



ACUMEN

**Assessing Quality and Improving Value of Breast
Cancer Screening Technical Expert Panel
Meeting #1**

January 2023

Acumen, LLC
500 Airport Blvd., Suite 365
Burlingame, CA 94010

PROJECT OVERVIEW

Acumen, LLC has been awarded a grant by the Gordon and Betty Moore Foundation (grant number GBMF11507) to develop measures that can assess the quality and improve the value of breast cancer screening. Acumen's measure development approach involves convening a Technical Expert Panel (TEP), composed of clinical experts, as well as patients, caregivers, and patient advocates, to contribute direction and thoughtful input during measure development. The measure development content is solely the responsibility of the authors and does not necessarily represent the views of the Moore Foundation.

The application of clinically accepted metrics in breast image quality and interpretation have been associated with improved outcomes. This project will link current practice standards to policy initiatives by creating a cohesive set of measures based on the Breast Imaging Reporting and Data System (BI-RADS) Atlas®. This will allow radiology practitioners to assess performance, value, and outcomes; engage in national policy; and work towards achieving a national standard for diagnostic excellence.

Acumen is developing four clinician-level measures – 3 quality measures and 1 episode-based cost measure – to assess performance of breast imaging teams using Medicare administrative claims. Together, these could provide a cohesive set of measures for a MIPS (Merit-based Incentive Payment System) Value Pathway (MVP) to reward diagnostic excellence, currently infeasible without outcome or cost measures. Accordingly, we plan to submit the measures through the Centers for Medicare & Medicaid Services (CMS) pre-rulemaking process for consideration for use in MIPS, with the intention of them filling critical measurement gaps now present.

We convened a TEP to provide input on the specifications of the measures. This is part of the measure development process, as defined by the CMS Measurement Management System (MMS) Blueprint, to gather expert clinical input and individual experience of person and family partners. Acumen held a call for nominations between December 14, 2022 and January 16, 2023. Our team notified interested parties via email, and collected nominations through an online survey. A panel of 14 TEP members was finalized mid-January 2023 to provide a balanced and diverse set of perspectives. This includes clinicians with expertise directly relevant to breast cancer, as well as Person and Family Partners (PFPs), people with lived experience of breast cancer screening, diagnosis, and treatment. The TEP met virtually on January 20, 2023, and will meet another two times throughout 2023 to further discuss measure specifications and review empirical testing results.

TABLE OF CONTENTS

- Project Overviewii**
- 1 Overview 1**
- 2 Summary of Sessions and Discussion 2**
 - 2.1 Measure Introduction..... 2
 - 2.2 Quality Measures: Denominator and Numerator..... 3
 - 2.2.1 Denominator Inclusions: Patient Cohort..... 3
 - 2.2.2 Denominator Inclusions: Service Codes..... 4
 - 2.2.3 Denominator Exclusions..... 5
 - 2.2.4 Numerators..... 6
 - 2.2.5 Key Takeaways..... 7
 - 2.3 Quality Measures: Scoring and Benchmarks 8
 - 2.3.1 Key Takeaways..... 9
 - 2.4 Cost Measure..... 9
 - 2.4.1 Episode Trigger 9
 - 2.4.2 Distinguishing Care through Cost Measure Performance 10
 - 2.4.3 Key Takeaways..... 11
- 3 Next Steps 12**
- Appendix A : List Of TEP Members and Acumen Project Team 13**
 - A.1 TEP Members: Clinical 13
 - A.2 TEP Members: Person and Family Partners 13
 - A.3 Acumen Project Team 14
- Appendix B : Meeting Materials 15**
- Appendix C : Literature Shared by TEP Panelists During Meeting..... 16**

LIST OF TABLES AND FIGURES

- Table 1: Services that Indicate Screening Mammography 4
- Table 2: Services that Indicate Biopsy that Could Follow Diagnostic Mammogram..... 4
- Table 3. Denominator Exclusion Diagnosis Codes 5
- Table 4. Services Indicating Diagnostic Workup After a Screening Mammogram (Recall Rate Numerator and PPV1 Denominator) 6
- Table 5. Diagnosis and Service Codes Identifying Breast Cancer (PPV1 and PPV3 Numerators) 6
- Table 6. Cost Measure Trigger Codes..... 10

1 OVERVIEW

This meeting summary document outlines the purpose, discussion, and recommendations from the Improving the Diagnostic Performance of Screening Tests for Breast Cancer TEP #1. The goals of the Improving the Diagnostic Performance of Screening Tests for Breast Cancer TEP meeting on January 20, 2023 were the following:

- (i) Provide TEP panelists with an overview of the project, quality measure framework, and episode-based cost measure construction methodology;
- (ii) Gather input from panelists on the patient population applicable for each measure, patient cohorts that may have distinct characteristics from the whole patient population, and identify clinically related services that can be used in quality and cost measure specifications.

