

Low Back Pain Service Assignment and Refinement (SAR) Meeting Summary

MACRA Episode-Based Cost Measures: Clinician Expert Workgroups

SAR Webinar, August 24, 2021

October 2021

Contents

| | |
|--|----------|
| Project Overview | 1 |
| Low Back Pain SAR Webinar, August 24, 2021 | 2 |
| 1. Person and Family Partner (PFP) Findings and Discussion..... | 2 |
| 2. Refinements to Draft Specifications..... | 3 |
| 2.1 Sharing Trigger Code Updates and Attribution Window Validity Testing | 3 |
| 2.2 Accounting for Patient Heterogeneity | 3 |
| 3. Assigning Services to the Episode Group..... | 6 |
| 3.1 Part D Costs..... | 6 |
| 3.2 Post-Acute Care..... | 7 |
| 3.3 Specific Services within Categories of Assigned Services Previously Discussed..... | 7 |
| 4. Next Steps..... | 8 |

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. This approach will be revisited for future Waves of development. 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) Major Depressive Disorder.

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. Then, Acumen convened the workgroups

¹ For information on measure development in Waves 3, refer to the [2020 Episode-Based Cost Measures Field Testing Wave 3 Measure Development Process](https://www.cms.gov/files/document/macra-cmft-ebcm-testing-wave-3-measure-development-process-2020.pdf) document (<https://www.cms.gov/files/document/macra-cmft-ebcm-testing-wave-3-measure-development-process-2020.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](https://www.cms.gov/files/document/wave-4-public-comment-summary.pdf) document (<https://www.cms.gov/files/document/wave-4-public-comment-summary.pdf>).

again for a Service Assignment and Refinement (SAR) Webinar to revisit the specifications recommended during the initial meeting and refine the measures prior to national field testing. For Wave 4, all workgroup meetings will be held virtually. The workgroups will convene for a third meeting to continue measure specification and refinement discussions after a national field test, currently slated for early 2022.

Low Back Pain SAR Webinar, August 24, 2021

The Low Back Pain workgroup met on August 24, 2021, to continue building out the specifications for the measure. The meeting was held via webinar and 19 of the 21 workgroup members attended.³ The webinar was facilitated by an Acumen moderator, Walter Park, and the workgroup chair, Dheeraj Mahajan. Person and Family Partners (PFPs) Lorraine Krug and Roger LaCoy presented findings from discussions with a broader group of PFPs. Members of the public also attended with a listen-only line for transparency into the measure development process.

This document summarizes the discussions from the 3.5-hour virtual meeting. Section 1 discusses the findings from PFPs about lived experience with low back pain. Section 2 describes refinements to the measure, particularly focusing on how to account for severe or complex low back pain and the high costs of spine surgery. Section 3 discusses assigning services to the measure. Section 4 summarizes the next steps in the measure development process. 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.

1. Person and Family Partner (PFP) Findings and Discussion

The 2 attending PFPs presented findings from focus groups with 4 PFPs with low back pain experience. These discussions built on the earlier discussions prior to the June 2021 workgroup webinars and focused on the types of care that they received.

PFPs identified a range of care provided by different members of their care team at various points during their low back pain experience. Primary care clinicians were often the first point of contact who managed more stable pain. Specialists, such as orthopedic surgeons, were identified as being more involved in referrals to imaging, while physical therapists and physiatrists treated patients who didn't need surgery. Other specialists, like cardiologists and pain specialists, provided medication and pain management care.

The PFPs also discussed services that weren't helpful in managing low back pain and common roadblocks to receiving care. Two PFPs received injections, citing no benefit. In terms of barriers to care, PFPs noted that accessing imaging results and the amount of paperwork required to navigate the Medicare cap on physical therapy sessions was difficult and a source of stress. One PFP highlighted that some common information about should be made available, e.g., a holistic approach to care; physical and occupational therapy; nutrition; and support for anxiety related to their condition.

