Comparative Effectiveness Review

Comparative Effectiveness of Core Needle and Open Surgical Biopsy for the Diagnosis of Breast Lesions

September 22, 2008

Revised Draft

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. 290-02-0019

Prepared by:
ECRI Institute

Investigators
Wendy Bruening, Ph.D.
Karen Schoelles, M.D., S.M.
Jonathon Treadwell, Ph.D.
Jason Launders MSc
Joann Fontanarosa, Ph.D.
Kelley Tipton, MPH


**Disclaimer**

This draft Comparative Effectiveness Review is distributed solely for the purpose of pre-release peer review. It has not been otherwise disseminated by AHRQ. It does not represent, and should not be construed to represent, an AHRQ determination or policy.
Question 1. In women with a palpable or non-palpable breast abnormality what is the accuracy of different types of large-core breast biopsy compared with open biopsy for diagnosis?

Evidence Base ................................................................. 21

Accuracy of Core-Needle Biopsy ........................................ 21

Freehand Core-Needle Biopsies ........................................ 21

Ultrasound Guided Automated Gun Core-Needle Biopsies ........ 21
  Patient and breast lesion factors ........................................ 22
  Biopsy procedure factors .................................................. 22
  Clinician and facility factors .............................................. 22

Stereotactically-guided Automated Gun Core-Needle Biopsies ... 22
  Patient and breast lesion factors ........................................ 23
  Biopsy procedure factors .................................................. 23
  Clinician and facility factors .............................................. 23

Ultrasound-guided Vacuum-assisted Core-needle Biopsies .......... 23
  Patient and breast lesion factors ........................................ 24
  Biopsy procedure factors .................................................. 24
  Clinician and facility factors .............................................. 24

Stereotactically-guided Vacuum-assisted Core-needle Biopsies .... 24
  Patient and breast lesion factors ........................................ 24
  Biopsy procedure factors .................................................. 24
  Clinician and facility factors .............................................. 24

MRI Guided Core-Needle Biopsies .................................... 25

Perforated Compression Grid Guided Core-Needle Biopsies ........ 25

Multiple Methods .............................................................. 25
  Patient and breast lesion factors ........................................ 25
  Biopsy procedure factors .................................................. 25
  Clinician and facility factors .............................................. 26

Question 2. In women with a palpable or non-palpable breast abnormality what are the harms associated with large core breast biopsy compared to the open biopsy technique in the diagnosis of breast cancer?

Use of Pain Medications .................................................. 38

Bruising, Bleeding, and Hematomas .................................... 38
Infections ...................................................................................................................................38
Pain ............................................................................................................................................39
Vasovagal Reactions..................................................................................................................39
Time to Recovery.......................................................................................................................39
Impact of Biopsy Procedure on Usual Activities.................................................................39
Impact of Biopsy Procedure on Subsequent Mammographic Procedures ..................39
Miscellaneous Reported Harms .................................................................................................40
Dissemination of cancerous cells during the biopsy procedure .................................................40

Question 3. How do open biopsy and various large-core 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? .........................................................42

Relative Costs ............................................................................................................................42
Spared Surgical Procedures .......................................................................................................43
Procedure Preference .................................................................................................................43
Cosmetic Results ........................................................................................................................43
Physician Experience .................................................................................................................44
Availability of a Qualified Pathologist ......................................................................................44
Availability of Equipment .........................................................................................................44
Resource Usage ..........................................................................................................................44
Procedure Duration Time ...........................................................................................................44
Wait Time for Test Results ........................................................................................................45

Conclusions ....................................................................................................................................50

Key Question 1. In women with a palpable or non-palpable breast abnormality what is the accuracy of different types of large core breast biopsy compared with open biopsy for diagnosis? .........................................................................................................................50

Key Question 2. In women with a palpable or non-palpable breast abnormality what are the harms associated with different types of large core breast biopsy compared with open biopsy for diagnosis? .................................................................................................................50

Other Conclusions ......................................................................................................................51

Summary Tables ............................................................................................................................52