The meeting was held virtually and attended by 12 of the 14 TEP members. The webinar was facilitated by the moderator, David Moore, and the TEP Chair/Co-Principal Investigator, David Seidenwurm. Appendix A provides the list of TEP members and the Acumen project team.

2 SUMMARY OF SESSIONS AND DISCUSSION

This section is organized based on the meeting sessions and describes panel member discussions and recommendations. Section 2.1 summarizes the measure overview and policy context of MIPS, presented during the webinar. Section 2.2 covers the quality measures' denominator and numerator discussions. Section 2.3 summarizes scoring and benchmarks for the quality measures. Section 2.4 summarizes the cost measure construction discussion.

2.1 Measure Introduction

Acumen provided an overview of the policy context for this project and how the measures meet CMS needs and priorities. The measures are being developed with the intent of submitting them to the Merit-based Incentive Payment System (MIPS), a program mandated by statute which adjusts payment to clinicians based on their performance across four categories: Quality, Cost, Improvement Activities, and Promoting Interoperability. Clinicians choose how to participate in MIPS; in 2023, they will be able to choose to participate through MIPS Value Pathways (MVPs). This participation option has a smaller set of measures and activities relevant to a clinical topic or specialty to provide a more connected assessment of the value of care.

The clinical topic of breast cancer screening is the focus of this measure development project as it represents a current gap in MIPS. Radiologists, as non-patient facing clinicians, have been identified by CMS as a priority for measure needs. Radiologists do have quality measures, albeit only process measures. Also, there are no cost measures applicable to diagnostic radiologists: in 2023, there are 23 episode-based cost measures and 2 population-based cost measures in MIPS but none are specific to radiology care. The Chair commented that even though the costs of radiology services are included in MIPS measures, radiologists themselves are most often not attributed.

Acumen presented the measure concepts for the four mammography measures that would provide more meaningful ways to assess the performance of radiologists than the measures currently available in MIPS:

- Recall Rate: Percentage of screening mammograms followed by diagnostic work-up
- Positive Predictive Value (PPV) 1: Percentage of screening mammograms followed by diagnostic work-up that lead to cancer diagnosis
- Positive Predictive Value (PPV) 3: Positive biopsy rate followed by cancer diagnosis
- Episode-based cost measure: Risk-adjusted cost for a screening mammography episode

Each measure will be constructed using administrative claims data. This method is advantageous because there is no additional reporting burden for clinicians and it removes the selection bias in quality measure reporting. There is also some precedent in CMS programs for using claims data to assess breast cancer screening: the Hospital Outpatient Quality Reporting (HOQR) program uses a claims-based breast cancer screening recall rate measure (OP-39). Since CMS considers alignment across programs and measures, Acumen's clinician-level Recall Rate measure will use the same definitions as OP-39 wherever possible.

2.2 Quality Measures: Denominator and Numerator

The 3 quality measures for breast cancer screening are constructed using a defined patient cohort (denominator value) and an outcome among the patient cohort (numerator value). A clinician's performance is then evaluated based on the resulting value. Section 2.2.1 summarizes the denominator patient cohort inclusions. Section 2.2.2 the denominator service code inclusions. Section 2.2.3 summarizes denominator exclusions. Section 2.2.4 summarizes numerator considerations. Lastly, Section 2.2.5 summarizes key takeaways from the above sections.

