2024 MIPS Peer-Reviewed Journal Article Requirement Template

Section 101(c)(1) of the Medicare Access and CHIP Reauthorization Act of 2015 (MACRA) requires submission of new measures for publication in applicable specialty-appropriate, peer-reviewed journals prior to implementing in the Merit-based Incentive Payment System (MIPS). Such measures will be submitted by the Centers for Medicare & Medicaid Services (CMS), to a journal(s), before including any new measure on the MIPS Quality Measures List. The measure submitter shall provide the required information for article submission under the MACRA per the MIPS Annual Call for Quality Measures submission process.

Interested parties submitting measures for consideration through the MIPS Annual Call for Quality Measures must complete the required information by the CMS Annual Call for Measures deadline (8 p.m. ET on May 10, 2024). Some of the information requested below may be listed in specific fields in the CMS Measures Under Consideration (MUC) Entry/Review Information Tool (MERIT); however, to ensure that CMS has all of the necessary information and avoid delays in the evaluation of your submission, please fully complete this form as an attached Word document. The information in MERIT must be consistent with the information below, including the following, but not limited to:

- **Breast Cancer Screening**
- **Affordability and Efficiency**

> **Measure Steward:** Acumen, LLC
> **Measure Developer:** Acumen, LLC
> **Description:** The Breast Cancer Screening episode-based cost measure evaluates a clinician's average risk-adjusted cost to Medicare for care provision in women 40 years of age or older who received a screening mammogram during an episode of care. The measure score is the clinician's risk-adjusted cost for the episode group averaged across all episodes attributed to the clinician. This measure includes costs of services that are clinically related to the attributed clinician's role in managing care during each episode starting from the screening mammogram that opens, or "triggers," the episode through 12 months after the trigger or the next screening mammogram.

I. **Statement**
   - *Background (Why is this measure important?)*

   Women have a 1 in 8 chance of developing breast cancer during their life.[1] Breast cancer accounts for around 30% of all new cancers for women each year. Breast cancer found during screening, before symptoms appear, is less likely to spread, including beyond the breast (metastasis). Early detection makes it easier to treat breast cancer successfully, with a better prognosis for the patient. Screening mammography reduces breast cancer mortality by an estimated 20%-35% in women aged 50-69 years.[2] As such, early detection is one of the most important strategies for preventing deaths from breast cancer, the second leading cause of cancer death in women in the United States.[1]

   However, two challenges with screening mammography are false negatives and false positives. Screening mammograms miss an estimated one in eight breast cancers; research indicates that

---

[1] American Cancer Society. "American cancer society recommendations for the early detection of breast cancer." ACS Breast Cancer Early Detection Recommendations (2022).
[2] Elmore JG, Armstrong K, Lehman CD, Fletcher SW. Screening for breast cancer. JAMA. 2005;293(10):1245-1256. doi:10.1001/jama.293.10.1245.

1.3%-45% of missed cancers were visible on mammograms.[3,4] False positives are common; around half of all women getting annual mammograms over a 10-year period will have a false positive finding, which can result in unnecessary testing and patient anxiety.[5] To balance the risks of false negatives and false positives, organizations have defined key metrics and developed acceptable ranges for breast cancer screening and diagnosis. Use of these has been associated with improvements in breast imaging programs.[6]

- *Environmental scan (Are there existing measures in this area?)*

Based on a search of the CMS Measure Inventory Tool (CMIT), there is one cost/resource use measure currently in use in the MIPS cost performance category that may be relevant to this measure. The Lumpectomy, Partial Mastectomy, Simple Mastectomy episode-based cost measure (COST_LPMSM_1) and this cost measure may share a focus of services that are clinically related to breast cancer. Additionally, we found three process MIPS quality measures that may be relevant to this measure. These quality measures (listed in Table 1 below) may include metrics focused on similar patient cohorts or clinically related to the care provided for the episode group. While these measures may pertain to breast cancer services, they do not capture the costs of services attributed to a clinician following a patient's screening mammogram.

