**National Quality Forum—Evidence (subcriterion 1a)**

**Measure Number** (*if previously endorsed*)**:** 2372

**Measure Title**: Breast Cancer Screening

**IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here:** Click here to enter composite measure #/ title

**Date of Submission**: 4/16/2018

|  |
| --- |
| **Instructions**  *Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.*  *Complete* ***EITHER 1a.2, 1a.3 or 1a.4*** *as applicable for the type of measure and evidence.*  *For composite performance measures:*  *A separate evidence form is required for each component measure unless several components were studied together.*  *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*   * All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed. * If you are unable to check a box, please highlight or shade the box for your response. * Contact NQF staff regarding questions. Check for resources at [Submitting Standards webpage](http://www.qualityforum.org/Measuring_Performance/Submitting_Standards.aspx). |

|  |
| --- |
| **Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the evidence for this measure meets NQF’s evaluation criteria.**   1a. Evidence to Support the Measure Focus The measure focus is evidence-based, demonstrated as follows:   * Outcome: **3** Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias. * Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence **4** that the measured intermediate clinical outcome leads to a desired health outcome. * Process: **5** a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence **4** that the measured process leads to a desired health outcome. * Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence **4**  that the measured structure leads to a desired health outcome. * Efficiency: **6** evidence not required for the resource use component. * For measures derived from patient reports, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful. * Process measures incorporating Appropriate Use Criteria: See NQF’s guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.   **Notes**  **3.** Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.  **4.** The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation [(GRADE) guidelines](http://www.gradeworkinggroup.org) and/or modified GRADE.  **5.** Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.  **6.** Measures of efficiency combine the concepts of resource use and quality (see NQF’s [Measurement Framework: Evaluating Efficiency Across Episodes of Care](http://www.qualityforum.org/Publications/2010/01/Measurement_Framework__Evaluating_Efficiency_Across_Patient-Focused_Episodes_of_Care.aspx); [AQA Principles of Efficiency Measures](http://www.aqaalliance.org/files/PrinciplesofEfficiencyMeasurementApril2006.doc)). |

**1a.1.This is a measure of**: (*should be consistent with type of measure entered in De.1*)

Outcome

Outcome: Click here to name the health outcome

Patient-reported outcome (PRO): Click here to name the PRO

*PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors.* (*A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)*

Intermediate clinical outcome (*e.g., lab value*): Click here to name the intermediate outcome

Process: Breast Cancer Screening

Appropriate use measure: Click here to name what is being measured

Structure: Click here to name the structure

Composite: Click here to name what is being measured

**1a.2** **LOGIC MODEL** Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient’s health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

2018 Submission

Females age 50-74 years >> screening for breast cancer is performed >> abnormal screening result >> evaluation and follow-up >> early detection and treatment >> improved length and/or quality of life

2014 Submission

Females age 50-74 years >> screening for breast cancer is performed >> results are evaluated >> results are positive for breast cancer >> treatment given >> improved length and/or quality of life

**1a.3** **Value and Meaningfulness:**  **IF** this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

**\*\*RESPOND TO ONLY ONE SECTION BELOW -EITHER 1a.2, 1a.3 or 1a.4) \*\***

**1a.2** **FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.**

**1a.3.****SYSTEMATIC REVIEW(SR) OF THE EVIDENCE (for intermediate outcome, PROCESS, or STRUCTURE PERFORMANCE measures, including those that are instrument-based) If the evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.**

**What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)**

☐ Clinical Practice Guideline recommendation (with evidence review)

