

Draft Comparative Effectiveness Review

Number XX

**Core Needle and Open Surgical Biopsy for Diagnosis
of Breast Lesions**

An Update to the 2009 Report

Prepared for:

Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. [To be provided post peer review]

Prepared by:

[To be provided post peer review]

Investigators:

[To be provided post peer review]

AHRQ Publication No.

[To be provided post peer review]

This report is based on research conducted by an Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. xxx-xxxx-xxxxx). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decisionmakers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

This document is in the public domain and may be used and reprinted without special permission. Citation of the source is appreciated.

Suggested citation: [To be provided post peer review]

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of systematic reviews to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. These reviews provide comprehensive, science-based information on common, costly medical conditions, and new health care technologies and strategies.

Systematic reviews are the building blocks underlying evidence-based practice; they focus attention on the strength and limits of evidence from research studies about the effectiveness and safety of a clinical intervention. In the context of developing recommendations for practice, systematic reviews can help clarify whether assertions about the value of the intervention are based on strong evidence from clinical studies. For more information about AHRQ EPC systematic reviews, see www.effectivehealthcare.ahrq.gov/reference/purpose.cfm

AHRQ expects that these systematic reviews will be helpful to health plans, providers, purchasers, government programs, and the health care system as a whole. Transparency and stakeholder input are essential to the Effective Health Care Program. Please visit the web site (www.effectivehealthcare.ahrq.gov) to see draft research questions and reports or to join an e-mail list to learn about new program products and opportunities for input.

We welcome comments on this systematic review. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by email to epc@ahrq.hhs.gov.

Richard G. Kronick, Ph.D.
Director
Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.
Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.
Director
Evidence-based Practice Program
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Elisabeth Kato, M.D., M.S.
Task Order Officer
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Acknowledgments

[To be provided post peer review]

Key Informants

[To be provided post peer review]

Technical Expert Panel

[To be provided post peer review]

Peer Reviewers

[To be provided post peer review]

Core Needle and Open Surgical Biopsy for Diagnosis of Breast Lesions

An Update to the 2009 Report

Abstract

Objective: Core-needle biopsy and open surgical biopsy are the most frequently used procedures for diagnosis of suspicious breast lesions. An AHRQ evidence report on the comparative effectiveness and adverse events of breast biopsy methods was completed in 2009. The availability of additional studies and the uncertainties surrounding newer biopsy techniques prompted an update of that report.

Study Eligibility Criteria: We searched eight electronic databases (last search on May 7, 2013) for English-language full-text reports of prospective or retrospective cohort studies of women not previously diagnosed with breast cancer who were undergoing biopsy for diagnosis of a breast lesion.

Study Appraisal and Synthesis Methods: A single investigator extracted data from each study; quantitative results and intervention descriptions were verified by a second reviewer. We assessed the strength and applicability of the evidence following the processes described in the AHRQ Methods Guide. We performed Bayesian meta-analyses to estimate summary test performance and performed indirect comparisons to assess the relative effectiveness of alternative core-needle biopsy methods. Statistical models accounted for between-study heterogeneity.

Results: 151 studies of moderate to high risk of bias provided information on the test performance of alternative core-needle biopsy techniques. Open biopsy continues to be considered the “gold” standard diagnostic procedure, and we found no new studies investigating its test performance. For women at average risk of cancer, both ultrasound- and stereotactically guided biopsies had average sensitivities higher than 0.97 and average specificities ranging from 0.92 to 0.99; freehand biopsy methods had average sensitivity of 0.91 and specificity of 0.98. Differences among core-needle biopsy methods other than freehand did not exceed ± 0.1 , regardless of extraction technique (automated or vacuum). However, evidence on the test performance of MRI-guided biopsy (4 studies) was insufficient to draw conclusions. Comparisons of test performance between women at average and high baseline risk of cancer did not indicate an association but were imprecise. 135 studies contributed information on potential harms of different core-needle biopsy techniques. Overall, core-needle biopsy had a lower risk of complications than open surgical biopsy; however information on the latter was sparse. The absolute incidence of adverse events was low and the incidence of severe complications was less than 1 percent for all techniques. Vacuum-assisted procedures appeared to be associated with increased bleeding and hematoma formation; biopsies performed with patients seated upright appeared to be associated with increased risk of vasovagal reactions. Harms were reported inconsistently, raising concerns about selective outcome reporting. We found 10 reports of patients developing tumors at the site of prior core-needle biopsies. We found information on only a few patient-relevant and resource-related outcomes. Based on 41 studies, core-needle biopsy obviated the need for surgical procedures in about 75 percent of women. Meta-analysis of 9 studies reporting the number of surgical procedures required after biopsy suggested that the odds of

requiring only one procedure were more than 13 times higher among women receiving core-needle biopsy, as compared to those receiving open surgical biopsy, although this observation may be confounded by indication.

Limitations: Information about study- or population-level characteristics did not allow the identification of modifiers of test performance, adverse events, or clinical outcomes. Studies reported adverse events incompletely, and did not provide details of their outcome ascertainment methods.

Conclusions: A large body of evidence suggests that core-needle biopsy procedures have sensitivity and specificity at or near that of open biopsy procedures, and are associated with fewer adverse events. Image-guided core needle biopsy approaches appear to have similar test performance and safety profiles for women at average risk of breast cancer, although freehand procedures have lower sensitivity, and vacuum-assisted procedures appear to have a higher risk of bleeding. The strength of our conclusions about comparative test performance is generally low, because of concerns about the risk of bias of included studies, incomplete reporting, and the reliance on indirect comparisons. There were insufficient data to draw conclusions for MRI-guided biopsy or women at high baseline risk of cancer. Harms were reported inconsistently, raising concerns about selective outcome and analysis reporting. Women diagnosed with breast cancer by core-needle biopsy were more likely to have their cancer treated with a single surgical procedure, compared with women diagnosed by open surgical biopsy.

PROSPERO registration number: CRD42013004381.

Contents

- Abstract..... v
- Executive Summary..... 1
 - Background..... 1
 - Methods 2
 - External Stakeholder Input 2
 - Study Eligibility Criteria 2
 - Literature Search and Study Selection 2
 - Data Abstraction and Management 3
 - Assessment of Risk of Bias 3
 - Data Synthesis 3
 - Grading the Strength of Evidence 4
- Results 5
 - Test Performance of Open Surgical Biopsy 5
 - Test Performance of Core-Needle Biopsy Methods 5
 - Comparative Test Performance 7
 - Factors that Affect Test Performance 7
 - Contextualizing the Results of Test Performance Meta-analyses 10
- Adverse Events of Open Biopsy..... 12
- Adverse Events of Core-needle Biopsy..... 12
- Factors that Affect the Development of Adverse Events 12
- Discussion..... 13
 - Key Findings and Assessment of the Strength of Evidence 13
 - Test performance and comparative test performance 13
 - Underestimation rates 15
 - Adverse Events and Additional Surgeries After Biopsy 15
 - Applicability of Review Findings 17
 - Limitations of the Evidence Base 17
 - Limitations of This Review 17
 - Future Research Needs 18
- Conclusions 18
- References 19
- Background..... 1
 - Breast Cancer Epidemiology and Clinical Diagnosis 1
 - Original Evidence Report and Rationale for the Update 2
 - Key Questions 2
- Methods 3
 - AHRQ TOO..... 3

| | |
|--|----|
| External Stakeholder Input | 4 |
| Key Questions | 4 |
| Analytic Framework | 4 |
| Scope of the Review | 4 |
| Populations and Conditions of Interest..... | 4 |
| Interventions | 5 |
| Comparators (reference standard and comparator index tests) | 5 |
| Outcomes | 5 |
| Timing | 6 |
| Setting..... | 6 |
| Study Design and Additional Criteria | 6 |
| Literature search and Abstract Screening..... | 8 |
| Study Selection and Eligibility Criteria..... | 9 |
| Results | 13 |
| Included Studies | 15 |
| Test Performance for Breast Cancer Diagnosis..... | 15 |
| Test Performance of Open Surgical Biopsy | 15 |
| Test Performance of Core-Needle Biopsy Methods..... | 15 |
| Contextualizing the Results of Test Performance Meta-analyses | 21 |
| Comparative Test Performance | 23 |
| Factors that Affect Test Performance..... | 25 |
| Within-study Evidence | 25 |
| Meta-regression analyses..... | 29 |
| Risk Of Bias Assessment for Studies Addressing Key Question 1 | 29 |
| Adverse Events of Open Biopsy..... | 31 |
| Adverse Events of Core-needle Biopsy..... | 31 |
| Hematomas and Bleeding | 32 |
| Infections | 33 |
| Pain and Use of Pain Medications..... | 34 |
| Vasovagal Reactions | 34 |
| Impact of Biopsy Procedure on Usual Activities and Time to Recovery..... | 35 |
| Impact of Biopsy Procedure on Subsequent Mammographic Procedures | 35 |
| Miscellaneous Reported Adverse Events | 36 |
| Dissemination of Cancerous Cells During the Biopsy Procedure..... | 36 |
| Factors that modify the association of biopsy procedures with adverse events | 37 |
| Anxiety and Distress..... | 39 |
| Procedure Preference..... | 40 |
| Surgical Procedures Avoided | 40 |
| Cosmetic Results | 40 |
| Resource utilization and costs | 41 |
| Physician Experience..... | 41 |
| Availability of a Qualified Pathologist..... | 42 |

| | |
|--|----|
| Availability of Equipment/Utilization | 42 |
| Procedure Duration Time | 42 |
| Time to Complete Tumor Removal..... | 43 |
| Wait Time for Test Results..... | 43 |
| Recurrence Rates | 43 |
| Discussion..... | 44 |
| Key Findings and Assessment of the Strength of Evidence..... | 44 |
| Test performance and comparative test performance..... | 44 |
| Underestimation rates | 46 |
| Adverse Events and Additional Surgeries After Biopsy | 46 |
| Limitations of the Evidence Base..... | 47 |
| Strengths and Limitations of This Review | 48 |
| Applicability of Review Findings | 49 |
| Evidence Gaps and Ongoing Research..... | 49 |
| Future Research Needs | 50 |
| Conclusions | 50 |
| List of abbreviations | 52 |
| References | 53 |

List of Tables

| | |
|--|----|
| Table A: Summary estimates of test performance for alternative core-needle biopsy methods – women at average risk of cancer | 6 |
| Table B: Summary estimates of test performance for alternative core-needle biopsy methods – women at high risk of cancer..... | 7 |
| Table C: Summary estimates of underestimation rates for alternative core-needle biopsy methods – women at average risk of cancer | 7 |
| Table D: Differences in sensitivity between pairs of biopsy methods (meta-regression based indirect comparisons) | 9 |
| Table E: Differences in specificity between pairs of biopsy methods (meta-regression based indirect comparisons) | 9 |
| Table F: Strength of evidence about comparative test performance in women at average risk of breast cancer | 14 |
| Table G: Strength of evidence for underestimation rates in women at average risk of cancer | 15 |
| Table H: Strength of evidence assessment for adverse events of biopsy | 16 |
| Table 1: Definitions of diagnostic groups based on index and reference standard test results | 7 |
| Table 2: Summary of new evidence evaluated in this update* | 14 |
| Table 3: Summary estimates of test performance for alternative core-needle biopsy methods – women at average risk of cancer | 17 |
| Table 4: Summary estimates of test performance for alternative core-needle biopsy methods – women at high risk of cancer..... | 18 |
| Table 5: Differences in sensitivity between pairs of biopsy methods (meta-regression based indirect comparisons) | 24 |

| | |
|---|----|
| Table 6: Differences in specificity between pairs of biopsy methods (meta-regression based indirect comparisons) | 24 |
| Table 7: Studies evaluating factors that may affect test performance (20 studies, reporting on multiple factors each) | 26 |
| Table 8: Meta-regression analysis for test performance outcomes. | 29 |
| Table 9: Adverse events associated with core-needle biopsy for breast cancer diagnosis | 32 |
| Table 10: Core-needle biopsy procedures and rates of hematoma formation and bleeding..... | 33 |
| Table 11: Core-needle biopsy procedures and rates of infectious complications | 34 |
| Table 12: Core-needle biopsy procedures and rates of vasovagal reactions | 34 |
| Table 13: Studies evaluating factors that may affect the incidence of adverse events..... | 37 |
| Table 14: Strength of evidence about comparative test performance in women at average risk of breast cancer | 44 |
| Table 15: Strength of evidence for underestimation rates in women at average risk of cancer | 46 |
| Table 16: Strength of evidence assessment for adverse events of biopsy | 47 |
| Table 17: Evidence gaps for biopsy methods for the diagnosis of breast cancer | 49 |

List of Figures

| | |
|---|----|
| Figure A: Outcomes of testing in a hypothetical cohort of 1000 women | 11 |
| Figure 1. Analytic framework | 4 |
| Figure 2: Flow chart of included studies | 13 |
| Figure 3: Scatterplot of results in the receiver operating characteristic space and summary receiver operating characteristic curves of alternative core-needle biopsy methods for the diagnosis of breast cancer | 18 |
| Figure 4: Outcomes of testing in a hypothetical cohort of 1000 women | 22 |
| Figure 5: Positive predictive value of alternative test methods | 23 |

Appendix A. Search strategy

Appendix B. List of excluded studies

Appendix C. Assessment of the strength of evidence

Executive Summary

Background

Approximately one in eight U.S. women will develop breast cancer during her lifetime, and as of 2009 an estimated 2.7 million women had a current or past diagnosis of breast cancer.¹ Because the earliest stages of breast cancer are often asymptomatic, the process of breast cancer diagnosis is often initiated by detection of an abnormality through self-examination, physical examination by a clinician, or screening mammography. If the initial assessment suggests that the abnormality could be breast cancer, the woman is likely to be referred for a biopsy – a sampling of cells or tissue from the suspicious lesion. Among women screened annually for 10 years, approximately 50 percent will need additional imaging, and a large proportion will have biopsies.^{2,3}

There are currently three techniques for obtaining samples from suspicious breast lesions: fine-needle aspiration, biopsy with a hollow core needle, or open surgical retrieval of tissue. Fine-needle aspiration is generally considered less sensitive than core-needle and open biopsy methods,⁴ and is used less frequently. Lesion samples obtained by any of these methods are evaluated by pathologists and classified into histological categories with the primary goal of determining whether the lesion is benign or malignant. Because core-needle biopsy samples only part of the breast abnormality, there is the risk that a lesion will be classified as benign, high-risk, or non-invasive when invasive cancer is in fact present in unsampled areas. Open surgical biopsy samples most or all of the lesion, and is therefore considered to have a smaller risk of misdiagnosis. However, open procedures may carry a higher risk of complications, such as bleeding or infection, compared to core-needle biopsy.⁵

Alternative core-needle biopsy methods differ with respect to the use of imaging (e.g., stereotactic mammography; ultrasound; or magnetic resonance imaging, MRI), the use of vacuum to assist in tissue acquisition, the use of needles of varying diameter, and the numbers of samples taken. These and other factors may affect test performance and the rate of complications. For example, some biopsy procedures may retrieve larger amounts of tissue, improving test performance, but the retrieval of larger amounts of tissue may also result in more complications, such as bleeding. The impact of various aspects of biopsy technique and patient or lesion characteristics on test performance and safety is not clear.

In 2009, the ECRI Evidence-based Practice Center (EPC) conducted a comparative effectiveness review for core-needle versus open surgical biopsy on behalf of the Agency for Health Care Research and Quality (AHRQ).^{6,7} The original evidence report assessed the diagnostic test performance and adverse events of core-needle biopsy techniques compared to open surgical biopsy and evaluated differences between open biopsy and core-needle biopsy with regards to patient preferences, costs, availability, and other factors. The authors concluded that core-needle biopsies were almost as accurate as open surgical biopsies, had a lower risk of severe complications, and were associated with fewer subsequent surgical procedures.⁷

The publication of additional studies and changes in practice raised the concern that the conclusions of the original report may be out of date, particularly for the ductal carcinoma in situ (DCIS) underestimation rate of stereotactically guided vacuum-assisted core-needle biopsy, the performance of magnetic resonance imaging (MRI)-guided core-needle biopsy, and the performance of freehand automated device core-needle technology. New studies may also provide additional information allowing the exploration of heterogeneity for test performance

and safety outcomes. Therefore, an updated review of the published literature was considered necessary to synthesize all evidence on currently available methods for core-needle and open surgical breast lesion biopsy.

Methods

We performed a systematic review of the published scientific literature using established methodologies as outlined in the Agency for Healthcare Research and Quality's (AHRQ) *Methods Guide for Comparative Effectiveness Reviews*,⁸ hereafter referred to as the Methods Guide. We followed the reporting requirements of the "Preferred Reporting Items for Systematic Reviews and Meta-analyses" (PRISMA) statement.⁹ A full description of all review steps is included in the full report and the study protocol (PROSPERO registration number CRD42013005690).

External Stakeholder Input

We convened a 9-member Technical Expert Panel (TEP), including representatives of professional societies, experts in the diagnosis and treatment of breast cancer (including radiologists and surgeons), and a patient representative. The TEP provided input to help further refine the Key Questions and protocol, identify important issues, and define the parameters for the review of evidence.

Study Eligibility Criteria

We included only English-language full-text articles. Studies included for the assessment of diagnostic test performance (Key Question 1) met the following inclusion criteria: (1) enrolled women not previously diagnosed with breast cancer who received core-needle or open biopsy for initial diagnosis of possible breast cancer; (2) compared diagnoses on core-needle biopsy to a reference standard of open surgery or follow-up by clinical examination or imaging of at least 6 months; (3) reported or allowed the calculation of sensitivity, specificity, positive or negative predictive value; (4) were prospective or retrospective cohort studies (including randomized controlled trials); and (5) enrolled 10 or more patients and followed at least 50 percent of them to the completion of the study. In contrast to the original report, we did not restrict eligibility to studies including only women at average risk for breast cancer, because MRI-guided biopsy, which was identified as a topic of interest for this update, is used mainly in women at a higher-than-average risk for breast cancer. Studies included for the assessment of possible adverse events of core-needle biopsy (Key Question 2) or the assessment of patient-relevant outcomes, resource use and logistics, and availability of technology and relevant expertise (Key Question 3) were not required to compare diagnoses on core-needle biopsy to a reference standard of open surgery or clinical follow-up, or to contain extractable information on diagnostic test performance. Furthermore, for Key Question 2 we included any primary research articles, regardless of design, that addressed the dissemination of cancer cells by the biopsy procedure (i.e., seeding).

Literature Search and Study Selection

We searched MEDLINE®, EMBASE®, the Cochrane Central Register of Controlled Trials, the Cochrane Database of Systematic Reviews, the Database of Abstracts of Reviews of Effects, the Health Technology Assessment Database, the U.K. National Health Service

Economic Evaluation Database, the U.S. National Guideline Clearinghouse, and CINAHL.¹⁰ **Appendix A** describes the search strategy we employed which is based on an expansion of the search strategy used in the original report. We did not use a search filter for studies of diagnostic tests in order to increase search sensitivity.¹¹ We also searched for systematic reviews on the topic and used their lists of included studies to validate our search strategy and to make sure we identified all relevant studies.

To identify studies excluded from the original evidence report because they enrolled women at high risk for cancer, we rescreened both the set of abstracts screened for the original report and the full text of studies excluded from the original report because they included women at high risk for cancer. Titles and abstracts were manually screened in duplicate. A single reviewer screened each potentially eligible article in full-text to determine eligibility and a second reviewer examined all articles deemed relevant. Disagreements regarding article eligibility were resolved by consensus involving a third reviewer.

Data Abstraction and Management

Data were extracted using electronic forms and entered into the Systematic Review Data Repository (SRDR; <http://srdhr.ahrq.gov/>). We pilot-tested the forms on several studies extracted by multiple team members to ensure consistency in operational definitions. A single reviewer extracted data from each eligible study. A second reviewer verified extracted data and discrepancies were resolved by consensus including a third reviewer. We contacted authors (1) to clarify information reported in their papers and to verify suspected overlap between study populations in publications from the same group of investigators.

Assessment of Risk of Bias

We assessed the risk of bias for each individual study using the assessment methods detailed in the Methods Guide. We used elements from the Quality Assessment for Diagnostic Accuracy Studies instrument (QUADAS version 2), to assess risk of bias for studies of diagnostic test accuracy.¹²⁻¹⁵ We used items from the Newcastle-Ottawa scale,¹⁶ the Cochrane Risk of Bias tool,¹⁷ and the checklist proposed by Drummond et al.,^{18, 19} to assess nonrandomized cohort studies, randomized controlled trials, and studies of resource utilization and costs, respectively.^{18, 19}

Data Synthesis

We summarized included studies qualitatively and presented important features of the study populations, designs, tests used, outcomes, and results in summary tables. Statistical analyses were conducted using methods currently recommend for use in Comparative Effectiveness Reviews of diagnostic tests.^{20, 21}

For Key Question 1 we performed meta-analyses because studies were deemed sufficiently similar with respect to included populations, and the core-needle biopsy and reference standard tests they employed. We use a mixed effects binomial-bivariate normal regression model that accounted for different imaging methods (e.g. ultrasound, stereotactic mammography, MRI), the use of vacuum (yes vs. not), the baseline of risk of cancer of included patients (high versus average risk), and residual (unexplained) heterogeneity. This model allowed us to estimate the test performance of alternative diagnostic tests, and to perform indirect comparisons among them.²² Furthermore, it allowed us to model the correlation between

sensitivity and specificity and to derive meta-analytic receiver operating characteristic (ROC) curves.^{23, 24} A univariate mixed effects logistic regression (binomial-normal) model was used for the meta-analysis of DCIS and high risk lesion underestimation rates.²⁵ We used meta-regression analyses to evaluate the impact of risk of bias items and other study-level characteristics.^{26, 27}

For Key Question 2, we found that adverse events were inconsistently reported across studies and that the methods for ascertaining their occurrence were often not presented in adequate detail. For this reason we refrained from performing meta-analyses for these outcomes. Instead, we calculated descriptive statistics (medians, 25th and 75th percentiles, minimum and maximum values) across all studies and for specific test types. For Key Question 3, because of the heterogeneity of research designs and outcomes assessed, we were only able to perform a meta-analysis comparing core-needle and open surgical biopsies with respect to the number of patients who required one versus more than one surgical procedures for treatment, after the establishment of breast cancer diagnosis. This analysis used a univariate normal random effects model with a binomial within-study distribution.

All statistical analyses were performed using Bayesian methods; models were fit using Markov Chain Monte Carlo methods and non-informative prior distributions. Empirical evidence suggests that, when the number of studies is large, this approach produces results similar to those of maximum likelihood methods (which do not require the specification of priors).²⁸ Results were summarized as medians of posterior distributions with associated 95 percent central credibility intervals (CrIs).

Grading the Strength of Evidence

We followed the Methods Guide⁸ to evaluate the strength of the body of evidence for each Key Question with respect to the following domains: risk of bias, consistency, directness, precision, and reporting bias.^{8, 29} Generally, strength of evidence was downgraded when risk of bias was not low, in the presence of inconsistency, when evidence was indirect or imprecise, or when we suspected that results were affected by selective analysis or reporting.

We determined risk of bias (low, medium, or high) on the basis of the study design and the methodological quality. We assessed consistency on the basis of the direction and magnitude of results across studies. We considered the evidence to be indirect when we had to rely on comparisons of biopsy methods across different studies (i.e., indirect comparisons). We considered studies to be precise if the credible interval (CrI) was narrow enough for a clinically useful conclusion, and imprecise if the CrI was wide enough to include clinically distinct conclusions. The potential for reporting bias (“suspected” vs. “not suspected”) was evaluated with respect to publication, selective outcome reporting, and selective analysis reporting. We made qualitative dispositions rather than perform formal statistical tests to evaluate differences in the effect sizes between more precise (larger) and less precise (smaller) studies because such tests cannot distinguish between “true” heterogeneity between smaller and larger studies, other biases, and chance.^{30, 31} Therefore, instead of relying on statistical tests, we evaluated the reported results across studies qualitatively, on the basis of completeness of reporting, number of enrolled patients, and numbers of observed events. Judgment on the potential for selective outcome reporting bias was based on reporting patterns for each outcome of interest across studies. We acknowledge that both types of reporting bias are difficult to reliably detect on the basis of data available in published research studies. We believe that our searches (across multiple databases), combined with our plan for contacting test manufacturers (for additional data) and the authors of published studies (for data clarification) limited the impact of reporting

and publication bias on our results, to the extent possible.

Finally, we rated the body of evidence using four strength of evidence levels: high, moderate, low, and insufficient.⁸ These describe our level of confidence that the evidence reflects the true effect for the major comparisons of interest.

We qualitatively evaluated similarities and differences in study populations, diagnostic methods, and outcomes among study designs. We used these comparisons to inform our judgments on applicability of study findings to clinical practice.

Results

Key Question 1: In women with a palpable or non-palpable breast abnormality, what is the diagnostic test performance of different types of core-needle breast biopsy compared with open biopsy or with each other?

One hundred and fifty one studies, published between 1990 and 2013, provided information on test performance outcomes (44 new studies and 107 studies included in the original evidence report). Forty-seven studies were prospectively designed, and 58 were conducted in the U.S. Ten studies provided outcome information on more than one group of patients (typically undergoing biopsy with a different biopsy device). In statistical analyses, these groups were treated separately, leading to a total of 161 independent patient groups with information on 68,942 breast lesions.

Test Performance of Open Surgical Biopsy

Published information on the test performance of open surgical biopsy was limited. Neither the original report nor our updated searches identified any clinical studies of open surgical biopsy that met the inclusion criteria. However, research studies of needle biopsy methods and technical experts generally suggested that open surgical biopsy could be considered a “gold” standard test (i.e., a test without measurement error). One study reported that open surgical biopsy may miss one to two percent of breast cancers (i.e. sensitivity of 98% or greater). No studies provided information on underestimation rates for open surgical biopsy.

Test Performance of Core-Needle Biopsy Methods

A total of 151 studies contributed information to analyses of test performance of core-needle biopsy methods; 146 enrolled women at average risk and only five enrolled women at high risk of cancer. Studies varied by type of imaging guidance (stereotactic guidance, ultrasound guidance, MRI guidance, other guidance, or freehand), how the biopsy sample was extracted (automated or vacuum), and other factors (e.g., needle size). If studies included multiple cohorts of patients undergoing biopsy by different methods (e.g., some patients were biopsied with vacuum-assistance and others were not) but the study did not report the test performance of each method, these groups were treated together as ‘multiple methods’ in statistical analyses for that factor. 114 studies reported the use of a single form of imaging guidance (74 stereotactic; 35 ultrasound; 4 MRI), whereas seven used freehand methods and 29 used multiple methods in their study population. Fifty-one studies used vacuum-assisted methods to obtain the biopsy sample; 71 used automated methods; 28 used multiple methods; and one did not report adequate details. Needle size also varied across studies: 56 used 14G needles, nine used smaller needles, 42 used larger bores, and 44 studies did not report relevant information.

Reference standard tests also differed across studies: 26 used open biopsy on all included patients; 87 used mean or median followup of between six and 24 months for test negative patients, and 38 used mean or median followup of 24 months or more for test negative cases. Additional study details are available in the SRDR. Consistent with the findings of the original report, the risk of bias for this body of evidence was considered moderate to high, mainly due to concerns about spectrum bias, retrospective data collection, differential verification, and lack of information regarding the blinding of reference standard test assessors to the index test results.

Table A summarizes meta-analysis results for alternative diagnostic biopsy methods, together with information on the number of lesions evaluated with each method for women at average risk of cancer. Sensitivity estimates were higher than 0.90 and specificity estimates were higher than 0.92 for all methods. CrIs, particularly for ultrasound- and stereotactically-guided biopsy methods, were fairly precise, reflecting the large number of studies reporting information on the test performance of these methods. In contrast, results for MRI-guided methods were based on only two studies and were imprecise, particularly for sensitivity. **Table B** summarizes the same information for women deemed to be at high risk for cancer (e.g. due to genetic factors or strong family history). Information for this subgroup was limited (5 studies) and we did not find evidence to suggest that the test performance of breast biopsy methods was different between women at average and high risk of cancer. However, there was substantial uncertainty around the relative test performance estimates of the two groups. **Table C** summarizes the results of analyses of underestimation rates for women at average risk of breast cancer. Results were rather imprecise (CrI widths were wider than 10%) for all estimates except the underestimation rate for stereotactically guided, vacuum-assisted biopsy methods. Analyses of underestimation rates were not possible for women at high risk of cancer because of lack of data.

Table A: Summary estimates of test performance for alternative core-needle biopsy methods – women at average risk of cancer

| Biopsy method or device | N studies [N biopsies] for sensitivity & specificity | Sensitivity | Specificity |
|--|--|-------------------|-------------------|
| Freehand, automated | 10 [786] | 0.91 (0.80, 0.96) | 0.98 (0.95, 1.00) |
| US-guided, automated | 27 [16287] | 0.99 (0.98, 0.99) | 0.97 (0.95, 0.98) |
| US-guided, vacuum-assisted | 10 [1456] | 0.97 (0.89, 0.99) | 0.99 (0.97, 0.99) |
| Stereotactically guided, automated | 36 [9342] | 0.97 (0.95, 0.98) | 0.97 (0.96, 0.98) |
| Stereotactically guided, vacuum- assisted | 40 [14421] | 0.99 (0.98, 0.99) | 0.92 (0.89, 0.94) |
| MRI-guided, automated | 2 [89] | 0.90 (0.58, 0.99) | 0.98 (0.90, 1.00) |
| Multiple techniques | 28 [25391] | 0.98 (0.97, 0.99) | 0.95 (0.93, 0.97) |

All numbers are medians with 95% CrIs, unless otherwise stated. Summary results are shown when at least two studies were available. Results are not shown for three studies that used devices not belonging to any of the categories listed in the table (1 grid guidance, 2 unclear).

CrI = credible interval; DCIS = ductal carcinoma in situ; MRI = magnetic resonance imaging; N = number; NA = not applicable; US = ultrasound.

Table B: Summary estimates of test performance for alternative core-needle biopsy methods – women at high risk of cancer

| Biopsy method or device | N studies (N biopsies) for sensitivity and specificity | Sensitivity (95% CrI) | Specificity (95% CrI) |
|--|---|--------------------------|--------------------------|
| Stereotactically guided, automated | 1 [416] | 0.98 (0.93, 0.99) | 0.97 (0.83, 1.00) |
| Stereotactically guided, vacuum-assisted | 2 [311] | 0.94 (0.83, 0.98) | 0.99 (0.93, 1.00) |
| MRI-guided, automated | 2 [56] | 0.99 (0.93, 1.00) | 0.89 (0.58, 0.98) |

No studies provided information on the test performance of freehand or US-guided biopsy methods, or the use of multiple methods in populations of women at high risk of cancer. Results are based on bivariate model with risk group as a covariate.

CrI = credible interval; DCIS = ductal carcinoma in situ; MRI = magnetic resonance imaging; N = number; US = ultrasound.

Table C: Summary estimates of underestimation rates for alternative core-needle biopsy methods – women at average risk of cancer

| Biopsy method or device | N studies [N biopsies] for DCIS underestimation | DCIS underestimation probability | N studies [N biopsies] for high risk lesion underestimation | High risk lesion underestimation probability |
|--|---|----------------------------------|---|--|
| Freehand, automated | 0 [0] | NA | 1 [6] | NA |
| US-guided, automated | 14 [307] | 0.38 (0.25, 0.51) | 20 [502] | 0.22 (0.14, 0.34) |
| US-guided, vacuum-assisted | 4 [21] | 0.11 (0.01, 0.41) | 7 [16] | 0.09 (0.01, 0.30) |
| Stereotactically guided, automated | 17 [649] | 0.27 (0.18, 0.37) | 28 [353] | 0.47 (0.37, 0.58) |
| Stereotactically guided, vacuum-assisted | 33 [1803] | 0.11 (0.08, 0.14) | 37 [949] | 0.18 (0.13, 0.24) |
| MRI-guided, automated | 0 [0] | NA | 1 [1] | NA |
| Other or multiple techniques | 16 [573] | 0.22 (0.15, 0.31) | 22 [822] | 0.33 (0.24, 0.43) |

Analyses for underestimation were not possible for high risk women due to sparse data. CrI = credible interval; DCIS = ductal carcinoma in situ; MRI = magnetic resonance imaging; N = number; US = ultrasound.

Comparative Test Performance

To compare test performance across different biopsy methods we used indirect (meta-regression-based) comparisons. **Tables D** and **E** present comparisons between all possible pairs of tests for sensitivity and specificity, respectively. In general, differences among tests were relatively small: for example, differences in sensitivity or specificity never exceeded 0.1 (i.e., 10% absolute difference) and 95 percent CrIs often included the null value (i.e. 0, indicating no difference). One exception to this general pattern was the comparative performance of freehand biopsy against other techniques. With respect to sensitivity, freehand biopsy had worse performance compared to the other methods and the CrIs of between-test differences excluded 0 (i.e., no difference) when compared against ultrasound-guided automated device biopsy, stereotactically guided biopsy (both vacuum-assisted and automated), as well as studies that used multiple biopsy methods in their study population (without stratifying results by biopsy method). Another fairly consistent finding was that stereotactically guided, vacuum-assisted biopsy had lower specificity than all other methods except MRI (and CrIs did not include 0).

Factors that Affect Test Performance

We considered evidence on the impact of patient or study level-factors on test performance from two complementary sources: (1) within-study evidence (i.e. comparisons of test performance over levels of a factor within the patient population enrolled in a study) and (2) evidence from meta-regression analyses (that combine information across studies). Ideally, all

studies would consistently report comparisons of test performance across well-defined subgroups (e.g., by patient, or lesion characteristics). Such within-study comparisons are more informative than comparisons across studies: factors related to study setting are common for all patients within the same study and other patient differences can be addressed (at least to some extent) by appropriate analytic methods (e.g., regression adjustment). In the absence of such information, one has to rely on indirect (across-study) comparisons that are generally less convincing because they cannot account for all differences across included populations.