Table 1. MIPS Measures Potentially Relevant for the Breast Cancer Screening Episode-Based Cost Measure

Measure Title	Measure ID	Measure Description	Measure Type
Lumpectomy, Partial Mastectomy, Simple Mastectomy	01592	The Lumpectomy, Partial Mastectomy, Simple Mastectomy episode-based cost measure evaluates a clinician's risk-adjusted cost to Medicare for patients who undergo partial or total mastectomy for breast cancer during the performance period. The measure score is the clinician's risk-adjusted cost for the episode group averaged across all episodes attributed to the clinician. This procedural measure includes costs of services that are clinically related to the attributed clinician's role in managing care during each episode from 30 days prior to the clinical event that opens, or "triggers," the episode through 90 days after the trigger.	Cost/Resource Use
Breast Cancer Screening	00093	Percentage of women 40 - 74 years of age who had a mammogram to screen for breast cancer in the 27 months prior to the end of the measurement period.	Process
Appropriate Treatment for Patients with Stage I (T1c) – III HER2 Positive Breast Cancer	00073	Percentage of female patients aged 18 to 70 with stage I (T1c) - III HER2 positive breast cancer for whom appropriate treatment is initiated.	Process

---

[3] Ekpo EU, Alakhras M, Brennan P. Errors in mammography cannot be solved through technology alone. *Asian Pac J Cancer Prev.* 2018;19(2):291-301. doi:10.22034/APJCP.2018.19.2.291.

[4] Warren Burhenne LJ, Wood SA, D'Orsi CJ, Feig SA, Kopans DB, O'Shaughnessy KF, Sickles EA, Tabar L, Vyborny CJ, Castellino RA. Potential contribution of computer-aided detection to the sensitivity of screening mammography. Radiology. 2000 May;215(2):554-62.

[5] American Cancer Society. Limitations of mammograms. Accessed May 15, 2022. https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection/mammograms/limitations-of-mammograms.html

[6] Hussain S, Omar A, Shah BA. The breast imaging medical audit: what the radiologist needs to know. *Contemp Diagn Radiol.* 2021;44(8):1-6.

Measure Title	Measure ID	Measure Description	Measure Type
Sentinel Lymph Node Biopsy for Invasive Breast Cancer	000676	The percentage of clinically node negative (clinical stage T1N0M0 or T2N0M0) breast cancer patients before or after neoadjuvant systemic therapy, who undergo a sentinel lymph node (SLN) procedure.	Process

II. Gap Analysis

- *Provide evidence for the measure (What are the gaps and opportunities to improve care?)*

The Breast Cancer Screening episode-based cost measure assesses costs related to breast cancer screening, a current measurement gap in the MIPS cost performance category. Furthermore, an environmental scan of the literature identified three critical areas for improving care and reducing costs, including:
1. Improving screening and diagnostic accuracy
2. Incentivizing early cancer detection
3. Reducing unnecessary resource use

The estimated cost for mammography screening in the U.S. in 2010 was $7.8 billion.[7] While costs associated with appropriate use are not concerning, costs associated with excessive use (e.g., unnecessary repeat imaging) or delayed detection are a significant concern. For example, costs may be high for false-positives and increased follow-up visits.[8] It is estimated that national expenditure for these false-positive follow-ups cost around $4 billion a year.[9] Increasing the accuracy of screenings can lower the need for excessive and expensive follow up treatment. In addition, one study showed that the annual estimated cost for breast cancer screening for women ages 40-49 was $2.13 billion despite unclear benefits for women in this age group receiving screenings.[10]

Moreover, with approximately 6.1 million screening mammograms in Medicare Part B physician/supplier billing, a modest improvement among these radiologists to recall 20% fewer patients would result in up to 92,000 fewer recalls. This would represent roughly $12.7 million in savings (average Medicare allowed amount of $140).[11] A cost measure would incentivize reductions in the costs of care following screening mammography. Choices in diagnostic imaging would have to consider the appropriate modality; for example, ultrasound is cheaper than magnetic resonance imaging (MRI), but has higher rates of repeat usage.[11] The cost measure could also credit providers for early cancer detection, as cancer detected at an early stage is cheaper to treat than at later stages.[11]