☒ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

|  |  |
| --- | --- |
| **USPSTF Recommendation:**   * **Title** * **Author** * **Date** * **Citation, including page number** * **URL** | 2018 Submission  U.S. Preventive Services Task Force (USPSTF). 2016. Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Annals of Internal Medicine* 164(4) 279-296. doi: 10.7326/M15-2886.  2014 Submission  U.S. Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2009 Nov 17; 151(10):716-26, W-236.  **URL:** <http://www.uspreventiveservicestaskforce.org/uspstf/uspsbrca.htm> |
| Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR. | 2018 Submission  The USPSTF recommends biennial screening mammography for women aged 50 to 74 years. Grade: B Recommendation  The decision to start screening mammography in women prior to age 50 years should be an individual one. Women who place a higher value on the potential benefit than the potential harms may choose to begin biennial screening between the ages of 40 and 49 years. Grade: C Recommendation  The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of screening mammography in women aged 75 years or older. Grade: I Recommendation  The USPSTF concludes that the current evidence is insufficient to assess the benefits and harms of digital breast tomosynthesis (DBT) as a primary screening method for breast cancer. Grade: I Recommendation  The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of adjunctive screening for breast cancer using breast ultrasonography, magnetic resonance imaging, DBT, or other methods in women identified to have dense breasts on an otherwise negative screening mammogram. Grade: I Recommendation  2014 Submission  The USPSTF recommends biennial screening mammography for women aged 50 to 74 years. (B Recommendation)  The decision to start regular, biennial screening mammography before the age of 50 years should be an individual one and take patient context into account, including the patient’s values regarding specific benefits and harms. (C Recommendation)  The USPSTF concludes that the current evidence is insufficient to assess the additional benefits and harms of screening mammography in women 75 years or older. (I Statement)  The USPSTF recommends against teaching breast self-examination. (D Recommendation)  The USPSTF concludes that the current evidence is insufficient to assess the additional benefits and harms of clinical breast examination beyond screening mammography in women 40 years or older. (I Statement)  The USPSTF concludes that the current evidence is insufficient to assess the additional benefits and harms of either digital mammography or magnetic resonance imaging (MRI) instead of film mammography as screening modalities for breast cancer. (I Statement) |
| Grade assigned to the **evidence** associated with the recommendation with the definition of the grade | 2018 Submission  The USPSTF concludes with moderate certainty that the net benefit of screening mammography in women aged 50 to 74 years is moderate.  For a general population of women aged 40 to 49 years, there is moderate certainty that the net benefit of screening mammography in the general population of women.  For women age 75 years and older, there is insufficient evidence on mammography screening and the balance of benefits and harms cannot be determined.  The USPSTF concludes that the evidence on DBT as a primary screening modality for breast cancer is insufficient, and the balance of benefits and harms cannot be determined. The USPSTF concludes that the evidence on adjunctive screening for breast cancer using breast ultrasound, MRI, DBT, or other methods in women identified to have dense breasts on an otherwise negative screening mammogram is insufficient, and the balance of benefits and harms cannot be determined.  2014 Submission  In the analytic framework, the USPSTF addressed in Key Question 1a whether screening with mammography (film or digital) or MRI decrease breast cancer mortality among women age 40-49 years and 70 and older. For this question, the USPSTF used seven studies in their meta-analysis, all rated *Fair*.  The USPSTF grades the quality of the overall evidence for a service on a 3-point scale (good, fair, poor). A *Fair* rating meansevidence is sufficient to determine effects on health outcomes, but the strength of the evidence is limited by the number, quality, or consistency of the individual studies, generalizability to routine practice, or indirect nature of the evidence on health outcomes. |
| Provide all other grades and definitions from the evidence grading system | 2018 Submission  N/A  2014 Submission  *Good***:** Evidence includes consistent results from well-designed, well-conducted studies in representative populations that directly assess effects on health outcomes.  *Poor***:** Evidence is insufficient to assess the effects on health outcomes because of limited number or power of studies, important flaws in their design or conduct, gaps in the chain of evidence, or lack of information on important health outcomes. |
| Grade assigned to the **recommendation** with definition of the grade | 2018 Submission  Grade B: The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial.  Grade C: The USPSTF recommends selectively offering or providing this service to individual patients based on professional judgment and patient preferences. There is at least moderate certainty that the net benefit is small.  Grade I: The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.  2014 Submission  The measure is based on a guideline to screen women age 50-74 years biennially, which is a grade B recommendation (Grade B: The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial). |
| Provide all other grades and definitions from the recommendation grading system | 2018 Submission  Grade A: The USPSTF recommends the service. There is high certainty that the net benefit is substantial.  Grade D: The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits.  2014 Submission  A Recommendation: The USPSTF recommends the service. There is high certainty that the net benefit is substantial.  C Recommendation: Clinicians may provide this service to selected patients depending on individual circumstances. However, for most individuals without signs or symptoms there is likely to be only a small benefit from this service.  D Recommendation: The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits.  I Statement: The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined. |