Twenty studies provided information that allowed an evaluation of the impact of any factor on test performance. Specifically, 16 studies provided information on patient and lesion-related factors, 10 on procedural factors, and three on clinician and facility factors (some studies provided information on multiple factors). Of note, the majority of studies (131 of 151) did not allow investigation of the impact of any factors on test performance, raising concerns about selective analysis or reporting of results on modifiers of test performance. Among the 20 studies reporting relevant results, factors were coded inconsistently and details that would allow formal statistical testing were not available. Because of these reasons, within-study comparisons could not support conclusions regarding possible modifiers of test performance.

Meta-regression analyses were possible for the following factors: needle size, choice of reference standard, country where the study was performed, whether multiple centers contributed patients to a study, study design, and risk of bias. In general, test performance was not affected by the factors examined (i.e., CrIs included the null value), with the exception of higher sensitivity in studies conducted in the U.S. (vs. any other country) and higher specificity in studies using followup of 6 or more and 24 or more months (as compared to studies using surgical pathology results for all patients) and higher sensitivity in studies with a prospective design (as compared to studies with a retrospective design). These results must be interpreted with caution given that they reflect indirect comparisons across studies, which cannot be adjusted for other factors that vary across studies.

Overall, within-study analyses and meta-regression analyses were insufficient to confirm (or exclude) any single factor as a modifier of test performance.

Table D: Differences in sensitivity between pairs of biopsy methods (meta-regression based indirect comparisons)

| | | | | | | | |
|--|--------------------------|------------------------------|------------------------------------|------------------------------------|--|-----------------------|----------------------|
| | Freehand, automated | | | | | | |
| US-guided, automated | 0.08 (0.02, 0.19) | Ultrasound-guided, automated | | | | | |
| US-guided, vacuum-assisted | 0.06 (-0.03, 0.17) | -0.02 (-0.09, 0.01) | Ultrasound-guided, vacuum-assisted | | | | |
| Stereotactically guided, automated | 0.07 (0.01, 0.18) | -0.01 (-0.03, 0.00) | 0.01 (-0.02, 0.08) | Stereotactically guided, automated | | | |
| Stereotactically guided, vacuum-assisted | 0.08 (0.03, 0.19) | 0.00 (-0.01, 0.02) | 0.02 (-0.00, 0.10) | 0.02 (0.00, 0.03) | Stereotactically guided, vacuum-assisted | | |
| MRI-guided, automated | 0.00 (-0.33, 0.14) | -0.08 (-0.41, 0.00) | -0.06 (-0.38, 0.05) | -0.07 (-0.39, 0.02) | -0.09 (-0.41, 0.00) | MRI-guided, automated | |
| Multiple techniques* | 0.08 (0.02, 0.19) | 0.00 (-0.01, 0.01) | 0.02 (-0.01, 0.09) | 0.01 (-0.00, 0.03) | 0.00 (-0.02, 0.01) | 0.08 (0.00, 0.41) | Multiple techniques* |

* Populations not stratified by biopsy method.

All results are shown as medians of differences (95% CrI). Positive values denote that the method on the left-most column has higher sensitivity than the comparator (on the diagonal). CrIs that do not include the null value (0) are highlighted in bold. CrI = credible interval; MRI = magnetic resonance imaging.

Table E: Differences in specificity between pairs of biopsy methods (meta-regression based indirect comparisons)

| | | | | | | | |
|--|-----------------------------|------------------------------|------------------------------------|------------------------------------|--|-----------------------|----------------------|
| | Freehand, automated | | | | | | |
| US-guided, automated | -0.01 (-0.04, 0.02) | Ultrasound-guided, automated | | | | | |
| US-guided, vacuum-assisted | 0.00 (-0.02, 0.03) | 0.02 (-0.01, 0.04) | Ultrasound-guided, vacuum-assisted | | | | |
| Stereotactically guided, automated | -0.01 (-0.03, 0.02) | 0.00 (-0.02, 0.02) | -0.02 (-0.03, 0.01) | Stereotactically guided, automated | | | |
| Stereotactically guided, vacuum-assisted | -0.07 (-0.10, -0.03) | -0.05 (-0.09, -0.02) | -0.07 (-0.10, -0.04) | -0.05 (-0.09, -0.03) | Stereotactically guided, vacuum-assisted | | |
| MRI-guided, automated | 0.00 (-0.08, 0.03) | 0.01 (-0.07, 0.04) | 0.00 (-0.09, 0.02) | 0.01 (-0.07, 0.04) | 0.06 (-0.02, 0.10) | MRI-guided, automated | |
| Multiple techniques* | -0.03 (-0.06, 0.01) | -0.02 (-0.04, 0.01) | -0.03 (-0.06, -0.01) | -0.02 (-0.04, 0.01) | 0.04 (0.01, 0.07) | -0.03 (-0.06, 0.06) | Multiple techniques* |

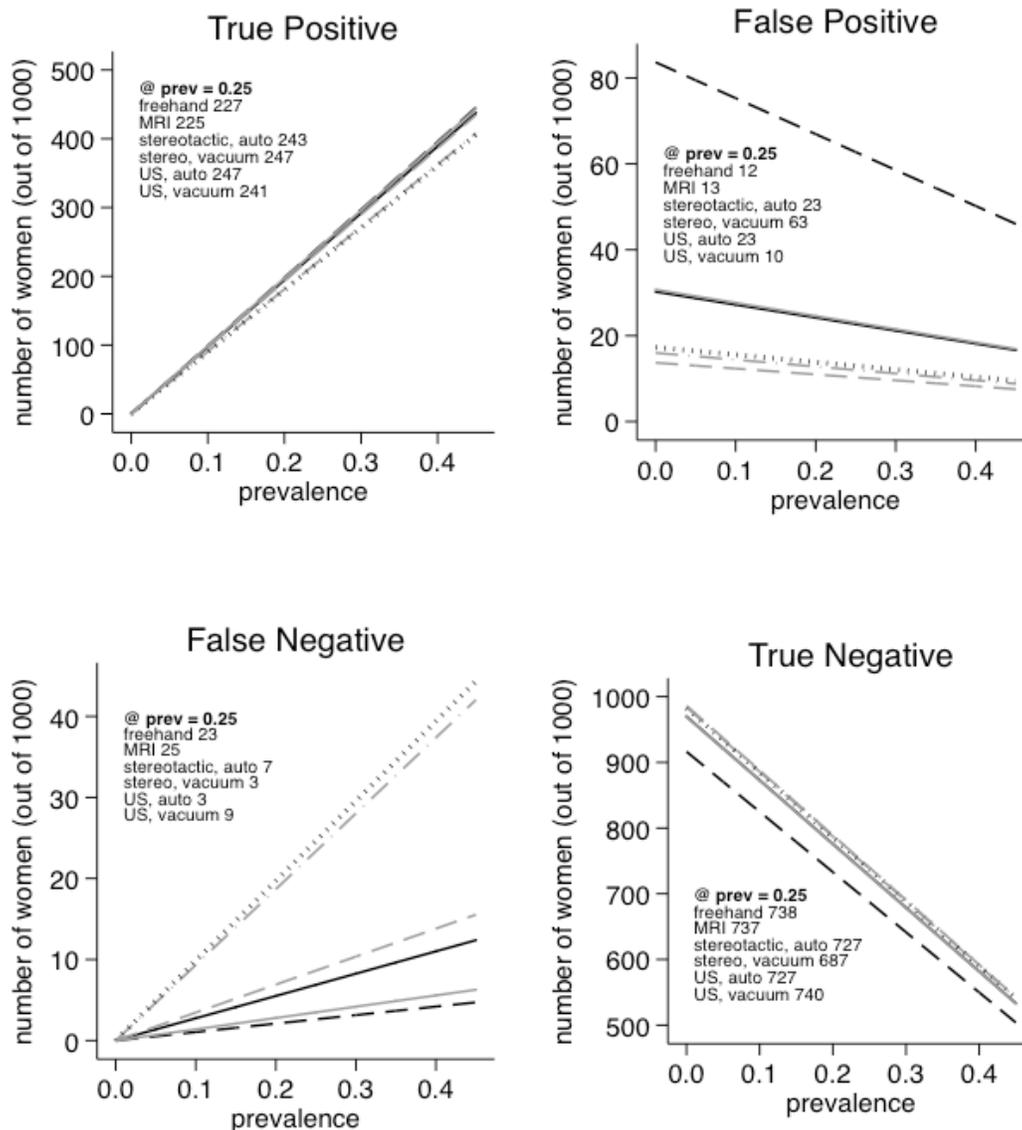
* Populations not stratified by biopsy method.

Positive values denote that the method on the left-most column has higher specificity than the comparator (on the diagonal). CrIs that do not include the null value (0) are highlighted in bold. CrI = credible interval; MRI = magnetic resonance imaging.

Contextualizing the Results of Test Performance Meta-analyses

To contextualize the results of the test performance meta-analyses presented in the preceding sections we evaluated the impact of testing in a hypothetical cohort of 1000 women, under alternative scenarios for disease prevalence. The results are presented in **Figure A**. In populations with low cancer prevalence, the number of cases where treatment may be delayed on the basis of biopsy results (i.e., false negative biopsies) is expected to be small (e.g., for all ultrasound or stereotactically guided biopsy methods less than five out of 1000 women, if prevalence is 10 percent or less). As prevalence increases the number of false negative results increases for all biopsy methods, but more rapidly for MRI-guided and freehand methods, which had the lowest sensitivity. The number of false positive cases declines with increasing prevalence. Automated device (both stereotactically and ultrasound-guided) biopsy methods have comparable results (approximately 20-30 false positive results in the range examined). MRI- or ultrasound-guided, vacuum assisted, and freehand methods appear to do best (less than 20 false positives in the range of prevalence examined). Stereotactically guided, vacuum-assisted methods appear to produce the most false positive results (more than 40 per 1000 women over the range of prevalence examined). **Figure A** also presents numerical results for a prevalence of 0.25, which is approximately the prevalence of breast cancer among women referred for breast biopsy in the U.S.

Figure A: Outcomes of testing in a hypothetical cohort of 1000 women



Different lines represent different test modalities: black solid = stereotactically guided, automated; grey solid = US guided, automated; black dashed = stereotactically guided, vacuum-assisted; grey dashed = US guided, vacuum assisted; black dotted = MRI guided, automated; grey dash-dot = freehand, automated.

Key Question 2. In women with a palpable or nonpalpable breast abnormality, what are the adverse events (harms) associated with different types of core-needle breast biopsy compared with open biopsy for diagnosis?

We synthesized information on adverse events from a total of 135 studies (63 new studies and 72 from the original evidence report) reporting on at least one of the outcomes relevant to Key Question 2: 2 for open biopsy, 113 for core-needle biopsy, and 20 on the dissemination of cancerous cells during core-needle biopsy. Overall, studies were considered to be of moderate to

high risk of bias. Selective outcome reporting was considered likely for all adverse events examined, because of the large proportion of studies with unclear or missing data.

Adverse Events of Open Biopsy

Very few studies reported information about complications occurring in association with open surgical biopsy procedures. One study reported results from a series of 425 wire-localized open biopsy procedures and reported that 10.2 percent were complicated by vasovagal reactions. Another study reported that 6.3 percent of open surgical biopsies were complicated by infections. A third study reported that 2.1 percent of open biopsy procedures were complicated by the development of an abscess, but zero abscesses complicated 234 ultrasound-guided vacuum-assisted core-needle procedures. One study reported that 4 of 100 surgical biopsies required repeat biopsy, compared to 2 of 100 vacuum-assisted core-needle biopsies.

Adverse Events of Core-needle Biopsy

We identified 133 studies reporting information on at least one of the adverse events of interest following core-needle biopsy (20 reported information related to the dissemination of cancerous cells during biopsy). Overall, core-needle biopsy appeared to have a lower risk of complications than open surgical biopsy; however, direct comparative information was sparse. The incidence of severe complications with core-needle biopsy was less than one percent. The incidence of all adverse events was low: in more than 50 percent of studies reporting information on hematomas, bleeding, vasovagal reactions, and infections, the percentage of patients experiencing each of the aforementioned outcomes was less than 2 percent; in 75 percent of studies the event rate was less than 1 percent for infections, less than 5 percent for bleeding and vasovagal reactions, and less than 8 percent for hematoma formation. Use of vacuum assistance was associated with a greater rate of bleeding and hematoma formation.

Of 14 studies that used histopathology to demonstrate dissemination of cells by core-needle biopsy procedures (nine cohort and five case series or case reports), the percentage of needle tracks reported to contain displaced cancerous cells ranged from 0 to 69 percent. The clinical significance of these findings is unclear; tumor development on the biopsy needle track is extremely rare.

Factors that Affect the Development of Adverse Events

Four studies provided information on patient and lesion-related factors, seven studies provided information on procedural factors, and one study provided information on clinician and facility factors. The vast majority of studies reporting on adverse events from core-needle biopsy did not allow investigation of the impact of factors on adverse events and no individual factor was evaluated by more than five of the total included studies, raising concerns regarding selective outcome and analysis reporting. No studies reported information on factors that affect the development of adverse events from open biopsy. We did not perform meta-regression analyses because studies reported information on adverse events inconsistently and because data were missing from more than half of the studies for all adverse events. Studies suggested that vacuum-assisted biopsy methods led to increased bleeding and performing biopsies with patients seated upright was associated with increased incidence of vasovagal reactions; however, results were reported in a way that precluded quantitation of the relative risk.

Key Question 3. How do open biopsy and various core-needle techniques differ in terms of patient preference, availability, costs, availability of qualified pathologist interpretations, and other factors that may influence choice of a particular technique?

We reviewed a total of 127 studies for Key Question 3 (41 new studies and 86 studies from the original report). Generally, the evidence supported the conclusions of the original report that core-needle biopsy costs less than open surgical biopsy, consumes fewer resources, and is preferred by patients. In addition, utilization of core needle biopsy has grown consistently since the mid-1990s. Studies reported that women were generally satisfied with the cosmetic results of core-needle procedures, but tended to feel intense anxiety just before and during the procedure, which may be partially ameliorated with the use of medication, relaxation and empathy techniques, or hypnosis. Core-needle biopsy obviated the need for surgical procedures in about 75 percent of women. Nine studies reported comparisons against open surgical biopsy with respect to the number of patients requiring only one surgical procedure (vs. more than one) after cancer diagnosis. Meta-analysis of these studies suggested that the odds of requiring only one surgical procedure were more than 13 times higher among women receiving core-needle biopsy; odds ratio = 13.4 (95% CrI, 5.6 to 43.4). This result should be interpreted with caution because of the possibility of confounding by indication.

Discussion

Key Findings and Assessment of the Strength of Evidence

In this update of the 2009 Comparative Effectiveness Review on breast biopsy methods we synthesized evidence from a total of 319 studies (111 new studies and 208 from the original report). We found few studies providing information on the test performance of open surgical biopsy. In contrast, the evidence base on core-needle biopsy methods now includes a large number of studies reporting on almost 70,000 breast lesions. Tables E-G summarize our assessment of the strength of evidence. Following the original evidence report, and in view of the paucity of evidence on open surgical biopsy, we refrained from rating the strength of evidence for this technique for all Key Questions. For Key Questions 1 and 2, we assessed the strength of evidence by integrating our (subjective) judgments on the risk of bias of included studies, the consistency of their findings, the directness of the available data, and the precision of quantitative results. For Key Question 3 we only rated the strength of evidence for the outcome of additional surgical procedures required after biopsy. We did not rate the strength of evidence for other Key Question 3 outcomes because of the diversity of designs employed and outcomes addressed. Please see the Methods section for a detailed discussion of our approach to rating the strength of evidence.

Test performance and comparative test performance

Among women at average risk of cancer, core-needle biopsy using ultrasound or stereotactic guidance had average sensitivities ranging from 0.97 to 0.99 and average specificities ranging from 0.92 to 0.99. Freehand biopsy methods appeared to have lower average sensitivity (0.91) compared to other methods, but similar specificity. Stereotactically guided vacuum-assisted techniques were associated with lower specificity compared to other biopsy methods. Although these results were consistent across studies and (in many cases) fairly

precise, they were derived from indirect comparisons across studies of moderate to high risk of bias. MRI-guided biopsies were evaluated in only four studies with small sample sizes, leading to substantial uncertainty around estimates of test performance. **Table F** summarizes our assessment of the strength of evidence for comparisons among alternative biopsy methods in women at average risk of cancer. Of note, we rated the strength of evidence on *comparative* test performance, whereas the original report considered *absolute* test performance; for this reason, and for this subset of outcomes, our ratings are not directly comparable with those of the original report.

There were few studies of women at high risk of cancer; however, statistical comparisons of test performance between women at low and high risk of breast cancer did not identify a difference. Because the number of available studies was small, comparisons of test performance between low and high risk women had substantial uncertainty and results were not sufficient to support definitive conclusions. Evidence on modifiers of test performance was also sparse, for all biopsy methods, raising concerns about selective outcome and analysis reporting.

Table F: Strength of evidence about comparative test performance in women at average risk of breast cancer

| Outcome | Comparison or biopsy method | Overall Rating | Key Findings and Comments |
|---|--|----------------|---|
| Comparison of test performance among alternative biopsy methods | Freehand vs. ultrasound-guided, automated | Low | – Difference in sensitivity: 0.08 (0.02 to 0.19) [ultrasound-guided, automated better] – Difference in specificity: -0.01 (-0.04, 0.02) [no difference] |
| | Freehand vs. stereotactically guided, automated | Low | – Difference in sensitivity: 0.07 (0.01 to 0.18) [stereotactically guided, automated better] – Difference in specificity: -0.01 (-0.03 to 0.02) [no difference] |
| | Freehand vs. ultrasound-guided, vacuum-assisted | Low | – Difference in sensitivity: 0.06 (-0.03 to 0.17) [ultrasound-guided, vacuum-assisted better] – Difference in specificity: 0.00 (-0.02 to -0.03) [no difference] |
| | Freehand vs. stereotactically guided, vacuum-assisted | Low | – Difference in sensitivity: 0.08 (0.03 to 0.19) [stereotactically guided, vacuum-assisted better] – Difference in specificity: -0.07 (-0.10 to -0.03) [freehand better] |
| | Stereotactically guided, vacuum assisted vs. automated | Low | – Difference in sensitivity: 0.02 (0.00 to 0.03) [vacuum-assisted better] – Difference in specificity: -0.05 (-0.09 to -0.03) [automated better] |
| | Stereotactically guided vacuum assisted vs. ultrasound-guided, automated | Low | – Difference in sensitivity: 0.00 (-0.01 to 0.02) [no difference] – Difference in specificity: -0.05 (-0.09 to -0.02) [ultrasound-guided, automated better] |
| | Stereotactically guided vacuum assisted vs. ultrasound-guided, vacuum-assisted | Low | – Difference in sensitivity: 0.02 (0.00 to 0.10) [no difference] – Difference in specificity: -0.07 (-0.10 to -0.04) [ultrasound-guided vacuum-assisted better] |
| | Other comparisons between biopsy techniques | Low | – There were no differences in sensitivity and specificity for ultrasound-guided automated vs. vacuum-assisted methods. – There were no differences in sensitivity and specificity for stereotactically guided automated vs. ultrasound-guided methods. – The Crls for all comparisons included zero and were fairly precise. |
| | MRI vs. any other device | Insufficient | – Only 4 small studies were available – Differences in sensitivity and specificity comparing MRI with other biopsy methods had Crls intervals that included 0 but were imprecise |

| Outcome | Comparison or biopsy method | Overall Rating | Key Findings and Comments |
|--|-----------------------------|----------------|---|
| Modifiers of test performance for women at average and high risk of breast cancer; | All biopsy methods | Insufficient | <ul style="list-style-type: none"> – Few studies provided within sample information for each modifier of interest; meta-regression results rely on cross-study comparisons so consistency of effects cannot be assessed – Within-study (direct) evidence was sparse; between study evidence relied on indirect comparisons across studies – In meta-regression analyses Crls were wide and extreme odds ratio values were often observed because sensitivity and specificity for all tests were very close to 1 (see Results for additional details) |

Crls = credible interval; MRI = magnetic resonance imaging.

Underestimation rates

Underestimation rates varied among alternative biopsy methods and were often imprecisely estimated because of the relatively small number of lesions contributing data for these analyses. In general, underestimation was less common with stereotactically guided vacuum-assisted biopsy methods, as compared to stereotactically or ultrasound-guided automated methods. Our assessment of the strength of evidence for this outcome is summarized in **Table G**.

Table G: Strength of evidence for underestimation rates in women at average risk of cancer

| Outcome | Comparison or biopsy method | Overall Rating | Key Findings and Comments |
|---------------------------------------|--|----------------|---|
| DCIS underestimation | Stereotactically guided, automated | Low | – Average underestimation probability: 0.27 (0.18 to 0.37) [17 studies] |
| | Stereotactically guided, vacuum-assisted | Low | – Average underestimation probability: 0.11 (0.08 to 0.14) [33 studies] |
| | Ultrasound-guided, automated | Low | – Average underestimation probability: 0.38 (0.25 to 0.51) [14 studies] |
| | Other biopsy methods | Insufficient | Few or no available studies with very small numbers of lesions. |
| High risk lesion underestimation rate | Stereotactically guided, automated | Low | – Average underestimation probability: 0.47 (0.37 to 0.58) [28 studies] |
| | Stereotactically guided, vacuum-assisted | Low | – Average underestimation probability: 0.18 (0.13 to 0.24) [37 studies] |
| | Ultrasound-guided, automated | Low | – Average underestimation probability: 0.22 (0.14 to 0.34) [20 studies] |
| | Other biopsy methods | Insufficient | Few or no available studies with very small numbers of lesions. |

DCIS = ductal carcinoma in situ.

Adverse Events and Additional Surgeries After Biopsy

In general, adverse events were reported inconsistently, raising concerns about selective outcome and analysis reporting. Few studies provided information on the harms of open surgical biopsy. Core-needle biopsy was only infrequently associated with serious adverse events. Comparisons between open and core-needle biopsy are based on indirect comparisons and expert opinion, with limited empirical evidence. Open biopsy appeared to be associated with an increased incidence of adverse events (including serious adverse events) compared to core-needle biopsy. Our assessment of the strength of evidence for adverse events is summarized in **Table H**.

Among core-needle biopsy methods, vacuum-assisted methods appeared to be associated with increased bleeding. Sitting upright during the biopsy procedure was associated with more vasovagal reactions. Information about the dissemination of cancer cells during the biopsy procedure was provided by a small number of studies with various designs. Studies reported that women were generally satisfied with the cosmetic results of core-needle procedures.

Women diagnosed with breast cancer by core-needle biopsy were able to have their cancer treated with a single surgical procedure, more often than women diagnosed by open surgical biopsy. Although the magnitude of this association was large (the ratio of the odds was approximately 13), women and their physicians are likely to choose biopsy methods on the basis of factors (e.g., lesion location, or characteristics of the lesion on imaging) that may also be associated with the need for additional surgeries. Because such selection would lead to confounding by indication, we rated the strength of evidence for this association as moderate. .

Table H: Strength of evidence assessment for adverse events of biopsy

| Outcomes | Comparison | Overall Rating | Key findings |
|--|---|----------------|--|
| Bleeding, including bleeding events that require treatment | Comparisons among alternative core-needle biopsy methods | Low | <ul style="list-style-type: none"> – Median %: 0.76 (25th perc. = 0.22; 75th perc = 3.97) – Potential for selective outcome and analysis reporting – Few studies reported bleeding requiring treatment; the event rate was low (<0.40 perc.) in those studies |
| Hematoma formation | Comparisons among alternative core-needle biopsy methods | Low | <ul style="list-style-type: none"> – Median %: 1.19 (25th perc. = 0.05; 75th perc = 7.20) – Potential for selective outcome and analysis reporting |
| Infectious complications | Comparisons among alternative core-needle biopsy methods | Low | <ul style="list-style-type: none"> – Median %: 0.05 (25th perc. = 0.00; 75th perc = 0.63) – Potential for selective outcome and analysis reporting |
| Vasovagal reactions: | Comparisons among alternative core-needle biopsy methods | Low | <ul style="list-style-type: none"> – Median %: 1.78 (25th perc. = 0.43; 75th perc = 4.29) – Potential for selective outcome and analysis reporting |
| Pain and severe pain | Comparisons among alternative core-needle biopsy methods | Low | 25 studies of a wide variety of biopsy methods reported information about patient pain during the procedure (pain was assessed heterogeneously across studies). |
| Other adverse events | Comparisons among alternative core-needle biopsy methods | Insufficient | <ul style="list-style-type: none"> – Most events were reported by a single study precluding assessment of consistency – Individual studies did not provide adequate information for precise estimation of the event rate) – Only informal indirect comparisons among biopsy methods were possible – Potential for selective outcome and analysis reporting |
| Modifiers of adverse events – vasovagal reactions | Sitting upright during the biopsy procedure | Low | <ul style="list-style-type: none"> – Vasovagal reactions were more common among patients sitting during the biopsy procedure – Results were reported in few studies (11 studies; 8 from the original evidence report and 3 from this update) – Potential for selective outcome and analysis reporting |
| Modifiers of adverse events – bleeding | Vacuum-assisted versus non-vacuum assisted biopsy methods | Low | <ul style="list-style-type: none"> – Vacuum-assisted procedures were generally associated with increased rates of bleeding and hematoma formation – Bleeding events were generally uncommon – Comparisons among biopsy methods were based on informal indirect comparisons (across studies) – Potential for selective outcome and analysis reporting |
| All other modifiers of | Comparisons among alternative core-needle biopsy | Insufficient | – Most factors assessed by a single study limiting our ability to assess consistency |

Applicability of Review Findings

The existing evidence base on core-needle biopsy of breast lesions in women at average risk of cancer appears to be applicable to clinical practice in the U.S. Studies enrolled patients with an average age similar to that of women undergoing breast biopsy in the U.S., and for indications that represent the most prevalent indications in U.S. clinical practice (i.e. mammographic findings of suspicious lesions). Almost all of the studies were carried out in either the U.S. or in industrialized European or Asian countries where core-biopsy methods are likely sufficiently similar to those used in the U.S. The applicability of our findings to women at high risk of breast cancer may be uncertain because we found few studies explicitly reporting on groups of patients at high baseline risk of breast cancer and comparisons of test performance between subgroups of women produced imprecise results.

Limitations of the Evidence Base

We believe that the evidence regarding the performance of core-needle biopsy for diagnosis of breast lesions is limited in the following ways: (1) published evidence on the test performance and adverse events of open surgical biopsy was sparse; (2) available studies were at moderate to high risk of bias and information on patient selection criteria, patient or lesion characteristics, adverse events, or patient-relevant outcomes was often missing or inconsistently reported, and pathology results were not reported with adequate granularity; (3) studies typically used lesions (or biopsy procedures) as the unit of analysis, instead of patients, reporting results in a way that did not allow for the correlation to be accounted for in our statistical analyses; (4) studies provided limited information to assess the impact of various patient-, lesion-, procedure-, or system- related factors on the outcomes of breast biopsy; (5) we found very few studies on MRI-guided biopsy for women at average or high risk of cancer; (6) there is limited information on the comparative effectiveness of alternative biopsy methods on patient-relevant outcomes, resource use and logistics, and availability of technology and expertise for different core-needle biopsy techniques.

Limitations of This Review

Our work has several limitations, which – to a large extent – reflect the limitations of the underlying evidence base. Studies were deemed to be of moderate to high risk of bias because of characteristics related to their design and conduct, limiting our ability to draw strong conclusions. Information for several outcomes of interest was often missing. Studies did not provide adequate information about study- or population level characteristics that could be modifiers of test performance, adverse events, or clinical outcomes. Thus, our ability to explore between-study heterogeneity was limited. Further, because we relied on published information, we were unable to evaluate the impact of patient- or lesion-level factors on outcomes of interest. We did not include studies published in languages other than English; however, given the very large number of studies from diverse geographic locations included in the review, we believe that the addition of non-English language studies would not affect our conclusions.

Future Research Needs

There is now a large body of evidence to suggest that stereotactic and US guided core-needle techniques have comparable sensitivity to each other and to open biopsy. The next focus of research should be biopsy under MRI guidance, which is a new technique that is likely to come into wider use. The data is not yet adequate to define its advantages or disadvantages of MRI guided biopsy compared with alternative techniques. Studies should be powered to achieve adequate precision (i.e., produce confidence intervals or CrIs that are narrow enough to allow clinically meaningful conclusions), have a prospective design, enroll patients across multiple centers, and use standardized histological classification systems for pathological classification.³²

³³ For all biopsy methods, additional well-designed and fully reported prospective cohort studies are needed, primarily for addressing questions about the impact of patient-, lesion-, procedure-, or system-level factors on test performance, adverse events, and patient-relevant outcomes. This would help resolve uncertainties regarding effect modification (e.g., over patient and lesion factors) that cannot be resolved with the currently available data. Such studies could be conducted at relatively low cost, and large-scale databases of prospectively-collected observational data on breast biopsy procedures and outcomes could be used to evaluate the comparative effectiveness of alternative biopsy methods with respect to short and long term outcomes, and potential modifying factors. In all future studies, baseline risk of cancer development should be characterized using consistent and widely accepted criteria to allow appropriate subgroup analyses. We believe that a randomized comparison of alternative biopsy methods would not be fruitful because existing studies indicate that biopsy procedures have sensitivities and specificities that are fairly similar and also close to 1. Additional information is also needed to identify factors that may influence the rate of adverse events of specific biopsy methods. Future research needs to be reported in accordance with recent reporting guidelines (e.g., STAndards for the Reporting of Diagnostic accuracy studies; <http://www.stard-statement.org/>), for progress to be made on these questions.

Conclusions

A large body of evidence suggests that imaging-guided core-needle biopsy procedures have sensitivity and specificity close to that of open biopsy procedures, and are associated with fewer adverse events. Imaging-guided core needle biopsy approaches appear to have similar test performance and safety profiles for women at average risk of breast cancer, although freehand procedures have lower sensitivity, and vacuum-assisted procedures appear to have a higher risk of bleeding. The strength of conclusions about comparative test performance was generally low, because of concerns about the risk of bias of included studies, incomplete reporting, and the reliance on indirect comparisons. There were insufficient data to draw conclusions for MRI-guided biopsy or women at high baseline risk of cancer. Harms were reported inconsistently, raising concerns about selective outcome and analysis reporting. Women diagnosed with breast cancer by core-needle biopsy were more likely to have their cancer treated with a single surgical procedure, compared with women diagnosed by open surgical biopsy.

References

1. Bendifallah S, Chabbert-Buffet N, Maurin N, et al. Predictive factors for breast cancer in patients diagnosed with ductal intraepithelial neoplasia, grade 1B. *Anticancer research*. 2012 Aug;32(8):3571-9. PMID: 22843948.
2. Hubbard RA, Kerlikowske K, Flowers CI, et al. Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: a cohort study. *Ann Intern Med*. 2011 Oct 18;155(8):481-92. PMID: 22007042.
3. Elmore JG, Barton MB, Mocerri VM, et al. Ten-year risk of false positive screening mammograms and clinical breast examinations. *N Engl J Med*. 1998 Apr 16;338(16):1089-96. PMID: 9545356.
4. Yu YH, Wei W, Liu JL. Diagnostic value of fine-needle aspiration biopsy for breast mass: a systematic review and meta-analysis. *BMC Cancer*. 2012;12:41. PMID: 22277164.
5. Bernardi D, Borsato G, Pellegrini M, et al. On the diagnostic accuracy of stereotactic vacuum-assisted biopsy of nonpalpable breast abnormalities. Results in a consecutive series of 769 procedures performed at the Trento Department of Breast Diagnosis. *Tumori*. 2012 Jan-Feb;98(1):113-8. PMID: 22495711.
6. Bruening W, Schoelles K, Treadwell J, et al. Comparative Effectiveness of Core-Needle and Open Surgical Biopsy for the Diagnosis of Breast Lesions. Comparative Effectiveness Review No. 19 (Prepared by the ECRI Institute Evidence-based Practice Center under Contract No. 290-02-0019). AHRQ Publication No. 10-EHC007-EF. Rockville, MD: Agency for Healthcare Research and Quality; December 2009. Available at <http://www.ncbi.nlm.nih.gov/books/NBK45220/pdf/TOC/pdf>.
7. Bruening W, Fontanarosa J, Tipton K, et al. Systematic review: comparative effectiveness of core-needle and open surgical biopsy to diagnose breast lesions. *Ann Intern Med*. 2010 Feb 16;152(4):238-46. PMID: 20008742.
8. AHRQ. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. 2012; Available from: <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=318>.
9. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg*. 2010;8(5):336-41. PMID: 20171303.
10. Whiting P, Westwood M, Burke M, et al. Systematic reviews of test accuracy should search a range of databases to identify primary studies. *J Clin Epidemiol*. 2008 Apr;61(4):357-64. PMID: 18313560.
11. Leeflang MM, Scholten RJ, Rutjes AW, et al. Use of methodological search filters to identify diagnostic accuracy studies can lead to the omission of relevant studies. *J Clin Epidemiol*. 2006 Mar;59(3):234-40. PMID: 16488353.
12. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011 Oct 18;155(8):529-36. PMID: 22007046.
13. Whiting P, Rutjes AW, Reitsma JB, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*. 2003 Nov 10;3:25. PMID: 14606960.
14. Whiting P, Rutjes AW, Dinnes J, et al. Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technol Assess*. 2004 Jun;8(25):iii, 1-234. PMID: 15193208.