2.2.1 Denominator Inclusions: Patient Cohort

In the draft measure specifications, the denominators are representative of the patient cohort. These values are determined based on the initial patient population, or all patients that should be considered for inclusion in the given measure:

- Recall Rate: Patients who had a screening mammogram and digital breast tomosynthesis (DBT)
- PPV1: Patients who had diagnostic workup after a screening mammogram
- PPV3: Patients who had a biopsy from a diagnostic mammogram

A central issue to defining the patient population is to determine the appropriate screening age range to use, given that clinical guidelines are inconsistent. Acumen reviewed various US and international guidelines for appropriate age ranges to start and end screening mammograms. While the panel acknowledged the variation in guidelines for the appropriate age ranges for screening mammogram, there was general consensus that the measures should begin to capture patients over 40. A PFP commented that although they were high risk for breast cancer and required screening at an age earlier than 40, they did not think that high risk patients like them were the standard; they felt high risk patients should be excluded from the cohort. A panelist commented that mammograms for patients under 40 are not considered *screening* mammograms, but *diagnostic* mammograms, reiterating the sentiment of the PFP's comment to exclude those under 40. As for an upper limit, some panelists preferred cutting off the cohort for patients over 74 years of age. Other panelists preferred to keep an open upper limit, since breast cancer rates can spike at later ages.

The group briefly considered the issue of using all-gender data for the measure denominators. Generally they agreed the measure cohort should only include females. Acumen proposed to test the impact of this inclusion criteria.

2.2.2 Denominator Inclusions: Service Codes

After defining the general patient cohorts of interest, the measures use services to identify the individuals from that cohort who have received the type of care of interest to build out the denominator. These services are identified through Current Procedural Terminology (CPT) and Healthcare Common Procedure Coding System (HCPCS) codes. The panel generally agreed on the set of codes listed in Table 1, below, which would be used to define the Recall Rate denominator. The TEP noted that these codes should only be the ones that indicate screening mammogram, not diagnostic mammogram. A panelist asked that MRI scan of both breasts without contrast (77047) be removed because it is not done for screening. One panelist requested that Acumen look into how to differentiate between screening and diagnostic ultrasounds. Another panelist suggested that 2-D vs. 3-D ultrasounds be accounted for separately in future codes lists. The TEP Chair noted that tomosynthesis (G0279) can be considered separately to yield better results.

Table 1: Services that Indicate Screening Mammography

CPT/HCPCS	Description
G0279	Diagnostic digital breast tomosynthesis, unilateral or bilateral (list separately in addition to 77065 or 77066)
77063	Screening 3D breast mammography
77067	Screening mammography
76641, 76642	Ultrasound of breast (complete, limited)

The PPV3 Positive Biopsy Rate measure denominator is based on patients who have a biopsy from a diagnostic mammogram. The TEP reviewed an initial list of codes from Acumen and discussed which should be removed, if they are typically unrelated to normal diagnostic cases. For example, the panel determined that codes for fluoroscopic guidance for needle placement (77002) and review by radiologist for CT guidance for needle placement (77012) were not medically relevant to normal diagnostic mammogram cases and should not be part of the denominator. Table 2 includes a list of codes that the TEP tentatively agreed on to indicate biopsy from a diagnostic mammogram.

Table 2: Services that Indicate Biopsy that Could Follow Diagnostic Mammogram

CPT/HCPCS	Description
19081-19086	Biopsy of breast, and placement of locating device
10006-10012	Fine needle aspiration biopsy
77021	Review by radiologist for MRI guidance for needle placement

CPT/HCPCS	Description
76942	Ultrasonic guidance for needle placement
19000, 19001	Needle biopsy of breast
19290, 19291	Pre-surgical insertion of needle wire to localize breast growth

The TEP noted that these codes would still need to be tested. Multiple panelists questioned if the pre-surgical procedure from codes 19290 and 19291, was within the intended scope of the measure, given that this procedure required a surgeon and is far from a normal screening mammogram or biopsy. Additionally, Acumen will review breast biopsy (19081-19086) and pre-surgical insertion of needle wire to localize breast growth (19290, 19291) codes individually to determine which are appropriate to use.

The TEP discussed how to define the timing of how soon after a diagnostic mammogram a biopsy takes place to be included in the denominator. A panelist noted that every practice is different and timing could range from same-day to months later. The panel came to agreement on a window of biopsy occurring somewhere within 45 or 90 day windows after the diagnostic mammogram, since these are standard timeframes based on their experience. The panelists were satisfied with the idea of age range stratifications, barring the future beta testing does not reveal a better option.