PFPs discussed the importance of cross-specialty coordination and highlighted this as an area for improvement. In general, PFPs experienced a lack of communication between different

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>).

specialties, unless they were part of the same healthcare system. For instance, 2 PFPs had vascular issues causing low back pain but had no involvement from vascular specialists coordinating with other clinicians early in their care. For surgery, only 1 of 3 PFPs had clinicians other than the orthopedic surgeon discuss the decision to have surgery.

2. Refinements to Draft Specifications

This session provided a feedback loop from the previous meeting about defining the scope of the measure (Section 2.1) and a detailed discussion of how to account for specific patient cohorts (Section 2.2).

2.1 Sharing Trigger Code Updates and Attribution Window Validity Testing

The first portion of this session provided updates regarding trigger codes and testing for the attribution window as a feedback loop from the workgroup's discussions and votes from the June 2021 meeting. The workgroup reached consensus on minor updates to the trigger codes; otherwise, the set of codes that identify the start of a clinician-patient relationship remained the same. Acumen provided empirical testing results to show how the workgroup's recommendation to use a 120-day attribution window reflects clinician-patient relationships for low back pain. The majority of episodes (77%) are reaffirmed, meaning that there's evidence of continued care for low back pain beyond 120 days. The episode costs are normalized so that episodes of different duration can be compared to each other; the mean observed cost for reaffirmed and non-reaffirmed episodes is almost identical (\$3,611 vs \$3,614), confirming the workgroup's intuition that care for time-limited and long-lasting low back pain could both be captured within the measure.

2.2 Accounting for Patient Heterogeneity

The workgroup spent the majority of the session discussing how to account for patient cohort heterogeneity. 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 cohorts⁴
- (ii) Defining covariates in the risk adjustment model⁵
- (iii) Identifying measure exclusions⁶
- (iv) Monitoring certain sub-populations for further testing⁷

⁴ Sub-grouping is a method that's intended for when we 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 up 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.

⁷ Monitoring for further testing is an option for flagging certain sub-populations that the workgroup may revisit later during measure development upon review of further data. This approach is best used when the workgroup requests additional data or information on a sub-population to discuss the appropriate method for meaningful clinical comparison.

2.2.1 Patients with a Spinal Surgery during the Episode

The workgroup reviewed analyses and discussed approaches for accounting for the high costs of spinal surgery. Patients who receive spinal surgery could represent a cohort with distinct pathologies relative to more common forms of low back pain; however, this hasn't been empirically tested. By definition, episodes with spinal surgery have higher observed and risk-adjusted cost than the total pool of low back pain episodes. As such, it's important to mitigate the impact of the high costs of spinal surgery to avoid the potential to disincentivize necessary surgical care.

The workgroup discussed 2 main approaches for how to account for these high-cost episodes:

- Stratify the measure into sub-groups based on whether or not an episode has a spine surgery during it
- Excluding episodes with spine surgery

The first approach effectively creates 2 measures: one for episodes where a patient receives a spine surgery, and one for episodes where a patient doesn't. For each, episodes are only compared with each other. This means that high-cost episodes with spine surgery are compared only with other high-cost episodes with spine surgery. This approach neither rewards nor penalizes clinicians for the occurrence or absence of spine surgery during an episode.

The second approach simply excludes episodes with a spine surgery. Members in favor of this approach noted that these patients are clinically distinct from other patients who don't have spine surgery; on the other hand, members opposed to this approach noted that an exclusion assumes that surgery is warranted in each case. An exclusion could also limit the utility of the cost measure in a value assessment, such as a potential future MIPS Value Pathway (MVP). This is because it would leave a gap in the data for cost to be considered alongside quality outcomes, and one would be unable to distinguish between cases where the spine surgery improved patient outcomes (which could be high-value care) and where the spine surgery didn't (which could be low-value care).

The third approach that the workgroup discussed briefly was to identify and risk adjust for characteristics associated with spinal surgery. This would involve empirically testing to identify effective predictors of spinal surgery which could then be added to the risk adjustment model. The benefit of this approach is to capture variation associated with the use of surgery in low back pain care. If there are cases where spine surgery occurs without the conditions or characteristics identified as being associated with spine surgery, then this would likely be reflected as a higher observed over expected cost for that episode.

