Table A-2. Effectiveness of noninvasive diagnostic tests for breast abnormalities

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Original Key Questions/Conclusions | | | Updated Key Questions/Conclusions (2011-2012) | 2009 Prediction | | Concordance | | |
| Effectiveness of Noninvasive Diagnostic Tests for Breast Abnormalities (Original report date - Feb 2006[3](#_ENREF_3) and Update report date - Feb 2012[4](#_ENREF_4)) | | | | | | | | |
| Key Question 1 - For the following diagnostic tests as applied to the breast (positron emission tomography (PET) scanning, scintimammography (SC), magnetic resonance imaging (MRI), and ultrasonography (US)) what are the sensitivity and specificity of the tests for diagnosis of breast cancer in women presenting with: a) An abnormal mammogram, overall and by BIRADS classification or other relevant clinical classification (e.g., presence or absence of calcification, well circumscribed lesions, etc.)b) A palpable breast abnormalityc) What percentage of women in the studies in this question were age 65 or older, and do sensitivity and specificity vary by older vs. younger than age 65? | | Key Question 1 - What is the accuracy (expressed as sensitivity, specificity, predictive values, and likelihood ratios) of noninvasive tests for diagnosis of breast cancer in women referred for further evaluation after identification of a possible breast abnormality on routine screening (mammography and/or clinical or self-detection of a palpable lesion)? The noninvasive tests to be evaluated are: Ultrasound (conventional B-mode, color Doppler, power Doppler, tissue harmonics, and tomography)Magnetic resonance imaging (MRI) with breast-specific coils and gadolinium-based contrast agents, with or without computer-aided diagnosis (CADx) Positron emission tomography (PET) with 18-fluorodeoxyglucose (FDG) as the tracer, with or without concurrent computed tomography (CT) scans Scintimammography (SMM) with technetium-99m sestamibi (MIBI) as the tracer, including Breast Specific Gamma Imaging (BSGI) | | |  | |  | |
| To place the tests’ accuracy information into perspective, an average woman in the U.S. who has an abnormal mammogram requiring a biopsy for evaluation has approximately a 20-percent risk of cancer. For women at this average level of risk of cancer after an abnormal mammogram, based upon the tests' negative likelihood ratios: \* For every 1,000 women who had a negative PET scan, about 924 women would have avoided an unnecessary biopsy, but 76 women would have missed cancers. \* For every 1,000 women who had a negative scintimammogram, about 907 women would have avoided an unnecessary biopsy, but 93 women would have missed cancers. (These numbers are for nonpalpable lesions only; numbers could not be calculated for all lesions.) \* For every 1,000 women who had a negative MRI, about 962 women would have avoided an unnecessary biopsy, but 38 women would have missed cancers. \* For every 1,000 women who had a negative US, about 950 women For every 1,000 women who had a negative US, about 950 women would have avoided an unnecessary biopsy, but 50 women would have missed cancers. | This CER is an update of a CER finalized in 2006.7 The updated results are, in general, very similar to the findings of the 2006 report. For MRI, in 2006 we found that the sensitivity was 92.5 percent and the specificity was 75.5 percent; the updated evidence base supported estimates of 91.7 percent sensitivity and 77.5 percent specificity. In both reports, MRI was found to be less sensitive (approximately 85%) for evaluation of microcalcifications than for evaluation of lesions in general. For PET, in 2006 we found that the sensitivity was 82.2 percent and the specificity was 78.3 percent; the updated evidence base supported estimates of 83.0 percent sensitivity and 74.0 percent specificity. In the updated report we attempted to evaluate the accuracy of PET/CT, but only one study that met the inclusion criteria was identified. For scintimammography, the updated evidence base identified a sensitivity of 84.7 percent, much higher than the sensitivity estimate from 2006 of 68.7 percent. Specificity was estimated at 84.8 percent in 2006, and at 77.0 percent in the update; however, the confidence intervals around the updated estimate of specificity are wide. It is possible that improvements in the technology in the last few years improved the sensitivity of the technique. For ultrasound, in 2006 we evaluated a relatively small set of studies of B-mode grayscale ultrasound, and estimated a sensitivity of 86.1 percent and a specificity of 66.4 percent. The update included a significantly expanded evidence base on B-mode grayscale ultrasound, and identified a sensitivity of 92.4 percent and specificity of 75.8 percent. In the update we included numerous other types of ultrasound, including power and color Doppler ultrasound, that were not studied in the 2006 report.  