| Body of evidence:   * Quantity – how many studies? * Quality – what type of studies? | 2018 Submission  The evidence report supporting this guideline outlines the quantity and quality of evidence (Nelson et al 2016).  Key Question 1: What is the effectiveness of routine mammography screening in reducing breast cancer–specific and all–cause mortality, and how does it differ by age, risk factor, and screening interval?   * 65 studies (including 8 RCTs) of fair quality assessed breast cancer mortality in relationship to screening   Key Question 2: What is the effectiveness of routine mammography screening in reducing the incidence of advanced breast cancer and treatment-related morbidity, and how does it differ by age, risk factors, and screening interval?   * 20 studies (including 8 RCTs) of fair quality assessed incidence of advanced breast cancer in relationship to screening   Key Question 3: How does the effectiveness of routine breast cancer screening in reducing breast cancer–specific and all-cause mortality vary by different screening modality?   * No studies of tomosynthesis, ultrasound, or MRI addressed this question.   Key Question 4: How does the effectiveness of routine breast cancer screening in reducing the incidence of advanced breast cancer and treatment-related morbidity vary by different screening modality?   * 2 case studies of unknown quality compared digital mammography versus tomosynthesis and digital mammography reported detection rates by cancer stage using various categories of cancer staging.   Key Question 5: What are the harms of routine mammography screening, and how do they differ by age, risk factor, and screening interval?   * 53 studies (including meta-analyses, reviews, modeling studies, observational studies and a surveillance analysis of good and fair quality) assessed overdiagnosis, impact of false-positive and false-negative screening results on women, radiation exposure and incidence of pain, discomfort or distress after screening.   Key Question 6: How do the harms of routine breast cancer screening vary by different screening modality?   * 6 observational studies of fair or unknown quality compared false-positive recall rates of screening for breast cancer using mammography and tomosynthesis, or clinical breast exam compared with mammography alone.   Nelson HD, Cantor A, Humphrey L, Fu R, Pappas M, Daeges M, Griffin J. Screening for Breast Cancer: A Systematic Review to Update the 2009 U.S. Preventive Services Task Force Recommendation. Evidence Synthesis No. 124. AHRQ Publication No. 14-05201-EF-1. Rockville, MD: Agency for Healthcare Research and Quality; 2016.  2014 Submission  The 2009 evidence review included a meta-analysis of 7 randomized controlled trials of mammography screening. A modeling study estimated the benefits/harms of screening of different screening scenarios. The USPSTF Quality Rating for studies used in the meta analysis was fair. |
| Estimates of benefit and consistency across studies | 2018 Submission  The USPSTF found adequate evidence that mammography screening reduces breast cancer mortality in women aged 40 to 74 years. The number of breast cancer deaths averted increases with age; women aged 40 to 49 years benefit the least and women aged 60 to 69 years benefit the most. Age is the most important risk factor for breast cancer, and the increased benefit observed with age is at least partly due to the increase in risk. Women aged 40 to 49 years who have a first-degree relative with breast cancer have a risk for breast cancer similar to that of women aged 50 to 59 years without a family history. Direct evidence about the benefits of screening mammography in women aged 75 years or older is lacking.  The USPSTF found inadequate evidence on the benefits and harms of DBT as a primary screening method for breast cancer. Similarly, the USPSTF found inadequate evidence on the benefits and harms of adjunctive screening for breast cancer using breast ultrasonography, MRI, DBT, or other methods in women identified to have dense breasts on an otherwise negative screening mammogram. In both cases, while there is some information about the accuracy of these methods, there is no information on the effects of their use on health outcomes, such as breast cancer incidence, mortality, or overdiagnosis rates.  2014 Submission  There is convincing evidence that screening with film mammography reduces breast cancer mortality, with a greater absolute reduction for women aged 50 to 74 years than for women aged 40 to 49 years. The strongest evidence for the greatest benefit is among women aged 60 to 69 years. Among women 75 years or older, evidence of benefits of mammography is lacking. |
| What harms were identified? | 2018 Submission  The USPSTF found adequate evidence that screening for breast cancer with mammography results in harms for women aged 40 to 74 years. The most important harm is the diagnosis and treatment of noninvasive and invasive breast cancer that would otherwise not have become a threat to a woman’s health, or even apparent, during her lifetime (that is, overdiagnosis and overtreatment). False-positive results are common and lead to unnecessary and sometimes invasive follow-up testing, with the potential for psychological harms (such as anxiety). False-negative results (that is, missed cancer) also occur and may provide false reassurance. Radiation-induced breast cancer and resulting death can also occur, although the number of both of these events is predicted to be low.  The USPSTF found inadequate evidence on the benefits and harms of DBT as a primary screening method for breast cancer. Similarly, the USPSTF found inadequate evidence on the benefits and harms of adjunctive screening for breast cancer using breast ultrasonography, MRI, DBT, or other methods in women identified to have dense breasts on an otherwise negative screening mammogram. In both cases, while there is some information about the accuracy of these methods, there is no information on the effects of their use on health outcomes, such as breast cancer incidence, mortality, or overdiagnosis rates.  2014 Submission  Harms associated with screening for breast cancer include unnecessary imaging tests and biopsies in women without cancer and psychological harms and inconvenience due to false-positive screening results. Additional harms include treatment of cancer that would not become clinically apparent during a woman's lifetime (overdiagnosis) and the harms of unnecessary earlier treatment of breast cancer that would have become clinically apparent but would not have shortened a woman's life. Radiation exposure (from radiologic tests), although a minor concern, is also a consideration. The USPSTF determined that adequate evidence suggests that the overall harms associated with mammography are moderate for every age group considered. However, false-positive results are more common for women aged 40 to 49 years, whereas overdiagnosis is a greater concern for women in the older age groups.  Results from randomized controlled trials show that screening mammography can help reduce the number of deaths from breast cancer, especially for those over age 50. The USPSTF noted with moderate certainty that the net benefits of screening mammography in women aged 50 to 74 years were at least moderate, and that the greatest benefits were seen in women aged 60 to 69 years. Thus, the harms did not outweigh the benefits in this age group. |
| Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR? | 2018 Submission  We are not aware of any published studies since the systematic review that would impact the recommendation. There are other clinical guidelines that recommend the use of digital breast tomosynthesis as a primary screening method for breast cancer, which we summarize in the section below.  2014 Submission  N/A |