15. Whiting PF, Weswood ME, Rutjes AW, et al. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Med Res Methodol*. 2006 Mar 6;6:9. PMID: 16519814.
16. Wells G, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. [29 April 2013]; Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
17. Higgins JP, Altman DG, Gotzsche PC, et al; Cochrane Bias Methods Group/ Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011 Oct 18;343:d5928. PMID: 22008217.
18. Houssami N, Ciatto S, Ellis I, et al. Underestimation of malignancy of breast core-needle biopsy: concepts and precise overall and category-specific estimates. *Cancer*. 2007 Feb 1;109(3):487-95. PMID: 17186530.
19. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ*. 1996 Aug 3;313(7052):275-83. PMID: 8704542.
20. Trikalinos TA, Balion CM. Chapter 9: options for summarizing medical test performance in the absence of a "gold standard". *J Gen Intern Med*. 2012 Jun;27 Suppl 1:S67-75. PMID: 22648677.
21. Trikalinos TA, Balion CM, Coleman CI, et al. Chapter 8: meta-analysis of test performance when there is a "gold standard". *J Gen Intern Med*. 2012 Jun;27 Suppl 1:S56-66. PMID: 22648676.
22. Hamza TH, van Houwelingen HC, Stijnen T. The binomial distribution of meta-analysis was preferred to model within-study variability. *J Clin Epidemiol*. 2008 Jan;61(1):41-51. PMID: 18083461.
23. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med*. 2001 Oct 15;20(19):2865-84. PMID: 11568945.
24. Arends LR, Hamza TH, van Houwelingen JC, et al. Bivariate random effects meta-analysis of ROC curves. *Med Decis Making*. 2008 Sep-Oct;28(5):621-38. PMID: 18591542.
25. Stijnen T, Hamza TH, Ozdemir P. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Stat Med*. 2010 Dec 20;29(29):3046-67. PMID: 20827667.
26. Hamza TH, van Houwelingen HC, Heijnenbrok-Kal MH, et al. Associating explanatory variables with summary receiver operating characteristic curves in diagnostic meta-analysis. *J Clin Epidemiol*. 2009 Dec;62(12):1284-91. PMID: 19398297.
27. van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Stat Med*. 2002 Feb 28;21(4):589-624. PMID: 11836738.
28. Dahabreh IJ, Trikalinos TA, Lau J, et al. *An Empirical Assessment of Bivariate Methods for Meta-Analysis of Test Accuracy*. Rockville (MD)2012.
29. Singh S, Chang SM, Matchar DB, et al. Grading a body of evidence on diagnostic tests. In: Chang SM, Matchar DB, Smetana GW, Umscheid CA, eds. *Methods Guide for Medical Test Reviews*. AHRQ Publication No. 12-EHC017. Rockville, MD: Agency for Healthcare Research and Quality; June 2012:chapter 7. Available at <http://www.ncbi.nlm.nih.gov/books/NBK98241/>.

30. Lau J, Ioannidis JP, Terrin N, et al. The case of the misleading funnel plot. *BMJ*. 2006 Sep 16;333(7568):597-600. PMID: 16974018.
31. Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*. 2011;343:d4002. PMID: 21784880.
32. Compton CC. Reporting on cancer specimens: case summaries and background documentation: A publication of the College of American Pathologists Cancer Committee; 2003.
33. Ellis IO, Humphreys S, Michell M, et al. Best Practice No 179. Guidelines for breast needle core biopsy handling and reporting in breast screening assessment. *Journal of clinical pathology*. 2004 Sep;57(9):897-902. PMID: 15333647.

Background

Breast Cancer Epidemiology and Clinical Diagnosis

Among women in the United States (U.S), breast cancer is the second most common malignancy (after skin cancer), and the second most common cause of cancer death (after lung cancer).¹ Approximately one in eight women in the U.S. will develop breast cancer during their lifetime, and as of 2009 an estimated 2.7 million women had a current or past diagnosis of breast cancer.² The American Cancer Society estimates that 232,340 new cases of invasive breast cancer and 64,640 new cases of non-invasive breast cancer will be diagnosed in 2013, and 39,620 women will die of breast cancer.³

During the earliest stages of breast cancer, there are usually no symptoms. The process of breast cancer diagnosis is often initiated by detection of an abnormality through self-examination, physical examination by a clinician, or screening mammography. Data from the Behavioral Risk Factor Surveillance System show that, in 2010, 75.4 percent of U.S. women aged ≥ 40 years and 79.7 percent of women aged 50 to 74 years reported having a mammogram within the past 2 years.⁴ If initial assessment suggests that the abnormality may be breast cancer, the woman may be referred for a biopsy, which is a sampling of cells or tissue from the suspicious lesion. Among women screened annually for 10 years, approximately 50 percent will need additional imaging, a large proportion will have biopsies.^{5,6} Over a million women have breast biopsies each year in the U.S. There are currently three techniques for obtaining samples from suspicious breast lesions: fine-needle aspiration, biopsy with a hollow core needle, or open surgical retrieval of tissue. Fine-needle aspiration, which retrieves a sample of cells, is generally considered less sensitive than both core-needle and open biopsy methods and will not be discussed in this report.⁷ Core-needle biopsy, which retrieves a sample of tissue, and open surgical procedures are therefore the most frequently used procedures.

Samples obtained by any of these methods are evaluated by pathologists and classified into histological categories with the primary goal of determining whether the lesion is benign or malignant. Because core-needle biopsy often samples only part of the breast abnormality, there is the risk that a lesion will be classified as benign or as high-risk (e.g., atypical ductal hyperplasia, ADH) or non-invasive (e.g., ductal carcinoma in situ, DCIS) when invasive cancer is in fact present in unsampled areas. In contrast, open surgical biopsy often samples most or all of the lesion, and it is thought that there is a smaller risk of misdiagnosis. However, while open surgical biopsy methods are considered to be the most accurate, they also appear to carry a higher risk of complications, such as bleeding or infection, compared to core-needle biopsy.⁸ Therefore, if core-needle biopsy is also highly accurate, women and their clinicians may prefer some type of core-needle biopsy to open surgical biopsy.

Core-needle biopsy may be carried out using a range of techniques. If the breast lesion to be biopsied is not palpable, an imaging method (i.e., stereotactic mammography, ultrasound, or magnetic resonance imaging (MRI)) may be used to locate the lesion. The biopsy may be carried out with needles of varying diameter, and one or more samples of tissue may be taken. Sometimes a vacuum device is used to assist in removing the tissue sample through the needle. It is thought that these and other variations in how core-needle biopsy is carried out may affect the accuracy and rate of complications of the biopsy. However, the impact of aspects of biopsy technique on test performance and safety are not clear.

Original Evidence Report and Rationale for the Update

In 2009, the ECRI Evidence-based Practice Center (EPC) conducted a comparative effectiveness review for core-needle versus open surgical biopsy.^{9, 10} The original report provided a detailed description of the technical aspects of alternative biopsy methods and we have not repeated this information here. The original report assessed the diagnostic test performance and adverse events of multiple core-needle biopsy techniques and tools, compared to open surgical biopsy, and also evaluated differences between open biopsy and core-needle biopsy with regards to patient preference, costs, availability, and other factors. The key conclusions were that core-needle biopsies were almost as accurate as open surgical biopsies, had a lower risk of severe complications, and were associated with fewer subsequent surgical procedures.¹⁰ The need for update of the 2009 report was assessed in 2010 by the RAND EPC.¹¹ Several high-impact general medical and specialty journals were searched, a panel of experts in the field was consulted, and an overall assessment of the need to update the report was produced. The conclusion of the update Surveillance Report was that additional studies and changes in practice render some conclusions of the original report possibly out of date. Specifically, the Surveillance Report noted the following:

- New studies are available regarding
 - the DCIS underestimation rate of stereotactic vacuum-assisted core-needle biopsy
 - test performance of MRI-guided core-needle biopsy
 - test performance of freehand automated device core-needle technology
- New studies on the test performance of core-needle biopsy may allow the exploration of heterogeneity for test performance or harm outcomes

On the basis of the Surveillance Report findings, an updated review of the published literature was considered necessary to synthesize all evidence on currently available methods for core-needle and open surgical breast biopsy.

Key Questions

To determine the Key Questions and study selection criteria (population, intervention, comparator, outcome, timing and setting; PICOTS) for this update, we began by considering the criteria used in the original Evidence Report. On the basis of input from clinical experts during the development of our protocol, we made minor revisions to the Key Questions and study eligibility criteria to clarify the focus of the updated review. We specified the following three Key Questions to guide the conduct of the update:

Key Question 1: In women with a palpable or nonpalpable breast abnormality, what is the test performance of different types of core-needle breast biopsy compared with open biopsy for diagnosis?

- a) What factors associated with the patient and her breast abnormality influence the test performance of different types of core-needle breast biopsy compared with open biopsy for diagnosis of a breast abnormality?
- b) What factors associated with the procedure itself influence the test performance of different types of core-needle breast biopsy compared with open biopsy for diagnosis of a breast abnormality?

- c) What clinician and facility factors influence the test performance of core-needle breast biopsy compared with open biopsy for diagnosis of a breast abnormality?

Key Question 2: In women with a palpable or nonpalpable breast abnormality, what are the adverse events (harms) associated with different types of core-needle breast biopsy compared with open biopsy for diagnosis?

- a) What factors associated with the patient and her breast abnormality influence the adverse events of core-needle breast biopsy compared with the open biopsy technique in the diagnosis of a breast abnormality?
- b) What factors associated with the procedure itself influence the adverse events of core-needle breast biopsy compared with the open biopsy technique in the diagnosis of a breast abnormality?
- c) What clinician and facility factors influence the adverse events of core-needle breast biopsy compared with the open biopsy technique in the diagnosis of a breast abnormality?

Key Question 3: How do open biopsy and various core-needle techniques differ in terms of patient preference, availability, costs, availability of qualified pathologist interpretations, and other factors that may influence choice of a particular technique?

Methods

This report updates a previously completed Comparative Effectiveness Review on core needle and open surgical biopsy methods for the diagnosis of breast cancer.¹² To update the report we performed a systematic review of the published scientific literature using established methodologies as outlined in the Agency for Healthcare Research and Quality’s (AHRQ) *Methods Guide for Comparative Effectiveness Reviews*, which is available at: <http://effectivehealthcare.ahrq.gov>.¹³ The main sections in this chapter reflect the elements of the protocol that guided this review. We have followed the reporting requirements of the “Preferred Reporting Items for Systematic Reviews and Meta-analyses” (PRISMA) checklist.¹⁴ All key methodological decisions were made *a priori*. The protocol was developed with input from external clinical and methodological experts, in consultation with the AHRQ task order officer (TOO), and was posted online to solicit additional public comments. Its PROSPERO registration number is CRD42013005690.

AHRQ TOO

The AHRQ TOO was responsible for overseeing all aspects of this project. The TOO facilitated a common understanding among all parties involved in the project, resolved ambiguities, and fielded all Evidence-based Practice Center (EPC) queries regarding the scope and processes of the project. The TOO and other staff at AHRQ helped to establish the Key Questions and protocol and reviewed the report for consistency, clarity, and to ensure that it conforms to AHRQ standards.

External Stakeholder Input

A new panel of experts was convened to form the Technical Expert Panel (TEP). The TEP included representatives of professional societies, experts in the diagnosis and treatment of breast cancer (including radiologists and surgeons), and a patient representative. The TEP provided input to help further refine the Key Questions and protocol, identify important issues, and define the parameters for the review of evidence. Discussions among the EPC, TOO, and the TEP occurred during a series of teleconferences and via email.

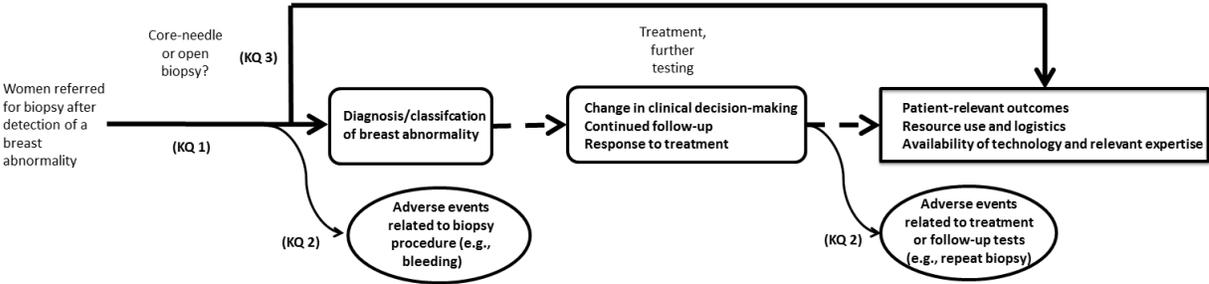
Key Questions

The final Key Questions are listed at the end of the Background section. The refinement of the Key Questions took into account the patient populations, interventions, comparators, outcomes, and study designs that are clinically relevant for core needle biopsies.

Analytic Framework

We used an analytic framework (Figure 1) that maps the Key Questions within the context of populations, interventions, comparators, and outcomes of interest. The framework was adapted from that used in the original 2009 CER. It depicts the chain of logic that links the test performance of core needle biopsy for the diagnosis of breast abnormalities (Key Question 1) with patient-relevant outcomes (Key Question 3) and adverse events of testing (Key Question 2).

Figure 1. Analytic framework



KQ = Key Question.

Scope of the Review

Populations and Conditions of Interest

The population of interest for all Key Questions was women who have been referred for biopsy for the diagnosis of primary breast cancer (including multifocal and bilateral disease) following self-examination, physical examination, or screening mammography. Studies carried out in women who had been previously diagnosed with breast cancer and were being examined for recurrence or to assess the extent of disease (staging) were excluded. The original report excluded studies carried out in women at high risk of breast cancer; however, MRI-guided biopsy is used mainly in this subset of patients. For this reason, following extensive discussions

with the TEP, we decided to broaden the scope of the review to include studies carried out in women at high baseline risk of breast cancer (e.g., on the basis of BRCA genetic testing or family history of breast cancer).

Interventions

For all Key Questions, the interventions of interest were core-needle and open biopsy done to evaluate whether a breast lesion is malignant. Other uses of biopsy techniques (e.g., use of biopsy to examine the sentinel lymph nodes in women with an established diagnosis of breast cancer) were not considered. Studies were required to have used biopsy instrumentation that is currently commercially available, as studies of discontinued devices are not applicable to current practice.

Comparators (reference standard and comparator index tests)

For test performance outcomes (Key Question 1) the reference standard was either open surgical biopsy, or follow-up by clinical examination and/or mammography for at least six months. The diagnostic performance of each core biopsy technique (each index test) was quantified versus the reference standard. Most assessments of diagnostic performance quantify the sensitivity and the specificity of each index test – here each needle core biopsy technique. Sensitivity and specificity are probabilities conditional on true disease status, and are noncomparative in nature. The reference standard is used in their definition, and is not a “comparator test”. The comparative diagnostic performance of alternative needle core biopsy techniques was also evaluated. For adverse events and patient-relevant outcomes (outcomes other than diagnostic performance; Key Questions 2 and 3) the comparators of interest were: open surgical biopsy, follow-up by clinical examination and/or mammography for at least six months, or alternative core-needle biopsy methods (e.g., stereotactic mammography versus ultrasound to locate the breast lesion; use versus non-use of vacuum-assistance to extract tissue samples).

Outcomes

For Key Question 1, the outcome of interest was test performance, as assessed by sensitivity (proportion of cancers detected by the reference standard that are also detected by core needle biopsy); specificity (proportion of negative findings according to core needle biopsy that were classified as negative by the reference standard; equal to one minus the false negative rate); underestimation rate for high risk lesions (most often atypical ductal hyperplasia, ADH), defined as the proportion of core needle biopsy findings of high risk lesions that are found to be malignant according to the reference standard); and underestimation rate for ductal carcinoma in situ (DCIS), defined as the proportion of core needle biopsy findings of DCIS that are found to be invasive according to the reference standard.

For Key Question 2 we looked for the following outcomes: rate of inconclusive biopsy findings (e.g. inadequate sampling of lesion); repeat biopsy rate; subsequent false positive and false negative rates on mammography; dissemination of cancerous cells along the needle track; and patient-centered outcomes (including bruising, bleeding or hematomas, pain, use of pain medication, infections, fainting or near fainting, time to recover). Because adverse events were not consistently defined across studies, we accepted the definitions used in the individual studies (when available).

For Key Question 3, we considered patient-relevant outcomes [patient preferences for specific procedures, cosmetic results, quality of life, anxiety and other psychological outcomes, time to complete tumor removal (for women with cancer), recurrence rate (for women with cancer, including local, regional, and distant recurrence), cancer-free survival and overall survival]; resource use and logistics [costs, resource utilization other than cost (number of additional surgical procedures, procedural time), subsequent surgical procedures, wait time for test results]; and availability of technology and relevant expertise [physician experience, availability of equipment, availability of (qualified) pathologists to evaluate biopsy samples].

Timing

We required that the duration of clinical and/or mammography follow-up was at least six months in studies where open surgical biopsy was not performed.

Setting

Studies in all geographic locations and care settings were evaluated, including general hospitals, academic medical centers, and ambulatory surgical centers, among others.

Study Design and Additional Criteria

We required that studies had been published in peer-reviewed journals as full articles. For all Key Questions, studies were required to have been published in English. Restricting included studies to those published in English, which was also an inclusion criterion in the original review, was deemed unlikely to bias the results of the review and avoids the resource-intensive translation of research articles published in languages other than English.

For Key Question 1 eligible studies were prospective or retrospective cohort studies or randomized controlled trials. Retrospective case studies (“case series”¹⁵) and other studies sampling patients on the basis of outcomes (e.g. diagnostic case-control studies, or studies selecting cases on the basis of specific histological findings) were excluded. Empirical evidence from meta-epidemiological studies suggests that diagnostic case-control studies may overestimate test performance.^{16, 17} Studies were required to report information on the sensitivity, specificity, positive or negative predictive value of tests, or to include data that allow the calculation of one or more of these outcomes. Specifically, studies needed to provide adequate information to reconstruct 2×2 tables of test performance of the index against the reference standard. **Table 1** illustrates how index and reference standard results were used to construct such 2×2 tables.

Table 1: Definitions of diagnostic groups based on index and reference standard test results

| | | Reference standard results (open surgery or followup) | |
|---|--|--|----------------|
| | | <i>Malignant (invasive or in situ)</i> | <i>Benign</i> |
| 'Core-needle biopsy results (index test) | <i>Malignant (invasive or in situ)</i> | considered TP | considered TP* |
| | <i>High risk lesion (e.g., ADH)</i> | considered TP | considered FP |
| | <i>Benign</i> | considered FN | considered TN |

*Some study authors specifically stated that diagnoses of malignancy on core-needle biopsy were assumed to be correct, whether or not a tumor was observed upon surgical excision. The original version of this review also classified all diagnoses of malignancy on core-needle biopsy as true positives.

ADH = atypical ductal hyperplasia; FN = false negative; FP = false positive; TN = true negative; TP = true positive.

Non-comparative studies of test performance (i.e. studies of a single index test) were required to have enrolled at least 10 participants per arm or per comparison group. This inclusion criterion was intended to reduce the risk of bias from non-representative participants in small studies. Further, smaller studies do not produce precise estimates of test performance and as such are unlikely to substantially affect results. Studies were also required to have followed at least fifty percent of participants to completion. This criterion was intended to reduce the risk of bias from high rates of attrition.

Key Question 2 was addressed by extracting harm-related information for core-needle biopsy and open surgical biopsy from studies meeting the criteria for Key Question 1. In addition, we included studies that met all other selection criteria for Key Question 1 except for the use of a reference standard and the reporting of information on test performance outcomes. This allowed us to consider additional sources of evidence that assess adverse events. Finally, for this Key Question, we also reviewed primary research articles, regardless of design (i.e., case reports and case series, case-control studies, cohort studies, randomized trials), that address the dissemination of cancer cells by the biopsy procedure, a relatively rare harm that is specific to core biopsy.

The original report did not use formal criteria for study selection for Key Question 3.¹⁸ Based on the findings of the original report, we used the same PICOTS criteria described above and considered the following study designs:

- Randomized controlled trials, cohort studies, and cross-sectional studies on patient preferences, cosmetic results of biopsy procedures, physician experience (including studies of the “learning curve” for different biopsy methods and tools).
- Cost studies, including cost-minimization and cost-consequence analyses, were used to obtain information on resource utilization and unit costs. Given the large variability of cost information among different jurisdictions, we only considered studies conducted in the U.S. setting and published after 2004.¹⁹
- Cost-effectiveness/cost-utility analyses based on primary trials²⁰ of breast biopsy interventions were used to obtain information on unit costs and resource utilization. Specifically, we considered the components of cost and resource use but did not use cost-effectiveness ratios or other summary measures of cost-effectiveness/utility. As for cost studies, we only considered primary cost-effectiveness/-utility studies conducted in the US setting and published after 2004.¹⁹ We did not use model-based cost-effectiveness results.

- Studies of pathologist qualifications for interpreting core-needle biopsy results; including interlaboratory initiatives to standardize diagnostic criteria (e.g., proficiency testing) or minimal competency requirements.
- Surveys of the availability of equipment for obtaining core-needle biopsies and of qualified pathologists to examine biopsy samples.

Literature search and Abstract Screening

We searched MEDLINE®, EMBASE®, the Cochrane Central Register of Controlled Trials (CENTRAL), the Cochrane Database of Systematic Reviews, the Database of Abstracts of Reviews of Effects (DARE), the Health Technology Assessment Database (HTA), the U.K. National Health Service Economic Evaluation Database (NHS EED), the U.S. National Guideline Clearinghouse (NGC), and the Cumulative Index to Nursing and Allied Health Literature (CINAHL®).²¹ **Appendix A** describes the search strategy we employed which is a revision and expansion of the search strategy used in the original report. Of note, the original report used a search filter for studies of diagnostic tests to increase search specificity; this is a reasonable approach given the large volume of literature on studies on diagnostic biopsy methods for breast cancer. Because this update covered a short time period (from 2009 to 2013) we opted to not use this filter, in order to increase search sensitivity.²² Our searches covered the time period from six months before the most recent search date in the original report, to ensure adequate overlap.

To identify studies excluded from the original report because they enrolled women at high risk for cancer, the set of abstracts screened for the original report was obtained and rescreened for potentially eligible studies of high-risk women. In addition, the list of studies excluded from the original report following full text review was checked to identify studies excluded because they included women at high risk for cancer. We also performed a search for systematic reviews on the topic and used their reference lists of included studies to validate our search strategy and to make sure we identified all relevant studies.

All reviewers screened a common set of 200 abstracts (in 2 pilot rounds, each with 100 abstracts), and discussed discrepancies, in order to standardize screening practices and ensure understanding of screening criteria. The remaining citations were split into non-overlapping sets, each screened by two reviewers independently. Discrepancies were resolved by consensus involving a third investigator.

We asked the TEP to provide citations of potentially relevant articles. Additional studies were identified through the perusal of reference lists of eligible studies, published clinical practice guidelines, relevant narrative and systematic reviews, Scientific Information Packages from manufacturers, and a search of U.S. Food and Drug Administration databases. All articles identified through these sources were screened for eligibility against the same criteria as for articles identified through literature searches. We sent the final list of included studies to the TEP to ensure that no key publications had been missed.

Following submission of the draft report (in December 2013), an updated literature search (using the same search strategy) will be conducted. Abstract and full-text screening of citations retrieved by this search will be performed as described above. Any additional studies that meet the eligibility criteria, including those that are identified through the peer review and comment process, will be added to the final report.

Study Selection and Eligibility Criteria

Potentially eligible citations were obtained in full text and reviewed for eligibility on the basis of the predefined inclusion criteria. A single reviewer screened each potentially eligible article in full-text to determine eligibility; reviewers were instructed to be inclusive. A second reviewer verified all relevant articles. Disagreements regarding article eligibility were resolved by consensus involving a third reviewer. **Appendix B** lists all the studies excluded after full-text screening and the reason for exclusion.

Data Abstraction and Management

Data was extracted using electronic forms and entered into the Systematic Review Data Repository (SRDR; <http://srdhr.ahrq.gov/>). The basic elements and design of these forms is similar to those we have used for other reviews of diagnostic tests and includes elements that address population characteristics, sample size, study design, descriptions of the index and reference standard tests of interest, analytic details, and outcome data. Prior to data extraction, forms were customized to capture all elements relevant to the Key Questions. We used separate sections in the extraction forms for Key Questions related to short-term outcomes, including classification of breast abnormalities, intermediate outcomes (such as clear surgical margins), patient-relevant outcomes (such as quality of life), and factors affecting (modifying) test performance. We pilot-tested the forms on several studies extracted by multiple team members to ensure consistency in operational definitions.

A single reviewer extracted data from each eligible study. The extracted data was reviewed and confirmed by at least one other team member (data verification). Disagreements were resolved by consensus including a third reviewer. We contacted authors (1) to clarify information reported in the papers that is hard to interpret (e.g., inconsistencies between tables and text); and (2) to verify suspected overlap between study populations in publications from the same group of investigators. (NB: The author contact process will be completed during the peer review of this Draft Report and any changes will be reflected in the Final Report).

Assessment of the Risk of Bias of Individual Studies

We assessed the risk of bias for each individual study using the assessment methods detailed in the AHRQ Methods Guide for Effectiveness and Comparative Effectiveness Review hereafter referred to as the Methods Guide. We used elements from the Quality Assessment for Diagnostic Accuracy Studies instrument (QUADAS version 2), to assess the risk of bias (methodological quality or internal validity) of the diagnostic test studies included in the review (these studies comprise the majority of the available studies).²³⁻²⁶ The tool assesses four domains of risk of bias related to patient selection, index test, reference standard test, and patient flow and timing. For studies of other designs we used appropriate sets of items to assess risk of bias or methodological “quality”: for nonrandomized cohort studies we used items from the Newcastle-Ottawa scale,²⁷ for randomized controlled trials we used items from the Cochrane Risk of Bias tool,²⁸ and for studies of resource utilization and costs we used items from the checklist proposed by Drummond et al.^{29,30}

We assessed and reported methodological quality items (as “Yes”, “No”, or “Unclear/Not Reported”) for each eligible study. We then rated each study as being of low, intermediate, or high risk of bias on the basis of adherence to accepted methodological principles. Generally, studies with low risk of bias have the following features: lowest likelihood of confounding due to

comparison to a randomized controlled group; a clear description of the population, setting, interventions, and comparison groups; appropriate measurement of outcomes; appropriate statistical and analytic methods and reporting; no reporting inconsistencies; clear reporting of dropouts and a low dropout rate; and no other apparent sources of bias. Studies with moderate risk of bias are susceptible to some bias but not sufficiently to invalidate results. They do not meet all the criteria for low risk of bias owing to some deficiencies, but none are likely to introduce major bias. Studies with moderate risk of bias may not be randomized or may be missing information, making it difficult to assess limitations and potential problems. Studies with high risk of bias are those with indications of bias that may invalidate the reported findings (e.g., observational studies not adjusting for any confounders, studies using historical controls, or studies with very high dropout rates). These studies have serious errors in design, analysis, or reporting and contain discrepancies in reporting or have large amounts of missing information. We discuss the handling of high risk of bias studies in evidence synthesis in the following sections. Studies of different designs were graded within the context of their study design.

Data Synthesis

We summarized included studies qualitatively and presented important features of the study populations, designs, tests used, outcomes, and results in summary tables. Population characteristics of interest included age, race/ethnicity, and palpability of lesion. Design characteristics included methods of population selection and sampling, and follow-up duration. Test characteristics included imaging-guided versus not imaging-guided, and vacuum-assisted versus not vacuum-assisted methods. We looked for information on test performance, adverse events, patient preferences, and resource utilization including costs.

Statistical analyses were conducted using methods currently recommend for use in Comparative Effectiveness Reviews of diagnostic tests.^{31, 32} For all outcomes we assessed heterogeneity graphically (e.g. by inspecting a scatterplot of studies in the receiver operating characteristic, ROC, space) and by examining the posterior distribution of between-study variance parameters.

For Key Question 1 we performed meta-analysis on studies that were deemed sufficiently similar. Based on the technical characteristics of the different tests, and the findings of the original Evidence Report, we developed a mixed effects binomial-bivariate normal regression model that accounted for different imaging methods (e.g. US, stereotactic mammography, MRI), the use of vacuum (yes vs. not), the baseline of risk of cancer of included patients (high versus average risk), and residual (unexplained) heterogeneity.³³⁻³⁵ This model allowed us to estimate the test performance of alternative diagnostic tests, and perform indirect comparisons among them.³³ Furthermore, it allowed us to model the correlation between sensitivity and specificity and to derive meta-analytic ROC curves using the methods proposed by Rutter & Gatsonis and Arends et al.^{34, 35} A univariate mixed effects logistic regression (binomial-normal) model was used for the meta-analysis of DCIS and high risk lesion underestimation rates.³⁶

We performed meta-regression analyses (e.g. to evaluate the impact of study risk of bias items, or the effect of other study-level characteristics) by extending the model to include additional appropriately coded terms in the regression equations.^{37, 38} Such analyses were planned for patient and breast lesion factors (e.g., age, density of breast tissue, microcalcifications, and palpability of the lesions), biopsy procedure factors (e.g., needle size, imaging guidance, vacuum extraction, and number of samples), clinician and facility-related factors (e.g., training of the operator, country where the study was conducted), and risk of bias

items. We performed additional sensitivity analyses (e.g., leave-one-out meta-analysis and comparisons of studies added in the update versus studies included in the original report).³⁹

For Key Question 2, we found that adverse events were inconsistently reported (across studies) and that the methods for ascertaining their occurrence were often not presented in adequate detail. For this reason we refrained from performing meta-analyses for these outcomes. Instead, we calculated descriptive statistics (medians, 25th and 75th percentiles, minimum and maximum values) across all studies and for specific test types. For Key Question 3, because of the heterogeneity of research designs and outcomes assessed, for all outcomes except the number of surgical procedures, we did not perform meta-analysis but instead chose to summarize the data qualitatively. We performed a meta-analysis comparing core-needle and open surgical biopsies with respect to the number of patients who required one versus more than one surgical procedures for treatment, after the establishment of breast cancer diagnosis. This analysis used a standard univariate normal random effects model with a binomial distribution for the within-study likelihood of each biopsy group (core-needle vs. open).

All statistical analyses were performed using Bayesian methods; models were fit using Markov Chain Monte Carlo methods and non-informative prior distributions.⁴⁰ Theory and extensive empirical evidence suggests that, when the number of studies is large, this approach produces results similar to those of maximum likelihood methods (which do not require the specification of priors).⁴¹ Results were summarized as medians of posterior distributions with associated 95 percent central credible intervals.

Grading the Strength of Evidence for Individual Comparisons and Outcomes

We followed the Methods Guide⁴² to evaluate the strength of the body of evidence for each Key Question with respect to the following domains: risk of bias, consistency, directness, precision, and reporting bias.^{42, 43} Generally, strength of evidence was downgraded when risk of bias was not low, in the presence of inconsistency, when evidence was indirect or imprecise, or when we suspected that results were affected by selective analysis or reporting.

We determined risk of bias (low, medium, or high) on the basis of the study design and the methodological quality. We assessed consistency on the basis of the direction and magnitude of results across studies. We considered the evidence to be indirect when we had to rely on comparisons of biopsy methods across different studies (i.e., indirect comparisons). We considered studies to be precise if the credible interval (CrI) was narrow enough for a clinically useful conclusion, and imprecise if the CrI was wide enough to include clinically distinct conclusions. The potential for reporting bias (“suspected” vs. “not suspected”) was evaluated with respect to publication, selective outcome reporting, and selective analysis reporting. We made qualitative dispositions rather than perform formal statistical tests to evaluate differences in the effect sizes between more precise (larger) and less precise (smaller) studies because such tests cannot distinguish between “true” heterogeneity between smaller and larger studies, other biases, and chance.^{44, 45} Therefore, instead of relying on statistical tests, we evaluated the reported results across studies qualitatively, on the basis of completeness of reporting, number of enrolled patients, and numbers of observed events. Judgment on the potential for selective outcome reporting bias was based on reporting patterns for each outcome of interest across studies. We acknowledge that both types of reporting bias are difficult to reliably detect on the basis of data available in published research studies. We believe that our searches (across multiple databases), combined with our plan for contacting test manufacturers (for additional

data) and the authors of published studies (for data clarification) limited the impact of reporting and publication bias on our results, to the extent possible.

Finally, we rated the body of evidence using four strength of evidence levels: high, moderate, low, and insufficient.⁴² These describe our level of confidence that the evidence reflects the true effect for the major comparisons of interest.

Assessing Applicability

We followed the Methods Guide⁴² in evaluating the applicability of included studies to patient populations of interest. Applicability to the population of interest was also judged separately on the basis of patient characteristics (e.g., age may affect test performance because the consistency of the breast tissue changes over time), method by which suspicion is established (e.g., mammography vs. other methods may affect test performance through spectrum effects), baseline risk of cancer (“average risk” vs. “high risk” women may affect estimated test performance because of differences in diagnostic algorithms), outcomes (e.g., prevalence of breast cancers diagnosed upon biopsy may also be a marker of spectrum effects), and setting of care (because differences in patient populations, diagnostic algorithms, and available technologies may affect test results).