2.2.3 Denominator Exclusions

The TEP discussed whether there are patient health characteristics that should be the basis for recall rate denominator exclusions. This would mean that patients are excluded from the measure before considering the numerator, and would be appropriate to use if they are reliably coded in claims data and show statistically significant distortion in the measure results. This would have to be further determined from measure testing. The preliminary set of denominator exclusion codes are shown in Table 3, below.

Table 3. Denominator Exclusion Diagnosis Codes

Health Characteristic	Diagnosis Codes
Genetic mutations BRCA1 and BRCA2	Z15.01 Genetic susceptibility to malignant neoplasm of breast
Prior breast cancer	Z85.3 Personal history of malignant neoplasm of breast
	C50 Malignant neoplasm of breast
	D05 Carcinoma in situ of the breast

The panel considered these codes to have the most probative value in determining whether a patient has higher than normal risk for breast cancer. Panelists noted that these listed prior breast cancer codes (Z85.3, C50, D05, and Z15.01) are more indicative of cancer risk for the measure than the diagnosis code for family history of breast cancer (Z80.3). Family history can mean anything from a mother with breast cancer to paternal aunt with breast cancer, which clinically is a very different level of risk. As such, panelists were not in favor of using Z80.3 as a

denominator exclusion. Also, the panel affirmed that genetic susceptibility – like BRCA carrier status (Z15.01) – should be excluded from measure denominators because these patients receive different screening care, such as MRIs with contrast.

2.2.4 Numerators

The measure numerators represent the outcome being accessed amongst the population defined by the denominator criteria and codes. The numerator is then divided by the denominator to create a ratio for each clinician/group.

The TEP discussed how to identify diagnostic workup after a screening mammogram. These services would be used to define both the Recall Rate numerator and the PPV1 denominator. Panelists noted that the definition should not include any magnetic resonance imaging without contrast codes (C8904, C8905, C8907, C8907) since these are not clinically relevant. Additionally, they agreed that pet imaging (G0252) and scintimammography (S8080) should not be used in the numerators, as these are more advanced types of screening that are often used as supplemental screenings for patients with dense breasts. In rural and low-resource areas, dense breast patients often do not have access to advanced screening technology; not all facilities have high tech services such as 3-D mammography or tomosynthesis machines. Instead, facilities sometimes have to do repeated 2-D mammography to compensate. This should not be included in the Recall Rate measure because these repeat scans are not necessarily “recalls” but rather repeats to ensure accurate imaging of the dense breast patient. The TEP noted that the issue of repeat scans is not unique to the rural and low-resource areas.

Table 4. Services Indicating Diagnostic Workup After a Screening Mammogram (Recall Rate Numerator and PPV1 Denominator)

CPT/ HCPCS	Description
G0279	Diagnostic digital breast tomosynthesis, unilateral or bilateral (list separately in addition to 77065 or 77066)
0638T	CT of both breasts before and after contrast with 3D rendering
C8903	Magnetic resonance imaging with contrast, breast; unilateral
C8906	Magnetic resonance imaging with contrast, breast; bilateral
76641, 76642	Ultrasound of breast (complete, limited)

The TEP also considered how to identify the numerators for PPV1 and PPV3, which is cancer detection. The most straightforward is to use diagnosis codes C50 and D05. Panelists also considered services that are typical once cancer has been detected. Table 5 lists these codes.

Table 5. Diagnosis and Service Codes Identifying Breast Cancer (PPV1 and PPV3 Numerators)

Code Type	Code	Description
ICD-10-DGN	C50	Malignant neoplasm of breast
	D05	Carcinoma in situ of breast

Code Type	Code	Description
CPT/HCPCS	19083-19086	Biopsy of breast and placement of locating device
	19125, 19126	Removal of growth of breast identified by x-ray marker
	19304	Removal to tumor and breast tissue, accessed beneath the skin
	00404 ,00406, 00550	Anesthesia for removal of breast
	19281-19288	Placement of breast localization devices accessed through the skin with mammographic/stereotactic/ultrasonic/MRI guidance
	19294	Preparation of tumor cavity and placement of radiation therapy applicator into breast
	19296-19298	Insertion of catheter into breast for radiation therapy
	19301-19303	Partial breast removal; Total breast removal
	19305-19307	Removal of breast, lymph nodes, and muscle
	M1080, M1081	Radiation therapy for breast cancer under the radiation oncology model
	0301T, 0581T	Destruction or reduction of malignant breast tumor (microwave/ freezing)
	81519, 81162-81167, 81211-81217, 81432, 81433, 81518-81521	Test for detecting genes associated with breast cancer; Gene analysis (breast cancer)
	0102U, 0104U, 0129U, 0131U, 0153U	Gene analysis hereditary breast cancer