Lastly, although the proposed cost measure is not intended to capture global costs of cancer care, delays in detection of breast cancer are expected to contribute to higher costs of treating cancer once detected. One review of the literature on breast cancer treatment costs found the stage at initial diagnosis to be an important determinant of resource use. For instance, cancers diagnosed at stage 2 were determined to have treatment costs that were 32 percent higher than cancers

[7] O'Donoghue, Cristina, Martin Eklund, Elissa M. Ozanne, and Laura J. Esserman. "Aggregate cost of mammography screening in the United States: comparison of current practice and advocated guidelines." *Annals of internal medicine* 160, no. 3 (2014): 145-153.
[8] Morris, Elizabeth, Stephen A. Feig, Madeline Drexler, and Constance Lehman. "Implications of overdiagnosis: impact on screening mammography practices." *Population health management* 18, no. S1 (2015): S-3.
[9] Ong, Mei-Sing, and Kenneth D. Mandl. "National expenditure for false-positive mammograms and breast cancer overdiagnoses estimated at $4 billion a year." *Health affairs* 34, no. 4 (2015): 576-583.
[10] Kunst, Natalia, Jessica B. Long, Xiao Xu, Susan H. Busch, Kelly A. Kyanko, Ilana B. Richman, and Cary P. Gross. "Use and costs of breast cancer screening for women in their 40s in a US population with private insurance." *JAMA internal medicine* 180, no. 5 (2020): 799-801.
[11] Centers for Medicare & Medicaid Services. Medicare provider utilization and payment data: physician and other practitioners. Accessed May 15, 2022. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Charge-Data/Physician-and-Other-Supplier

diagnosed at stage 1. Breast cancers detected in stage 3 and stage 4 were, respectively, found to cost 95% and 109% more than cancers detected in stage 1.[12]

- *Expected outcome (patient care/patient health improvements, cost savings).*

There are a number of strategies or interventions that have been shown to improve screening and diagnostic accuracy for patients and incentivize early cancer detection, which have the potential to help reduce unnecessary resource use.

Improving screening and diagnostic accuracy
Clinicians can improve their screening interpretation performance. For example, providing proper training for clinicians can improve accuracy in interpretations and reduce unnecessary follow-up care. Offering programs early in training for radiologists could lead to improved performance rates.[13] A randomized, controlled trial showed that radiologists having a self-paced DVD containing mammography interpretation content overall improved their mammography interpretive performance.[14]

Another opportunity for performance improvement involves gaining better access to technology for mammograms. Clinicians choose to use the traditional 2-D imaging process or a newer 3-D imaging process – also known as tomosynthesis. Preliminary findings suggest clinicians utilizing tomosynthesis seem to have improved recall rates and more accurate cancer detection.[15] In addition, tomosynthesis offers less compression on the breast and less discomfort, which is a known deterrent for women considering getting a mammogram.[16] Studies examining the use of tomosynthesis have shown the method can potentially produce better imaging results, with fewer false positives/negatives.[17,18,19] Clinicians switching to tomosynthesis could potentially improve patient care by minimizing discomfort and increasing accuracy of breast cancer detection.

Incentivizing early cancer detection

Early detected breast tumors are generally small and easier to successfully treat.[20] Common treatments include simple surgeries and chemotherapy; late or metastatic breast cancer, on the other hand, can require more invasive surgeries, more prolonged chemotherapy, and targeted drug

[12] Sun L, Legood R, Santos-Silva I, et al. Global Treatment Costs of Breast Cancer by Stage: A Systematic Review. PLos One, 2018: 13(11): 30207993.

[13] Miglioretti, Diana L., Charlotte C. Gard, Patricia A. Carney, Tracy L. Onega, Diana SM Buist, Edward A. Sickles, Karla Kerlikowske et al. "When radiologists perform best: the learning curve in screening mammogram interpretation." *Radiology* 253, no. 3 (2009): 632.