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**1a.4 OTHER SOURCE OF EVIDENCE**

*If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.*

2018 Submission

The National Comprehensive Cancer Network (NCCN) publishes statements of evidence and expert consensus of currently accepted approaches to treatment. The American College of Radiology (ACR) publishes Appropriateness Criteria® for radiology procedures. NCCN’s 2017 breast cancer screening clinical practice guideline and ACR’s 2017 appropriateness criteria for breast cancer screening recommend the use of digital breast tomosynthesis for primary screening of breast cancer.

2014 Submission

N/A

**1a.4.1** **Briefly SYNTHESIZE the evidence that supports the measure.** A list of references without a summary is not acceptable.

2018 Submission

NCCN states: “Combined use of digital mammography (two-dimensional, 2D) in conjunction with digital breast tomosynthesis (DBT) appears to improve cancer detection and reduce false-positive call-back rates. Tomosynthesis allows acquisition of three-dimensional (3D) data using a moving x-ray and digital detector. These data are reconstructed using computer algorithms to generate thin sections of images. The combined use of 2D and DBT results in double the radiation exposure compared with mammography alone. However, this increase in radiation dose falls below limits of radiation set by the U.S. Food and Drug Administration for standard mammography. The radiation dose can be minimized by newer tomosynthesis techniques that create a synthetic 2D image, which may obviate the need for a conventional digital image.”

ACR states: “Digital breast tomosynthesis (DBT) can address some of the limitations encountered with standard mammographic views. In addition to planar images, DBT allows for creation and viewing of thin-section reconstructed images that may decrease the lesion-masking effect of overlapping normal tissue and reveal the true nature of potential false-positive findings without the need for recall. Several studies confirm that in a screening setting, the cancer detection rate is increased with use of DBT compared with 2-D mammography alone. Additionally, the rate of recall for benign findings (false-positives) can be decreased. Some authors found these advantages to be especially pronounced in women under age 50, in those with dense breasts, and with lesion types including spiculated masses and asymmetries. Interpretation time for DBT images is greater than for standard mammography. Additionally, dose is increased if standard 2-D images are obtained in addition to DBT images. However, synthesized reconstructed images (a virtual planar image created from the tomographic dataset) may replace the need for a 2-D correlative view; current data suggest that these synthetic images perform as well as standard full-field digital images. DBT is almost always performed as part of an examination that also includes digital mammography. The digital mammography part of the examination may be in the form of traditional projection mammography or synthesized image from the DBT data.”

2014 Submission

N/A

**1a.4.2 What process was used to identify the evidence?**

2018 Submission

NCCN: The development of the NCCN Guidelines is an ongoing and iterative process, which is based on a critical review of the best available evidence and derivation of recommendations by a multidisciplinary panel of experts in the field of cancer. Prior to the annual update of the Guidelines, an electronic search of the PubMed database, provided by the U.S. National Library of Medicine, is performed to obtain key literature published since the previous Guidelines update. The PubMed database was chosen as it remains the most widely used resource for medical literature and indexes only peer-reviewed biomedical literature. Articles from additional sources (e.g., e-publications ahead of print, meeting abstracts) deemed as relevant to the Guidelines may be included in the literature review process.

ACR: Appropriateness criteria are based on expert consensus and evidence review. A literature search was conducted in December 2015 and updated on March 2016 to identify additional evidence published since the ACR Appropriateness Criteria® Breast Cancer Screening topic was finalized. 379 articles were found. Twenty-four articles were added to the bibliography. The remaining articles were not used due to either poor study design, the articles were not relevant or generalizable to the topic, the results were unclear, misinterpreted, or biased, or the articles were already cited in the original bibliography. The author added 27 citations from bibliographies, websites, or books that were not found in the new literature search.

2014 Submission

N/A

**1a.4.3.** **Provide the citation(s) for the evidence.**

2018 Submission

National Comprehensive Cancer Network (NCCN). 2017. “Breast Cancer Screening and Diagnosis.” (April 13, 2018). Guideline available at: <https://www.nccn.org/professionals/physician_gls/pdf/breast-screening.pdf>

American College of Radiology (ACR). 2017. “*ACR Appropriateness Criteria®: Breast Cancer Screening*.” (April 13, 2018). Guideline available at: <https://acsearch.acr.org/docs/70910/Narrative/>

2014 Submission

N/A