A few panelists debated and decided to remove stereotactic body radiation therapy since it is supposed to be used for metastatic cancer (77373, 77435). A panelist also asked that insertion of metal clip during breast biopsy or aspiration (19295) also be removed since it is done for all breast biopsies not just cancers. Panelists also decided that genic blood testing codes should be removed, however, genetic tissue tests should be included.

A panelist noted that services for the placement of breast localization devices accessed through the skin with mammographic/stereotactic/ultrasonic/MRI guidance (19281-19288) can be done to remove a variety of things (i.e. Benign papilloma), not just cancer. Acumen will need to determine how to discern this in the data, or exclude. The panel also agreed that certain codes for genetic testing that are not directly related to breast cancer should not be used to identify the breast cancer since testing done near screening may have nothing to do with the screening (0009U, 0067U, 0155U, 0172U, 0177U, 0220U).

2.2.5 Key Takeaways

- The measure populations should be limited to females.
- Recall Rate and PPV1 measures should include patients from age 40 to align with lower limit of most screening guidelines. There should not be an upper limit to the patient cohort.

- 2D and 3E ultrasounds and tomosynthesis can have different efficacy in detecting cancer, which may affect recall and cancer detection rates, depending on the incidence screening values that are used.
- Recall Rate measure exclusions should be restricted to personal patient health characteristics (BRCA carrier status and prior breast cancer).
- The TEP is subject to adjust recommendations for measure numerators and denominators once testing data provides further evidence for evaluation.

2.3 Quality Measures: Scoring and Benchmarks

Each of the quality measures generates a score for clinicians where the numerator is divided by the denominator. This score should provide a meaningful reflection of performance and be able to distinguish between good and poor care. To do so, these rates can be compared against clinically accepted benchmarks.

Most of the panelists generally agreed with using the benchmarks of the American College of Radiology (ACR). ACR's benchmarks provide the widest range of variability. The panel believed that this is beneficial to imitate, or to mimic, actual variability in a given set of patients. The TEP also recommended examining performance with age-group segmentation, especially for the Recall Rate measure. The panel suggested that breast cancer screening incidence be separated from prevalence within claims data, to improve the accuracy of the measures for new and overall cancer diagnosis. One strategy that could be considered is using a lookback to identify patient with prior breast cancer screening services.

However, the panel was concerned about the impact of applying these benchmarks – which are intended to serve as guidelines – to clinicians who care for riskier populations. Some panelists noted that the rates from each measure are not a direct indicator of performance if patient risk is not considered when interpreting the rates. One suggestion to address this was to create a composite measure where outcomes could be weighted according to importance: specifically, high detection rates are more meaningful metrics for performance than high recall rate. PFPs however did emphasize that recall has impacts that are important to patients and families as it leads to high levels of anxiety, particularly where results are delayed. In addition, the TEP Chair noted that a composite measure could be difficult to create, and may stray from the original intention of the separate measures.

The TEP discussed which risk factors would need to be considered and tested throughout development to guard against the risk of penalizing clinicians who care for higher-risk patients. The panel noted that social risk factors (SRFs) such as race and socioeconomic status (SES) can impact the outcomes being measured, such as through screening compliance and access to screening services. These, as well as health characteristics like dense breast tissue, should be accounted for. A panelist commented that the field of diagnostic radiology has had external

scrutiny which has resulted in clinician hesitancy in providing breast screening services. They thought that the fear of litigation and additional radiologist attributed-responsibility towards cost/quality measures, could result in reduced care provision. Another panelist affirmed a similar sentiment regarding potentially high recall rates for radiologists who perform supplemental screening for dense breast patients, which is now recommended by some states and the FDA, despite mixed data on benefits. The panel agreed that these measures should not exacerbate issues with access to screening and related care.