[14] Geller, Berta M., Andy Bogart, Patricia A. Carney, Edward A. Sickles, R. A. Smith, Barbara Monsees, Lawrence W. Bassett et al. "Educational interventions to improve screening mammography interpretation: a randomized, controlled trial." *AJR. American journal of roentgenology* 202, no. 6 (2014): W586.

[15] Lowry, Kathryn P., Rebecca Yates Coley, Diana L. Miglioretti, Karla Kerlikowske, Louise M. Henderson, Tracy Onega, Brian L. Sprague et al. "Screening performance of digital breast tomosynthesis vs digital mammography in community practice by patient age, screening round, and breast density." *JAMA Network open* 3, no. 7 (2020): e2011792-e2011792.

[16] Miller, Dawn, Vicki Livingstone, and G. Peter Herbison. "Interventions for relieving the pain and discomfort of screening mammography." *Cochrane Database of Systematic Reviews* 1 (2008).

[17] DePolo, Jamie. "3D Mammograms More Effective Than 2D Mammograms in Women 65 and Older." Accessed October 18, 2022. https://www.breastcancer.org/research-news/3d-mammos-better-than-2d-for-65-and-up.

[18] Houssami, Nehmat, Petra Macaskill, Daniela Bernardi, Francesca Caumo, Marco Pellegrini, Silvia Brunelli, Paola Tuttobene et al. "Breast screening using 2D-mammography or integrating digital breast tomosynthesis (3D-mammography) for single-reading or double-reading–evidence to guide future screening strategies." *European Journal of Cancer* 50, no. 10 (2014): 1799-1807.

[19] Aragon, Lourdes Noemi Santos, and Dafne Soto-Trujillo. "Effectiveness of Tomosynthesis Versus Digital Mammography in the Diagnosis of Suspicious Lesions for Breast Cancer in an Asymptomatic Population." *Cureus* 13, no. 3 (2021).

[20] Sumbaly, Ronak, N. Vishnusri, and S. Jeyalatha. "Diagnosis of breast cancer using decision tree data mining technique." *International Journal of Computer Applications* 98, no. 10 (2014).

therapy which can become costly.[21] Therefore, screening and early detection decreases spending costs associated with cancer treatment.[22]

Women may not undergo screening due to lack of information on how and when to get mammogram screenings. Studies have shown that sending out reminders, notices, or supplemental follow-up materials increased the number of patients receiving mammogram screenings.[23,24] Patients have also expressed significant levels of anxiety before, during, and after undergoing mammograms. Clinicians can implement multiple strategies to lower anxiety levels in patients and thereby increase participation in breast cancer screenings.[25] These efforts can be simple, such as providing education about screening mammography prior to the appointment,[26] minimizing wait times,[27] and offering relaxation techniques.[28] Lower anxiety levels not only improve patients' mental health, but can also improve the patient-clinician relationship to further increase participation in future mammogram screenings.

Reducing unnecessary resource use

By improving screening and diagnostic accuracy for patients and incentivizing early cancer detection using the strategies outlined above, clinicians can reduce unnecessary resource use associated to diagnostic testing and costs associated with delayed treatment of breast cancer.

- *Recommendation for the measure (Is it based on a study, consensus opinion, USPSTF recommendation etc.?)*

The American Cancer Society (ACS) states that early breast cancer detection and quality cancer treatment are two of the most important strategies for preventing deaths from breast cancer. ACS recommends that for women at average breast cancer risk, women between 40 and 44 have the option to start screening with a mammogram every year, women between 45 and 54 should get mammograms every year, and women 55 and older can switch to a mammogram every other year, or they can choose to continue yearly. Additionally, screening should continue as long as a woman is in good health and is expected to live at least 10 more years. ACS' breast cancer screening guidelines consider having had either a 2D or 3D mammogram as being in line with current screening recommendations.[29]