Previously Published Systematic Reviews ................................................................................57
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 4. Discussion</td>
<td>59</td>
</tr>
<tr>
<td>Conclusions</td>
<td>59</td>
</tr>
<tr>
<td>Limitations of the Evidence Base</td>
<td>59</td>
</tr>
<tr>
<td>Applicability</td>
<td>59</td>
</tr>
<tr>
<td>Possible Impact of Key Assumptions on the Conclusions</td>
<td>60</td>
</tr>
<tr>
<td>Correlation with Findings from Prior Systematic Reviews</td>
<td>61</td>
</tr>
<tr>
<td>Future Research Needed</td>
<td>61</td>
</tr>
<tr>
<td>References and Included Studies</td>
<td>62</td>
</tr>
<tr>
<td>List of Acronyms/Abbreviations</td>
<td>73</td>
</tr>
</tbody>
</table>
Tables
Table 1. Summary of key accuracy findings (Key Question 1) .................................................. 5
Table 2. Summary of key harms findings (Key Question 2) .................................................. 5
Table 3. Summary of impact of various factors on accuracy and harms ................................ 6
Table 4. Studies addressing Key Questions 1 and 2 .......................................................... 27
Table 5. Dissemination of cancerous cells during biopsy procedures ............................... 41
Table 7. Economic considerations .................................................................................... 46
Table 8. Key Question 3: other outcomes ........................................................................ 48
Table 9. Summary of accuracy by type of biopsy procedure .......................................... 52
Table 10. Summary of harms complicating core-needle biopsies .................................... 53
Table 11. Summary of the impact of factors on accuracy and harms ............................... 54
Table 12. Summary of economic aspects of core-needle biopsy ...................................... 56
Table 13. Summary of patient perspectives on choice of biopsy method ....................... 56
Table 14. Summary of clinician and facility factors related to core-needle biopsy ....... 56

Figures
Figure 1. Sensitivity of different types of biopsy ............................................................. 7
Figure 2. Negative likelihood ratios of different types of biopsy ...................................... 7
Figure 3. DCIS underestimation rates of different types of biopsy ................................. 8
Figure 4. ADH underestimation rates of different types of biopsy .................................. 8
Figure 5. Analytical framework ....................................................................................... 15

Appendices
Appendix A. Technical experts and peer reviewers
Appendix B. Methods of identifying the literature
Appendix C. Excluded studies
Appendix D. Data abstraction forms
Appendix E. Evidence tables
Appendix F. Data analysis
Executive Summary

Background

Breast cancer is the second most common malignancy of women, with over 180,000 new cases diagnosed each year. Survival rates depend on the stage of disease at diagnosis. Women diagnosed with early stages of breast cancer have a five-year survival rate of 100%. However, early breast cancer is asymptomatic, and the only way to detect it is by population-wide screening programs.

Mammography uses x-rays to examine the breast for calcifications, masses, or other abnormal structures. Currently most professional organizations recommend that all women older than fifty years of age receive a yearly mammogram. Some professional organizations recommend that routine breast cancer screening begin earlier, though x-ray mammography screening is less effective in younger women. Most experts believe that annual x-ray mammographic screening of all women who are between the age of 50 and 70 can reduce mortality from breast cancer.

The American College of Radiology has created a standardized system for reporting the results of mammography, the Breast Imaging Reporting and Data System (BIRADS). There are seven categories of assessment and recommendation:

0 Assessment is incomplete, and additional imaging evaluation is needed.
1 Negative. There is no appreciable abnormality to report.
2 Benign finding. Benign finding such as benign calcifications, intramammary lymph nodes and calcified fibroadenomas.
3 Probably benign finding. An abnormality that has a high probability of being benign.
4 Suspicious abnormality. Biopsy should be considered.
5 Highly suggestive of malignancy. Biopsy is very strongly recommended.
6 Confirmed diagnosis of malignancy.

After identification of an abnormality on screening mammography, women typically undergo additional imaging studies (diagnostic mammography, ultrasound, magnetic resonance imaging (MRI)) and a physical examination. If these studies suggest the abnormality may be malignant, a biopsy of the suspicious area may be recommended. Biopsy material may be obtained by fine-needle aspiration, core-needle biopsy, or open surgical procedures.