The probability that a woman actually has cancer (invasive or in situ) even after a finding of “benign” on MRI depends on her probability of having cancer before undergoing the test. Bayes’ theorem and the summary likelihood ratios indicate that if a woman with an estimated 5 to 10 percent chance of having cancer undergoes MRI and has a finding of “benign” she will then have an estimated 1 percent chance of having cancer; a woman with an estimated 20 percent chance of having cancer who has a finding of “benign” on MRI will then have an estimated 3 percent chance of having cancer; and a woman with an estimated 50 percent chance of having cancer who has a finding of “benign” on MRI will then have an estimated 10 percent chance of having cancer.  The probability that a woman actually does have cancer (invasive or in situ) even after a finding of “benign” on PET depends on her probability of having cancer before undergoing the test. Bayes’ theorem and the summary likelihood ratios indicate that if a woman with an estimated 5 percent chance of having cancer undergoes PET and has a finding of “benign” she will then have an estimated 1 percent chance of having cancer; a woman with an estimated 20 percent chance of having cancer who has a finding of “benign” on PET will then have an estimated 6 percent chance of having cancer; and a woman with an estimated 50 percent chance of having cancer who has a finding of “benign” on PET will then have an estimated 19 percent chance of having cancer.  The probability that a woman actually does have cancer (invasive or in situ) even after a finding of “benign” on scintimammography depends on her probability of having cancer before undergoing the test. Bayes’ theorem and the summary likelihood ratios indicate that if a woman with an estimated 5 percent chance of having cancer undergoes scintimammography and has a finding of “benign” she will then have an estimated 1 percent chance of having cancer; a woman with an estimated 20 percent chance of having cancer who has a finding of “benign” on scintimammography will then have an estimated 5 percent chance of having cancer; and a woman with an estimated 50 percent chance of having cancer who has a finding of “benign” on scintimammography will then have an estimated 17 percent chance of having cancer.  The probability that a woman actually does have cancer (invasive or in situ) even after a finding of “benign” on ultrasound depends on her probability of having cancer before undergoing the test. Bayes’ theorem and the summary likelihood ratios indicate that if a woman with an estimated 5 to 10 percent chance of having cancer undergoes B-mode grayscale ultrasound and has a finding of “benign” she will then have an estimated 1 percent chance of having cancer; a woman with an estimated 20 percent chance of having cancer who has a finding of “benign” on B-mode grayscale ultrasound will then have an estimated 2 percent chance of having cancer; and a woman with an estimated 50 percent chance of having cancer who has a finding of “benign” on B-mode | | | | It is difficult to estimate whether this conclusion is still valid or not, since it consists of calculations made based on operating characteristics of the test with “average” level of cancer risk. Using data from the Peters, 2008 meta-analysis, the “missed cancers” number would be 20, not 38. Therefore, this conclusion is possibly out of date, although probably modestly so. | | | Good |
| Although all of the technologies evaluated could reduce the need for biopsy in women with an abnormal mammogram who do not have cancer, each would miss some cancers. | | There was a great deal of heterogeneity (I2 = 93%) in the reported data. We were unable to identify with meta-regression any study- related characteristics that explained this heterogeneity, such as consecutive enrollment of patients, blinding of the diagnostic test reader to patient history/other clinical information, and use of the gold standard (biopsy) as the reference standard. | | | Conclusion is still valid and this portion of the CER does not need updating. | | Good | |
| Key Question 2 - For women with relevant demographic risk factors (e.g., age, family history) and clinical risk factors (e.g., BIRADS status or morphologic characteristics of the lesion), what are the positive and negative predictive values of the above diagnostic tests? | | Key Question 2 - Are there demographic (e.g., age) and clinical risk factors (e.g., morphologic characteristics of the lesion) that affect the accuracy of the tests considered in Key Question 1? | | |  | |  | |