---

[21] American Cancer Society. Treatment of Breast Cancer by Stage. Accessed November 14, 2022.
https://www.cancer.org/cancer/breast-cancer/treatment/treatment-of-breast-cancer-by-stage.html
[22] Gross, Cary P., Jessica B. Long, Joseph S. Ross, Maysa M. Abu-Khalaf, Rong Wang, Brigid K. Killelea, Heather T. Gold, Anees B. Chagpar, and Xiaomei Ma. "The cost of breast cancer screening in the Medicare population." *JAMA internal medicine* 173, no. 3 (2013): 220-226.
[23] Buist, Diana SM, Hongyuan Gao, Melissa L. Anderson, Tracy Onega, Susan Brandzel, Melissa A. Rabelhofer, Susan Carol Bradford, and Erin J. Aiello Bowles. "Breast cancer screening outreach effectiveness: Mammogram-specific reminders vs. comprehensive preventive services birthday letters." *Preventive medicine* 102 (2017): 49-58.
[24] Baron, Roy C., Stephanie Melillo, Barbara K. Rimer, Ralph J. Coates, Jon Kerner, Nancy Habarta, Sajal Chattopadhyay et al. "Intervention to increase recommendation and delivery of screening for breast, cervical, and colorectal cancers by healthcare providers: a systematic review of provider reminders." *American journal of preventive medicine* 38, no. 1 (2010): 110-117.
[25] Pai, Vidya R., and Murray Rebner. "How to minimize patient anxiety from screening mammography." *Journal of Breast Imaging* 3, no. 5 (2021): 603-606.
[26] Lee, Jiyon, Lara A. Hardesty, Nathan M. Kunzler, and Andrew B. Rosenkrantz. "Direct interactive public education by breast radiologists about screening mammography: impact on anxiety and empowerment." *Journal of the American College of Radiology* 13, no. 1 (2016): 12-20.
[27] Randel S. Mammograms: reducing patient anxiety. Radiol Technol 2016;87(6):707–709.
[28] Zavotsky, Kathleen Evanovich. "The effects of music on pain and anxiety during screening mammography." *Number 3/June 2014* 18, no. 3 (1969): E45-E49.
[29] American Cancer Society. "American cancer society recommendations for the early detection of breast cancer." ACS Breast Cancer Early Detection Recommendations (2022).

III. Reliability/Validity

- *What testing has been performed at the level of implementation? (MIPS requires full measure testing at the individual clinician level (and may also need to be tested at the group level) for MIPS Clinical Quality Measures (CQMs) and Electronic Clinical Quality Measures (eCQMs) collection types. Administrative claims measures tested at the group level require a reliability threshold to be implemented at the group level.)*

  Please provide testing results including the N value, Bonnie test case results, correlation coefficient and any other pertinent information or values to be considered.

  o *Reliability Testing Results at the accountable entity level*

Reliability evaluates a measure's ability to consistently differentiate the performance of one clinician from another. The signal-to-noise ratio is used to estimate reliability, which indicates how much of the variation in the measure score is explained by differences among clinicians' performance (i.e., signal) instead of differences within each clinician's performance (i.e., noise). Specifically, noise is the variation from one episode to another during the performance period for a particular clinician.

Table 2 shows reliability metrics at various testing volume thresholds. While higher thresholds yield higher reliability results, it is at the cost of further reducing the number of clinicians and clinician groups eligible for the measure, which would reduce the potential impact of the measure. We used a 20-episode volume threshold (bolded in the table below). If the measure is implemented in the Merit-based Incentive Payment System (MIPS) in the future, CMS will establish a case minimum through notice-and-comment rulemaking.