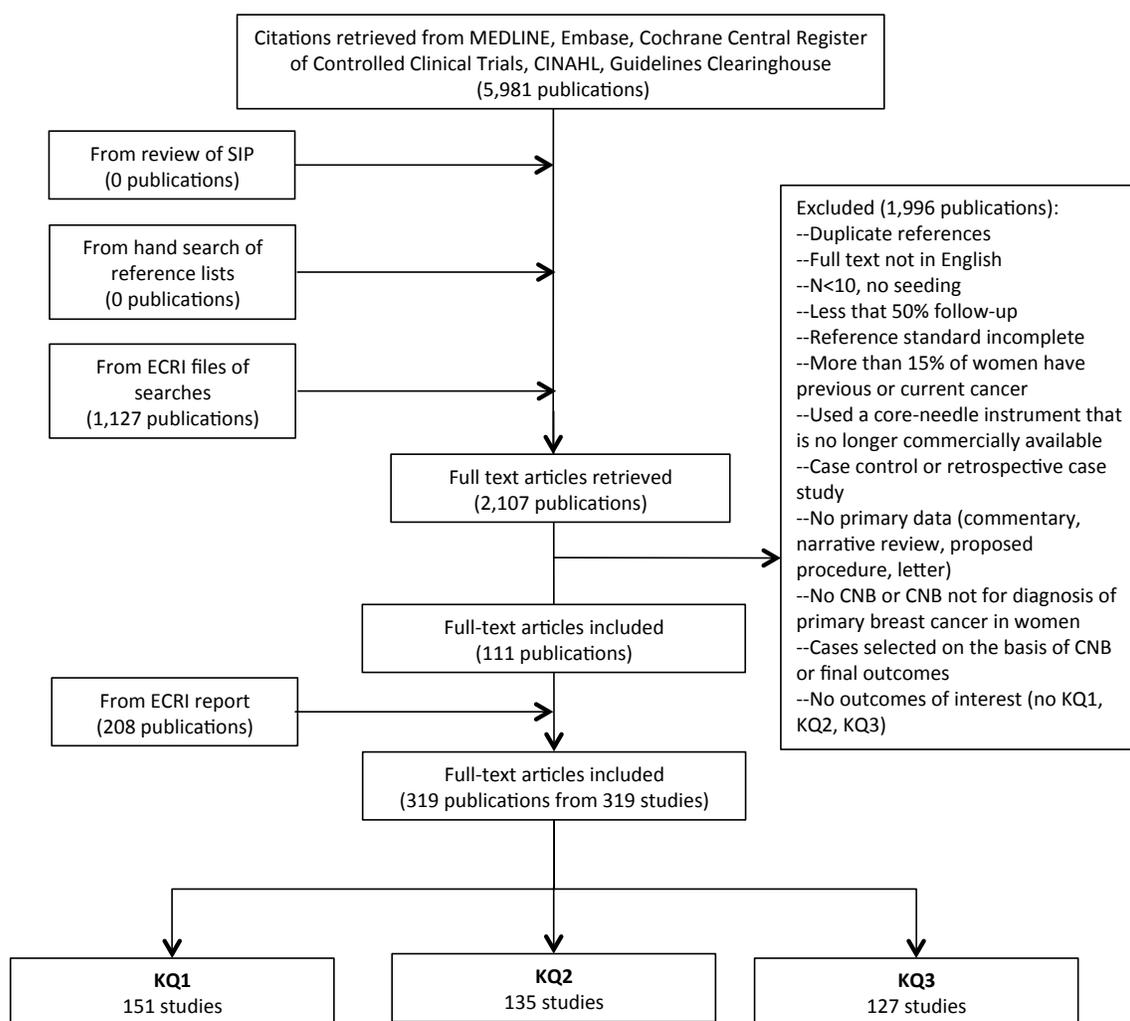
Peer Review

The initial draft report was pre-reviewed by the TOO and an AHRQ Associate Editor (a senior member of another EPC). Following revisions, the draft report was sent to invited peer reviewers and was simultaneously uploaded to the AHRQ Web site where it was available for public comment for 30 days. All reviewer comments (both invited and from the public) were collated and individually addressed. The revised report and the EPC’s responses to invited and public reviewers’ comments were again reviewed by the TOO and Associate Editor prior to completion of the report. The authors of the report had final discretion as to how the report was revised based on the reviewer comments, with oversight by the TOO and Associate Editor.

Results

Our literature searches identified 7,108 potentially relevant articles. After review of the abstracts, the full-length articles of 2,107 of these studies were obtained and examined in full text. Of these, 111 articles (reporting the results of 111 studies) were considered eligible for inclusion in the updated review. **Figure 2** presents the literature flow and **Table 2** summarizes the additions to the original report, separately by Key Question.

Figure 2: Flow chart of included studies



CNB = core-needle biopsy; KQ = Key Question; N = number of patients; SIP = Scientific Information Packet.

Table 2: Summary of new evidence evaluated in this update*

| Key Question | Studies included in the original report | Studies identified by the updating process | Total number of studies synthesized in this report |
|--|--|---|---|
| <i>Key Question 1: What is the test performance of different types of core-needle breast biopsy compared with open biopsy in the diagnosis of breast cancer?</i> | 107 | 44 | 151 |
| <i>Key Question 2: What are the adverse events (harms) associated with core-needle breast biopsy compared to the open biopsy in the diagnosis of breast cancer?</i> | 72 | 63 | 135 |
| <i>Key Question 3: How do open biopsy and various core-needle techniques differ in terms of patient preference, availability, costs, availability of qualified pathologist interpretations, and other factors that may influence choice of a particular technique?</i> | 86 | 41 | 127 |

*Some studies address multiple Key Questions

Key Question 1: In women with a palpable or non-palpable breast abnormality, what is the test performance of different types of core-needle breast biopsy compared with open biopsy for diagnosis?

Included Studies

Forty-four new studies identified by this update met the inclusion criteria for Key Question 1. We synthesized these studies with the 107 studies identified by the original evidence report, for a total of 151 studies providing information on test performance outcomes. Studies had been published between 1990 and 2013. Forty-seven studies were prospectively designed, and 58 were conducted in the United States. Ten studies provided information on more than one group of patients (typically undergoing biopsy with a different biopsy device). In statistical analyses these groups were treated as separate strata, leading to a total of 161 complete 2x2 tables of diagnostic test results, with information on 68,942 breast lesions.

Test Performance for Breast Cancer Diagnosis

Test Performance of Open Surgical Biopsy

Neither the original report nor our updated searches identified any clinical studies of open surgical biopsy that met our inclusion criteria. Research studies of needle biopsy methods and technical experts generally suggested that open surgical biopsy could be considered a “gold” standard test (i.e. a test without measurement error). One study identified by the original evidence report, provided information on the test performance of open surgical biopsy, using published literature and primary patient data (patient charts) from patients evaluated at a single medical center. Based on a re-review of archived open biopsy material by a second pathologist, patient chart review, study of cases with benign results on biopsy after suspicious mammography results, and expert opinion, the authors concluded that open surgical biopsy may miss one to two percent of breast cancers (i.e. sensitivity of 98% or greater). The original evidence report and our update did not identify any information on underestimation rates for open surgical biopsy. Because open surgical biopsy samples the entire target lesion or a large part of it, in theory underestimation should not occur.

Test Performance of Core-Needle Biopsy Methods

A total of 151 studies contributed information to analyses of test performance of core-needle biopsy methods.⁴⁶⁻¹⁹⁶ Five studies enrolled women at high-risk of cancer development and 146 enrolled women at average risk. The studies reported on a variety of biopsy techniques: 114 reported on the use of a single form of imaging guidance (74 stereotactic; 35 ultrasound; 4 MRI) whereas seven used freehand methods, and 29 used multiple methods, including freehand techniques in some cases (and did not report test performance results separately by each method); 51 studies used vacuum-assisted methods to obtain the biopsy sample; 71 used automated methods; 28 used multiple methods; and 1 did not report adequate details. Needle size also varied across studies: of the 107 studies reporting information on this aspect of the biopsy procedure, 56 used 14G needles, nine used smaller and 42 used larger bores; 44 studies did not report relevant information. Reference standard tests also differed across studies: 26 used open biopsy on all included patients; 87 used mean or median followup of between 6 and 24 months

for test negative patients, and 38 used mean or median followup of 24 months or more for test negative cases. Additional details about the designs of included studies, their selection criteria, enrolled patients, biopsy methods and results, are publically available in the SRDR. Consistent with the findings of the original report, the overall risk of bias was considered moderate to high, mainly due to concerns about spectrum bias, retrospective data collection, differential verification, and lack of information regarding the blinding of reference standard test assessors to the index test results. Additional results from our risk of bias assessment are provided at the end of this section.

Table 3 summarizes the results for alternative diagnostic biopsy methods, together with information on the number of lesions evaluated by each test and summary test performance information, for women at average risk of cancer. **Table 4** summarizes the same information for women deemed to be at high-risk for cancer (e.g. due to genetic factors or strong family history). Figure 3 presents individual study estimates and meta-analytic results in the ROC space for both groups of women. These plots indicate that results were fairly homogeneous across studies for each test and that test sensitivity and specificity were close to 1 (studies cluster at the top left corner of the space).

Key findings with respect to test performance (sensitivity, specificity, and positive and negative likelihood ratios¹) and underestimation rates are summarized narratively below. As mentioned, only five studies reported results on the test performance of various biopsy methods for breast cancer diagnosis in high risk women. Of these studies, only two reported information on underestimation rates (both for high risk lesions; none for DCIS); for this reason, we did not include studies of high risk women in statistical analyses of underestimation; instead we summarized their results narratively.

¹ To aid in the interpretation of likelihood ratios we remind readers that these statistics can be used to convert pre-test probabilities to post-test probabilities. For example, before testing, assume that a patient has probability of

disease $pre\text{-test } p = 0.1$ and $pre\text{-test odds} = \frac{pre\text{-test } p}{1 - pre\text{-test } p} = \frac{0.1}{0.9} = 0.11$. If the diagnostic test has a positive

likelihood ratio (LR^+) of 15 then the post-test odds are $post\text{-test odds} = pre\text{-test odds} \times LR^+ = 0.11 \times 15 = 1.67$.

This corresponds to a post-test probability of $post\text{-test } p = \frac{post\text{-test odds}}{post\text{-test odds} + 1} = \frac{1.67}{1.67 + 1} = 0.625$ (i.e. the post-test

probability is approximately 6 times greater than the pre-test value). If the test results had been negative and the test had a negative likelihood ratio (LR^-) of 0.1, the post-tests odds would be

$post\text{-test odds} = pre\text{-test odds} \times LR^- = 0.11 \times 0.1 = 0.011$, which corresponds to $post\text{-test } p = \frac{0.011}{0.011 + 1} = 0.011$

(i.e. approximately the post-test probability is approximately 10 times lower than the pre-test value). As a rule of thumb, $LR^+ > 10$ and $LR^- < 0.1$ are generally considered clinically meaningful.

Table 3: Summary estimates of test performance for alternative core-needle biopsy methods – women at average risk of cancer

| Biopsy method or device | N studies [N biopsies] for sensitivity & specificity | Sensitivity | Specificity | N studies [N biopsies] for DCIS underestimation | DCIS underestimation probability | N studies [N biopsies] for high risk lesion underestimation | High risk lesion underestimation probability |
|--|--|-------------------|-------------------|---|----------------------------------|---|--|
| Freehand, automated | 10 [786] | 0.91 (0.80, 0.96) | 0.98 (0.95, 1.00) | 0 [0] | NA | 1 [6] | NA |
| US-guided, automated | 27 [16287] | 0.99 (0.98, 0.99) | 0.97 (0.95, 0.98) | 14 [307] | 0.38 (0.25, 0.51) | 20 [502] | 0.22 (0.14, 0.34) |
| US-guided, vacuum-assisted | 10 [1456] | 0.97 (0.89, 0.99) | 0.99 (0.97, 0.99) | 4 [21] | 0.11 (0.01, 0.41) | 7 [16] | 0.09 (0.01, 0.30) |
| Stereotactically guided, automated | 36 [9342] | 0.97 (0.95, 0.98) | 0.97 (0.96, 0.98) | 17 [649] | 0.27 (0.18, 0.37) | 28 [353] | 0.47 (0.37, 0.58) |
| Stereotactically guided, vacuum-assisted | 40 [14421] | 0.99 (0.98, 0.99) | 0.92 (0.89, 0.94) | 33 [1803] | 0.11 (0.08, 0.14) | 37 [949] | 0.18 (0.13, 0.24) |
| MRI-guided, automated | 2 [89] | 0.90 (0.58, 0.99) | 0.98 (0.90, 1.00) | 0 [0] | NA | 1 [1] | NA |
| Multiple techniques | 28 [25391] | 0.98 (0.97, 0.99) | 0.95 (0.93, 0.97) | 16 [573] | 0.22 (0.15, 0.31) | 22 [822] | 0.33 (0.24, 0.43) |

All numbers are medians with 95% Crls, unless otherwise stated. 10 studies provided information multiple patient groups (cohorts) treated with alternative biopsy methods. Summary results are shown when at least two studies were available. Results are not shown for three studies that used devices not belonging to any of the categories listed in the table (1 grid guidance, 2 unclear) are not shown.

Crl = credible interval; DCIS = ductal carcinoma in situ; MRI = magnetic resonance imaging; N = number; NA = not applicable; US = ultrasound.

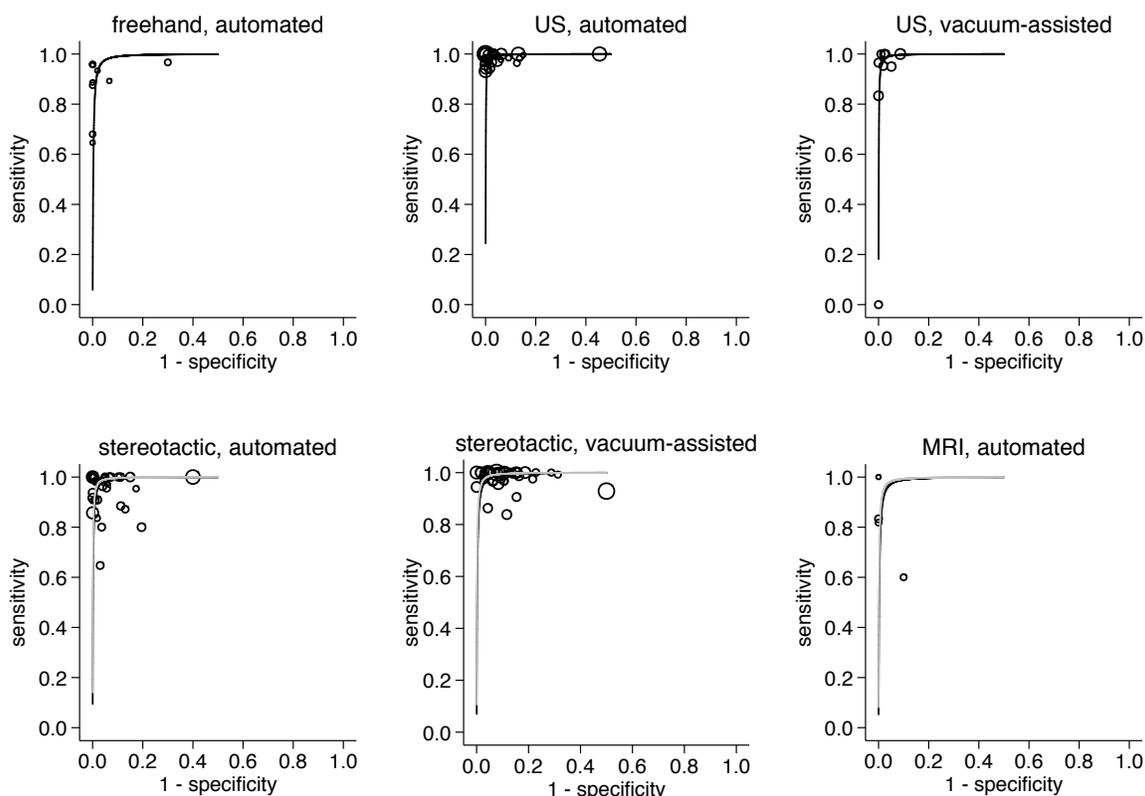
Table 4: Summary estimates of test performance for alternative core-needle biopsy methods – women at high risk of cancer

| Biopsy method or device | N studies (N biopsies) for sensitivity and specificity | Sensitivity (95% CrI) | Specificity (95% CrI) |
|--|---|--------------------------|--------------------------|
| Stereotactically guided, automated | 1 [416] | 0.98 (0.93, 0.99) | 0.97 (0.83, 1.00) |
| Stereotactically guided, vacuum-assisted | 2 [311] | 0.94 (0.83, 0.98) | 0.99 (0.93, 1.00) |
| MRI-guided, automated | 2 [56] | 0.99 (0.93, 1.00) | 0.89 (0.58, 0.98) |

No studies provided information on the test performance of freehand or US-guided biopsy methods, or the use of multiple methods in populations of women at high risk of cancer. Results are based on bivariate model with risk group as a covariate.

CrI = credible interval; DCIS = ductal carcinoma in situ; MRI = magnetic resonance imaging; N = number; US = ultrasound.

Figure 3: Scatterplot of results in the receiver operating characteristic space and summary receiver operating characteristic curves of alternative core-needle biopsy methods for the diagnosis of breast cancer



Solid black lines represent results for average risk women; solid gray lines represent results for high-risk women (when results were available). MRI = magnetic resonance imaging; US = ultrasound.

Freehand Core-Needle Biopsies

Women at average risk of cancer: Ten studies reported data on the accuracy of non-guided (i.e., freehand) core-needle biopsies performed with automated biopsy devices. The summary sensitivity was 0.91 (95% CrI, 0.80 to 0.96) and the summary specificity was 0.98 (95% CrI, 0.95 to 1.00), corresponding to a positive likelihood ratio of 56.2 (95% CrI, 18.4 to 206.5) and a negative likelihood ratio of 0.10 (95% CrI, 0.04 to 0.21). Only one study provided information on the high risk lesion underestimation rate (five cancers misclassified as high risk lesions among a total of six such lesions on core-needle biopsy). No studies provided information on the DCIS underestimation rate.

Women at high risk of cancer: No studies provided information on the test performance (sensitivity, specificity, or underestimation rates) of freehand core needle biopsy techniques in women at high risk of breast cancer.

Ultrasound-Guided Automated Device Core-Needle Biopsies

Women at average risk of cancer: Twenty-seven studies of 16,287 biopsies used ultrasound guidance and an automated biopsy device. The summary sensitivity was 0.99 (95% CrI, 0.98 to 0.99) and the summary specificity was 0.97 (95% CrI, 0.95 to 0.98), corresponding to a positive likelihood ratio of 32.1 (95% CrI, 20.2 to 52.8) and a negative likelihood ratio of 0.01 (95% CrI, 0.01 to 0.03). Fourteen studies provided information on the DCIS underestimation; the summary rate was 0.38 (95% CrI, 0.25 to 0.51). Twenty studies provided information on high risk lesion underestimation; the summary rate was 0.22 (95% CrI, 0.14 to 0.34).

Women at high risk of cancer: No studies provided information on the test performance (sensitivity, specificity, or underestimation rates) of ultrasound-guided automated core needle biopsy techniques in women at high risk of breast cancer.

Ultrasound-Guided Vacuum-Assisted Core-Needle Biopsies

Women at average risk of cancer: Ten studies of 1,456 biopsies used ultrasound guidance and a vacuum-assisted device to perform breast biopsies. The summary sensitivity was 0.97 (95% CrI, 0.89 to 0.99) and the summary specificity was 0.99 (95% CrI, 0.97 to 0.99), corresponding to a positive likelihood ratio of 70.0 (95% CrI, 29.5 to 179.0) and a negative likelihood ratio of 0.04 (95% CrI, 0.01 to 0.11). Four studies provided information on DCIS underestimation: the summary rate was 0.11 (95% CrI, 0.01 to 0.41). Seven studies provided information on high risk lesion underestimation: the summary rate was 0.09 (95% CrI, 0.01 to 0.30).

Women at high risk of cancer: No studies provided information on the test performance (sensitivity, specificity, or underestimation rates) of ultrasound-guided automated core needle biopsy techniques in women at high risk of breast cancer.

Stereotactically Guided Automated Device Core-Needle Biopsies

Women at average risk of cancer: Thirty-six studies of 9,342 biopsies used stereotactic guidance and an automated biopsy device. The summary sensitivity was 0.97 (95% CrI, 0.95 to 0.98) and the summary specificity was 0.97 (95% CrI, 0.96 to 0.98), corresponding to a positive likelihood ratio of 32.1 (95% CrI, 21.8 to 49.3) and a negative likelihood ratio of 0.03 (95% CrI, 0.02 to 0.05). Seventeen studies provided information on DCIS underestimation; the summary

rate was 0.27 (95% CrI, 0.18 to 0.37). Twenty-eight studies provided information on high risk lesion underestimation; the summary rate was 0.47 (95% CrI, 0.37 to 0.58).

Women at high risk of cancer: One study reported information on the test performance of stereotactically guided automated core-needle biopsy methods. Using results from the joint bivariate model, sensitivity was 0.98 (0.93 to 0.99) and specificity was 0.97 (0.83 to 1.00).

Stereotactically Guided Vacuum-Assisted Core-Needle Biopsies

Women at average risk of cancer: Forty studies of 14,421 biopsies used stereotactic guidance and a vacuum-assisted device to perform core-needle biopsies. The summary sensitivity was 0.99 (95% CrI, 0.98 to 0.99) and the summary specificity was 0.92 (95% CrI, 0.89 to 0.94), corresponding to a positive likelihood ratio of 11.8 (95% CrI, 8.68 to 16.2) and a negative likelihood ratio of 0.01 (95% CrI, 0.01 to 0.02). Thirty-three studies provided information on DCIS underestimation; the summary rate was 0.11 (95% CrI, 0.08 to 0.14). Thirty-seven studies provided information on high risk lesion underestimation; the summary underestimation rate was 0.18 (95% CrI, 0.13 to 0.24).

Women at high risk of cancer: Two studies provided information on the test performance of stereotactically guided vacuum assisted core-needle biopsies. The summary sensitivity was 0.94 (95% CrI 0.83 to 0.98) and summary specificity was 0.99 (0.93 to 1.00). One of the two studies also reported that two cancer cases were underestimated by the biopsy diagnosis, among a total of 17 high risk lesions (for an underestimation rate of 12%).

MRI-Guided Core-Needle Biopsies

Women at average risk of cancer: Two studies reported data on the accuracy of MRI-guided biopsies performed with automated biopsy devices. The summary sensitivity was 0.90 (95% CrI, 0.58 to 0.99) and the summary specificity was 0.98 (95% CrI, 0.90 to 1.00), corresponding to a positive likelihood ratio of 50.2 (95% CrI, 8.37 to 524.7) and a negative likelihood ratio of 0.10 (95% CrI, 0.01 to 0.43). None of the studies provided information on the DCIS underestimation rate. One study provided information on the a high risk lesion underestimation rate (one biopsy-detected high risk lesion was found to be malignant).

Women at high risk of cancer: Two studies provided information on the test performance of MRI-guided core-needle biopsies among women at high risk for cancer. The summary sensitivity was 0.99 (95% CrI 0.93 to 1.00) and summary specificity was 0.89 (0.58 to 0.98). One of the two studies also reported that no cancers developed in the two women considered to have high risk lesions on core-needle biopsy (i.e. no underestimation was observed in the study).

Populations Biopsied with Multiple Core-Needle Methods

Women at average risk of cancer: An additional 28 studies reported results from populations of women undergoing core-needle biopsy with diverse methods, without stratifying their results by biopsy method. In this heterogeneous group of studies, the summary sensitivity was 0.98 (95% CrI, 0.97 to 0.99) and the summary specificity was 0.95 (95% CrI, 0.93 to 0.97), corresponding to a positive likelihood ratio of 21.4 (95% CrI, 14.3 to 32.6) and a negative likelihood ratio of 0.02 (95% CrI, 0.01 to 0.03). Sixteen studies provided information on the DCIS underestimation; the summary DCIS underestimation rate was 0.22 (95% CrI, 0.15 to 0.31). Twenty-two studies provided information on high risk lesion underestimation; the summary underestimation rate was 0.33 (95% CrI, 0.24 to 0.43).

Women at high risk of cancer: No studies of high-risk women were included in this subgroup.

Contextualizing the Results of Test Performance Meta-analyses

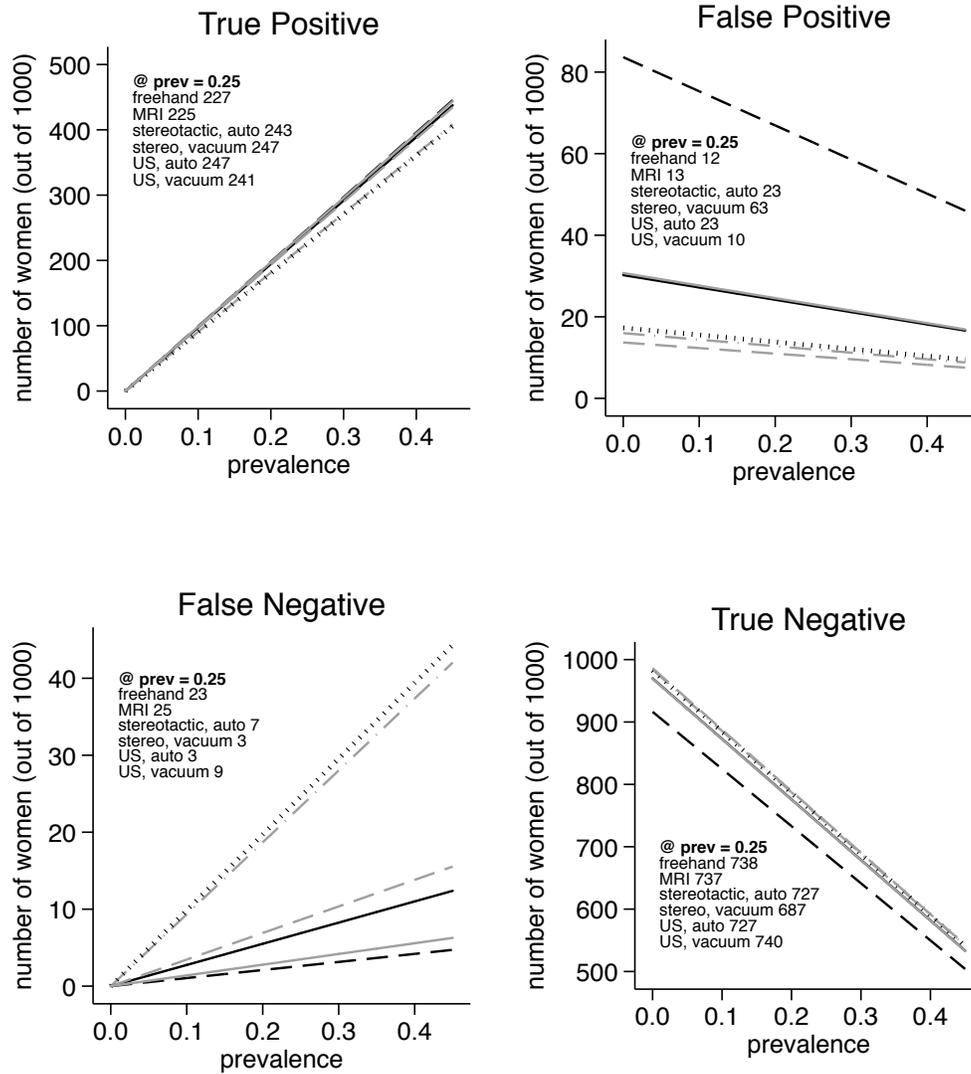
To contextualize the results of the test performance meta-analyses presented in the preceding sections, we evaluated the impact of testing in a hypothetical cohort of 1000 women, under alternative scenarios for disease prevalence. The results are presented in **Figure 4**. In low cancer prevalence settings, the number of false negative results (i.e. cases where treatment may be delayed on the basis of biopsy results) is expected to be small (e.g., for all ultrasound- or stereotactically guided biopsy methods, less than five out of 1000 women, if prevalence is 10 percent or less). As prevalence increases, the number of false negative results increases for all biopsy methods (and more rapidly for MRI-guided and freehand methods, which had the lowest summary sensitivity).

The number of false positive cases declines with increasing prevalence. Automated device biopsy methods (both stereotactically and ultrasound-guided) have comparable results (approximately 20-30 false positive results in the range examined). MRI- or ultrasound-guided vacuum-assisted, and freehand methods appear to perform somewhat better (less than 20 false positives in the range of prevalence examined). Stereotactically guided, vacuum-assisted methods appear to produce the most false positive results (more than 40 per 1000 women over the range of prevalence examined), which is consistent with them having the lowest summary specificity. **Figure 4** also presents numerical results for a prevalence of 0.25, which is approximately the prevalence of breast cancer among women referred for breast biopsy in the U.S.

To illustrate the dependence of the number of true positive results among patients who are test positive by breast biopsy on the prevalence of disease, we calculated positive predictive values over a range of prevalences for different biopsy methods (**Figure 5**). These results suggest that even in low breast cancer prevalence settings (of 5 to 10%) 70 to 80% of women who test positive will truly have breast cancer for all tests except stereotactically guided, vacuum-assisted biopsy. The latter test has a somewhat lower positive predictive value (approximately 60%) in low-prevalence settings, reflecting its lower specificity (compared to other tests). However, as the prevalence increases, the positive predictive value approaches 1 for all tests.

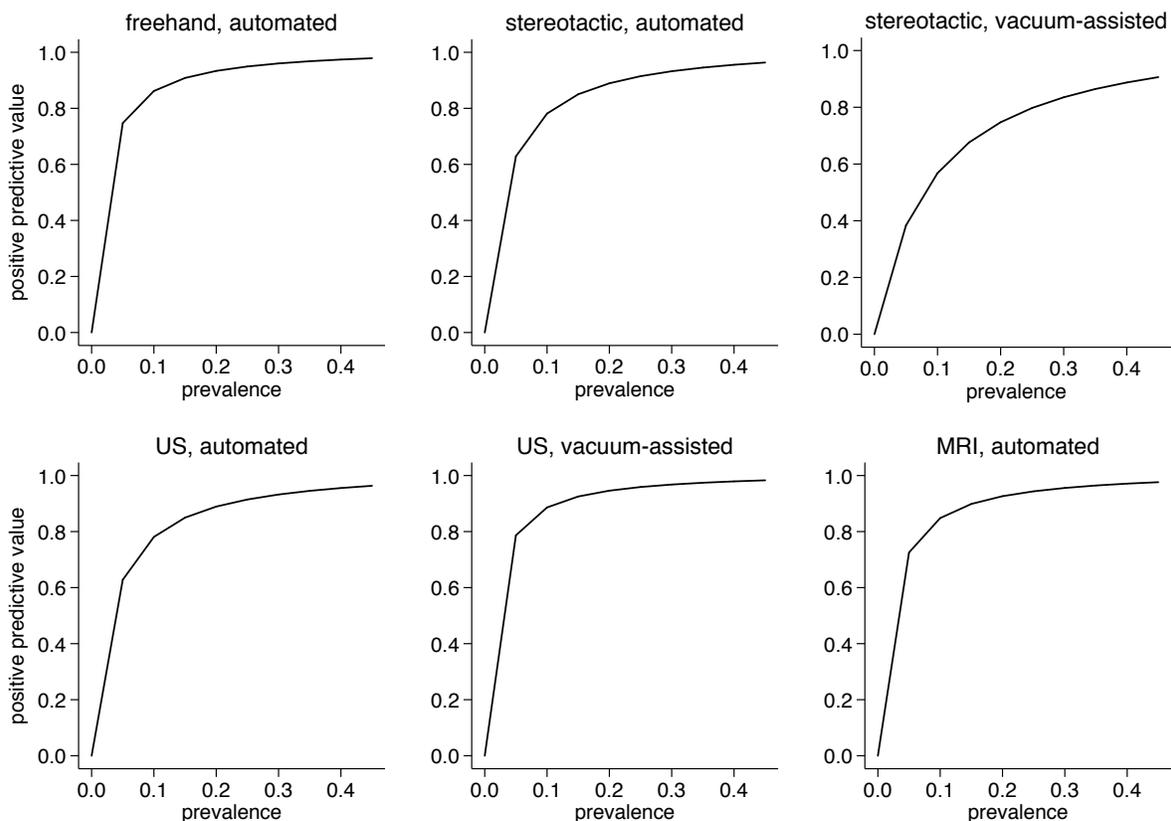
The above comparisons (test outcomes in a hypothetical cohort of known prevalence and positive predictive value calculations) can serve as aids for contextualizing the test performance meta-analysis results presented above. However, they do not reflect the uncertainty around the meta-analytic summary estimates. The following section presents the results of formal (indirect) comparisons among alternative core-needle biopsy methods.

Figure 4: Outcomes of testing in a hypothetical cohort of 1000 women



Different lines represent different test modalities: black solid = stereotactically guided, automated; gray solid = ultrasound-guided, automated; black dashed = stereotactically guided, vacuum-assisted; gray dashed = ultrasound-guided, vacuum-assisted; black dotted = MRI-guided, automated; gray dash-dot = freehand, automated.

Figure 5: Positive predictive value of alternative test methods



MRI = magnetic resonance imaging; US = ultrasound.

Comparative Test Performance

To compare test performance across different biopsy methods, we used indirect (meta-regression-based) comparisons. **Tables 5** and **6** present comparisons between all possible pairs of tests for sensitivity and specificity, respectively. In general, differences among tests were relatively small: for example, differences in sensitivity or specificity never exceeded 0.1 and 95 percent CrIs included the null value (i.e. 0, indicating no difference) in most cases.

One exception to this pattern was the comparative performance of freehand automated biopsy against other techniques. With respect to sensitivity, freehand biopsy had worse performance compared to the other methods, and the credible intervals of the pairwise between-test differences exclude the null value when compared against ultrasound-guided automated device biopsy, stereotactically guided biopsy (both vacuum-assisted and automated), as well as the catch-all category of studies using multiple (mixed) biopsy methods. Another fairly consistent pattern emerged in comparisons of techniques with respect to specificity: stereotactically guided, vacuum-assisted biopsy had lower specificity than all other methods except MRI-guided biopsy (and credible intervals did not include 0). Of note, the above results are based on indirect comparisons (across studies); as such they may be distorted by factors (confounders or effect modifiers) that vary across studies.

Table 5: Differences in sensitivity between pairs of biopsy methods (meta-regression based indirect comparisons)

| | | | | | | | |
|--|--------------------------|----------------------|----------------------------|------------------------------------|--|-----------------------|----------------------|
| | Freehand, automated | | | | | | |
| Ultrasound-guided, automated | 0.08 (0.02, 0.19) | US-guided, automated | | | | | |
| Ultrasound-guided, vacuum-assisted | 0.06 (-0.03, 0.17) | -0.02 (-0.09, 0.01) | US-guided, vacuum-assisted | | | | |
| Stereotactically guided, automated | 0.07 (0.01, 0.18) | -0.01 (-0.03, 0.00) | 0.01 (-0.02, 0.08) | Stereotactically guided, automated | | | |
| Stereotactically guided, vacuum-assisted | 0.08 (0.03, 0.19) | 0.00 (-0.01, 0.02) | 0.02 (-0.00, 0.10) | 0.02 (0.00, 0.03) | Stereotactically guided, vacuum-assisted | | |
| MRI-guided, automated | 0.00 (-0.33, 0.14) | -0.08 (-0.41, 0.00) | -0.06 (-0.38, 0.05) | -0.07 (-0.39, 0.02) | -0.09 (-0.41, 0.00) | MRI-guided, automated | |
| Multiple techniques* | 0.08 (0.02, 0.19) | 0.00 (-0.01, 0.01) | 0.02 (-0.01, 0.09) | 0.01 (-0.00, 0.03) | 0.00 (-0.02, 0.01) | 0.08 (0.00, 0.41) | Multiple techniques* |

* Populations not stratified by biopsy method.