2.3.1 Key Takeaways

- Timely diagnosis of cancer should be prioritized above all other measures of screening outcome performance since the intention of the measures is exactly to improve the breast cancer screening process as a whole.
- ACR's broad range of benchmarks may be better honed for the measures once age-stratifications can be evaluated from testing data. This will provide the measures with appropriate ranges of care quality to score clinicians on.
- Equity and inclusion need to remain at the forefront of the measure development to ensure patient access to screening services remains available. This aligns with the TEP's focus on not penalizing providers who treat potentially higher risk populations.

2.4 Cost Measure

An episode-based cost measure is calculated as the ratio of the observed episode cost to expected episode cost, averaged across all episodes attributed to a clinician. To be effective and to be accepted into MIPS, a cost measure is constructed using the steps below to ensure that it is consistent with other measures in the Cost performance category:

1. Trigger and define an episode
2. Attribute the episode to the clinician group and clinician(s)
3. Assign the cost of services clinically related to mammography
4. Exclude episodes
5. Calculate expected costs through risk adjustment
6. Calculate the measure score

Section 2.4.1 outlines the discussion on episode triggers. Section 2.4.2 summarizes the in-depth discussion on how good and poor performance can be evaluated using cost measures. Finally, section 2.4.3 outlines key takeaways from these sections.

2.4.1 Episode Trigger

The episode trigger is a key first component to cost measure construction because it serves multiple purposes: identifies patient cohort, identifies the attributed clinician, and denotes

the start of when care is going to be assessed. In the draft measure specifications, episodes are triggered when a clinician group, identified by their Taxpayer Identification Number (TIN), or individual clinician (identified by TIN-NPI) bills Medicare for a screening mammogram or screening DBT.

Table 6 provides a list of services that the TEP agreed could be used to trigger an episode. These align with the services in Table 1 for Recall Rate denominator. Similarly to the Recall rate cohort, the panel agreed that MRI without contrast (77047) should not be used to identify the patient cohort.

Table 6. Cost Measure Trigger Codes

CPT/ HCPCS	Description
G0279	Diagnostic digital breast tomosynthesis, unilateral or bilateral (list separately in addition to 77065 or 77066)
77063	Screening 3D breast mammography
77067	Screening mammography
76641, 76642	Ultrasound of breast

2.4.2 Distinguishing Care through Cost Measure Performance

The goal of the cost measure is to reflect the costs related to the role of the attributed clinician so that it can be used to inform practice changes to reduce costs without impacting quality. Like the quality measures, it must be able to distinguish good from poor care. When considering the role of the diagnostic radiologist in breast cancer screening, the care that indicates good performance is true positives and true negatives, with false positives being a better outcome than false negatives. One challenge is that the cost impact of these differ: true positives will involve higher costs (e.g., diagnosis, treatment) than the other outcomes. True and false negatives alike will have fewer costs associated with them. The cost measure can account for these scenarios through service assignment, or determining which costs should be included in the measure. Acumen’s recommendation was to include only the costs of cancer treatment for interval cancer. This is breast cancers that are diagnosed during the time between a regular screening mammogram that appears normal and the next screening mammogram. However, the panel made a final recommendation to shift away from including costs of cancer treatment and towards costs of diagnosis.

The panel discussed the potential unintended consequences with including the costs of cancer treatment in a cost measure. They were concerned about attribution as there are many other clinicians involved in patient care. Panelists also noted that radiologists may miss cancer detection due to a variety of reasons outside of their control. For instance, patient screening non-compliance or test machine accuracy could affect the attributed clinician’s ability to detect

cancer in a timely manner. The panel was in full agreement that cost of treatment should not be included if cancer was correctly diagnosed.

The panel also considered other challenges with distinguishing good from poor cost performance. One panelist started by reminding Acumen that low recall rates and high cancer detection do not always correlate and that the benchmark ranges discussed earlier could help point to appropriate ranges of both. Additionally, the TEP discussed the impact of guidelines around test coverage. For example, patients with dense breasts sometimes receive MRI screening instead of mammography, but there are generally strict guidelines for approval. A panelist commented that although screening MRI is very expensive, it is underutilized by patients who exhibit BRCA markers or have history of breast cancer. Another factor mentioned was that clinicians will often over-order diagnostic tests to avoid legal risk for concerned patients. Some panelists argued that the competing pressures to perform well on these potential measures as well as concerns over legal risk may have a confounding effect on a cost measure score and put undue pressure on diagnostic radiologists.