At the testing volume of 20 episodes, the reliability score for the Breast Cancer Screening cost measure is high, specifically 0.97 at the TIN level and 0.93 at the TIN-NPI level (Table 2). CMS generally considers 0.4 as the threshold indicating 'moderate' reliability and 0.7 indicating 'high' reliability, which is supported by previous work into reliability and the threshold was finalized in the 2022 Physician Fee Schedule final rule.[30,31]

Table 2. Sample Size, Reliability Score, and Proportion of Clinicians at Various Testing Volume Thresholds

Testing Volume Threshold	TIN		TIN-NPI	
	Number of TINs	Reliability Score	Number TIN-NPIs	Reliability Score
10	2,497	0.96	18,459	0.90
20	**2,441**	**0.97**	**16,289**	**0.93**
30	2,411	0.97	14,838	0.94
40	2,382	0.98	13,835	0.95
50	2,343	0.98	13,031	0.96

  o *Face Validity Testing Results, Clinician Sites*

Face validity testing was conducted for the cost measure. Out of eight total experts and patients/caregivers who voted on face validity, seven voted in agreement that the cost measure could differentiate good from poor quality care. There were six votes for agree (75%), one vote

---

[30] Mathematica, Inc., "Memorandum: Reporting Period and Reliability of AHRQ, CMS 30-Day and HAC Quality Measures – Revised," http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/hospital-value-based-purchasing/Downloads/HVBP_Measure_Reliability-.pdf.

[31] CMS, "Medicare Program; CY 2022 Payment Policies Under the Physician Fee Schedule and Other Changes to Part B Payment Policies; Medicare Shared Savings Program Requirements; Provider Enrollment Regulation Updates; and Provider and Supplier Prepayment and Post-Payment Medical Review Requirements," 86 FR 64996-66031.

for strongly agree (12.5%), and one vote for undecided (12.5%) about whether the cost measure could distinguish good from poor quality care.

   o *Empiric Validity Testing Results at the accountable entity level*

This measure is tested using a mediation analysis to demonstrate construct validity and a correlation analysis to demonstrate concurrent validity.

Table 3. Estimated Effect on Treatment Choices on the Measure Score

Service Categories	Coefficient in Thousands [95% Confidence Interval] (p-value)			
	TIN		TIN-NPI	
	Model 1: Mean O/E = Mean Cost of Treatment Choices + Mean Cost of Adverse Events	Model 2: Mean Cost of Adverse Events = Mean Cost of Treatment Choices	Model 1: Mean O/E = Mean Cost of Treatment Choices + Mean Cost of Adverse Events	Model 2: Mean Cost of Adverse Events = Mean Cost of Treatment Choices
Adverse Events	44.17 [-452.6, 540.96] (p=0.86)	-	206.89 [90.55, 323.24] (p<.01)	-
Outpatient Evaluation & Management Services	1054.5 [1040.9, 1068.1] (p<.01)	0.00 [0.00, 0.00] (p=0.65)	1069.5 [1063.4, 1075.6] (p<.01)	0.00 [0.00, 0.00] (p=0.60)
Major Procedures	496.63 [440.77, 552.49] (p<.01)	-0.02 [-0.03, -0.02] (p<.01)	150.21 [137.67, 162.75] (p<.01)	-0.01 [-0.01, -0.01] (p<.01)
Ambulatory/Minor Procedures	-272.8 [-306.2, -239.4] (p<.01)	0.00 [0.00, 0.00] (p=0.69)	-252.9 [-260.9, -245.0] (p<.01)	0.00 [0.00, 0.00] (p=0.84)
Laboratory, Pathology, and Other Tests	-614.7 [-709.6, -519.8] (p<.01)	0.03 [0.03, 0.04] (p<.01)	-71.10 [-91.53, -50.66] (p<.01)	0.01 [0.01, 0.01] (p<.01)
Imaging Services	1031.8 [1021.7, 1041.8] (p<.01)	0.00 [0.00, 0.00] (p=0.69)	1038.6 [1034.2, 1043.0] (p<.01)	0.00 [0.00, 0.00] (p=0.89)
Anesthesia Services	9761.0 [8678.8, 10843] (p<.01)	0.37 [0.29, 0.46] (p<.01)	-357.6 [-685.9, -29.38] (p=0.03)	0.32 [0.28, 0.37] (p<.01)