| In general, the higher a woman’s risk of cancer is before undergoing a noninvasive test, the higher is the risk that she has cancer even if the test is negative.  If a less than 2-percent risk of having breast cancer with a negative diagnostic test is considered an acceptable level of risk for a diagnostic test to reliably preclude biopsy, none of these tests was sufficiently accurate to replace biopsy for women at average risk of breast cancer. | | Two studies evaluated only patients with palpable breast masses,57,62 one study evaluated only patients with non-palpable breast masses,63 and one study evaluated only patients with microcalcifications detected on x-ray mammography.61 With so few studies reporting on each category, evidence-based conclusions are difficult to support. None of the studies reported outcomes by patient demographics or any other clinical risk factors that may have affected the accuracy of SMM.  For all of the technologies evaluated in this assessment, only women with a low suspicion of malignancy after standard-of-care workup might be expected to experience a change in management decisions as a result of additional noninvasive imaging. A woman with a ≤12 percent suspicion of malignancy who has benign findings on MRI could have her suspicion of malignancy drop below the 2 percent threshold, and therefore she might be assigned to short-interval imaging followup management rather than tissue sampling management; a woman with a 1 percent suspicion of malignancy who has benign findings on MRI could have her suspicion of malignancy drop to near 0 percent and therefore she might be assigned to return to normal screening rather than short-interval followup imaging.  Therefore, if the 2 percent threshold is chosen, the use of noninvasive imaging in addition to standard workup may be clinically useful for diagnostic purposes only for women with a low suspicion of malignancy. When choosing which noninvasive imaging technology to use for this purpose, diagnostic B-mode grayscale ultrasound and MRI appear to be more accurate than PET, scintimammography, or the other types of ultrasound (e.g., Doppler) that were evaluated in this comparative effectiveness review. | | | Conclusion is probably out of date and this portion of the CER may need updating based on new data on MRI and US. | | Good. The updated report does not use the same "average" categorization, but concludes that MRI and B-mode grayscale ultrasound are clinically useful in women with a low suspicion of malignancy. | |
| Key Question 3 - Are there other factors that affect the accuracy or acceptability of the tests considered in Questions 1 and 2? | | Key Question 3 - Are there other factors and considerations (e.g., safety, care setting, patient preferences, ease of access to care) that may affect the accuracy or acceptability of the tests considered in Key Questions 1 and 2? | | |  | |  | |
| Based on results for only nonpalpable lesions (usually detected by mammography), data were insufficient to estimate the accuracy of PET scanning, MRI, or US. Scintimammography was not sufficiently accurate to avoid biopsy in women at average risk as judged by the acceptability standard of less than a 2-percent risk of breast cancer with a negative diagnostic test. Based on results for only palpable lesions, data were insufficient to estimate the accuracy of PET scanning, MRI, ultrasound, and scintimammography. | | **MRI.** One study reported the accuracy of MRI images interpreted with and without a Computer Aided Diagnosis (CAD) software system.12 The study reported virtually no difference in either sensitivity (77.4% vs. 78.9%) or specificity (73.2% vs. 73.2%) with or without CAD assistance. **Positron Emission Tomography.** None of the seven studies on stand-alone PET scanning or the one study on PET with CT reported information that addressed this question. **Scintimammography.** None were identified.  **Ultrasound.** None were identified. | | | Conclusion is possibly out of date and this portion of the CER may need updating based on the new meta-analysis. It would need to be reviewed to assess whether data can be stratified. | | Fair. New data was scanned and did not materially effect the conclusions. | |

**References**

3. Bruening W, Launders J, Pinkney N, et al. Effectiveness of Noninvasive Diagnostic Tests for Breast Abnormalities. Comparative Effectiveness Review No. 2. (Prepared by ECRI Evidence-based Practice Center under Contract No. 290-02-0019.) Rockville, MD: Agency for Healthcare Research and Quality. February 2006. [www.effectivehealthcare.ahrq.gov/reports/final.cfm](http://www.effectivehealthcare.ahrq.gov/reports/final.cfm).

4. Bruening W, Uhl S, Fontanarosa J, et al. Noninvasive Diagnostic Tests for Breast Abnormalities: Update of a 2006 Review. Comparative Effectiveness Review No. 47. (Prepared by the ECRI Institute Evidence-based Practice Center under Contract No. 290-02-0019.) AHRQ Publication No. 12-EHC014-EF. Rockville, MD: Agency for Healthcare Research and Quality; February 2012. [www.effectivehealthcare.ahrq.gov/reports/final.cfm](http://www.effectivehealthcare.ahrq.gov/reports/final.cfm).