The TEP considered the question of whether a separation of patients with and without higher risk would address these concerns. A panelist commented that there are different calculators of cancer risk and clinicians are adept at discerning which calculators provide a favorable result for their patients to obtain insurance approval for tests that are usually only covered for high-risk patients. The TEP Chair recommended that age could also be used as a risk factor as well, though given the earlier conversations about age stratification, this may not be necessary.

Lastly, the panel discussed the topic of episode window length for the cost measure. The panel considered that a 1-year measure period is too short for cancer screening. They suggested a 15, 18, or 24-month timeframe of evaluation. Acumen explained that the MIPS program evaluates measures on an annual basis. The panel requested testing data analysis to determine if the measures can be properly discerned in a 12-month episode. Acumen will work to determine if a longer claims look-back period can be used to discern screening incidence and prevalence and provide better estimates of interval cancers.

2.4.3 Key Takeaways

- The measure needs to be tested empirically to guard against unintended consequences. There are many factors which could affect diagnosis that are outside of radiologists' control.
- Patients with higher screening care needs (e.g., with dense breasts) should be accounted for, so that clinicians are not penalized due to their patient case mix.
- The episode window needs to be long enough to discern between screening incidence and prevalence for patients with prior screening mammograms or none. This might

require episodes to be longer than 12 months (up to 24-month look back). Alternatively, Acumen will commence data testing to see if there is a way to explore a longer look-back period from the claims, rather than adjusting the whole measure.

3 NEXT STEPS

Acumen will use the TEP's input to build out draft measures using the specifications discussed. This will allow for alpha testing using administrative claims data. The TEP will convene another two times this year to iteratively review and refine the measures. At the end of the development process, Acumen intends to submit the final measures to CMS's MUC list: this process allows CMS to consider whether to propose to add the measures to the MIPS Quality and Cost performance categories through the notice-and-comment rulemaking process.

Please contact Acumen Mammography Measures Moore Support Team at mmg_measures_moore@acumenllc.com if you have any questions.

APPENDIX A: LIST OF TEP MEMBERS AND ACUMEN PROJECT TEAM

The *Assessing Quality and Improving Value of Breast Cancer Screening* Technical Expert Panel is made up of 14 members (11 clinical, 3 person and family partners). Of the 14 members, 12 were able to attend the January 20, 2023 meeting. Section A.1 lists the clinical members. Section A.2. Lists the person and family partner members. Section A.3 lists the 10 Acumen project team members for additional reference.

A.1 TEP Members: Clinical

- Megan Adamson, MD, MHS-CL, FAAFP
- Jose Bazan, MD, MS
- Stamatia Destounis, MD, FACR, FSBI, FAIUM
- Carolyn Dueñas, RN, Absent
- Sarah Eakin, MD, FCAP
- Sharad Goyal, MD, MS
- Cindy Lee, MD, FACMQ, FSBI
- Lauren Nicola, MD, Absent
- Lydia Pace, MD, MPH
- Barbara Spivak, MD
- Barbara Wexelman, MD, MBA

A.2 TEP Members: Person and Family Partners

- Rosie Bartel
- Nancy Farrar
- Barbara Kivowitz

A.3 Acumen Project Team

- David Moore, Moderator
- Rose Do, MD, Principal Investigator
- David Seidenwurm, MD, FACR, Co-Principal Investigator
- Heather Litvinoff, PT, MPH, Project Manager
- Sri Nagavarapu, PhD, Technical Analytic Advisor
- Lois Olinger, MCP, Senior Policy Advisor
- Laurie Feinberg, MD, MPH, Clinical Associate
- Joyce Lam, MPP, Research Manager
- Ken Tran, PhD, Senior Policy Associate
- Tatiana Valentine, MS, Data & Policy Analyst

APPENDIX B: MEETING MATERIALS

Prior to the meeting, TEP members were provided with the following materials:

- Agenda and Slides
- Measure Construction Reference Sheet
- TEP Draft Charter, which was ratified by panelists during the meeting
- Assessing Quality and Improving Value of Breast Cancer Screening: environmental scan and literature review, which provided a comprehensive overview of the importance of breast cancer screening for women’s health, the interest in measure development for screening, core elements of the proposed measures, and provided precedent for the measures.
- Background materials about MIPS and episode-based cost measures
 - CMS 2022 MIPS Overview Webinar Slides¹
 - JAMA Health Forum Paper “Development of Episode-Based Cost Measures for the US Medicare Merit-based Incentive Payment System”²