The mediation analysis estimates both the direct and indirect effect of treatment choices on the measure score. This analysis first estimates the correlation between treatment choices and the measure score while controlling for adverse outcomes. Then the correlation between treatment choices and related adverse outcomes is calculated to demonstrate the indirect effect. Generally, adverse outcomes are non-trigger inpatient hospitalizations, non-trigger emergency room visits, and post-acute care. The remaining service categories are typically considered treatment. The results show that spending on ambulatory/minor procedures, anesthesia, and laboratory testing is statistically associated with a better measure score. On the other hand, the main drivers of cost are major procedures, outpatient evaluation and management, and imaging services. While major procedures show a statistical association with lower costs of adverse events, the reduction in costs of adverse events is not enough to offset the cost of major procedures, which further reinforces the need for early detection to avoid high-intensity treatments. Even though outpatient evaluation and management and imagining are essential, the results indicate that they can be prone to overuse because more spending on these services does not offset the costs of adverse events.

The correlation analysis shows that the cost measure score is positively associated with the Breast Cancer Screening Recall Rate measure (TIN: r = 0.32, p-value <0.001; TIN-NPI: r = 0.27, p-value <0.001) and OP-39 Breast Cancer Screening Recall Rate (TIN: r = 0.23, p-value <0.001; TIN-NPI: r = 0.21, p-value <0.001), which is aligned with mediation analysis in suggesting that imaging services are necessary but can be prone to overuse. The cost measure score is negatively associated with Breast Cancer Screening with an Eventual Breast Cancer Diagnosis: PPV1 measure (TIN: r = -0.12, p-value <0.001; TIN-NPI: r = -0.13, p-value <0.001) and Use of Biopsy After Diagnostic Follow-up with an Eventual Breast Cancer Diagnosis: PPV3 (TIN: r = -0.15, p-value <0.001; TIN-NPI: r = -0.14, p-value <0.001), which are also aligned in with the mediation analysis in emphasizing the importance of cancer detection in reducing costs of delayed treatment or adverse events.

o   *Data Element/Patient Encounter Level Testing*

This is not applicable to the Breast Cancer Screening episode-based cost measure.

o   *Exclusion Frequency*

Exclusions specific to this measure are developed with input from the project's Technical Expert Panel (TEP). These exclusion criteria ensure that the reportable episode populations are more homogenous and comparable than all episodes meeting the triggering logic for the measure. The table below displays descriptive statistics of all episodes meeting the measure's triggering logic, excluded episodes, and final reportable episodes at both TIN and TIN-NPI levels.

Table 4: Statistics for Measure Exclusions

Exclusion Criteria	Episode Count	% of All Episodes	Mean Episode Observed Cost	Standard Deviation of Episode Observed Cost
All Episodes Meeting Trigger Logic	5,266,086	100.00%	$278	$558
Patient Death in Episode	54,245	1.03%	$309	$1,007
Outliers	93,773	1.78%	$752	$1,689
TIN Does Not Meet Case Minimum	8,123	0.15%	$259	$471
TIN-NPI Does Not Meet Case Minimum	72,030	1.37%	$237	$443
Episodes with Patients with History of Breast Cancer on or 365 Days Prior to Trigger	417,179	7.92%	$440	$1,017
Male Patients	226	<0.01%	$165	$150
Beneficiary Less Than 40 Years of Age	1,803	0.03%	$278	$403
Reportable Episodes - Group Reporting	**4,733,329**	**89.88%**	**$255**	**$436**
Reportable Episodes - Individual Reporting	**4,685,647**	**88.98%**	**$255**	**$436**

o *What were the minimum sample sizes used for reliability results?*

Please refer to table 2 for the breakdown of TINs and TIN-NPIs that meet the 10, 20, 30, 40, and 50 case volume thresholds used to assess reliability.