2.2.2 Patients with Spinal Disorders

The workgroup reviewed descriptive statistics for observed and risk-adjusted costs for patients with various spinal disorders. These represent sub-populations that the workgroup had identified in the previous meeting as being clinically distinct from patients with less complex forms of low back pain. Results indicated that the conditions examined had higher risk-adjusted costs than the overall patient cohort. Members generally agreed that, of the conditions examined, radiculopathy, spinal stenosis, and spondylolisthesis could be considered together as one sub-group. These have large numbers of episodes and similar risk-adjusted cost profiles. They're also clinically similar, so they could be coded similarly. This would result in one sub-group for patients with these conditions and a separate sub-group for patients without these conditions.

For the remaining spinal disorders analyzed, the workgroup was in favor of risk adjusting for 2 other conditions: spondylolysis and scoliosis and other spinal deformities. Members noted that these are distinct from radiculopathy, spinal stenosis, and spondylolisthesis, so they shouldn't be considered as part of the sub-grouping. However, they do have different cost profiles (even after risk adjustment), so they should be accounted for in this way.

The workgroup's review of the results for spondylosis suggested that this patient cohort doesn't appear as distinct from the overall patient cohort as the other conditions, so they stated that they didn't need to be adjusted.

2.2.3 Confirming Risk Adjustors and Exclusions where Consensus wasn't Previously Reached

The final portion of the session confirmed exclusions and risk adjustors where consensus had previously not been reached. Members reviewed results showing the counts of episodes and compared the observed and risk-adjusted episode costs. For exclusions, there was general agreement to exclude small groups of patients with distinct care pathways:

- Cauda equina syndrome
- Infection
- Osteoporotic compression fracture
- Myelopathy
- Trauma

A few members also suggested excluding cancer, although many types of cancer are accounted for through the standard hierarchical condition category (HCC) risk adjustors. The workgroup was also generally in agreement about implementing the following patient sub-populations as risk adjustors:

- Prior hospitalization for medical back problems
- Depression
- Osteoarthritis
- Spine surgery during a 1-year lookback period.

This final risk adjustor incorporated feedback from the June workgroup meeting to broaden the definition to identify other indicators of prior spine surgery (e.g., diagnosis codes).

Key Takeaways from Discussion and/or Polls for Addressing Sub-Populations for Meaningful Clinical Comparison:

- Members were split on how to account for the high cost of spine surgery during episodes.
 - A majority was in favor of stratifying the measure into sub-groups for episodes with and without spinal surgery, but they didn't reach the greater than 60% consensus threshold
 - A minority of members were in favor of excluding episodes with spine surgery, with a couple of members favoring the third approach of identifying and risk adjusting for conditions associated with the need for spinal surgery
 - As the vote was split, Acumen will proceed with field testing the measure with sub-groups for episodes with and without spine surgery (i.e., the preference of the majority)

- Members agreed to create further sub-groups based on radiculopathy, spinal stenosis, and spondylolisthesis. As such, the set of sub-groups for field testing will be:⁸
 - Episodes with spine surgery with radiculopathy, spinal stenosis, or spondylolisthesis
 - Episodes with spine surgery without radiculopathy, spinal stenosis, or spondylolisthesis
 - Episodes without spine surgery with radiculopathy, spinal stenosis, or spondylolisthesis
 - Episodes without spine surgery without radiculopathy, spinal stenosis, or spondylolisthesis
- The following risk adjustors were confirmed:
 - Spondylolysis
 - Scoliosis and other spinal deformities
 - Medical back problems hospitalization
 - Depression
 - Osteoarthritis
 - Spine surgery during 1-year lookback period
- The following exclusions were confirmed:
 - Cauda equina syndrome
 - Infection
 - Osteoporotic compression fracture
 - Myelopathy
 - Trauma

3. 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. This discussion focused on particular types of costs: Part D costs (Section 3.1), post-acute care (Section 3.2), and specific services within the broader categories of services that the workgroup had identified during the June meeting (Section 3.3).