¹ Centers for Medicare & Medicaid Services Quality Payment Program. Full Webinar Library. <https://qpp-cm-prod-content.s3.amazonaws.com/uploads/1925/2022%20MIPS%20Overview%20Webinar%20Slides.pdf>

² Duseja R, Andress J, Sandhu AT, et al. Development of Episode-Based Cost Measures for the US Medicare Merit-based Incentive Payment System. JAMA Health Forum. 2021;2(5):e210451. doi:[10.1001/jamahealthforum.2021.0451](https://doi.org/10.1001/jamahealthforum.2021.0451)

APPENDIX C: LITERATURE SHARED BY TEP PANELISTS DURING MEETING

Below is a list of references to literature shared by the panelists during the meeting. Additionally, abstract conclusion excerpts have been included to provide further context of the articles.

- Ganguli, I., Keating, N. L., Thakore, N., Lii, J., Raza, S., & Pace, L. E. (2022). Downstream Mammary and Extramammary Cascade Services and Spending Following Screening Breast Magnetic Resonance Imaging vs Mammography Among Commercially Insured Women. *JAMA network open*, 5(4), e227234.
<https://doi.org/10.1001/jamanetworkopen.2022.7234>
 - In this cohort study of commercially insured women, breast MRI was associated with more mammary and extramammary cascade events and spending relative to mammography.
- Hill, D. A., Haas, J. S., Wellman, R., Hubbard, R. A., Lee, C. I., Alford-Teaster, J., Wernli, K. J., Henderson, L. M., Stout, N. K., Tosteson, A. N. A., Kerlikowske, K., & Onega, T. (2018). Utilization of breast cancer screening with magnetic resonance imaging in community practice. *Journal of general internal medicine*, 33(3), 275–283.
<https://doi.org/10.1007/s11606-017-4224-6>
 - Utilization of screening MRI in community settings is not consistent with current professional guidelines and the goal of delivery of high-value care.
- Miles, R., Wan, F., Onega, T. L., Lenderink-Carpenter, A., O'Meara, E. S., Zhu, W., Henderson, L. M., Haas, J. S., Hill, D. A., Tosteson, A. N. A., Wernli, K. J., Alford-Teaster, J., Lee, J. M., Lehman, C. D., & Lee, C. I. (2018). Underutilization of Supplemental Magnetic Resonance Imaging Screening Among Patients at High Breast Cancer Risk. *Journal of women's health* (2002), 27(6), 748–754.
<https://doi.org/10.1089/jwh.2017.6623>
 - While nearly half of women at high lifetime breast cancer risk undergo routine screening mammography at a facility with on-site breast MRI availability, supplemental breast MRI remains widely underutilized among those who may benefit from earlier cancer detection.
- Morrell, B. L., Morrell, M. B., Ball, J. A., Ochoa, A. C., & Seewaldt, V. L. (2023). Disparities in the use of screening breast magnetic resonance imaging persist in Louisiana after the Affordable Care Act: A question of access, policy, institutional support, or

something else?. Cancer, 10.1002/cncr.34605. Advance online publication.

<https://doi.org/10.1002/cncr.34605>

- Since the adoption of the ACA in Louisiana, Black women have continued to have disproportionately high breast cancer mortality rates. This persistent disparity provides evidence that additional change is needed. This change should include exploring innovative ways to make advanced imaging technology such as breast MRI more accessible and expanding research to specifically address community and culturally specific barriers.
- Nattinger, A. B., Laud, P. W., Bajorunaite, R., Sparapani, R. A., & Freeman, J. L. (2004). An algorithm for the use of Medicare claims data to identify women with incident breast cancer. Health services research, 39(6 Pt 1), 1733–1749. <https://doi.org/10.1111/j.1475-6773.2004.00315.x>.
 - The new algorithm has better performance characteristics than previously proposed algorithms evaluating data from the Population-based Surveillance, Epidemiology, and End Results (SEER) Tumor Registry database. The ability to examine national patterns of breast cancer care using Medicare claims data would open new avenues for the assessment of quality of care.