All results are shown as medians of differences (95% CrI). Positive values denote that the method on the left-most column has higher sensitivity than the comparator (on the diagonal). CrIs that do not include the null value (0) are highlighted in bold. CrI = credible interval; MRI = magnetic resonance imaging.

Table 6: Differences in specificity between pairs of biopsy methods (meta-regression based indirect comparisons)

| | | | | | | | |
|--|-----------------------------|-----------------------------|-----------------------------|------------------------------------|--|-----------------------|----------------------|
| | Freehand, automated | | | | | | |
| Ultrasound-guided, automated | -0.01 (-0.04, 0.02) | US-guided, automated | | | | | |
| Ultrasound-guided, vacuum-assisted | 0.00 (-0.02, 0.03) | 0.02 (-0.01, 0.04) | US-guided, vacuum-assisted | | | | |
| Stereotactically guided, automated | -0.01 (-0.03, 0.02) | 0.00 (-0.02, 0.02) | -0.02 (-0.03, 0.01) | Stereotactically guided, automated | | | |
| Stereotactically guided, vacuum-assisted | -0.07 (-0.10, -0.03) | -0.05 (-0.09, -0.02) | -0.07 (-0.10, -0.04) | -0.05 (-0.09, -0.03) | Stereotactically guided, vacuum-assisted | | |
| MRI-guided, automated | 0.00 (-0.08, 0.03) | 0.01 (-0.07, 0.04) | -0.00 (-0.09, 0.02) | 0.01 (-0.07, 0.04) | 0.06 (-0.02, 0.10) | MRI-guided, automated | |
| Multiple techniques* | -0.03 (-0.06, 0.01) | -0.02 (-0.04, 0.01) | -0.03 (-0.06, -0.01) | -0.02 (-0.04, 0.01) | 0.04 (0.01, 0.07) | -0.03 (-0.06, 0.06) | Multiple techniques* |

* Populations not stratified by biopsy method.

All results are shown as medians of differences (95% CrI). Positive values denote that the method on the left-most column has higher specificity than the comparator (on the diagonal). CrIs that do not include the null value (0) are highlighted in bold. CrI = credible interval; MRI = magnetic resonance imaging.

Factors that Affect Test Performance

We considered evidence on the impact of patient or study level factors on test performance from two complementary sources: (1) within-study evidence (i.e. comparisons of test performance over levels of a factor within the patient population enrolled in a study) and (2) evidence from meta-regression analyses (that combine information across studies). Ideally, all studies would consistently report comparisons of test performance across well-defined subgroups (e.g., by patient, or lesion characteristics). Such within-study comparisons are more informative than comparisons across studies: factors related to study setting are common for all patients within the same study and other patient differences can be addressed (at least to some extent) by appropriate analytic methods (e.g., regression adjustment). In the absence of such information, one has to rely on indirect (across-study) comparisons that are generally less convincing because they cannot account for all differences across included populations. Overall, on the basis of both sources of information (within-study analyses and meta-regression analyses), we found that evidence was insufficient to support any specific factor as a modifier of test performance. Detailed results are presented below.

Within-study Evidence

Twenty studies (14 identified by the original review and six included in the current update) of 11,280 patients provided information on factors that affect test performance. Specifically, 16 studies provided information on patient and lesion-related factors, 10 on procedural factors, and three on clinician and facility factors (some studies provided information on multiple factors). The majority of studies (131 of 151) did not allow investigation of the impact of any factors on test performance. The 20 included studies were reported inconsistently and often lacked details necessary for formal statistical assessment of the impact of various factors on test performance. These findings raise concerns regarding selective analysis and outcome reporting with respect to modifiers of test performance. **Table 7** summarizes the findings of individual studies.

Table 7: Studies evaluating factors that may affect test performance (20 studies, reporting on multiple factors each)

| Author, year [PMID] | Biopsy method | Factors evaluated | Key findings |
|---|---|--|---|
| Patient and lesion factors | | | |
| Cusick et al., 1990 ⁷¹ [2183373] | Freehand | Lesion size | Smaller lesions (<2 cm in diameter) were more likely to be misdiagnosed. |
| Barreto et al. 1991 ⁵² [2044776] | Freehand | Lesion size Lesion location Patient age | Tumor size did not affect the accuracy of the procedure. All lesions in the study were > 2 cm in diameter. Lesions in the right breast were more likely to be misdiagnosed. Patient age was not related to accuracy. |
| Makkun et al. 2011 ¹²⁶ [no PMID] | Ultrasound-guided automated device | Lesion type | The accuracy of biopsy in palpable lesions was 100%, while the accuracy of biopsy in nonpalpable lesions was 79.16% |
| Povoski et al. 2011 ¹⁵⁶ [21835024] | Ultrasound-guided automated device | Size of lesion; BI-RADS classification; | There was no difference between the median size of the lesion in cases of false negative biopsies and the median size in all cases biopsied. Among women undergoing interval follow-up after biopsy, the rate of false negatives was 0% for lesions initially classified as BI-RADS 3, 0.6% for lesions classified BI-RADS 4, and 2.8% for lesions classified BI-RADS 5. |
| Wiratkapun et al. 2012 ¹⁸⁷ [22252182] | Ultrasound-guided automated device | Patient age; breast density; lesion type; BI-RADS classification; lesion location | There was no statistically significant relationship between underestimation and patient age, breast density, lesion size, lesion visibility on mammography, lesion type (pure mass vs. mass with calcification), lesion BI-RADS classification (4 vs. 5). Lesions in the lower outer quadrant of the breast were more often underestimated. There was a tendency for younger women with larger mass lesions located at the lower quadrants of the breast and with BI-RADS 5 lesions not seen on mammography to have underestimated lesions. |
| Dahlstrom et al. 1996 ⁷³ [8735717] | Stereotactically guided automated device | Lesion type | There was no difference in the number of cores needed for diagnosis of microcalcifications, densities, or stellate lesions. |
| Koskela et al. 2005 ¹¹³ [16020555] | Stereotactically guided automated device | Lesion type | There were zero false-negatives out of 97 procedures performed on lesions detected as masses on mammography, but 4 false-negatives out of 108 procedures performed on lesions with microcalcifications. |
| Walker et al. 1997 ¹⁸⁰ [no PMID] | Stereotactically guided automated device | Lesion type | The sensitivity of core-needle biopsy was much lower for microcalcifications than for any other type of lesion. |
| Lomoschitz et al. 2004 ¹²³ [15273332] | Stereotactically guided, vacuum- assisted | Lesion type | Biopsies were equally accurate for lesions with microcalcifications and lesions detected as masses on mammography. |
| Pfarl et al. 2002 ¹⁴⁹ [12438044] | Stereotactically guided, vacuum- assisted | Lesion type | Biopsies were equally accurate for lesions with microcalcifications and lesions detected as masses on mammography. |

| | | | |
|--|---|--------------------------|---|
| Reiner et al. 2009 ¹⁵⁸ [19565246] | Stereotactically guided, vacuum- assisted | Lesion type | The agreement rate between core-needle biopsy and surgery was higher in nonpapillary lesions than in papillary lesions, but this difference was not statistically significant. 5 of 6 cases of underestimation occurred in papillary lesions, and the one false negative occurred in a nonpapillary lesion. |
| Venkataraman et al. 2012 ¹⁹⁷ [22127375] | Stereotactically guided, vacuum- assisted | Patient age; lesion size | There was no correlation between patient age and upgrade. However, there was a positive correlation between size of the lesion and upgrade. |
| Abdsaleh et al. 2003 ⁴⁶ [12630998] | Multiple methods | Patient breast density | Technical failures were more likely to occur in women with very dense breast tissue. |
| Ciatto et al. 2007 ⁶⁵ [16823506] | Multiple methods | Lesion type | False negative results were 2.7% for palpable lesions, 2.2% for nonpalpable lesions, 2.3% for masses on mammography, 1.4% for distortions on mammography, and 2.5% for microcalcifications. |
| Cipolla et al. 2006 ⁶⁶ [16473738] | Multiple methods | Lesion type | Correspondence between core-needle biopsy and surgical biopsy results was 100% for palpable lesions but only 88% for nonpalpable lesions. |
| Fajardo et al. 2004 ⁸² [15035520] | Multiple methods | Lesion type | Sensitivity was 97.4% for biopsies of masses detected on mammography and 90.7% for biopsies of nonpalpable lesions and lesions with microcalcifications. |
| Procedural factors | | | |
| de Lucena et al. 2007 ⁷⁴ [17663457]. | Ultrasound-guided automated device | Number of cores | Taking >2 cores did not improve accuracy. Taking >2 cores did not reduce the rate of false negatives. The 6 tumors (out of 101) that were falsely diagnosed as benign by core-needle biopsy would not have been correctly diagnosed even if up to six cores were taken. |
| Fishman et al. 2003 ⁸⁴ [12601206] | Ultrasound-guided automated device | Number of cores | Taking >2 cores improved the accuracy of the biopsy, with 4 cores being the optimal number. 1 case of DCIS would have been missed if fewer than 4 cores had been taken; the other 13 tumors identified in the study would have been correctly diagnosed if only 2 cores had been taken. |
| Kirshenbaum et al. 2003 ¹¹¹ [484822] | Ultrasound-guided automated device | Number of cores | 1 core was diagnostic in 82.6% of cases, 2 cores in 90.5% of cases, 3 cores in 97.9% of cases, 4 cores in 98.9% of cases, and 5 cores in 100% of cases. 100% of malignant lesions were diagnosed after the fourth core. |
| Wiratkapun et al. 2012 ¹⁸⁷ [22252182] | Ultrasound-guided automated device | Number of cores | There was no statistically significant relationship between underestimation and number of biopsy cores. |
| Dahlstrom et al. 1996 ⁷³ [8735717] | Stereotactically guided automated device | Number of cores | One core was diagnostic in 71% of cases, two cores in 84% of cases, three cores in 90% of cases, four cores in 91% of cases, and five cores in 93% of cases. |
| Koskela et al. 2005 ¹¹³ [16020555] | Stereotactically guided automated device | Number of cores | Comment that more than three cores must be taken from lesions before an accurate diagnosis can be made. |
| Lomoschitz et al. 2004 ¹²³ [15273332] | Stereotactically guided, vacuum- assisted | Number of cores | 12 cores were necessary for accurate diagnosis, but taking >12 cores did not improve accuracy. |
| Venkataraman et al. 2012 ¹⁹⁷ [22127375] | Stereotactically guided, vacuum- assisted | Number of cores | There was no correlation between number of cores and upgrade. |

| | | | |
|--|--|---|--|
| Abdsaleh et al. 2003 ⁴⁶ [12630998] | Multiple methods | Number of cores; patient breast density | Taking 2 cores instead of one increased the accuracy of the procedure. Technical failures were more likely to occur in women with very dense breast tissue. |
| Helbich et al. 1997 ⁹⁶ [9169689] | Multiple methods | Patient position | Patients were randomly assigned to undergo stereotactic biopsies in either seated or prone position. The accuracy data were not reported separately for each group, but the authors commented that patient position did not affect the biopsy procedure. |
| <i>Clinician and facility factors</i> | | | |
| Barreto et al. 1991 ⁵² [2044776] | Freehand | Operator experience | Operator inexperience appeared to be related to misdiagnosis. |
| Pfarl et al, 2002 ¹⁴⁹ [12438044] | Stereotactically guided, vacuum-assisted | Operator experience | For 6 of the seven false-negatives, the biopsy had been performed by an operator who had previously performed fewer than 15 stereotactic-guided biopsies. |
| Ciatto et al. 2007 ⁶⁵ [16823506] | Multiple methods | Operator experience | Sensitivity of core-needle biopsies improved as the operators (radiologists) gained experience, from 88% in the first year of the study to 96% in the eighth year of the study. |

BI-RADS = Breast Imaging-Reporting and Data System; PMID = PubMed identification number.

Meta-regression analyses

Meta-regression analyses were possible for the following factors: needle size, choice of reference standard, country where the study was performed, whether multiple centers contributed patients to a study, study design, and risk of bias. All models accounted for the biopsy method used (i.e., imaging guidance method and type of device) and the population’s risk of cancer (average vs. high). **Table 8** summarizes the findings of the meta-regression analyses. The credible intervals for all examined factors included the null value (i.e. lack of difference in sensitivity or specificity), with the exception of increased sensitivity in studies conducted in the U.S. (vs. any other country); higher specificity in studies using followup of 6 or more and 24 or more months (as compared to studies using surgical pathology results for all patients); and higher sensitivity in studies with a prospective design (as compared to studies with a retrospective design). These results must be interpreted with caution, because they reflect indirect comparisons across studies. Furthermore, because these results represent odds ratios for test performance outcomes that are close to 1 (e.g. sensitivity and specificity for all tests were above 0.9), readers should keep in mind that small differences among subgroups of studies can result in (very) large odds ratio values. For example, if the summary sensitivities in two subgroups are 0.99 and 0.98, the relative odds for sensitivity are approximately $2 = (0.99/0.01)/(0.98/0.02)$.

Table 8: Meta-regression analysis for test performance outcomes.

| Modifier category | Potential modifier | Comparison | Relative odds for sensitivity (95% CrI) | Relative odds for specificity (95% CrI) |
|--------------------------------|---------------------------------------|-------------------------------|---|---|
| Biopsy procedure factors | Needle size | 14G vs <14G | 0.44 (0.11, 2.22) | 1.33 (0.51, 3.58) |
| | | >=15G vs. <14 | 0.28 (0.06, 1.41) | 0.39 (0.11, 1.31) |
| | | Unclear/NR vs. <14G | 0.75 (0.23, 2.74) | 2.02 (0.82, 4.34) |
| | Reference standard | 2yrs vs. open biopsy | 1.79 (0.72, 4.08) | 1.61 (0.81, 3.22) |
| | | 6mo vs. open biopsy | 1.60 (0.74, 3.44) | 2.02 (1.06, 3.88) |
| Clinician and facility factors | Country where the study was performed | U.S. vs. other countries | 1.89 (1.03, 3.27) | 1.02 (0.67, 1.60) |
| | Multicenter study | >=1 centers vs. single center | 1.30 (0.47, 3.22) | 1.16 (0.58, 2.24) |
| | Study design | Prospective vs. retrospective | 1.17 (0.60, 2.36) | 1.90 (1.19, 3.07) |
| | | Unclear/NR vs. retrospective | 0.82 (0.44, 1.62) | 2.06 (1.27, 3.23) |
| Study risk of bias | | Intermediate vs. high | 0.61 (0.19, 1.71) | 1.20 (0.63, 2.42) |
| | | Low vs. high | 0.69 (0.20, 1.97) | 1.51 (0.75, 3.58) |

Relative odds for sensitivity compare the odds of a positive test result among patients with cancer over the levels of the modifier. Relative odds for specificity compare the odds of a negative test among patients without cancer over the levels of the modifier. Both metrics are obtained from the bivariate meta-analysis model and are exponentiated coefficients from logistic regression; thus, they can be interpreted as odds ratios. Results were adjusted for biopsy technique and baseline risk of breast cancer (high vs. average). CrI = credible interval; NR = not reported.

Risk Of Bias Assessment for Studies Addressing Key Question 1

Overall, on the basis of 14 items related to risk of bias, we deemed 12 studies to be at low risk of bias, 106 to be at moderate risk of bias, and 33 to be at high risk of bias. Given our relatively strict selection criteria related to study design and completeness of followup, it is not

surprising that the majority of studies reported enrolling consecutive or randomly selected patients (67%), were successful in enrolling 85 percent of all eligible patients (65%), and reported complete data on at least 85 percent of all enrolled patients (70%). However, only 40 percent of studies were judged to be free of spectrum bias, 78 percent were conducted retrospectively (or did not report relevant information), and 83 percent did not apply a “gold” standard reference test on all patients. In most studies (85%), the index test was interpreted by readers blinded to the reference standard test results. However, the vast majority of studies (99%) either did not provide information on whether index test results were available to interpreters of the reference standard or reported that blinding was not used. Finally, information on the incorporation of clinical information in the interpretation of the index and reference standard tests was judged inadequate in the majority of studies (99% for both items).

Question 2: In women with a palpable or non-palpable breast abnormality, what are the adverse events (harms) associated with core-needle breast biopsy compared to the open biopsy technique in the diagnosis of breast cancer?

This section summarizes findings from a total of 135 studies (63 new studies and 72 from the original evidence report) reporting information on at least one of the outcomes relevant to Key Question 2 (2 for open biopsy only, 113 for core-needle biopsy, 20 on the dissemination of cancerous cells during core-needle biopsy). Overall, studies were considered to be of low to moderate risk of bias. Of note, 70 of the 151 core-needle biopsy studies included in Key Question 1 did not provide any information on adverse events, and thus do not allow us to determine whether any adverse events were observed (and not reported). As such, they are uninformative for this Key Question. Further, selective outcome reporting was considered likely for all adverse events examined, because of the large proportion of studies with unclear or missing data.

Adverse Events of Open Biopsy

Very few of the included studies reported information about complications associated with open surgical biopsy. The original evidence report reported findings from a study published in 1993 and a narrative review published in 2007. The study found that 10.2 percent of a series of 425 wire-localized open biopsy procedures were complicated by vasovagal reactions.¹⁹⁸ The narrative review reported that 2 to 10 percent of breast surgeries are complicated by hematoma formation, and that 3.8 percent are complicated by infections.¹⁹⁹ Our update identified three additional studies. One study, which was included in the original report to address Key Question 1, reported that 6.3 percent of open surgical biopsies were complicated by infections.²⁰⁰ A second study, which was also included in the original report to address Key Question 1, reported that 2.1 percent of open biopsy procedures were complicated by the development of an abscess, but none of the 234 ultrasound-guided vacuum-assisted core-needle procedures had abscess development.¹²⁸ Finally, one study reported that four of 100 surgical biopsies required repeat biopsy compared to two of 100 vacuum-assisted core-needle biopsies.¹⁵³

Adverse Events of Core-needle Biopsy

We identified 133 studies reporting information on at least one of the adverse events of interest following core-needle biopsy (63 new studies and 70 from the original evidence report). Of these studies, 20 reported information related to the dissemination of cancerous cells during the biopsy procedure, and 101 allowed for the calculation of event rates for hematomas, bleeding, vasovagal reactions, and infections. **Table 9** summarizes information for the incidence of these adverse events. Overall, their incidence was low: in more than 50 percent of studies reporting information, the percentage of patients experiencing each of the aforementioned outcomes was less than 2 percent; in 75 percent of studies the event rate was less than 1 percent for infections, less than 5 percent for bleeding and vasovagal reactions, and less than 8 percent for hematoma formation. Results for these outcomes, stratified by biopsy technique, are discussed below. Information on less commonly reported adverse events (including seeding) is summarized narratively in the following sections.

Table 9: Adverse events associated with core-needle biopsy for breast cancer diagnosis

| Outcome | Number of studies * | Number of procedures | Median % of procedures where an event was observed (25 th – 75 th percentile) | Minimum-maximum percentage of procedures where an event was observed |
|--------------------|---------------------|----------------------|---|--|
| Hematoma | 51 | 30,058 | 1.19 (0.05-7.20) | [0.00-100.00] |
| Bleeding | 42 | 19,787 | 0.76 (0.22-3.97) | [0.00-100.00] |
| Vasovagal reaction | 35 | 12,449 | 1.78 (0.43-4.29) | [0.00-10.90] |
| Infection | 35 | 23,522 | 0.05 (0.00-0.63) | [0.00-2.91] |

*Number of studies providing information on the outcome
max = maximum; min = minimum; perc. = percentile.

Hematomas and Bleeding

Fifty-one studies including 30,058 core-needle biopsy procedures reported information on hematoma formation. In 50 percent of these studies the event rate for hematomas was less than 1.2 percent, and in 75 percent the event rate was less than 8 percent. The highest rates of hematoma formation were observed in studies of vacuum-assisted procedures. For example, in 75 percent of studies of ultrasound-guided vacuum-assisted procedures, the event rate for hematomas was 16.4 percent or greater, while no hematomas were reported in two studies of ultrasound-guided biopsies without vacuum assistance. The median hematoma event rate for studies of stereotactic-guided vacuum-assisted biopsies was 1.55 percent, whereas the maximum event rate in studies of stereotactic-guided biopsies without vacuum-assistance was 1.25 percent. Due to incomplete (and potentially selective) reporting, these percentages should be interpreted with caution; however, vacuum-assisted procedures do appear to have a higher rate of hematoma formation than other core-needle biopsy methods. One study of ultrasound-guided vacuum-assisted biopsy identified for this update reported that in 1183 procedures one hematoma required surgical intervention.²⁰¹ The event rate of 0.085 percent reported in this individual study is the same as the event rate for hematomas requiring treatment calculated across 24 studies included in the original evidence report. No other newly identified studies reported information on the number of hematomas requiring treatment.

Forty-two studies of 19,787 core-needle biopsy procedures reported information on bleeding. In 50 percent of these studies the event rate for bleeding was less than 0.76 percent, and in 75 percent the event rate was less than 4 percent. In 25 percent of studies of stereotactic-guided vacuum-assisted procedures the event rate for bleeding was 3.75 percent or greater, while the maximum event rate reported in studies of stereotactic-guided biopsies without vacuum assistance was 1.29 percent. The highest event rate in studies of ultrasound-guided vacuum-assisted biopsy was just under 8 percent, while the single study of ultrasound-guided biopsy without vacuum assistance that contained information on bleeding reported an event rate of 5.26 percent. With the same caveats as for hematoma formation, vacuum-assisted procedures appeared to be associated with bleeding more often than non-vacuum-assisted procedures. Overall, bleeding was a rare complication. In addition to the studies reporting bleeding, we identified one study in which 19 percent of 1177 patients undergoing ultrasound guided vacuum-assisted biopsy were diagnosed with skin ecchymosis without hematoma.²⁰¹ One study of stereotactic-guided vacuum-assisted biopsy identified in our updated searches reported that of 485 women biopsied, one patient was observed in the hospital for one day due to persistent bleeding. The event rate of 0.21 percent in this study is consistent with the 0.34 percent of vacuum-assisted procedures reported in the previous report to be complicated by bleeding that

required treatment. No other newly identified studies reported information on bleeding events that required treatment.

Nine studies of various core needle techniques that were included in the original report specified that bruising occurred after core-needle biopsy procedures. Three of the nine reported that bruising was a common event, two reported that approximately 50 percent of patients had bruising, and four studies reported that 45 out of 976 patients (4.6%) had severe bruising. We identified one additional study of stereotactic-guided biopsy without vacuum assistance that reported information on bruising. The study noted that 1.2 percent of 200 patients reported tenderness, swelling or bruising at the biopsy site following the biopsy.⁷³ No other newly identified studies reported information on bruising.

Table 10 summarizes information on hematomas and bleeding stratified by biopsy technique.

Table 10: Core-needle biopsy procedures and rates of hematoma formation and bleeding

| Outcome | Biopsy technique | N studies (among 101 studies reporting adverse events) | Number of procedures | Median % of procedures where an event was observed (25 th – 75 th percentile) | Minimum-maximum percentage of procedures where an event was observed |
|--------------------|-------------------------------|--|----------------------|---|--|
| Hematoma formation | all devices | 51 | 30,058 | 1.19 (0.05-7.20) | 0.00-100.00 |
| | freehand, automated | 2 | 1487 | NA | 0.00-0.00 |
| | US, automated | 2 | 598 | NA | 0.00-0.00 |
| | US, vacuum-assisted | 5 | 1480 | 20.00 (16.40-26.47) | 0.99-36.27 |
| | stereotactic, automated | 5 | 1706 | 0.97 (0.77-1.00) | 0.00-1.25 |
| | stereotactic, vacuum-assisted | 21 | 12,145 | 1.55 (0.69-6.06) | 0.00-100.00 |
| | MRI, automated | 2 | 116 | NA | 1.33-4.88 |
| | other | 14 | 12,526 | 0.84 (0.00-7.20) | 0.00-79.12 |
| Bleeding | all devices | 42 | 19787 | 0.76 (0.22-3.97) | 0.00-100.00 |
| | freehand, automated | 3 | 1732 | NA | 0.14-3.97 |
| | US, automated | 1 | 190 | NA | 5.26 |
| | US, vacuum-assisted | 6 | 555 | 2.04 (0.70-4.95) | 0.00-7.84 |
| | stereotactic, automated | 4 | 951 | 0.24 (0.00-0.89) | 0.00-1.29 |
| | stereotactic, vacuum-assisted | 17 | 13,419 | 0.75 (0.30-3.75) | 0.14-26.94 |
| | MRI, automated | 0 | 0 | NA | NA |
| | other | 11 | 2940 | 1.07 (0.00-4.04) | 0.00-100.00 |

We only report the minimum and maximum percentage of events when fewer than three studies were available for a biopsy technique. When a single study reported information we simply list the percentage of procedures associated with complications. MRI = magnetic resonance imaging; NA = not applicable; US = ultrasound.

Infections

Across 35 studies, including 23,522 core-needle procedures, the median percentage of infectious complications was 0.05 percent. One study, identified by the original evidence report, reported that a patient developed an abscess that required surgical treatment in a series of 268 stereotactically guided vacuum-assisted procedures.¹⁷¹ We did not identify any new studies reporting information on the occurrence of abscesses. **Table 11** summarizes information on infections, stratified by biopsy technique.

Table 11: Core-needle biopsy procedures and rates of infectious complications

| Biopsy technique | N studies (among 101 studies reporting adverse events) | Number of procedures | Median % of procedures where an event was observed (25 th – 75 th percentile) | Minimum-maximum percentage of procedures where an event was observed |
|-------------------------------|--|----------------------|---|--|
| All devices | 35 | 23,522 | 0.05 (0.00-0.63) | 0.00-2.91 |
| Freehand, automated | 3 | 1637 | NA | 0.00-2.00 |
| US, automated | 4 | 1675 | 0.05 (0.00-0.92) | 0.00-1.74 |
| US, vacuum-assisted | 2 | 171 | NA | 0.00-1.98 |
| Stereotactic, automated | 9 | 2128 | 0.00 (0.00-0.66) | 0.00-2.91 |
| Stereotactic, vacuum-assisted | 8 | 4739 | 0.09 (0.00-0.30) | 0.00-0.89 |
| MRI, automated | 0 | 0 | NA | NA |
| Other techniques | 9 | 13,172 | 0.05 (0.00-0.15) | 0.00-2.20 |

We only report the minimum and maximum percentage of events when fewer than three studies were available for a biopsy technique. MRI = magnetic resonance imaging; NA = not applicable; US = ultrasound.

Pain and Use of Pain Medications

The original report identified three vacuum-assisted biopsy procedures reported to have been terminated after patients complained of severe pain, and we identified one study of 4086 stereotactic-guided vacuum-assisted procedures in which four biopsies were suspended due to pain.²⁰² No studies reported procedure termination due to patient complaints of pain in any other types of biopsy procedures. Twenty-five studies of a wide variety of biopsy methods reported information about patient pain during the procedure (pain was assessed heterogeneously and for that reason we did not calculate overall event rates).

Eleven studies reported information on the use of pain medications. One of these studies reported that 100 percent of patients were sent home with narcotics after an open biopsy procedure, and only three patients required narcotics after a core-needle procedure.¹⁰⁶ Twenty patients were reported to have required acetaminophen after a core-needle procedure.¹²⁸ Note that being sent home with a medication may not necessarily mean the patients required or used the medication.

Vasovagal Reactions

Thirty-five studies with 12,449 procedures reported information about the occurrence of vasovagal reactions (fainting or near-fainting) during core-needle biopsy. The median event rate in these studies was 1.78 percent, although one study reported an event rate of nearly 11 percent. More than 40 percent of the vasovagal reactions occurred in patients who were reported to have been positioned sitting upright for the biopsy procedure (many of the studies did not report patient position so the other 60 percent of vasovagal reactions could have occurred in patients positioned in a variety of positions, or could have occurred primarily in seated patients).

Table 12 summarizes information on vasovagal reactions, stratified by biopsy technique.

Table 12: Core-needle biopsy procedures and rates of vasovagal reactions

| Biopsy technique | N studies (among 101 studies reporting adverse events) | Number of procedures | Median % of procedures where an event was observed (25 th – 75 th percentile) | Minimum-maximum percentage of procedures where an event was observed |
|------------------|--|----------------------|---|--|
|------------------|--|----------------------|---|--|

| | | | | |
|-------------------------------|----|--------|------------------|------------|
| All devices | 35 | 12,449 | 1.78 (0.43-4.29) | 0.00-10.90 |
| Freehand, automated | 1 | 1431 | NA | 0.00 |
| US, automated | 2 | 235 | NA | 0.53-8.89 |
| US, vacuum-assisted | 2 | 766 | NA | 0.43-1.43 |
| Stereotactic, automated | 11 | 1785 | 1.94 (0.50-4.29) | 0.00-8.33 |
| Stereotactic, vacuum-assisted | 12 | 5779 | 2.89 (0.35-6.11) | 0.17-10.90 |
| MRI, automated | 0 | 0 | NA | NA |
| Other techniques | 7 | 2453 | 1.78 (0.99-2.20) | 0.00-3.47 |

We only report the minimum and maximum percentage of events when fewer than three studies were available for a biopsy technique. When a single study reported information we simply list the percentage of procedures associated with complications. MRI = magnetic resonance imaging; NA = not applicable; US = ultrasound.

Impact of Biopsy Procedure on Usual Activities and Time to Recovery

Three studies provided information on the impact of biopsy procedures on usual activities. The first study¹²⁸ reported that of 34 women undergoing ultrasound-guided vacuum-assisted breast biopsy, 16 (47%) women stated that the procedure did not interfere with usual activity, 14 (41%) stated that there was minor interference, and four (12%) felt that there was mild interference. The second study²⁰³ reported four cases in which the patient felt constrained in her daily life due to the procedure. The third study¹⁵³ reported vacuum-assisted biopsy results in less psychological/physical stress when compared to surgical procedures.

A single study provided information regarding time to recovery, measured by asking patients how long it had taken for them to return to their normal activities after the biopsy procedure.⁸⁵ This study reported that the average time of recovery was 3.5 days for open biopsy procedures and 1.5 days for stereotactically guided automated gun core-needle biopsy procedures.

Impact of Biopsy Procedure on Subsequent Mammographic Procedures

Five studies reported information about the impact of core-needle biopsies on subsequent mammographic examinations. Three studies reported on stereotactic-guided vacuum-assisted core-needle procedures. These studies enrolled 3,748 patients, of whom 3,345 (89.2%) were reported to have no mammographically visible scarring after the biopsy procedure. Only seven of the patients were reported to have scars that were potentially diagnostically confusing on subsequent mammographic procedures. In the fourth study, 91 patients underwent stereotactic- or ultrasound-guided vacuum-assisted core-needle biopsy. The researchers reported that at 6-month followup there was no evidence of scarring, architectural distortion, alterations of the skin, fat necrosis, or other changes that are frequently observed after surgical breast biopsy.²⁰³ In the fifth study, patients underwent mammography at 6 or 12 months, and the authors reported that mammograms showed structural distortions at the biopsy site in the 100 women who underwent surgical biopsy, and no sequelae in the 100 women who received vacuum-assisted core-needle biopsy.¹⁵³

Miscellaneous Reported Adverse Events

The original report identified eight studies with information on pneumothorax, seizures, vomiting, or acute inflammation, and we identified one additional study reporting vomiting and one additional study reporting inflammation. Four studies of 2,600 patients reported that four cases of pneumothorax, none of which required treatment, had occurred. None of these four studies used the same core-needle biopsy method. Two studies reported that one patient per study (out of 3,487 patients in total) had suffered a seizure during a stereotactic-guided vacuum-assisted procedure. One study of 268 patients undergoing stereotactic-guided vacuum-assisted biopsies reported that three patients developed acute inflammation at the biopsy site after the procedure. One study of 485 women undergoing stereotactic-guided vacuum-assisted biopsies reported that two patients developed signs of inflammation judged to be mastitis. Two studies reported that a patient vomited during the procedure; one of these studies was of 185 stereotactic-guided vacuum-assisted procedures and the second was of 236 vacuum-assisted procedures using either stereotactic or ultrasound guidance. We did not identify any new studies reporting any other significant adverse events associated with core biopsy procedures.

Dissemination of Cancerous Cells During the Biopsy Procedure

To address the potential dissemination of cancerous cells by breast biopsy we did not use the study-design evaluation criteria for Key Questions 1 and 2; instead, we considered any clinical study that addressed the topic (including case reports and case series). Full details of the included studies are available in SRDR.

We reviewed 14 studies that used histopathology to demonstrate dissemination of cells by core-needle biopsy procedures (four new studies and 10 studies included in the original report). Nine studies had a cohort design, and five were case series or case reports.

The percentage of needle tracks previously reported to contain displaced cancerous cells ranged from 0 to 65 percent. We identified a cohort study that reported that the percentage of ultrasound-guided biopsies with cancerous cells in the needle wash material ranged from 33 percent to 69 percent.²⁰⁴ The original report observed that the risk of finding displaced cancerous cells was increased by greater duration of the biopsy procedure,²⁰⁵ multiple needle passes,²⁰⁶ and a short interval between core-needle biopsy and surgical excision,²⁰⁷ while the risk was decreased by diagnosis of invasive lobular carcinoma²⁰⁶ and the use of vacuum-assisted core-needle biopsy.²⁰⁷ The incidence of positive cytological findings in needle wash material was also greater with multiple needle passes and automated device (versus vacuum-assisted) biopsy.²⁰⁴

Although the clinical significance of these displaced cancerous cells is debated,²⁰⁷ we found four case reports of patients developing tumors at the site of prior core-needle biopsies, which supplement the six case reports previously identified for this review.²⁰⁸⁻²¹² Four of these ten women were reported to have not received radiation therapy for the primary tumor; for the other six women it was not reported whether they had received radiation therapy.