- *Other Information*
  o *Is it risk adjusted? If so, how?*

The Breast Cancer Screening episode-based cost measure is a risk-adjusted measure. The risk adjustment model for this measure uses an ordinary least squares (OLS) regression model, which utilizes variables from the CMS Hierarchical Condition Code Version 24 (CMS-HCC V24) Risk Adjustment Model. This includes comorbidities captured by HCC codes that map with ICD-10-CM codes, interaction variables accounting for a range of comorbidities, patient age category, and patient disability status. Additional risk adjustors that are clinically relevant to this measure were developed with input from the TEP. This cost measure is further stratified into sub-groups (breast cancer detection and no breast cancer detection), and risk adjustment is performed separately for episodes within each sub-group of this measure to allow for comparisons within more clinically homogenous cohorts.

  o *What benchmarking information is available?*

This measure provides a score evaluating a clinician's risk-adjusted resource use as a dollar amount which can be compared with the scores for other clinicians, as well as relevant national averages.

  o *Collection Type: Specify the data collection type.*

This measure uses administrative Medicare claims data.

  o *Specify measure stage of development.*

This measure is fully developed.

  o *For Patient Reported Outcome Performance Measures:*
    - *The survey or tool has been tested and doesn't require modifications based on results?*
    - *Patient/encounter level testing for each critical data element doesn't require changes to the tool base on the results?*

This is not applicable to the Breast Cancer Screening episode-based cost measure.

IV. Endorsement

- *Provide the Consensus-Based Entity (CBE) (i.e., Partnership for Quality Measures (PQM)) endorsement status (and CBE ID) and/or other endorsing body. If the measure is only endorsed for paper records, please note endorsement for only the data source being submitted.*

This measure is not currently endorsed by the CBE and has never been submitted for endorsement.

V. Summary

- *Alignment with CMS Meaningful Measures Initiative or MACRA (if applicable).*

This cost measure aligns with the Cost goal of CMS's Meaningful Measures Initiative, and the domain of Affordability and Efficiency. Through this measure, we aim to improve accuracy in early breast cancer detection, minimize unnecessary testing costs, and reduce patient stress and discomfort.

- *Relevance to MIPS or other CMS programs.*

This measure would be proposed in future rulemaking for inclusion in the Cost performance category for MIPS. If finalized through rulemaking, the measure would assess clinician performance in the Cost performance category, and could count toward the overall MIPS final score. This measure will incentivize reductions in the costs of care following screening mammography and credit providers for early cancer detection.

- *Rationale: Use of measure for inclusion in program (specialty society, regional collaborative, other).*

The Breast Cancer Screening episode-based cost measure was selected for development because it fills a measurement and performance gap for radiologists in MIPS. This measure addresses unmet needs in MIPS, as none of the current clinician-level measures evaluate the outcome of screening mammograms, and goes beyond avoiding diagnostic errors to account for cost and patient experience. It was developed with extensive engagement from interested parties (i.e., clinicians, persons with lived experience, and the general public) through several mechanisms including a TEP, person and family engagement opportunities, and beta testing. This measure's development is aligned with episode-based cost measures currently used in the program.

- *Public reporting (if applicable).*

This is not applicable to the Breast Cancer Screening episode-based cost measure.

- *Preferable relevant peer-reviewed journal for publication.*

Journal of the American College of Radiology, Journal of the American Medical Association Oncology

- *Rationale as to how the measure correlates to existing cost measures and improvement activities, as applicable and feasible.*

The Lumpectomy, Partial Mastectomy, Simply Mastectomy (COST_LPMSM_1) episode-based cost measure correlates with this cost measure because they both focus on services that are clinically related to breast cancer. There are no improvement activities in MIPS specific to the breast cancer clinical area. However, there is an improvement activity related to adequate cancer screening, timely follow-up, and optimal patient care, Use of Computable Guidelines and Clinical Decision Support to Improve Adherence for Cervical Cancer Screening and Management Guidelines (IA_PM_23), which may correlate with the Breast Cancer Screening measure as it aims to improve adherence to clinical practice guidelines for cancer screening. Additionally, there are improvement activities related to cost management, Cost Display for Laboratory and Radiographic Orders (IA_PSPA_25) and Implementation of Analytic Capabilities to Manage Total Cost of Care for Practice Population (IA_PSPA_17), which may correlate with the Breast Cancer Screening measure as it aims to reduce costs of breast cancer episodes while also improving patient outcomes.