area for cost improvement, and therefore, important to measure.

During the meeting, the workgroup seemed generally in favor of including PAC, as this is an area of potential overuse and cost savings, but they requested clarification on what this includes. The measure currently includes the following:

- Skilled Nursing Facilities (SNF) stays when the qualifying inpatient stay is for a Medicare Severity Diagnosis Related Group (MS-DRG) that's assigned to the measure
- IRF and LTCH stays when preceded within 7 days by an inpatient stay with an MS-DRG that's assigned to the measure
- Home health based on the combination of the visit type and relevant diagnosis

Stays in PAC facilities are assigned based on whether there's a clinically related prior hospitalization to capture the downstream consequences of inpatient care. For example, if a patient is discharged earlier than they're ready, then this will be reflected in higher PAC costs.

2.4.2 Nutritional Services

After the previous webinar, a majority of workgroup members voted against including nutritional services in the measure, though this vote didn't meet the consensus threshold. Nutritional services weren't assigned to the episode group during field testing, and stakeholders were asked to provide feedback about potential inclusion of these services. Several stakeholders supported adding nutritional services with a few noting concerns about limited Medicare coverage.

During this webinar, workgroup members seemed generally opposed to including these services; one reason was that there's no evidence to support the effectiveness of these services. Additional analyses provided with the poll also showed a low frequency of Medicare

Part B physician/supplier claims for nutritional services and that the most common diagnosis was Type 2 Diabetes Mellitus.

Key Takeaways from Poll Results for Assigning Services to the Episode Group:

- Workgroup members didn't reach consensus on whether to assign LTCH services to the episode group, so we will continue to assign it (i.e., the default option from the draft measure's prior specifications).
- Workgroup members recommend not assigning nutritional services to the episode group.

2.5 Next Steps

In the last session, Acumen provided a wrap-up of the discussion and an overview of the next steps. After the meeting, Acumen distributed the PFTR Webinar Poll to gather input from members on the discussions held during the webinar about potential refinements. The poll also included a section for other general comments. Acumen will operationalize input for the measure specifications based on PFTR Webinar Poll results.

Please contact **Acumen MACRA Clinical Committee Support** at macra-clinical-committee-support@acumenllc.com if you have any questions. If you are interested in receiving updates about MACRA Episode-Based Cost Measures, please complete this Mailing List Sign-Up Form to be added to our mailing list.