The previous evidence report found four studies with 1,879 women that explored the risk of tumor recurrence following biopsy.²¹³⁻²¹⁶ Three of these four studies reported that women who did not have a preoperative needle biopsy had a higher rate of tumor recurrence than women who did receive a preoperative needle biopsy;²¹³⁻²¹⁵ the fourth study reported the opposite. We identified an additional cohort study, published in 2011, that reported no development of tumors along the needle track among more than a thousand women receiving a core-needle biopsy

diagnosis of cancer in early 2008 through 2009.²¹⁷ The majority of the women in the original four studies were treated with breast-conserving surgery and radiation therapy; the newly identified fifth study did not report whether women received radiation therapy.

The original evidence report found three studies with 3,103 women that investigated the risk of seeding the lymph nodes with cancerous cells after biopsy procedures.²¹⁸⁻²²⁰ Two of the three studies reported that the method of biopsy did not affect the rate of positive sentinel lymph nodes; the third study reported that the rate of metastases to the sentinel lymph node was higher in women who underwent some form of preoperative biopsy. We found two new studies examining the topic of epithelial displacement into lymph nodes after biopsy. One study described 15 cases of epithelial cell displaced into the lymph node subcapsular sinus in a series of axillary lymph node dissections taken approximately 2 weeks after either core-needle or open breast biopsy.²²¹ The authors stated that this was probably the result of mechanical transport of cells during biopsy and that the clinical implications are likely not significant. The second study examined epithelial displacement into lymphovascular spaces in the breast core needle biopsy specimens of seven women who were diagnosed with pure DCIS after core-needle biopsy and surgical excision.²²² These women did not have recurrences or metastases after 24 to 84 months followup. The authors suggest that because this epithelial displacement is seen in the initial core biopsy sample, the presence of tumor cell clusters in lymphovascular spaces may not reflect lymphovascular invasion.²²³

The original evidence report identified a case series report of 25 cases of false-positive sentinel lymph nodes, in which the false-positives appeared to be caused by displacement of benign epithelial cells during a biopsy procedure.²²³ Twelve of the false-positive cases had undergone core-needle biopsy prior to the sentinel lymph procedure, 12 had undergone wire-localization open biopsy, and one had undergone a fine-needle aspiration procedure. Findings of false-positive sentinel lymph nodes are clinically important because the findings are likely to lead to adverse events from unnecessary treatment. Because 22 of the 25 cases had intraductal papilloma at the biopsy site, the authors of the case series report suggested using caution when interpreting sentinel lymph node histopathology in cases where intraductal papilloma was noted during the initial biopsy procedure.

Factors that modify the association of biopsy procedures with adverse events

Due to the small number of studies providing information on any of the factors of interest and the poor reporting of adverse events across studies, we believe that the evidence is insufficient to establish any specific factor (other than patient positioning for vasovagal events and the use of vacuum for bleeding, as discussed in preceding sections) as a determinant of the rate of adverse events among women undergoing biopsy for breast cancer diagnosis. Information extracted from individual studies is summarized in **Table 13**.

Table 13: Studies evaluating factors that may affect the incidence of adverse events

| Author, year [PMID] | Biopsy technique | Factors evaluated | Key findings |
|---|-----------------------------------|-------------------|---|
| <i>Patient and lesion factors</i> | | | |
| Lin et al., 2000 ¹²² [no PMID] | Ultrasound guided vacuum-assisted | Breast density | Among 8 women with hematomas and pre-biopsy mammograms, 75% had breasts classified as dense. No patients with breasts classified as fatty developed |

| | | | |
|---|--|---|---|
| | | | hematomas. |
| Wang et al., 2012 ¹⁸¹ [21300503] | Ultrasound guided vacuum-assisted | Lesion size | No statistically significant difference was observed in mean lesion size for cases with and without hematoma. |
| Zografos et al. 2008 ¹⁹⁵ [18814132] | Stereotactic-guided vacuum-assisted | BI-RADS classification, patient age | There was no statistically significant association between hematoma formation and BI-RADS classification or patient age. |
| Frank et al., 2007 ²²⁴ [17661855] | Stereotactic-guided automated gun | Patient age | Pain was not associated with patient age (p=0.11). |
| Procedural factors | | | |
| McMahon et al. 1992 ¹³⁰ [1422715] | Freehand | Needle size | 18G core-needle procedure were associated with significantly less pain than 14G core-needle procedures, but there was no significant difference in pain between 14G and 16G procedures. |
| Wong and Hisham 2003 ¹⁸⁸ [484085] | Freehand | Needle size | No difference in the amount of pain experienced by patients undergoing a 14G core-needle procedure vs. a 16G core-needle procedure. |
| Zagouri et al., 2011 ²²⁵ [21709018] | Stereotactic-guided vacuum-assisted | Number of cores | In women who underwent additional sampling (96 cores vs. the standard 24-36), the rate of clinically significant hematomas doubled from 3.5% to 7.5%. |
| Frank et al., 2007 ²²⁴ [17661855] | Stereotactic-guided automated gun | Number of cores, duration of procedure | Pain was associated with the number of biopsy cores (p=0.032) and the duration of the procedure (p=0.046). |
| Schaefer et al., 2012 ²²⁶ [22381441] | Multiple methods | Size of needle; biopsy device | There were significantly higher rates of bleeding (p<0.001) and hematoma (p=0.029) in the Mammotome 8G than in the Mammotome 11G group. There were no significant differences in bleeding rates (p=0.799) or hematoma rates (p=0.596) between the ATEC 12G and the ATEC 9G group. There were no significant differences in bleeding or hematoma rates in the Mammotome 8G group and the ATEC 9G group, but there was less bleeding (p=0.015) and fewer hematomas (p=0.001) in the Mammotome 11G group than in the ATEC 12G group. |
| Seror et al., 2012 ²²⁷ [21310570] | Multiple methods | Size of needle/probe | There was no difference in pain with different probe sizes (12 mm, 15 mm, and 20 mm). |
| Szynglarewicz et al., 2011 ²²⁸ [21367573] | Multiple methods | Vacuum-assistance; biopsy device | Biopsy with an automated device was significantly more painful than biopsy with a vacuum-assisted hand-held device (p<0.01). |
| Clinician and facility factors | | | |
| Kirshenbaum et al., 2003 ¹¹¹ [12876040] | Multiple methods | Operator experience | The majority of vasovagal reactions occurred when inexperienced operators performed the biopsy procedures. |

PMID = PubMed identification number.

Key Question 3: How do open biopsy and various core-needle techniques differ in terms of patient preference, availability, costs, availability of qualified pathologist interpretations, and other factors that may influence choice of a particular technique?

We identified 41 new studies that addressed various aspects of KQ3. Together with the 86 studies included in the original evidence report, this section synthesizes evidence from 127 studies. Generally, our findings confirmed those of the original evidence report. In the following sub-sections, we first discuss aspects of diagnostic biopsy important to patients, followed by economic factors that may influence the choice of a particular technique, and then proceed to summarize information on other factors, including the availability of equipment, procedure duration time, time to complete tumor removal, wait time for test results, and recurrence rates. Because of the nature of this Key Question and the heterogeneity of the sources of information used to address each outcome of interest, we did not attempt to grade the strength of evidence for most outcomes considered for this Key Question (this is consistent with the original evidence report).

Anxiety and Distress

We identified 11 studies that looked at levels of anxiety and distress related to biopsy procedures. This outcome was not specifically examined in the original report, and we base our conclusions on the studies retrieved for this update. Overall, patients reported increased levels of anxiety and distress immediately before or during the procedure, and these levels were reduced after the procedure. One study reported mean anxiety levels just before the procedure to be well above normal on State Trait Anxiety Inventory (STAI) (mean 48; normal=35.9), Impact of Event Scale (mean 26; normal < 8.5), Center for Epidemiological Studies-Depression Scale (mean 16; normal 8), and *Perceived Stress Scale* (mean 19; normal 12.6).²²⁹ This was corroborated by a second study that reported participants prebiopsy STAI-S and STAI-T T scores were two standard deviations higher than the mean T score (T-score mean 50, SD 10).²³⁰ Yet another study reported that one procedure out of 602 could not be completed because of patient anxiety.¹²⁴ One study found greater anxiety in surgical biopsy patients than in those receiving core-needle vacuum-assisted biopsies.¹⁵³

Four studies, three of which were randomized controlled trials, looked at a range of options to ameliorate stress during core-needle biopsy procedures, with relaxation, medication, empathy, and hypnosis all showing reductions in stress either just before or during the procedure. One randomized controlled trial reported on stress levels in three groups of patients (those receiving usual care, relaxation, or medication to reduce anxiety). All three groups had preprocedural state anxiety levels that were significantly higher than normal and reported significant reductions in anxiety 24 hours after the procedure. Patients in the medication group reported significantly less anxiety during the procedure, when compared with the usual care and relaxation groups.²³¹ They also reported that there was no statistically significant difference in anxiety levels during the procedure for those who underwent stereotactically guided versus ultrasound-guided procedures.²³¹ A second randomized controlled trial looked at the use of empathy and hypnosis in relieving anxiety. The authors found that standard care patients experienced an increase in anxiety during the procedure, patients who were given empathy experienced no change in anxiety during the procedure, and patients receiving hypnosis experienced a decrease in anxiety during the procedure.²³² A final randomized controlled trial

reported that the main effect of an education intervention on anxiety was that those in the control group tended to have lower postconsultation anxiety than those in the education group.²³³

Procedure Preference

We found two studies that specifically addressed procedure preference^{234, 235} in addition to the 20 reported in the original evidence report. Both of the new studies reported a positive experience with core-needle biopsy, relative to surgical biopsy. One study reported that women who had previously experienced only core-needle or surgical biopsy were willing to wait a median of 3.2 weeks longer to avoid surgical than to avoid core-needle biopsy; while women who had experienced both were willing to wait 2.4 weeks longer to avoid surgical than to avoid core-needle biopsy.²³⁴ This supports the findings of the original report: the majority of studies reported core-needle biopsies to be preferable to open biopsies. However, a single study reported the reverse: a survey of 59 patients (20 open biopsy, 20 fine needle aspiration, and 19 core needle biopsy) from Detroit, Michigan in 1997 and 1998 found that 90 percent were satisfied with their open surgical biopsy compared to only 80 percent satisfied with a vacuum-assisted core-needle biopsy, though the authors reported that this difference was not statistically significant at the $p=0.05$ level.²³⁶ The original evidence report also noted that the majority of the studies reported such information as that the patients tolerated the procedure well or would recommend it to others in the future. One study reported that 99 percent of image-guided core-needle biopsy patients rated their overall experience as positive and 97 percent reported they would recommend the center to a family member or friend if they needed a biopsy.²³⁵ Another study reported that patients preferred the decubitus position to the prone position.²³⁷ Two studies reported that vacuum-assisted procedures were more comfortable than other types of core-needle biopsies.^{238, 239} Two other studies reported that patients lost less time to core-needle procedures than to open procedures.^{240, 241}

Surgical Procedures Avoided

We identified 10 new studies providing information on the number of surgical procedures avoided by the use of core needle biopsy methods for breast cancer diagnosis. Including the 31 studies considered by the original report, a total of 41 studies provide information on this outcome. In general, studies found that core-needle biopsy obviated the need for surgery for a substantial proportion of women, ranging from 29 to 87 percent. Of the 41 studies, nine reported comparisons against open surgical biopsy with respect to the number of patients requiring only one surgical procedure (vs. more than one). Meta-analysis of these studies suggested that the odds of requiring only one surgical procedure were more than 13 times higher among women receiving core-needle biopsy; odds ratio = 13.4 (95% CrI, 5.6 to 43.4). This result should be interpreted with caution because of the possibility of confounding by indication. Women may have been selected for a specific diagnostic approach on the basis of clinical or other factors, which may also be associated with the need for additional surgical interventions.

Cosmetic Results

We identified two new studies that addressed cosmetic results with core-needle or open biopsy. Both reported minimal scars that were acceptable to the patients. The original evidence report identified 10 other studies that included information on cosmetic results for vacuum-assisted core-needle biopsy, which reported that, overall, patients were satisfied with the

cosmetic results. Only one of the 10 studies included in the original report compared a group of patients undergoing core-needle biopsy to a group of patients undergoing open biopsy.²³⁶ This study reported a greater satisfaction with appearance of the breast 2 years after surgery in core-needle patients (95 percent very satisfied) than in open biopsy patients (25 percent very satisfied).²³⁶

Resource utilization and costs

We found two additional studies on relative costs of core-needle biopsy. The results below reflect a total of eight studies, including six studies identified in the original report. The original report concluded that the costs of surgical biopsy are considerably greater than those of core-needle biopsy. In this update we identified one study (2008) reporting average charges for core-needle biopsy at \$10,500 and excision biopsy at \$11,500.²⁴² The authors based their costs on the calculation of mean patient charges for initial diagnostic procedure and subsequent necessary surgeries, which were compared for patients undergoing biopsy for BI-RADS-5 lesions between 1998 and 2002. The authors recommend core needle biopsy as the initial diagnostic approach for highly suspicious lesions, based upon improved pathologic margins and fewer surgical procedures rather than significant costs savings.

Another study compared per-procedure costs of core-needle biopsy and fine needle biopsy. Based on reimbursements for facility fees, but excluding professional fees, the costs were \$477.92 versus \$166.34, respectively.²⁴³

The original evidence report reported on the relative costs of open surgical biopsy and various core-needle biopsy techniques in six studies. The studies reviewed factors, including purchase price of devices, personnel time and costs, the costs of processing and analyzing samples, patient volume, whether the device is used as a complementary procedure, and what mammography results determine the use of a core-needle biopsy technique. The original report also noted that MRI-guidance is the most expensive method of performing core-needle biopsies²⁴⁴ We did not find any new studies comparing the costs or cost-effectiveness of different core-needle or imaging techniques.

We did not identify any new studies for this domain. The two studies discussed in the original evidence report reported that vacuum-assisted procedures and procedures that required dedicated prone tables required more physician and room time.

Physician Experience

We identified three new studies²⁴⁵⁻²⁴⁸ which, together with the 10 studies included in the original report, support the conclusions that greater experience with particular devices improves accuracy, shortens procedure duration times, and leads to a decrease in the number of open biopsies. One study reported a trend that indicated that in a training program, the fellows were able to establish an accurate diagnosis with fewer core biopsy samples in their later cases (i.e. as the training progressed and they gained experience).²⁴⁶ A second study introduced a training program for breast lesion excision system biopsy, for which they reported that fellows who had previous experience in vacuum-assisted biopsy could perform the new procedure after four procedures (median), while those without previous exposure showed proficiency after nine procedures (median). This was compared to the 12 procedures required for a new user to become proficient with vacuum-assisted biopsy.²⁴⁸ A survey of 79 fellows who had graduated from

approved breast fellowships between 2005 and 2009 reported that many physicians feel poorly prepared to do ultrasound-guided (41 poorly prepared; 16 moderately prepared; 22 well prepared) or stereotactic (57 poorly prepared; 7 moderately prepared; 15 well prepared) core-needle biopsies.²⁴⁵

Availability of a Qualified Pathologist

We did not identify any new studies for this outcome. The two studies included in the original report showed conflicting results, with one reporting that whether the specimen was read by a local or central pathologist had little effect as agreement rates were very high,²⁴⁹ and the second reporting that the pathologist's lack of experience with the TruCut device explains its poor performance.²⁵⁰

Availability of Equipment/Utilization

The original report identified three studies reporting on the impact of equipment availability and utilization, to which we added four more for a total of seven. The original report concluded that wait times are longer for open procedures and dedicated prone biopsy tables. We found a randomized controlled trial that reported that patients who waited 4 days or more for a core-needle biopsy procedure were less satisfied than patients who waited 3 days or less ($p=0.007$).²³³ We did not find any new studies reporting the overall wait times for core-needle biopsies, or comparing wait times for core-needle vs. open biopsy procedures. Other studies looked at utilization rates of core-needle biopsies over time. One study reported that the non-operative diagnosis rates in core-needle biopsy had increased from 49 percent in 1995/96 to 87 percent in 2000/01 to 94 percent in 2005/06.²⁵¹ A second study reported that with a stable total patient population and constant number of open and needle-localized procedures, stereotactic breast biopsies had increased from 56 in 1995 to 68 in 1996, 118 in 1997, and 172 in 1998.²⁵² They further reported that diagnostic yield had increased in the stereotactic era.²⁵² A third study reported a similar increase in core-needle biopsy utilization between January 1992 and March 1998, with a corresponding decrease in open biopsies.²⁵³

Procedure Duration Time

We identified an additional 11 studies that reported results for procedure duration across various types of biopsy. When these studies are added to the 40 studies identified in the original evidence report, reported procedure times range between 3 and 128 minutes. This large range is probably the result of different definitions for procedure time. For example, one study reported times for “total procedure” (from signing of informed consent to end of preparation for next patient) as 26.7 minutes for ultrasound guided core biopsy and 47.5 minutes for stereotactic core biopsy; “room time” (from signing of informed consent to end of procedure) as 23.1 minutes for ultrasound guided core biopsy and 36.5 minutes for stereotactic core biopsy; and “physician time” (time radiologist located lesion to time enough samples had been obtained) as 12.3 minutes for ultrasound guided core biopsy and 18.6 minutes for stereotactic core biopsy.²²⁸

Mean procedure times for ultrasound-guided core-needle biopsies ranged from 3 to 60 minutes, while stereotactically guided core-needle procedures tended to take longer, with mean procedure times ranging from 10 to 100 minutes. Mean times for MRI guidance ranged from 8 to 70 minutes, with only one new study reporting mean times for MRI-guided procedures. The authors of that study reported a mean of 12 minutes and a range of 8 to 23 minutes.⁸³

Vacuum-assisted core biopsies had a reported mean or median duration of 3 to 70 minutes. We found no new results for open biopsy, which had a reported mean of 40 to 45 minutes, based on two studies included in the original evidence report. Again, these mean ranges may be artificially wide due to differences in definitions of procedure time across studies.

Time to Complete Tumor Removal

We identified seven studies that reported results for time in days from biopsy to surgery for tumor removal. There were no studies addressing this specific outcome reported in the original report. Overall times from biopsy to tumor removal ranged from 5 to 153 days. One study directly compared wait times for core-needle and surgical biopsies, reporting an average time from initial procedure to final surgical procedure for core-needle biopsy as 27 days and excisional biopsy as 22 days.²⁴² The rest of the studies gave results for core-needle biopsy only, with means ranging from 14 to 62 days and medians ranging from 11 to 83 days.

Wait Time for Test Results

We found five studies that discussed wait times for core-needle biopsy results.^{235, 254-258} There were two studies included in the original report, for a total of seven studies addressing this outcome. Overall, core-needle wait times ranged from 1 to 114 days, with most reported as between 1 and 1.3 days. The two studies in the original report that compared wait times after core-needle and open biopsies showed that wait times for core-needle biopsy results are shorter by an average of 7 to 10 days. One study reported that using a microwave processor to reduce wait times for test results reduced the average wait for results ($P < 0.001$).²⁵⁵ Another study assessed patient satisfaction with wait times and found that most participants (88 percent) thought the wait for test results (usually the day after the biopsy by phone) was reasonable.²³⁵

Recurrence Rates

We found five studies that discussed recurrence rates among core-needle biopsy patients. There were no studies addressing this specific outcome in the original report. One study reported no recurrence,²²⁷ one reported a single lesion recurring within 6 months (in 86 patients),²⁵⁹ one reported three cases of recurrent malignant lesions (in 420 patients),²⁶⁰ one reported two malignant lesions recurring in 405 patients,²⁶¹ and the last reported 9 lesions with mammographic progression requiring further intervention (in 270 patients).²⁵² A total of 15 recurrences were reported among 1344 patients.

Discussion

Key Findings and Assessment of the Strength of Evidence

In this update of the 2009 Comparative Effectiveness Review on breast biopsy methods we synthesized evidence from a total of 319 studies. We found few studies providing information on the test performance of open surgical biopsy. In contrast, the evidence base on core-needle biopsy methods now includes a large number of studies reporting on almost 70,000 breast lesions. The following subsections summarize our assessment of the strength of evidence. Following the original evidence report, and in view of the paucity of evidence on open surgical biopsy, we refrained from rating the strength of evidence for this technique for all Key Questions. For Key Questions 1 and 2, we assessed the strength of evidence by integrating our (subjective) judgments on the risk of bias of included studies, the consistency of their findings, the directness of the available data, and the precision of quantitative results. For Key Question 3 we only rated the strength of evidence for the outcome of additional surgical procedures required after biopsy. We did not rate the strength of evidence for other Key Question 3 outcomes because of the diversity of designs employed and outcomes addressed. Please see the Methods section for a detailed discussion of our approach to rating the strength of evidence. Details about the strength of evidence assessment are provided in **Appendix C**.

Test performance and comparative test performance

Among women at average risk of cancer, core-needle biopsy using ultrasound or stereotactic guidance had average sensitivities ranging from 0.97 to 0.99 and average specificities ranging from 0.92 to 0.99. Freehand biopsy methods appeared to have lower average sensitivity (0.91) compared to other methods, but similar specificity. Stereotactically guided vacuum-assisted techniques were associated with lower specificity compared to other biopsy methods. Although these results were consistent across studies and (in many cases) fairly precise, they were derived from indirect comparisons across studies of moderate to high risk of bias. MRI-guided biopsies were evaluated in only four studies with small sample sizes, leading to substantial uncertainty around estimates of test performance. **Table 14** summarizes our assessment of the strength of evidence for comparisons among alternative biopsy methods in women at average risk of cancer. Of note, we rated the strength of evidence on *comparative* test performance, whereas the original report considered *absolute* test performance; for this reason, and for this subset of outcomes, our ratings are not directly comparable with those of the original report.

There were few studies of women at high risk of cancer; however, statistical comparisons of test performance between women at low and high risk of breast cancer did not identify a difference. Because the number of available studies was small, comparisons of test performance between low and high risk women had substantial uncertainty and results were not sufficient to support definitive conclusions. Evidence on modifiers of test performance was also sparse, for all biopsy methods, raising concerns about selective outcome and analysis reporting.

Table 14: Strength of evidence about comparative test performance in women at average risk of breast cancer

| Outcome | Comparison or biopsy method | Overall Rating | Key Findings and Comments |
|--------------------|-----------------------------|----------------|--|
| Comparison of test | Freehand vs. ultrasound- | Low | – Difference in sensitivity: 0.08 (0.02 to 0.19) |

| Outcome | Comparison or biopsy method | Overall Rating | Key Findings and Comments |
|--|--|----------------|---|
| performance among alternative biopsy methods | guided, automated | | [ultrasound-guided, automated better] – Difference in specificity: -0.01 (-0.04, 0.02) [no difference] |
| | Freehand vs. stereotactically guided, automated | Low | – Difference in sensitivity: 0.07 (0.01 to 0.18) [stereotactically guided, automated better] – Difference in specificity: -0.01 (-0.03 to 0.02) [no difference] |
| | Freehand vs. ultrasound-guided, vacuum-assisted | Low | – Difference in sensitivity: 0.06 (-0.03 to 0.17) [ultrasound-guided, vacuum-assisted better] – Difference in specificity: 0.00 (-0.02 to -0.03) [no difference] |
| | Freehand vs. stereotactically guided, vacuum-assisted | Low | – Difference in sensitivity: 0.08 (0.03 to 0.19) [stereotactically guided, vacuum-assisted better] – Difference in specificity: -0.07 (-0.10 to -0.03) [freehand better] |
| | Stereotactically guided, vacuum assisted vs. automated | Low | – Difference in sensitivity: 0.02 (0.00 to 0.03) [vacuum-assisted better] – Difference in specificity: -0.05 (-0.09 to -0.03) [automated better] |
| | Stereotactically guided vacuum assisted vs. ultrasound-guided, automated | Low | – Difference in sensitivity: 0.00 (-0.01 to 0.02) [no difference] – Difference in specificity: -0.05 (-0.09 to -0.02) [ultrasound-guided, automated better] |
| | Stereotactically guided vacuum assisted vs. ultrasound-guided, vacuum-assisted | Low | – Difference in sensitivity: 0.02 (0.00 to 0.10) [no difference] – Difference in specificity: -0.07 (-0.10 to -0.04) [ultrasound-guided vacuum-assisted better] |
| | Other comparisons between biopsy techniques | Low | – There were no differences in sensitivity and specificity for ultrasound-guided automated vs. vacuum-assisted methods. – There were no differences in sensitivity and specificity for stereotactically guided automated vs. ultrasound-guided methods. – The CrIs for all comparisons included zero and were fairly precise. |
| | MRI vs. any other device | Insufficient | – Only 4 small studies were available – Differences in sensitivity and specificity comparing MRI with other biopsy methods had CrIs intervals that included 0 but were imprecise |
| Modifiers of test performance for women at average and high risk of breast cancer; | All biopsy methods | Insufficient | – Few studies provided within sample information for each modifier of interest; meta-regression results rely on cross-study comparisons so consistency of effects cannot be assessed – Within-study (direct) evidence was sparse; between study evidence relied on indirect comparisons across studies – In meta-regression analyses CrIs were wide and extreme odds ratio values were often observed because sensitivity and specificity for all tests were very close to 1 (see Results for additional details) |

CrIs = credible interval; MRI = magnetic resonance imaging.

Underestimation rates

Underestimation rates varied among alternative biopsy methods and were often imprecisely estimated because of the relatively small number of lesions contributing data for these analyses. In general, underestimation was less common with stereotactically guided vacuum-assisted biopsy methods, as compared to stereotactically or ultrasound-guided automated methods. Our assessment of the strength of evidence for this outcome is summarized in **Table 15**.

Table 15: Strength of evidence for underestimation rates in women at average risk of cancer

| Outcome | Comparison or biopsy method | Overall Rating | Key Findings and Comments |
|---------------------------------------|--|----------------|---|
| DCIS underestimation | Stereotactically guided, automated | Low | – Average underestimation probability: 0.27 (0.18 to 0.37) [17 studies] |
| | Stereotactically guided, vacuum-assisted | Low | – Average underestimation probability: 0.11 (0.08 to 0.14) [33 studies] |
| | Ultrasound-guided, automated | Low | – Average underestimation probability: 0.38 (0.25 to 0.51) [14 studies] |
| | Other biopsy methods | Insufficient | Few or no available studies with very small numbers of lesions. |
| High risk lesion underestimation rate | Stereotactically guided, automated | Low | – Average underestimation probability: 0.47 (0.37 to 0.58) [28 studies] |
| | Stereotactically guided, vacuum-assisted | Low | – Average underestimation probability: 0.18 (0.13 to 0.24) [37 studies] |
| | Ultrasound-guided, automated | Low | – Average underestimation probability: 0.22 (0.14 to 0.34) [20 studies] |
| | Other biopsy methods | Insufficient | Few or no available studies with very small numbers of lesions. |

DCIS = ductal carcinoma in situ.

Adverse Events and Additional Surgeries After Biopsy

In general, adverse events were reported inconsistently, raising concerns about selective outcome and analysis reporting. Few studies provided information on the harms of open surgical biopsy. Core-needle biopsy was only infrequently associated with serious adverse events. Comparisons between open and core-needle biopsy are based on indirect comparisons and expert opinion, with limited empirical evidence. Open biopsy appeared to be associated with an increased incidence of adverse events (including serious adverse events) compared to core-needle biopsy. Our assessment of the strength of evidence for adverse events is summarized in **Table 16**.

Among core-needle biopsy methods, vacuum-assisted methods appeared to be associated with increased bleeding. Sitting upright during the biopsy procedure was associated with more vasovagal reactions. Information about the dissemination of cancer cells during the biopsy procedure was provided by a small number of studies with various designs. Studies reported that women were generally satisfied with the cosmetic results of core-needle procedures.

Women diagnosed with breast cancer by core-needle biopsy were able to have their cancer treated with a single surgical procedure, more often than women diagnosed by open surgical biopsy. Although the magnitude of this association was large (the ratio of the odds was approximately 13), women and their physicians are likely to choose biopsy methods on the basis of factors (e.g., lesion location, or characteristics of the lesion on imaging) that may also be

associated with the need for additional surgeries. Because such selection would lead to confounding by indication, we rated the strength of evidence for this association as moderate. .

Table 16: Strength of evidence assessment for adverse events of biopsy

| Outcomes | Comparison | Overall Rating | Key findings |
|--|---|----------------|--|
| Bleeding, including bleeding events that require treatment | Comparisons among alternative core-needle biopsy methods | Low | <ul style="list-style-type: none"> – Median %: 0.76 (25th perc. = 0.22; 75th perc = 3.97) – Potential for selective outcome and analysis reporting – Few studies reported bleeding requiring treatment; the event rate was low (<0.40 perc.) in those studies |
| Hematoma formation | Comparisons among alternative core-needle biopsy methods | Low | <ul style="list-style-type: none"> – Median %: 1.19 (25th perc. = 0.05; 75th perc = 7.20) – Potential for selective outcome and analysis reporting |
| Infectious complications | Comparisons among alternative core-needle biopsy methods | Low | <ul style="list-style-type: none"> – Median %: 0.05 (25th perc. = 0.00; 75th perc = 0.63) – Potential for selective outcome and analysis reporting |
| Vasovagal reactions: | Comparisons among alternative core-needle biopsy methods | Low | <ul style="list-style-type: none"> – Median %: 1.78 (25th perc. = 0.43; 75th perc = 4.29) – Potential for selective outcome and analysis reporting |
| Pain and severe pain | Comparisons among alternative core-needle biopsy methods | Low | 25 studies of a wide variety of biopsy methods reported information about patient pain during the procedure (pain was assessed heterogeneously across studies). |
| Other adverse events | Comparisons among alternative core-needle biopsy methods | Insufficient | <ul style="list-style-type: none"> – Most events were reported by a single study precluding assessment of consistency – Individual studies did not provide adequate information for precise estimation of the event rate) – Only informal indirect comparisons among biopsy methods were possible – Potential for selective outcome and analysis reporting |
| Modifiers of adverse events – vasovagal reactions | Sitting upright during the biopsy procedure | Low | <ul style="list-style-type: none"> – Vasovagal reactions were more common among patients sitting during the biopsy procedure – Results were reported in few studies (11 studies; 8 from the original evidence report and 3 from this update) – Potential for selective outcome and analysis reporting |
| Modifiers of adverse events – bleeding | Vacuum-assisted versus non-vacuum assisted biopsy methods | Low | <ul style="list-style-type: none"> – Vacuum-assisted procedures were generally associated with increased rates of bleeding and hematoma formation – Bleeding events were generally uncommon – Comparisons among biopsy methods were based on informal indirect comparisons (across studies) – Potential for selective outcome and analysis reporting |
| All other modifiers of adverse events | Comparisons among alternative core-needle biopsy methods | Insufficient | <ul style="list-style-type: none"> – Most factors assessed by a single study limiting our ability to assess consistency – Potential for selective outcome and analysis reporting. – Within-study comparisons provided direct evidence |

perc. = percentile.

Limitations of the Evidence Base

We believe that the evidence regarding the performance of core-needle biopsy for diagnosis of breast lesions is limited in the following ways:

- Published evidence on the test performance and adverse events of open surgical biopsy was sparse.
- Available studies, particularly for Key Questions 1 and 2, were at moderate to high risk of

bias and the publications we reviewed did not follow the Standards for Reporting of Diagnostic Accuracy (STARD) guidelines.²⁶² Information on patient selection criteria, patient or lesion characteristics (e.g., granular reporting of pathology results), was often missing or inconsistently reported. Information on adverse events and patient-relevant outcomes was often incomplete, potentially selectively reported. Studies did not use standardized definitions and ascertainment methods for adverse events. Pathology results were not reported with adequate granularity in the majority of cases.

- Studies typically used lesions (or biopsy procedures) as the unit of analysis, instead of patients. This way, patients with multiple lesions contributed multiple observations to the analyses. Lesions belonging to the same patient are likely to have similar characteristics (i.e. they are correlated). Unfortunately, studies reported results in a way that did not allow for the correlation to be accounted for in our statistical models. As such, our analyses (and those of the original report) assume independence among lesions. If the correlation among lesions in the same patient is high (positive and close to one) individual study and meta-analytic results will underestimate uncertainty and may also be biased (the direction of bias is unpredictable). However, unless each patient contributes large numbers of lesions that are highly correlated, the underestimation of uncertainty will not be large. Further, bias is unlikely unless patients contributing large numbers of lesions also have lesions that are substantially harder (or easier) to diagnose compared to those of other patients. Without additional data on the test performance on individual lesions within patients it is not possible to ascertain the impact of this factors on our results.
- Studies provided limited information to assess the impact of various patient-, lesion-, procedure-, or system- related factors on the outcomes of breast biopsy. For example, the impact of patient age, breast density, lesion type, training and experience of the operators, and error rates of pathologists who read the samples, on test performance, adverse events, or clinical outcomes could not be assessed.
- We found very few studies on MRI-guided biopsy for women at average or high risk of cancer. Because MRI-guided biopsy is likely reserved for diagnostically challenging cases and may be available in specialized care settings indirect (i.e. across studies) comparisons between MRI-guided and other biopsy procedures may be confounded by factors unrelated to the diagnostic value of the tests compared.
- There is limited information on the comparative effectiveness of alternative biopsy methods on patient-relevant outcomes, resource use and logistics, and availability of technology and expertise for different core-needle biopsy techniques.

Strengths and Limitations of This Review

We conducted an up-to-date review of the benefits and risks of breast biopsy methods for breast cancer diagnosis, with respect to test performance, underestimation rates, adverse events, and patient-relevant outcomes. Previous reviews on this topic have focused on special patients populations (e.g., patients with non-palpable lesions²⁶³), selected outcomes (e.g. DCIS underestimation²⁶⁴ or seeding²⁶⁵), or biopsy methods (e.g., ultrasound-guided biopsy²⁶⁶). However, our work has several limitations, which – to a large extent – reflect the limitations of the underlying evidence base. Studies were deemed to be of moderate to high risk of bias because of characteristics related to their design and conduct, limiting our ability to draw strong conclusions. Information for several outcomes of interest was not reported from all available studies (e.g., underestimation rates, adverse events) raising concerns about selective outcome and

analysis reporting. Information on study- or population level characteristics that could be modifiers of test performance, adverse events, or clinical outcomes, was inadequate. Thus, our ability to explore between-study heterogeneity was limited. Further, because we relied on published information and did not have access to individual patient data, we were unable to evaluate the impact of patient- or lesion-level factors on outcomes of interest.