3.1 Part D Costs

Acumen summarized guidance from the project’s overarching technical expert panel (TEP), which noted that episode-based cost measures should consider Part D costs on a case-by-case basis. There are some conditions where it will be more important than others to include Part D costs to ensure the full picture of provider performance is part of the measure. The measure must also account for the challenge that not all patients are enrolled in Part D. One of the key challenges to including Part D is the lack of specificity in drugs used to treat low back pain versus other conditions. Also, many of the commonly prescribed Part D drugs are low-cost and may not represent much opportunity for cost improvement.

The workgroup considered these factors as well as the TEP’s guidance and agreed that the measure should include Part D costs, as capturing variation in their utilization outweighed the challenges. Members agreed that the following Part D drugs should be included:

⁸ The Low Back Pain measure will also sub-group by Part D enrollment since the workgroup was in favor of including Part D services in the measure.

- Opioids
- Nonsteroidal anti-inflammatory agents (NSAIDs)
- Central muscle relaxants
- Oral Glucocorticosteroids
- Anti-depressants

The workgroup noted that Proton Pump Inhibitors (PPI) used to treat common side effects of NSAIDs and antibiotics shouldn't be included, as they aren't relevant to the treatment of low back pain.

3.2 Post-Acute Care

The workgroup reviewed an analysis showing the frequency and cost of common services after an episode is triggered. Members discussed whether post-acute care (PAC) services should be included in the measure, noting that these are high-cost events that could adversely impact measure scores. Other members noted that these were important services to capture adverse outcomes, and that service assignment could ensure relatedness to low back pain treatment (e.g., by requiring a relevant diagnosis code to be present for the service to be assigned to the measure). The workgroup reached consensus about including inpatient rehabilitation facility (IRF) and home health care services, but the workgroup was closely split on whether or not to include long-term care hospital (LTCH) services.

3.3 Specific Services within Categories of Assigned Services Previously Discussed

Acumen identified specific services for discussion that would help build out the set of assigned services from earlier discussions. Members voted on these through a post-meeting poll, and they're listed below.

Key Takeaways from Discussion and/or Polls for Assigning Services to the Episode Group:

- Members recommended to include Part D costs in the measure, specifically:
 - Opioids
 - NSAIDs
 - Central muscle relaxants
 - Oral glucocorticosteroids
 - Anti-depressants
- The workgroup discussed the inclusion of PAC services and recommended including IRF and home health services that are related to low back pain
- Members recommended including the following services within particular categories of assigned services:
 - Behavioral health services: outpatient psychotherapy and depression screening
 - Drug tests: outpatient testing used for patients on opioids
 - Walking aids: services such as wheelchairs, crutches, canes, walkers
- Members recommended not including the following services:
 - Hospitalizations for behavioral health conditions/disorders
 - General wound care (e.g., dressings, gauze, bandages)
 - Gastrointestinal conditions: hospitalizations for gastrointestinal hemorrhage, peptic ulcer, inflammatory bowel disease
- The workgroup was split on whether to include nutritional services, with a majority believing that they shouldn't be included, though this was short of the greater than 60% consensus threshold
 - As such, the services won't be included for field testing; the decision can be revisited afterwards

4. 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 SAR Webinar Poll with a recording of the webinar to formally gather input from the meeting. The poll was open for one week and was structured to summarize discussion to reflect where there appeared to be verbal consensus; it included empirical testing results (where relevant) so that members could refer to this information when responding to the survey. The survey included comment boxes to provide additional thoughts. Based on National Quality Forum practices, the threshold for support was greater than 60% consensus among poll responses.

Acumen will operationalize input for the measure specifications based on SAR Webinar Poll results and will prepare specifications and related materials for the upcoming national field testing. The workgroup is slated to convene for a Post-Field Test Refinement (PFTR) Webinar in March 2022.

Please contact **Acumen MACRA Clinical Committee Support** at macra-clinical-committee-support@acumenllc.com if you have any questions. If you're 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.