Applicability of Review Findings

The existing evidence base on core-needle biopsy of breast lesions in women at average risk of cancer appears to be applicable to clinical practice in the U.S. Studies enrolled patients with an average age similar to that of women undergoing breast biopsy in the U.S., and for indications that represent the most prevalent indications in U.S. clinical practice (i.e. mammographic findings of suspicious lesions). While fewer than half of the studies in this review were conducted in the United States, almost all were carried out in either the U.S. or in industrialized European or Asian countries where core-biopsy methods are likely sufficiently similar to those used in the U.S. However, the applicability of our findings to women at high risk of breast cancer may be limited because we found few studies explicitly reporting on groups of patients at high baseline risk of breast cancer on the basis of factors such as genetic testing, or family history of disease. Of note, this may be an instance of incomplete reporting rather than a true characterization of the baseline risk of included populations (i.e. some high risk populations may have been misclassified as “average risk”).

Evidence Gaps and Ongoing Research

Table 17 summarizes the evidence gaps with regards to the Key Questions of diagnostic test performance and adverse events. A search on ClinicalTrials.gov for randomized trials comparing alternative biopsy methods did not identify trials examining biopsy techniques for breast cancer diagnosis (last search: Dec 5, 2013; 141 records retrieved).

Table 17: Evidence gaps for biopsy methods for the diagnosis of breast cancer

| Key Question | Category | Evidence Gap |
|--|-------------------------------|---|
| Comparative effectiveness of core-needle biopsy and open surgical biopsy | General | Limited information on the diagnostic test performance of open surgical biopsy was available. However, expert opinion and research studies consider open biopsies to have negligible measurement error. |
| | Population | Limited information for women specified to be at high baseline risk of breast cancer. |
| | Interventions & Comparators | Limited information on MRI-guided biopsy methods (all patient populations). For other biopsy methods a large body of evidence was available; however studies were at moderate to high risk of bias and poorly reported. |
| | Outcomes | Information on underestimation rates was relatively limited. Pathology results were not reported using consistent or sufficiently granular classification schemes. |
| | Modifiers of test performance | Optimal core-needle biopsy method for specific subgroups of patients, lesion characteristics. |
| Adverse events of core-needle biopsy and open surgical biopsy | General | Information for adverse events of interest was incompletely and (potentially) selectively reported. |
| | Interventions & Comparators | Evidence comparing the adverse events of open and alternative core-needle biopsy methods was limited. |
| | Outcomes | Limited information was available for key adverse events of interest. Reporting in existing studies was inconsistent and potentially selective. Outcome ascertainment was not standardized. |

| | | |
|--|-----------------------------|---|
| | Modifiers of adverse events | Information on factors that affect the incidence of adverse events is sparse. Unclear what subgroups of patients and lesions may be most likely to experience adverse events |
| Patient-relevant and resource-related outcomes | General | Comparative effectiveness information among alternative biopsy techniques (both open and core-needle) was very sparse and indirect. Comparisons between methods are susceptible to confounding and selection bias |
| | Population | Evidence is limited both for women at average and high risk of breast cancer. |
| | Outcomes | The balance of benefits and risks associated with alternative breast biopsy with respect to clinical outcomes, quality of life, and resource use has not been comprehensively assessed. |

MRI = magnetic resonance imaging.

Future Research Needs

- Studies of test performance are needed to evaluate MRI-guided biopsy methods. Ideally, these studies will be large (powered to achieve adequate precision), prospectively designed, multicenter investigations enrolling patients representative of those seen in clinical practice. Studies should use standardized histological classification systems for pathological classification.^{267, 268} The reference standard for test negative cases should be regular monitoring for an adequate period of time (e.g., 2 years).
- Although a large number of studies were available for other core-needle biopsy methods we believe that additional well-designed and fully reported prospective cohort studies are needed, primarily for addressing questions about the impact of patient-, lesion-, procedure-, or system-level factors on test performance, adverse events, and patient-relevant outcomes. Given that a large number of core-needle biopsies are performed annually in diverse settings, such studies could be conducted at relatively low cost.
- Large-scale databases of prospectively-collected observational data on breast biopsy procedures and outcomes could be used to evaluate the comparative effectiveness of alternative biopsy methods with respect to short and long term outcomes, and potential modifying factors. Such studies would need to collect detailed information on baseline factors that may be associated with both the choice of biopsy method and the outcomes of interest, to adjust for potential confounding factors
- In all future studies, baseline risk of cancer development should be characterized using consistent and widely accepted criteria to allow appropriate subgroup analyses.
- We believe that a randomized comparison of alternative biopsy methods is unlikely to be fruitful because existing studies indicate that biopsy procedures have sensitivities and specificities that are fairly similar and close to 1. Under these conditions randomized trials comparing alternative biopsy methods would need to attain very large sample sizes to allow reliable comparisons between tests.
- Additional information is also needed to define what patient and lesion factors may correspond with accuracy or adverse events of specific techniques. Future research needs to be better reported for progress to be made on these questions.

Conclusions

A large body of evidence suggests that core-needle biopsy procedures have sensitivity and specificity at or near that of open biopsy procedures, and are associated with fewer adverse events. Image-guided core needle biopsy approaches appear to have similar test performance and safety profiles for women at average risk of breast cancer, although freehand procedures have

lower sensitivity, and vacuum-assisted procedures appear to have a higher risk of bleeding. The strength of conclusions about comparative test performance was generally low, because of concerns about the risk of bias of included studies, incomplete reporting, and the reliance on indirect comparisons. There were insufficient data to draw conclusions for MRI-guided biopsy or women at high baseline risk of cancer. Harms were reported inconsistently, raising concerns about selective outcome and analysis reporting. Women diagnosed with breast cancer by core-needle biopsy were more likely to have their cancer treated with a single surgical procedure, compared with women diagnosed by open surgical biopsy.

List of abbreviations

| | |
|--------|--|
| ADH | Atypical ductal hyperplasia |
| AHRQ | Agency for healthcare Research and Quality |
| CrI | Credibility interval |
| DCIS | Ductal carcinoma in situ |
| EPC | Evidence-based Practice Center |
| FN | False negative |
| FP | False positive |
| MRI | Magnetic resonance imaging |
| PICOTS | Populations-Interventions-Comparators-Outcomes-Timing-Setting |
| PRISMA | Preferred Reporting Items for Systematic Reviews and Meta-Analyses |
| QUADAS | Quality Assessment of Diagnostic Accuracy Studies |
| ROC | Received operating characteristic |
| SRDR | Systematic Review Data Repository |
| STARD | Standards for Reporting of Diagnostic Accuracy |
| TEP | Technical expert panel |
| TOO | Task Order Officer |
| TN | True negative |
| TP | True positive |

References

1. Breast Cancer Facts & Figures 2011-2012. American Cancer Society, Atlanta, Georgia.; [April 25, 2013]; Available from: <http://www.cancer.org/acs/groups/content/@epidemiologysurveillance/documents/document/acspc-030975.pdf>.
2. SEER Stat Fact Sheets: Breast. Surveillance Research Program. National Cancer Institute.; [April 25, 2013]; Available from: <http://seer.cancer.gov/statfacts/html/breast.html - incidence-mortality>.
3. Breast Cancer : What are the key statistics about breast cancer? : American Cancer Society, Atlanta, Georgia.; [April 25, 2013]; Available from: <http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-key-statistics>.
4. Miller JW, King JB, Joseph DA, et al. Breast Cancer Screening Among Adult Women — Behavioral Risk Factor Surveillance System, United States, 2010. *Morbidity and Mortality Weekly Report (MMWR)* 2012;61(2):46-50. PMID:
5. Hubbard RA, Kerlikowske K, Flowers CI, et al. Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: a cohort study. *Annals of internal medicine* 2011 Oct 18;155(8):481-92. PMID: 22007042
6. Elmore JG, Barton MB, Mocerri VM, et al. Ten-year risk of false positive screening mammograms and clinical breast examinations. *The New England journal of medicine* 1998 Apr 16;338(16):1089-96. PMID: 9545356
7. Yu YH, Wei W, Liu JL. Diagnostic value of fine-needle aspiration biopsy for breast mass: a systematic review and meta-analysis. *BMC cancer* 2012;12:41. PMID: 22277164
8. Having a Breast Biopsy: A guide for women and their families. The Agency for Healthcare Research and Quality (AHRQ). [April 25, 2013]; Available from: http://effectivehealthcare.ahrq.gov/ehc/products/17/407/core_needle_consumer_guide.pdf.
9. Bruening W SK, Treadwell J, Launders J, Fontanarosa J, Tipton K. . Comparative Effectiveness of Core-Needle and Open Surgical Biopsy for the Diagnosis of Breast Lesions. Comparative Effectiveness Review No. 19. (Prepared by ECRI Institute Evidence-based Practice Center under Contract No. 290-02-0019.) Rockville, MD: Agency for Healthcare Research and Quality. December 2009. Available at: <http://www.effectivehealthcare.ahrq.gov/reports/final.cfm> December 2009.
10. Bruening W, Fontanarosa J, Tipton K, et al. Systematic review: comparative effectiveness of core-needle and open surgical biopsy to diagnose breast lesions. *Annals of internal medicine* 2010 Feb 16;152(4):238-46. PMID: 20008742
11. AHRQ Comparative Effectiveness Review Surveillance Program. Comparative Effectiveness of Core Needle Biopsy and Open Surgical Biopsy for Diagnosis of Breast Lesions 2011.
12. Agency for Healthcare Research and Quality. Having a breast biopsy: A guide for women and their families. AHRQ Publication No 10-EHC007-A. Rockville, MD: Agency for

- Healthcare Research and Quality; April 2010. Available at http://www.effectivehealthcare.ahrq.gov/ehc/products/17/407/core_needle_consumer_guide.pdf.
13. Agency for Healthcare Research and Quality. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No 10(12)-EHCO63-EF. Rockville, MD: Agency for Healthcare Research and Quality; April 2012. Available at http://www.effectivehealthcare.ahrq.gov/ehc/products/60/318/MethodsGuide_Prepublishing-Draft_20120523.pdf.
 14. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg* 2010;8(5):336-41. PMID: 20171303
 15. Dekkers OM, Egger M, Altman DG, et al. Distinguishing case series from cohort studies. *Ann Intern Med* 2012 Jan 3;156(1 Pt 1):37-40. PMID: 22213493
 16. Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA : the journal of the American Medical Association* 1999 Sep 15;282(11):1061-6. PMID: 10493205
 17. Rutjes AW, Reitsma JB, Di Nisio M, et al. Evidence of bias and variation in diagnostic accuracy studies. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne* 2006 Feb 14;174(4):469-76. PMID: 16477057
 18. Maglione M, Motala A, Shanman R, et al; AHRQ Comparative Effectiveness Review Surveillance Program. Surveillance Report: Comparative Effectiveness of Core Needle Biopsy and Open Surgical Biopsy for Diagnosis of Breast Lesions. Rockville, MD: Agency for Healthcare Research and Quality; December 2011 and July 2012. Available at http://www.effectivehealthcare.ahrq.gov/ehc/products/17/370/NeedleBiopsy_SurveillanceAssessment_20120919.pdf.
 19. Drummond M, Barbieri M, Cook J, et al. Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force report. *Value Health* 2009 Jun;12(4):409-18. PMID: 19900249
 20. Ramsey S, Willke R, Briggs A, et al. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. *Value Health* 2005 Sep-Oct;8(5):521-33. PMID: 16176491
 21. Whiting P, Westwood M, Burke M, et al. Systematic reviews of test accuracy should search a range of databases to identify primary studies. *J Clin Epidemiol* 2008 Apr;61(4):357-64. PMID: 18313560
 22. Leeflang MM, Scholten RJ, Rutjes AW, et al. Use of methodological search filters to identify diagnostic accuracy studies can lead to the omission of relevant studies. *Journal of clinical epidemiology* 2006 Mar;59(3):234-40. PMID: 16488353
 23. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011 Oct 18;155(8):529-36. PMID: 22007046
 24. Whiting P, Rutjes AW, Reitsma JB, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC medical research methodology* 2003 Nov 10;3:25. PMID: 14606960
 25. Whiting P, Rutjes AW, Dinnes J, et al. Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technol Assess* 2004 Jun;8(25):iii, 1-234. PMID: 15193208

26. Whiting PF, Weswood ME, Rutjes AW, et al. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC medical research methodology* 2006 Mar 6;6:9. PMID: 16519814
27. Wells G, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. [29 April 2013]; Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
28. Higgins JP, Altman DG, Gotzsche PC, et al; Cochrane Bias Methods Group/ Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011 Oct 18;343:d5928. PMID: 22008217
29. Keskek M, Tez M, Yoldas O, et al. Receiver operating characteristic analysis of leukocyte counts in operations for suspected appendicitis. *The American journal of emergency medicine* 2008 Sep;26(7):769-72. PMID: 18774040
30. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the *BMJ*. The *BMJ* Economic Evaluation Working Party. *BMJ* 1996 Aug 3;313(7052):275-83. PMID: 8704542
31. Trikalinos TA, Balion CM. Chapter 9: options for summarizing medical test performance in the absence of a "gold standard". *J Gen Intern Med* 2012 Jun;27 Suppl 1:S67-75. PMID: 22648677
32. Trikalinos TA, Balion CM, Coleman CI, et al. Chapter 8: meta-analysis of test performance when there is a "gold standard". *J Gen Intern Med* 2012 Jun;27 Suppl 1:S56-66. PMID: 22648676
33. Hamza TH, van Houwelingen HC, Stijnen T. The binomial distribution of meta-analysis was preferred to model within-study variability. *J Clin Epidemiol* 2008 Jan;61(1):41-51. PMID: 18083461
34. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2001 Oct 15;20(19):2865-84. PMID: 11568945
35. Arends LR, Hamza TH, van Houwelingen JC, et al. Bivariate random effects meta-analysis of ROC curves. *Med Decis Making* 2008 Sep-Oct;28(5):621-38. PMID: 18591542
36. Stijnen T, Hamza TH, Ozdemir P. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Stat Med* 2010 Dec 20;29(29):3046-67. PMID: 20827667
37. Hamza TH, van Houwelingen HC, Heijnenbroek-Kal MH, et al. Associating explanatory variables with summary receiver operating characteristic curves in diagnostic meta-analysis. *J Clin Epidemiol* 2009 Dec;62(12):1284-91. PMID: 19398297
38. van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Stat Med* 2002 Feb 28;21(4):589-624. PMID: 11836738
39. Olkin I. Diagnostic statistical procedures in medical meta-analyses. *Stat Med* 1999 Sep 15-30;18(17-18):2331-41. PMID: 10474143
40. Plummer M. JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003); March 20-22, 2003; Vienna, Austria.
41. Dahabreh IJ, Trikalinos TA, Lau J, et al. An Empirical Assessment of Bivariate Methods for Meta-Analysis of Test Accuracy. Rockville (MD)2012.

42. AHRQ. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. 2012; Available from: <http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=318>.
43. Singh S, Chang SM, Matchar DB, et al. Grading a body of evidence on diagnostic tests. In: Chang SM, Matchar DB, Smetana GW, Umscheid CA, editors. Methods Guide for Medical Test Reviews. AHRQ Publication No. 12-EHC017. Rockville, MD: Agency for Healthcare Research and Quality; June 2012:chapter 7. Available at <http://www.ncbi.nlm.nih.gov/books/NBK98241/>.
44. Lau J, Ioannidis JP, Terrin N, et al. The case of the misleading funnel plot. *BMJ* 2006 Sep 16;333(7568):597-600. PMID: 16974018
45. Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011;343:d4002. PMID: 21784880
46. Abdsaleh S, Azavedo E, Lindgren PG. Semiautomatic core biopsy. A modified biopsy technique in breast diseases. *Acta Radiol* 2003 Jan;44(1):47-51. PMID: 12630998
47. Ahmed ME, Ahmad I, Akhtar S. Ultrasound guided fine needle aspiration cytology versus core biopsy in the preoperative assessment of non-palpable breast lesions. *Journal of Ayub Medical College, Abbottabad : JAMC* 2010 Apr-Jun;22(2):138-42. PMID: 21702288
48. Akita A, Tanimoto A, Jinno H, et al. The clinical value of bilateral breast MR imaging: is it worth performing on patients showing suspicious microcalcifications on mammography? *Eur Radiol* 2009 Sep;19(9):2089-96. PMID: 19350244
49. Alonso-Bartolome P, Vega-Bolivar A, Torres-Tabanera M, et al. Sonographically guided 11-G directional vacuum-assisted breast biopsy as an alternative to surgical excision: utility and cost study in probably benign lesions. *Acta Radiol* 2004 Jul;45(4):390-6. PMID: 15323390
50. Ambrogetti D, Bianchi S, Ciatto S. Accuracy of percutaneous core biopsy of isolated breast microcalcifications identified by mammography. Experience with a vacuum-assisted large-core biopsy device. *La Radiologia medica* 2003 Oct;106(4):313-9. PMID: 14612823
51. Apesteguia L, Mellado M, Saenz J, et al. Vacuum-assisted breast biopsy on digital stereotaxic table of nonpalpable lesions non-recognisable by ultrasonography. *Eur Radiol* 2002 Mar;12(3):638-45. PMID: 11870480
52. Barreto V, Hamed H, Griffiths AB, et al. Automatic needle biopsy in the diagnosis of early breast cancer. *Eur J Surg Oncol* 1991 Jun;17(3):237-9. PMID: 2044776
53. Bauer RL, Sung J, Eckhert KH, Jr., et al. Comparison of histologic diagnosis between stereotactic core needle biopsy and open surgical biopsy. *Ann Surg Oncol* 1997 Jun;4(4):316-20. PMID: 9181231
54. Beck RM, Gotz L, Heywang-Kobrunner SH. Stereotaxic vacuum core breast biopsy--experience of 560 patients. *Swiss Surg* 2000;6(3):108-10. PMID: 10894010
55. Becker L, Taves D, McCurdy L, et al. Stereotactic core biopsy of breast microcalcifications: comparison of film versus digital mammography, both using an add-on unit. *AJR Am J Roentgenol* 2001 Dec;177(6):1451-7. PMID: 11717106
56. Bernardi D, Borsato G, Pellegrini M, et al. On the diagnostic accuracy of stereotactic vacuum-assisted biopsy of nonpalpable breast abnormalities. Results in a consecutive series of 769 procedures performed at the Trento Department of Breast Diagnosis. *Tumori* 2012 Jan-Feb;98(1):113-8. PMID: 22495711

57. Bolivar AV, Alonso-Bartolome P, Garcia EO, et al. Ultrasound-guided core needle biopsy of non-palpable breast lesions: a prospective analysis in 204 cases. *Acta Radiol* 2005 Nov;46(7):690-5. PMID: 16372687
58. Brancato B, Scialpi M, Pusiol T, et al. Needle core biopsy should replace fine needle aspiration cytology in ultrasound-guided sampling of breast lesions. *Pathologica* 2011 Apr;103(2):52. PMID: 21797145
59. Brenner RJ, Bassett LW, Fajardo LL, et al. Stereotactic core-needle breast biopsy: a multi-institutional prospective trial. *Radiology* 2001 Mar;218(3):866-72. PMID: 11230668
60. Britton PD, Flower CD, Freeman AH, et al. Changing to core biopsy in an NHS breast screening unit. *Clin Radiol* 1997 Oct;52(10):764-7. PMID: 9366536
61. Burbank F. Stereotactic breast biopsy of atypical ductal hyperplasia and ductal carcinoma in situ lesions: improved accuracy with directional, vacuum-assisted biopsy. *Radiology* 1997 Mar;202(3):843-7. PMID: 9051043
62. Cangiarella J, Waisman J, Symmans WF, et al. Mammotome core biopsy for mammary microcalcification: analysis of 160 biopsies from 142 women with surgical and radiologic followup. *Cancer* 2001 Jan 1;91(1):173-7. PMID: 11148574
63. Caruso ML, Gabrieli G, Marzullo G, et al. Core Biopsy as Alternative to Fine-Needle Aspiration Biopsy in Diagnosis of Breast Tumors. *Oncologist* 1998;3(1):45-9. PMID: 10388083
64. Chapellier C, Balu-Maestro C, Amoretti N, et al. Vacuum-assisted breast biopsies. Experience at the Antoine Lacassagne Cancer Center (Nice, France). *Clin Imaging* 2006 Mar-Apr;30(2):99-107. PMID: 16500540
65. Ciatto S, Houssami N, Ambrogetti D, et al. Accuracy and underestimation of malignancy of breast core needle biopsy: the Florence experience of over 4000 consecutive biopsies. *Breast Cancer Res Treat* 2007 Mar;101(3):291-7. PMID: 16823506
66. Cipolla C, Fricano S, Vieni S, et al. Validity of needle core biopsy in the histological characterisation of mammary lesions. *Breast* 2006 Feb;15(1):76-80. PMID: 16473738
67. Cornelis A, Verjans M, Van den Bosch T, et al. Efficacy and safety of direct and frontal macrobiopsies in breast cancer. *Eur J Cancer Prev* 2009 Aug;18(4):280-4. PMID: 19352188
68. Cross MJ, Evans WP, Peters GN, et al. Stereotactic breast biopsy as an alternative to open excisional biopsy. *Ann Surg Oncol* 1995 May;2(3):195-200. PMID: 7641014
69. Crowe Jr JP, Patrick RJ, Rybicki LA, et al. Does ultrasound core breast biopsy predict histologic finding on excisional biopsy? *The American Journal of Surgery* 2003;186:397-99. PMID:
70. Crystal P, Koretz M, Shcharynsky S, et al. Accuracy of sonographically guided 14-gauge core-needle biopsy: results of 715 consecutive breast biopsies with at least two-year follow-up of benign lesions. *J Clin Ultrasound* 2005 Feb;33(2):47-52. PMID: 15674836
71. Cusick JD, Dotan J, Jaecks RD, et al. The role of Tru-Cut needle biopsy in the diagnosis of carcinoma of the breast. *Surg Gynecol Obstet* 1990 May;170(5):407-10. PMID: 2183373
72. Dahlstrom JE, Jain S. Histological correlation of mammographically detected microcalcifications in stereotactic core biopsies. *Pathology (Phila)* 2001 Nov;33(4):444-8. PMID: 11827410
73. Dahlstrom JE, Jain S, Sutton T, et al. Diagnostic accuracy of stereotactic core biopsy in a mammographic breast cancer screening programme. *Histopathology* 1996 May;28(5):421-7. PMID: 8735717

74. de Lucena CE, Dos Santos Junior JL, de Lima Resende CA, et al. Ultrasound-guided core needle biopsy of breast masses: How many cores are necessary to diagnose cancer? *J Clin Ultrasound* 2007 Sep;35(7):363-6. PMID: 17663457
75. Delle Chiaie L, Terinde R. Three-dimensional ultrasound-validated large-core needle biopsy: is it a reliable method for the histological assessment of breast lesions? *Ultrasound Obstet Gynecol* 2004 Apr;23(4):393-7. PMID: 15065192
76. Dhillon MS, Bradley SA, England DW. Mammotome biopsy: impact on preoperative diagnosis rate. *Clin Radiol* 2006 Mar;61(3):276-81. PMID: 16488210
77. Dillon MF, Hill AD, Quinn CM, et al. The accuracy of ultrasound, stereotactic, and clinical core biopsies in the diagnosis of breast cancer, with an analysis of false-negative cases. *Ann Surg* 2005 Nov;242(5):701-7. PMID: 16244544
78. Doyle AJ, Collins JP, Forkert CD. Decubitus stereotactic core biopsy of the breast: technique and experience. *AJR Am J Roentgenol* 1999 Mar;172(3):688-90. PMID: 10063861
79. Doyle AJ, Murray KA, Nelson EW, et al. Selective use of image-guided large-core needle biopsy of the breast: accuracy and cost-effectiveness. *AJR Am J Roentgenol* 1995 Aug;165(2):281-4. PMID: 7618540
80. Elvecrog EL, Lechner MC, Nelson MT. Nonpalpable breast lesions: correlation of stereotaxic large-core needle biopsy and surgical biopsy results. *Radiology* 1993 Aug;188(2):453-5. PMID: 8327696
81. Faizi KS, Hameed T, Azim K, et al. Diagnostic accuracy of Trucut biopsy in the detection of breast lumps. *Pakistan Journal of Medical and Health Sciences* 2012;6(4):979-82. PMID:
82. Fajardo LL, Pisano ED, Caudry DJ, et al. Stereotactic and sonographic large-core biopsy of nonpalpable breast lesions: results of the Radiologic Diagnostic Oncology Group V study. *Acad Radiol* 2004 Mar;11(3):293-308. PMID: 15035520
83. Fischbach F, Eggemann H, Bunke J, et al. MR-guided freehand biopsy of breast lesions in a 1.0-T open MR imager with a near-real-time interactive platform: preliminary experience. *Radiology* 2012 Nov;265(2):359-70. PMID: 22923721
84. Fishman JE, Milikowski C, Ramsinghani R, et al. US-guided core-needle biopsy of the breast: how many specimens are necessary? *Radiology* 2003 Mar;226(3):779-82. PMID: 12601206
85. Frazee RC, Roberts JW, Symmonds RE, et al. Open versus stereotactic breast biopsy. *Am J Surg* 1996 Nov;172(5):491-3; discussion 4-5. PMID: 8942551
86. Fuhrman G, Cederbom G, Champagne J, et al. Stereotactic core needle breast biopsy is an accurate diagnostic technique to assess nonpalpable mammographic abnormalities. *J La State Med Soc* 1996 Apr;148(4):167-70. PMID: 8935619
87. Fuhrman GM, Cederbom GJ, Bolton JS, et al. Image-guided core-needle breast biopsy is an accurate technique to evaluate patients with nonpalpable imaging abnormalities. *Ann Surg* 1998 Jun;227(6):932-9. PMID: 9637557
88. Georgian-Smith D, D'Orsi C, Morris E, et al. Stereotactic biopsy of the breast using an upright unit, a vacuum-suction needle, and a lateral arm-support system. *AJR Am J Roentgenol* 2002 Apr;178(4):1017-24. PMID: 11906893
89. Gisvold JJ, Goellner JR, Grant CS, et al. Breast biopsy: a comparative study of stereotaxically guided core and excisional techniques. *AJR Am J Roentgenol* 1994 Apr;162(4):815-20. PMID: 8140997

90. Gruber R, Jaromi S, Rudas M, et al. Histologic work-up of non-palpable breast lesions classified as probably benign at initial mammography and/or ultrasound (BI-RADS category 3). *Eur J Radiol* 2013 Mar;82(3):398-403. PMID: 22429299
91. Gruber R, Walter E, Helbich TH. Impact of stereotactic 11-g vacuum-assisted breast biopsy on cost of diagnosis in Austria. *Eur J Radiol* 2011 Jan;77(1):131-6. PMID: 19853395
92. Hahn SY, Shin JH, Han BK, et al. Sonographically-guided vacuum-assisted biopsy with digital mammography-guided skin marking of suspicious breast microcalcifications: comparison of outcomes with stereotactic biopsy in Asian women. *Acta Radiol* 2011 Feb 1;52(1):29-34. PMID: 21498322
93. Hamed H, De Freitas R, Rasbridge S, et al. A prospective randomized study of two gauges of biopsy-cut needle in diagnosis of early breast cancer. *Breast* 1995;4(2):135-6. PMID: 855240
94. Han BK, Choe YH, Ko YH, et al. Stereotactic core-needle biopsy of non-mass calcifications: outcome and accuracy at long-term follow-up. *Korean journal of radiology : official journal of the Korean Radiological Society* 2003 Oct-Dec;4(4):217-23. PMID: 14726638
95. Head JF, Haynes AE, Elliott MC, et al. Stereotaxic localization and core needle biopsy of nonpalpable breast lesions: two-year follow-up of a prospective study. *Am Surg* 1996 Dec;62(12):1018-23. PMID: 8955240
96. Helbich TH, Mayr W, Schick S, et al. Coaxial technique: approach to breast core biopsies. *Radiology* 1997 Jun;203(3):684-90. PMID: 9169689
97. Heywang-Kobrunner SH, Heinig A, Hellerhoff K, et al. Use of ultrasound-guided percutaneous vacuum-assisted breast biopsy for selected difficult indications. *The breast journal* 2009 Jul-Aug;15(4):348-56. PMID: 19500104
98. Heywang-Kobrunner SH, Schaumloffel U, Viehweg P, et al. Minimally invasive stereotaxic vacuum core breast biopsy. *Eur Radiol* 1998;8(3):377-85. PMID: 9510569
99. Huang PC, Cheung YC, Lo YF, et al. A comparison of spring-loaded and vacuum-assisted techniques for stereotactic breast biopsy of impalpable microcalcification lesions: experience at Chang Gung Memorial Hospital at Linkou. *Chang Gung medical journal* 2011 Jan-Feb;34(1):75-83. PMID: 21392477
100. Huang YC, Ho SY, Liao CY, et al. Stereotactic 11-gauge vacuum-assisted breast biopsy: Experience of 420 cases with microcalcification on mammography of non-palpable lesions. *J Radiol Sci* 2011;36(4):195-202. PMID: 21392477
101. Ioffe OB, Berg WA, Silverberg SG, et al. Mammographic-histopathologic correlation of large-core needle biopsies of the breast. *Mod Pathol* 1998 Aug;11(8):721-7. PMID: 9720499
102. Jackman RJ, Lamm RL. Stereotactic histologic biopsy in breasts with implants. *Radiology* 2002 Jan;222(1):157-64. PMID: 11756720
103. Jackman RJ, Marzoni Jr FA, Rosenberg J. False-negative diagnoses at stereotactic vacuum-assisted needle breast biopsy: long-term follow-up of 1,280 lesions and review of the literature. *AJR Am J Roentgenol* 2009;192(2):341-51. PMID: 1920499
104. Jackman RJ, Nowels KW, Rodriguez-Soto J, et al. Stereotactic, automated, large-core needle biopsy of nonpalpable breast lesions: false-negative and histologic underestimation rates after long-term follow-up. *Radiology* 1999 Mar;210(3):799-805. PMID: 10207484
105. Jales RM, Sarian LO, Peralta CF, et al. Complex breast masses: assessment of malignant potential based on cyst diameter. *J Ultrasound Med* 2012 Apr;31(4):581-7. PMID: 22441915
106. Johnson AT, Henry-Tillman RS, Smith LF, et al. Percutaneous excisional breast biopsy. *Am J Surg* 2002 Dec;184(6):550-4; discussion 4. PMID: 12488164

107. Kettritz U, Morack G, Decker T. Stereotactic vacuum-assisted breast biopsies in 500 women with microcalcifications: radiological and pathological correlations. *Eur J Radiol* 2005 Aug;55(2):270-6. PMID: 16036159
108. Kettritz U, Rotter K, Schreer I, et al. Stereotactic vacuum-assisted breast biopsy in 2874 patients: a multicenter study. *Cancer* 2004 Jan 15;100(2):245-51. PMID: 14716757
109. Khattar SC, Torp-Pedersen S, Horn T, et al. Ultrasound-guided biopsy of palpable breast masses. *Eur J Ultrasound* 1997;6(1):1-7. PMID:
110. Kirshenbaum K, Keppke A, Hou K, et al. Reassessing specimen number and diagnostic yield of ultrasound guided breast core biopsy. *The breast journal* 2012 Sep;18(5):464-9. PMID: 22775971
111. Kirshenbaum KJ, Voruganti T, Overbeeke C, et al. Stereotactic core needle biopsy of nonpalpable breast lesions using a conventional mammography unit with an add-on device. *AJR Am J Roentgenol* 2003 Aug;181(2):527-31. PMID: 12876040
112. Kirwan SE, Denton ER, Nash RM, et al. Multiple 14G stereotactic core biopsies in the diagnosis of mammographically detected stellate lesions of the breast. *Clin Radiol* 2000 Oct;55(10):763-6. PMID: 11052877
113. Koskela AK, Sudah M, Berg MH, et al. Add-on device for stereotactic core-needle breast biopsy: how many biopsy specimens are needed for a reliable diagnosis? *Radiology* 2005 Sep;236(3):801-9. PMID: 16020555
114. Lai JT, Burrowes P, MacGregor JH. Diagnostic accuracy of a stereotactically guided vacuum-assisted large-core breast biopsy program in Canada. *Can Assoc Radiol J* 2001 Aug;52(4):223-7. PMID: 11512293
115. Latosinsky S, Cornell D, Bear HD, et al. Evaluation of stereotactic core needle biopsy (SCNB) of the breast at a single institution. *Breast Cancer Res Treat* 2000 Apr;60(3):277-83. PMID: 10930116
116. Levin MF, Papoff WJ, Doan L, et al. Stereotaxic percutaneous core biopsy versus surgical biopsy of nonpalpable breast lesions using a standard mammographic table with an add-on device. *Can Assoc Radiol J* 2001 Feb;52(1):29-32. PMID: 11247262
117. Liberman L, Dershaw DD, Glassman JR, et al. Analysis of cancers not diagnosed at stereotactic core breast biopsy. *Radiology* 1997 Apr;203(1):151-7. PMID: 9122384
118. Liberman L, Ernberg LA, Heerd A, et al. Palpable breast masses: is there a role for percutaneous imaging-guided core biopsy? *AJR Am J Roentgenol* 2000 Sep;175(3):779-87. PMID: 10954467
119. Liberman L, Feng TL, Dershaw DD, et al. US-guided core breast biopsy: use and cost-effectiveness. *Radiology* 1998 Sep;208(3):717-23. PMID: 9722851
120. Liberman L, Kaplan JB, Morris EA, et al. To excise or to sample the mammographic target: what is the goal of stereotactic 11-gauge vacuum-assisted breast biopsy? *AJR Am J Roentgenol* 2002 Sep;179(3):679-83. PMID: 12185043
121. Liberman L, Sama MP. Cost-effectiveness of stereotactic 11-gauge directional vacuum-assisted breast biopsy. *AJR Am J Roentgenol* 2000 Jul;175(1):53-8. PMID: 10882245
122. Lin WC, Hsu GC, Yu CP, et al. Value of sonographically guided needle sampling of cystic versus solid components in the diagnosis of complex cystic breast masses. *Acta Radiol* 2009 Jul;50(6):595-601. PMID: 19452336
123. Lomoschitz FM, Helbich TH, Rudas M, et al. Stereotactic 11-gauge vacuum-assisted breast biopsy: influence of number of specimens on diagnostic accuracy. *Radiology* 2004 Sep;232(3):897-903. PMID: 15273332

124. Luparia A, Durando M, Campanino P, et al. Efficacy and cost-effectiveness of stereotactic vacuum-assisted core biopsy of nonpalpable breast lesions: analysis of 602 biopsies performed over 5 years. *La Radiologia medica* 2011 Apr;116(3):477-88. PMID: 21225359
125. Mainiero MB, Philpotts LE, Lee CH, et al. Stereotaxic core needle biopsy of breast microcalcifications: correlation of target accuracy and diagnosis with lesion size. *Radiology* 1996 Mar;198(3):665-9. PMID: 8628852
126. Makkun S, Prueksadee J, Chayakulkheeree J, et al. The accuracy of ultrasound guided 14-gauge core needle breast biopsy: Correlation with surgical excision or long term follow-up. *Asian Pacific Journal of Tropical Disease* 2011;222-6. PMID:
127. Makoske T, Preletz R, Riley L, et al. Long-term outcomes of stereotactic breast biopsies. *Am Surg* 2000 Dec;66(12):1104-8; discussion 8-9. PMID: 11149580
128. March DE, Coughlin BF, Barham RB, et al. Breast masses: removal of all US evidence during biopsy by using a handheld vacuum-assisted device--initial experience. *Radiology* 2003 May;227(2):549-55. PMID: 12676972
129. Margolin FR, Leung JW, Jacobs RP, et al. Percutaneous imaging-guided core breast biopsy: 5 years' experience in a community hospital. *AJR Am J Roentgenol* 2001 Sep;177(3):559-64. PMID: 11517047
130. McMahan AJ, Lutfy AM, Matthew A, et al. Needle core biopsy of the breast with a spring-loaded device. *Br J Surg* 1992 Oct;79(10):1042-5. PMID: 1422715
131. Meloni GB, Becchere MP, Soro D, et al. Percutaneous vacuum-assisted core breast biopsy with upright stereotactic equipment. Indications, limitations and results. *Acta Radiol* 2002 Nov;43(6):575-8. PMID: 12485254
132. Meyer JE, Christian RL, Lester SC, et al. Evaluation of nonpalpable solid breast masses with stereotaxic large-needle core biopsy using a dedicated unit. *AJR Am J Roentgenol* 1996 Jul;167(1):179-82. PMID: 8659367
133. Meyer JE, Smith DN, Lester SC, et al. Large-core needle biopsy of nonpalpable breast lesions. *JAMA* 1999 May 5;281(17):1638-41. PMID: 10235159
134. Mikhail RA, Nathan RC, Weiss M, et al. Stereotactic core needle biopsy of mammographic breast lesions as a viable alternative to surgical biopsy. *Ann Surg Oncol* 1994 Sep;1(5):363-7. PMID: 7850536
135. Morris EA, Liberman L, Trevisan SG, et al. Histologic heterogeneity of masses at percutaneous breast biopsy. *The breast journal* 2002 Jul-Aug;8(4):187-91. PMID: 12100109
136. Mueller-Holzner E, Frede T, Daniaux M, et al. Ultrasound-guided core needle biopsy of the breast: does frozen section give an accurate diagnosis? *Breast Cancer Res Treat* 2007 Dec;106(3):399-406. PMID: 17318378
137. Nagar S, Iacco A, Riggs T, et al. An analysis of fine needle aspiration versus core needle biopsy in clinically palpable breast lesions: a report on the predictive values and a cost comparison. *Am J Surg* 2012 Aug;204(2):193-8. PMID: 22464444
138. Nguyen M, McCombs MM, Ghandehari S, et al. An update on core needle biopsy for radiologically detected breast lesions. *Cancer* 1996 Dec 1;78(11):2340-5. PMID: 8941004
139. Ohsumi S, Taira N, Takabatake D, et al. Breast biopsy for mammographically detected nonpalpable lesions using a vacuum-assisted biopsy device (Mammotome) and upright-type stereotactic mammography unit without a digital imaging system: experience of 500 biopsies. *Breast Cancer* 2012 Apr 5. PMID: 22477266

140. Ohsumi S, Takashima S, Aogi K, et al. Breast biopsy for mammographically detected non-palpable lesions using a vacuum-assisted biopsy device (Mammotome) and an upright-type stereotactic mammography unit. *Jpn J Clin Oncol* 2001 Nov;31(11):527-31. PMID: 11773259
141. Parker SH, Burbank F, Jackman RJ, et al. Percutaneous large-core breast biopsy: a multi-institutional study. *Radiology* 1994 Nov;193(2):359-64. PMID: 7972743
142. Parker SH, Jobe WE, Dennis MA, et al. US-guided automated large-core breast biopsy. *Radiology* 1993 May;187(2):507-11. PMID: 8475299
143. Parker SH, Lovin JD, Jobe WE, et al. Stereotactic breast biopsy with a biopsy gun. *Radiology* 1990 Sep;176(3):741-7. PMID: 2167501
144. Parkin E, Hindocha S, Subar D, et al. An initial experience with rapid microwave processing in the one-stop breast clinic. *World J Surg* 2010 Dec;34(12):3036-41. PMID: 20703464
145. Perez-Fuentes JA, Longobardi IR, Acosta VF, et al. Sonographically guided directional vacuum-assisted breast biopsy: preliminary experience in Venezuela. *AJR Am J Roentgenol* 2001 Dec;177(6):1459-63. PMID: 11717107
146. Peters N, Hoorntje L, Mali W, et al. Diagnostic performance of stereotactic large core needle biopsy for nonpalpable breast lesions in routine clinical practice. *Int J Cancer* 2008 Jan 15;122(2):468-71. PMID: 17935136
147. Peters NHGM, Meeuwis C, Bakker CJG, et al. Feasibility of MRI-guided large-core-needle biopsy of suspicious breast lesions at 3T. *Eur Radiol* 2009;19:1639-44. PMID: 19111111
148. Pettine S, Place R, Babu S, et al. Stereotactic breast biopsy is accurate, minimally invasive, and cost effective. *Am J Surg* 1996 May;171(5):474-6. PMID: 8651388
149. Pfarl G, Helbich TH, Riedl CC, et al. Stereotactic 11-gauge vacuum-assisted breast biopsy: a validation study. *AJR Am J Roentgenol* 2002 Dec;179(6):1503-7. PMID: 12438044
150. Pfleiderer SO, Brunzlow H, Schulz-Wendtland R, et al. Two-year follow-up of stereotactically guided 9-G breast biopsy: a multicenter evaluation of a self-contained vacuum-assisted device. *Clin Imaging* 2009 Sep-Oct;33(5):343-7. PMID: 19712812
151. Pfleiderer SO, Reichenbach JR, Azhari T, et al. A manipulator system for 14-gauge large core breast biopsies inside a high-field whole-body MR scanner. *J Magn Reson Imaging* 2003 Apr;17(4):493-8. PMID: 12655591
152. Philpotts LE, Hooley RJ, Lee CH. Comparison of automated versus vacuum-assisted biopsy methods for sonographically guided core biopsy of the breast. *AJR Am J Roentgenol* 2003 Feb;180(2):347-51. PMID: 12540431
153. Pistolese CA, Ciarrapico A, Perretta T, et al. Cost-effectiveness of two breast biopsy procedures: surgical biopsy versus vacuum-assisted biopsy. *La Radiologia medica* 2012 Jun;117(4):539-57. PMID: 22020428
154. Pistolese CA, Ciarrapico AM, Della Gatta F, et al. Cost-effectiveness analysis of two vacuum-assisted breast biopsy systems: Mammotome and Vacora. *La Radiologia medica* 2009 Aug;114(5):743-56. PMID: 19484585
155. Pitre B, Baron PL, Baron LF, et al. Stereotactic core biopsy of the breast: results of one-year follow-up of 101 patients. *Am Surg* 1997 Dec;63(12):1124-7. PMID: 9393264
156. Pivoski SP, Jimenez RE, Wang WP. Ultrasound-guided diagnostic breast biopsy methodology: retrospective comparison of the 8-gauge vacuum-assisted biopsy approach versus the spring-loaded 14-gauge core biopsy approach. *World journal of surgical oncology* 2011;9:87. PMID: 21835024

157. Puglisi F, Pertoldi B, Ramello M, et al. Diagnostic accuracy of perforated compression grid approach for mammographically guided core needle biopsy of breast lesions. *Cancer Lett* 1999 Nov 15;146(2):181-8. PMID: 10656624
158. Reiner CS, Helbich TH, Rudas M, et al. Can galactography-guided stereotactic, 11-gauge, vacuum-assisted breast biopsy of intraductal lesions serve as an alternative to surgical biopsy? *Eur Radiol* 2009 Dec;19(12):2878-85. PMID: 19565246
159. Rosenblatt R, Fineberg SA, Sparano JA, et al. Stereotactic core needle biopsy of multiple sites in the breast: efficacy and effect on patient care. *Radiology* 1996 Oct;201(1):67-70. PMID: 8816522
160. Sauer G, Deissler H, Strunz K, et al. Ultrasound-guided large-core needle biopsies of breast lesions: analysis of 962 cases to determine the number of samples for reliable tumour classification. *Br J Cancer* 2005 Jan 31;92(2):231-5. PMID: 15611793
161. Schueller G, Jaromi S, Ponhold L, et al. US-guided 14-gauge core-needle breast biopsy: results of a validation study in 1352 cases. *Radiology* 2008 Aug;248(2):406-13. PMID: 18641246
162. Schulz-Wendtland R, Kramer S, Lang N, et al. Ultrasonic guided micro biopsy in mammary diagnosis: indications, technique and results. *Anticancer Res* 1998 May-Jun;18(3C):2145-6. PMID: 9703772
163. Scopa CD, Koukouras D, Spiliotis J, et al. Comparison of fine needle aspiration and Tru-Cut biopsy of palpable mammary lesions. *Cancer Detect Prev* 1996;20(6):620-4. PMID: 8939348
164. Sim LS, Kei PL. Upright stereotactic vacuum-assisted needle biopsy of suspicious breast microcalcifications. *Journal of medical imaging and radiation oncology* 2008 Aug;52(4):358-64. PMID: 18811759
165. Smith DN, Rosenfield Darling ML, Meyer JE, et al. The utility of ultrasonographically guided large-core needle biopsy: results from 500 consecutive breast biopsies. *J Ultrasound Med* 2001 Jan;20(1):43-9. PMID: 11149527
166. Smith LF, Henry-Tillman R, Mancino AT, et al. Magnetic resonance imaging-guided core needle biopsy and needle localized excision of occult breast lesions. *Am J Surg* 2001 Oct;182(4):414-8. PMID: 11720683
167. Smyth AT, Cederbom GJ. Core biopsy of breast lesions. *J La State Med Soc* 1994 Nov;146(11):499-501. PMID: 7806950
168. Soo MS, Ghate S, DeLong D. Stereotactic biopsy of noncalcified breast lesions: utility of vacuum-assisted technique compared to multipass automated gun technique. *Clin Imaging* 1999 Nov-Dec;23(6):347-52. PMID: 10899415
169. Stolier AJ. Stereotactic breast biopsy: a surgical series. *J Am Coll Surg* 1997 Sep;185(3):224-8. PMID: 9291397
170. Sutton S, Dahlstrom JE, Jain S. Stereotactic large-gauge core biopsy: its role in the diagnosis of non-palpable mammographic abnormalities presenting to a screening service. *Australas Radiol* 1997 May;41(2):103-8. PMID: 9153803
171. Tonegutti M, Girardi V. Stereotactic vacuum-assisted breast biopsy in 268 nonpalpable lesions. *La Radiologia medica* 2008 Feb;113(1):65-75. PMID: 18338128
172. Toth D, Sebo E, Sarkadi L, et al. Role of core needle biopsy in the treatment of radial scar. *Breast* 2012 Dec;21(6):761-3. PMID: 22397896
173. Uematsu T, Kasami M, Yuen S. A cluster of microcalcifications: women with high risk for breast cancer versus other women. *Breast Cancer* 2009;16(4):307-14. PMID: 19350360

174. Uematsu T, Yuen S, Kasami M, et al. Dynamic contrast-enhanced MR imaging in screening detected microcalcification lesions of the breast: is there any value? *Breast Cancer Res Treat* 2007 Jul;103(3):269-81. PMID: 17063274
175. Vag T, Pfleiderer SO, Bottcher J, et al. Ultrasound-guided breast biopsy using a 10-gauge self-contained vacuum-assisted device. *Eur Radiol* 2007 Dec;17(12):3100-2. PMID: 17639409
176. Vega A, Arrizabalaga R, Garijo F, et al. Nonpalpable breast lesion. Stereotaxic core needle aspiration biopsy with a single pass. *Acta Radiol* 1995 Mar;36(2):117-21. PMID: 7710788
177. Vega Bolivar A, Ortega Garcia E, Garijo Ayensa F. Stereotaxic core needle aspiration biopsy with multiple passes in nonpalpable breast lesions. *Acta Radiol* 1998 Jul;39(4):389-94. PMID: 9685825
178. Venkataraman S, Dialani V, Gilmore HL, et al. Stereotactic core biopsy: Comparison of 11 gauge with 8 gauge vacuum assisted breast biopsy. *Eur J Radiol* 2012 Oct;81(10):2613-9. PMID: 22127375
179. Verkooijen HM. Diagnostic accuracy of stereotactic large-core needle biopsy for nonpalpable breast disease: results of a multicenter prospective study with 95% surgical confirmation. *Int J Cancer* 2002 Jun 20;99(6):853-9. PMID: 12115488
180. Walker TM. Impalpable breast lesions: Stereotactic core biopsy with an 'add-on' unit. *Breast* 1997 June;6(3):126-31. PMID:
181. Wang HY, Jiang YX, Zhu QL, et al. Differentiation of benign and malignant breast lesions: a comparison between automatically generated breast volume scans and handheld ultrasound examinations. *Eur J Radiol* 2012 Nov;81(11):3190-200. PMID: 22386134
182. Ward SE, Taves DH, McCurdy LI. Stereotaxic core needle biopsy of breast microcalcifications obtained using a standard mammography table with an add-on unit. *Can Assoc Radiol J* 2000 Feb;51(1):10-5. PMID: 10711288
183. Weber WP, Zanetti R, Langer I, et al. Mammotome: less invasive than ABBI with similar accuracy for early breast cancer detection. *World J Surg* 2005 Apr;29(4):495-9. PMID: 15770379
184. Welle GJ, Clark M, Loos S, et al. Stereotactic breast biopsy: recumbent biopsy using add-on upright equipment. *AJR Am J Roentgenol* 2000 Jul;175(1):59-63. PMID: 10882246
185. White RR, Halperin TJ, Olson JA, Jr., et al. Impact of core-needle breast biopsy on the surgical management of mammographic abnormalities. *Ann Surg* 2001 Jun;233(6):769-77. PMID: 11371735
186. Whitman GJ, Kopans DB, McCarthy KA, et al. Coaxial core needle biopsy under mammographic guidance: indications and applications. *AJR Am J Roentgenol* 1998 Jul;171(1):67-70. PMID: 9648766
187. Wiratkapun C, Treesit T, Wibulpolprasert B, et al. Diagnostic accuracy of ultrasonography-guided core needle biopsy for breast lesions. *Singapore Med J* 2012 Jan;53(1):40-5. PMID: 22252182
188. Wong TE, Hisham AN. Core needle biopsy of palpable breast lump: the influence of needle size. *Med J Malaysia* 2003 Aug;58(3):399-404. PMID: 14750380
189. Wunderbaldinger P, Helbich TH, Partik B, et al. First experience with a new dedicated ultrasound system for computer-guided large-core breast biopsy. *Eur Radiol* 2001;11(12):2460-4. PMID: 11734940

190. Y.K. W, Y.M. H, A.S. C, et al. Management of breast fibroadenomas by ultrasound-guided vacuum-assisted biopsy- Three years experience. *Tzu Chi Med J* 2005;17(6):405-8. PMID:
191. Yeoman LJ, Michell MJ, Humphreys S, et al. Radiographically guided fine needle aspiration cytology and core biopsy in the assessment of impalpable breast lesions. *The Breast* 1996;5:41-7. PMID:
192. Yeow KM, Lo YF, Wang CS, et al. Ultrasound-guided core needle biopsy as an initial diagnostic test for palpable breast masses. *J Vasc Interv Radiol* 2001 Nov;12(11):1313-7. PMID: 11698631
193. Youk JH, Kim EK, Kim MJ, et al. Sonographically guided 14-gauge core needle biopsy of breast masses: a review of 2,420 cases with long-term follow-up. *AJR Am J Roentgenol* 2008 Jan;190(1):202-7. PMID: 18094312
194. Zannis VJ, Aliano KM. The evolving practice pattern of the breast surgeon with disappearance of open biopsy for nonpalpable lesions. *Am J Surg* 1998 Dec;176(6):525-8. PMID: 9926783
195. Zografos G, Zagouri F, Sergentanis TN, et al. Vacuum-assisted breast biopsy in nonpalpable solid breast lesions without microcalcifications: the Greek experience. *Diagn Interv Radiol* 2008 Sep;14(3):127-30. PMID: 18814132
196. Zuiani C, Mazzarella F, Londero V, et al. Stereotactic vacuum-assisted breast biopsy: results, follow-up and correlation with radiological suspicion. *La Radiologia medica* 2007 Mar;112(2):304-17. PMID: 17361368
197. Venkataraman S, Dialani V, Gilmore HL, et al. Stereotactic core biopsy: Comparison of 11 gauge with 8 gauge vacuum assisted breast biopsy. *Eur J Radiol* 2012 Oct;81(10):2613-9. PMID: 22127375
198. Rissanen TJ, Makarainen HP, Mattila SI, et al. Wire localized biopsy of breast lesions: a review of 425 cases found in screening or clinical mammography. *Clin Radiol* 1993 Jan;47(1):14-22. PMID: 8428412
199. Vitug AF, Newman LA. Complications in breast surgery. *The Surgical clinics of North America* 2007 Apr;87(2):431-51, x. PMID: 17498536
200. Zannis VJ, Aliano KM. The evolving practice pattern of the breast surgeon with disappearance of open biopsy for nonpalpable lesions. *American journal of surgery* 1998 Dec;176(6):525-8. PMID: 9926783
201. Kibil W, Hodorowicz-Zaniewska D, Kulig J. Mammotome biopsy under ultrasound control in the diagnostics and treatment of nodular breast lesions - own experience. *Polski przeglad chirurgiczny* 2012 May 1;84(5):242-6. PMID: 22763299
202. Penco S, Rizzo S, Bozzini AC, et al. Stereotactic vacuum-assisted breast biopsy is not a therapeutic procedure even when all mammographically found calcifications are removed: analysis of 4,086 procedures. *AJR Am J Roentgenol* 2010 Nov;195(5):1255-60. PMID: 20966337
203. Huber S, Wagner M, Medl M, et al. Benign breast lesions: minimally invasive vacuum-assisted biopsy with 11-gauge needles patient acceptance and effect on follow-up imaging findings. *Radiology* 2003 Mar;226(3):783-90. PMID: 12616021
204. Uematsu T, Kasami M. The use of positive core wash cytology to estimate potential risk of needle tract seeding of breast cancer: directional vacuum-assisted biopsy versus automated core needle biopsy. *Breast Cancer* 2010;17(1):61-7. PMID: 19360459

205. Michalopoulos NV, Zagouri F, Sergentanis TN, et al. Needle tract seeding after vacuum-assisted breast biopsy. *Acta Radiol* 2008 Apr;49(3):267-70. PMID: 18365811
206. Uematsu T, Kasami M. Risk of needle tract seeding of breast cancer: cytological results derived from core wash material. *Breast Cancer Res Treat* 2008 Jul;110(1):51-5. PMID: 17674195
207. Diaz LK, Wiley EL, Venta LA. Are malignant cells displaced by large-gauge needle core biopsy of the breast? *AJR Am J Roentgenol* 1999 Nov;173(5):1303-13. PMID: 10541110
208. Kwo S, Grotting JC. Does stereotactic core needle biopsy increase the risk of local recurrence of invasive breast cancer? *The breast journal* 2006 May-Jun;12(3):191-3. PMID: 16684313
209. Ishizuna K, Ota D, Okamoto J, et al. A case of mucinous carcinoma of the breast in which needle tract seeding was diagnosed by preoperative diagnostic imaging. *Breast Cancer* 2011 Oct;18(4):324-7. PMID: 19701680
210. Uriburu JL, Vuoto HD, Cogorno L, et al. Local recurrence of breast cancer after skin-sparing mastectomy following core needle biopsy: case reports and review of the literature. *The breast journal* 2006 May-Jun;12(3):194-8. PMID: 16684314
211. Chao C, Torosian MH, Boraas MC, et al. Local recurrence of breast cancer in the stereotactic core needle biopsy site: case reports and review of the literature. *The breast journal* 2001 Mar-Apr;7(2):124-7. PMID: 11328321
212. Brenner RJ, Gordon LM. Malignant seeding following percutaneous breast biopsy: documentation with comprehensive imaging and clinical implications. *The breast journal* 2011 Nov-Dec;17(6):651-6. PMID: 21906209
213. Fitzal F, Sporn EP, Draxler W, et al. Preoperative core needle biopsy does not increase local recurrence rate in breast cancer patients. *Breast Cancer Res Treat* 2006 May;97(1):9-15. PMID: 16502019
214. Chen AM, Haffty BG, Lee CH. Local recurrence of breast cancer after breast conservation therapy in patients examined by means of stereotactic core-needle biopsy. *Radiology* 2002 Dec;225(3):707-12. PMID: 12461249
215. Knight R, Horiuchi K, Parker SH, et al. Risk of needle-track seeding after diagnostic image-guided core needle biopsy in breast cancer. *JSLs* 2002 Jul-Sep;6(3):207-9. PMID: 12166757
216. King TA, Hayes DH, Cederbom GJ, et al. Biopsy technique has no impact on local recurrence after breast-conserving therapy. *The breast journal* 2001 Jan-Feb;7(1):19-24. PMID: 11348411
217. Wei X, Li Y, Zhang S, et al. Experience in large-core needle biopsy in the diagnosis of 1431 breast lesions. *Med Oncol* 2011 Jun;28(2):429-33. PMID: 20339957
218. Hansen NM, Ye X, Grube BJ, et al. Manipulation of the primary breast tumor and the incidence of sentinel node metastases from invasive breast cancer. *Arch Surg* 2004 Jun;139(6):634-9; discussion 9-40. PMID: 15197090
219. Newman EL, Kahn A, Diehl KM, et al. Does the method of biopsy affect the incidence of sentinel lymph node metastases? *The breast journal* 2006 Jan-Feb;12(1):53-7. PMID: 16409587
220. Peters-Engl C, Konstantiniuk P, Tausch C, et al. The impact of preoperative breast biopsy on the risk of sentinel lymph node metastases: analysis of 2502 cases from the Austrian Sentinel Node Biopsy Study Group. *Br J Cancer* 2004 Nov 15;91(10):1782-6. PMID: 15477859
221. Carter BA, Jensen RA, Simpson JF, et al. Benign transport of breast epithelium into axillary lymph nodes after biopsy. *Am J Clin Pathol* 2000 Feb;113(2):259-65. PMID: 10664628

222. Koo JS, Jung WH, Kim H. Epithelial displacement into the lymphovascular space can be seen in breast core needle biopsy specimens. *Am J Clin Pathol* 2010 May;133(5):781-7. PMID: 20395526
223. Bleiweiss IJ, Nagi CS, Jaffer S. Axillary sentinel lymph nodes can be falsely positive due to iatrogenic displacement and transport of benign epithelial cells in patients with breast carcinoma. *J Clin Oncol* 2006 May 1;24(13):2013-8. PMID: 16606970
224. Frank LS, Frank JL, March D, et al. Does therapeutic touch ease the discomfort or distress of patients undergoing stereotactic core breast biopsy? A randomized clinical trial. *Pain Med* 2007 Jul-Aug;8(5):419-24. PMID: 17661855
225. Zagouri F, Gounaris A, Liakou P, et al. Vacuum-assisted breast biopsy: more cores, more hematomas? *In Vivo* 2011 Jul-Aug;25(4):703-5. PMID: 21709018
226. Schaefer FK, Order BM, Eckmann-Scholz C, et al. Interventional bleeding, hematoma and scar-formation after vacuum-biopsy under stereotactic guidance: Mammotome((R))-system 11 g/8 g vs. ATEC((R))-system 12 g/9 g. *Eur J Radiol* 2012 May;81(5):e739-45. PMID: 22381441
227. Seror JY, Lesieur B, Scheuer-Niro B, et al. Predictive factors for complete excision and underestimation of one-pass en bloc excision of non-palpable breast lesions with the Intact((R)) breast lesion excision system. *Eur J Radiol* 2012 Apr;81(4):719-24. PMID: 21310570
228. Szynglarewicz B, Kasprzak P, Kornafel J, et al. Duration time of vacuum-assisted biopsy for nonpalpable breast masses: comparison between stereotactic and ultrasound-guided procedure. *Tumori* 2011 Jul-Aug;97(4):517-21. PMID: 21989443
229. Flory N, Lang EV. Distress in the radiology waiting room. *Radiology* 2011 Jul;260(1):166-73. PMID: 21474702
230. Novy DM, Price M, Huynh PT, et al. Percutaneous core biopsy of the breast: correlates of anxiety. *Acad Radiol* 2001 Jun;8(6):467-72. PMID: 11394538
231. Bugbee ME, Wellisch DK, Arnott IM, et al. Breast core-needle biopsy: clinical trial of relaxation technique versus medication versus no intervention for anxiety reduction. *Radiology* 2005 Jan;234(1):73-8. PMID: 15564386
232. Lang EV, Berbaum KS, Faintuch S, et al. Adjunctive self-hypnotic relaxation for outpatient medical procedures: a prospective randomized trial with women undergoing large core breast biopsy. *Pain* 2006 Dec 15;126(1-3):155-64. PMID: 16959427
233. Walker MS, Farria D, Schmidt M, et al. Educational intervention for women undergoing image-guided breast biopsy: results of a randomized clinical trial. *Cancer control : journal of the Moffitt Cancer Center* 2007 Oct;14(4):380-7. PMID: 17914338
234. Swan JS, Lawrence WF, Roy J. Process utility in breast biopsy. *Med Decis Making* 2006 Jul-Aug;26(4):347-59. PMID: 16855124
235. Brandon CJ, Mullan PB. Patients' perception of care during image-guided breast biopsy in a rural community breast center: communication matters. *J Cancer Educ* 2011 Mar;26(1):156-60. PMID: 21188664
236. Chun K, Velanovich V. Patient-perceived cosmesis and satisfaction after breast biopsy: comparison of stereotactic incisional, excisional, and wire-localized biopsy techniques. *Surgery* 2002 May;131(5):497-501. PMID: 12019401
237. Welle GJ, Clark M, Loos S, et al. Stereotactic breast biopsy: recumbent biopsy using add-on upright equipment. *AJR American journal of roentgenology* 2000 Jul;175(1):59-63. PMID: 10882246

238. Hui JY, Chan LK, Chan RL, et al. Prone table stereotactic breast biopsy. *Hong Kong medical journal = Xianggang yi xue za zhi / Hong Kong Academy of Medicine* 2002 Dec;8(6):447-51. PMID: 12459602
239. Duchesne N, Parker SH, Lechner MC, et al. Multicenter evaluation of a new ultrasound-guided biopsy device: Improved ergonomics, sampling and rebiopsy rates. *The breast journal* 2007 Jan-Feb;13(1):36-43. PMID: 17214791
240. Geller BM, Oppenheimer RG, Mickey RM, et al. Patient perceptions of breast biopsy procedures for screen-detected lesions. *American journal of obstetrics and gynecology* 2004 Apr;190(4):1063-9. PMID: 15118643
241. Alonso-Bartolome P, Vega-Bolivar A, Torres-Tabanera M, et al. Sonographically guided 11-G directional vacuum-assisted breast biopsy as an alternative to surgical excision: utility and cost study in probably benign lesions. *Acta Radiol* 2004 Jul;45(4):390-6. PMID: 15323390
242. Wolf R, Quan G, Calhoun K, et al. Efficiency of Core Biopsy for BI-RADS-5 Breast Lesions. *The breast journal* 2008 Sep-Oct;14(5):471-5. PMID: 18821933
243. Nagar S, Iacco A, Riggs T, et al. An analysis of fine needle aspiration versus core needle biopsy in clinically palpable breast lesions: a report on the predictive values and a cost comparison. *Am J Surg* 2012 Aug;204(2):193-8. PMID: 22464444
244. Launders JEI, Plymouth Meeting, PA). CER Breast Biopsy Project. August 12, 2008. PMID:
245. Sclafani LM, Bleznak A, Kelly T, et al. Training a new generation of breast surgeons: are we succeeding? *Ann Surg Oncol* 2012 Jun;19(6):1856-61. PMID: 22219063
246. Brenner RJ, Fajardo L, Fisher PR, et al. Percutaneous core biopsy of the breast: effect of operator experience and number of samples on diagnostic accuracy. *AJR Am J Roentgenol* 1996 Feb;166(2):341-6. PMID: 8553943
247. Holmes DR, Silverstein MJ. A minimally invasive breast biopsy clinic: an innovative way to teach breast fellows how to perform breast ultrasound and ultrasound-guided breast procedures. *American journal of surgery* 2006 Oct;192(4):439-43. PMID: 16978945
248. Michalopoulos NV, Maniou I, Zografos GC. Breast lesion excision system biopsy: the learning curve. *AJR Am J Roentgenol* 2012 Nov;199(5):W667. PMID: 23096217
249. Collins LC, Connolly JL, Page DL, et al. Diagnostic agreement in the evaluation of image-guided breast core needle biopsies: results from a randomized clinical trial. *The American journal of surgical pathology* 2004 Jan;28(1):126-31. PMID: 14707874
250. Gukas ID, Nwana EJ, Ihezue CH, et al. Tru-cut biopsy of palpable breast lesions: a practical option for pre-operative diagnosis in developing countries. *The Central African journal of medicine* 2000 May;46(5):127-30. PMID: 11210334
251. Wallis MG, Cheung S, Kearins O, et al. Non-operative diagnosis--effect on repeat-operation rates in the UK breast screening programme. *Eur Radiol* 2009 Feb;19(2):318-23. PMID: 18751983
252. Hyser MJ, Vanuno D, Mallesh A, et al. Changing patterns of care for occult breast lesions in a community teaching hospital. *Am Surg* 2000 May;66(5):438-42; discussion 42-3. PMID: 10824743
253. Rubin E, Mennemeyer ST, Desmond RA, et al. Reducing the cost of diagnosis of breast carcinoma: impact of ultrasound and imaging-guided biopsies on a clinical breast practice. *Cancer* 2001 Jan 15;91(2):324-32. PMID: 11180078

254. de Roos MA, Pijnappel RM, Groote AD, et al. Ductal carcinoma in situ presenting as microcalcifications: the effect of stereotactic large-core needle biopsy on surgical therapy. *Breast* 2004 Dec;13(6):461-7. PMID: 15563852
255. Parkin E, Hindocha S, Subar D, et al. An initial experience with rapid microwave processing in the one-stop breast clinic. *World J Surg* 2010 Dec;34(12):3036-41. PMID: 20703464
256. Landercasper J, Ellis RL, Mathiason MA, et al. A community breast center report card determined by participation in the national quality measures for breast centers program. *The breast journal* 2010 Sep-Oct;16(5):472-80. PMID: 20722650
257. Landercasper J, Linebarger JH, Ellis RL, et al. A quality review of the timeliness of breast cancer diagnosis and treatment in an integrated breast center. *J Am Coll Surg* 2010 Apr;210(4):449-55. PMID: 20347737
258. Kulkarni D, Irvine T, Reyes RJ. The use of core biopsy imprint cytology in the 'one-stop' breast clinic. *European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology* 2009 Oct;35(10):1037-40. PMID: 19268520
259. He Q, Fan X, Guan Y, et al. Percutaneous excisional biopsy of impalpable breast lesions under ultrasound visualization. *Breast* 2008 Dec;17(6):666-70. PMID: 18835778
260. Huang YC, HS, Liao CY., Hsueh C, Lee KW, Chen WL, Wen CY, Wu HK. Stereotactic 11-gauge Vacuum-assisted Breast Biopsy: Experience of 420 Cases with Microcalcification on Mammography of Non-palpable Lesions. *J Radiol Sci* 2011;36(4):195-202. PMID:
261. Lee CH, Egglin TK, Philpotts L, et al. Cost-effectiveness of stereotactic core needle biopsy: analysis by means of mammographic findings. *Radiology* 1997 Mar;202(3):849-54. PMID: 9051045
262. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Ann Intern Med* 2003 Jan 7;138(1):40-4. PMID: 12513043
263. Verkooijen HM, Peeters PH, Buskens E, et al. Diagnostic accuracy of large-core needle biopsy for nonpalpable breast disease: a meta-analysis. *Br J Cancer* 2000 Mar;82(5):1017-21. PMID: 10737383
264. Brennan ME, Turner RM, Ciatto S, et al. Ductal carcinoma in situ at core-needle biopsy: meta-analysis of underestimation and predictors of invasive breast cancer. *Radiology* 2011 Jul;260(1):119-28. PMID: 21493791
265. Liebens F, Carly B, Cusumano P, et al. Breast cancer seeding associated with core needle biopsies: a systematic review. *Maturitas* 2009 Feb 20;62(2):113-23. PMID: 19167175
266. Schueller G, Schueller-Weidekamm C, Helbich TH. Accuracy of ultrasound-guided, large-core needle breast biopsy. *Eur Radiol* 2008 Sep;18(9):1761-73. PMID: 18414872
267. Compton CC. Reporting on cancer specimens: case summaries and background documentation: A publication of the College of American Pathologists Cancer Committee; 2003.
268. Ellis IO, Humphreys S, Michell M, et al. Best Practice No 179. Guidelines for breast needle core biopsy handling and reporting in breast screening assessment. *Journal of clinical pathology* 2004 Sep;57(9):897-902. PMID: 15333647