Open surgical biopsy involves removing a sample of tissue from the suspicious area through an open incision. To aid in location of a non-palpable lesion, it may be marked with a wire, dye, or carbon particles using an imaging method (mammography, ultrasound, MRI) to guide placement of the marker. The procedure may be performed under general anesthesia, sedation plus local anesthesia, or local anesthesia only. The surgeon may attempt to remove the entire lesion during the biopsy procedure (excisional biopsy) if the lesion is fairly small. After removing the tissue sample, the incision is closed with sutures.

Open surgical biopsy is the “gold standard” method of evaluating a suspicious breast lesion. While generally considered safe, it is a surgical procedure that, like all surgeries, places the patient at risk of experiencing morbidities and, in rare cases, mortality. Only 20 to 30% of women who undergo breast biopsy procedures are diagnosed with cancer. Exposing large
numbers of women who do not have cancer to invasive surgical procedures may be considered to be an undesirable medical practice. A less invasive method for evaluation of suspicious breast lesions would be preferable if it were sufficiently accurate.

A core-needle biopsy is a procedure that involves removing small samples of breast tissue through a hollow core needle inserted through the skin. Basic core-needle biopsy uses a special 11-, 14-, or 16-gauge needle (the smaller the gauge the larger the diameter of the needle). The suspicious lesion may be located by palpation or by imaging (stereotactic mammography, ultrasound, MRI). The procedure is usually performed under local anesthesia. Multiple core-needle samples may be taken from the suspicious area.

A variant on core-needle biopsy is vacuum-assisted biopsy. After locating the suspicious area by stereotactic mammography or ultrasound, the probe of the device is inserted into the suspicious area. The device uses vacuum suction to help remove tissue samples. Multiple samples may be taken from the suspicious area, but, unlike with traditional core-needle biopsy, the device need only be inserted through the skin once.

Medical indications may direct the preference of one type of breast biopsy procedure over another such as size and location of the lesion, imaging characteristics of the lesion, and likelihood of eventual surgical excision. However, other factors such as patient preferences, access, and practice and referral patterns also influence decisions about which procedure should be performed.

The large number of possible methods of performing breast biopsy can be bewildering to patients and healthcare providers alike. Which method to choose? Is a particular method clearly superior, or does the method of choice depend upon individual patient characteristics? We have performed a systematic review intended to evaluate the accuracy of different methods of performing breast biopsy, and to explore what factor(s) may impact the accuracy and possible harms of different methods of performing breast biopsy.

The chain of evidence linking better patient outcomes to the use of open biopsy after detection of a breast abnormality is firmly established. There is no need for randomized controlled trials with patient-oriented outcomes to demonstrate that patients may benefit from other types of breast biopsy. Establishing that a type of breast biopsy is safer than open surgical biopsy while being almost as accurate as open surgical biopsy is sufficient to justify its use.

Studies of diagnostic test performance compare the results of the experimental test to a reference test. The reference test is intended to measure the “true” disease status of each patient. For the diagnosis of breast cancer the “gold standard” reference test is open surgical biopsy. However, a difficulty with the use of the “gold standard” in large cohort studies of screening-detected breast abnormalities is the questionable ethics of subjecting all women with probably benign lesions to open surgery. Therefore we have chosen to use a combination of follow-up and open surgical biopsy as the reference standard for our analyses.

In our analysis of biopsy accuracy we focused on measures that evaluate the extent of false-negative errors (cancers falsely diagnosed as benign): sensitivity, negative likelihood ratio, and negative predictive value. Sensitivity is expressed as a percentage. A biopsy method with a sensitivity close to 100% will miss very few cancers. A negative likelihood ratio can be used to calculate an individual woman’s risk of having a malignancy following a “benign” diagnosis on breast biopsy. In general, the smaller the negative likelihood ratio the more accurate the
diagnostic test is in predicting the absence of disease. However, each individual woman’s post-test risk varies by her pre-test risk of malignancy. Negative predictive value applies to specific populations of women and can be used to predict how many women in that particular population do not have a malignancy following a “benign” diagnosis on core-needle biopsy. Negative predictive values vary by the prevalence of disease in each specific population and should not be applied to other populations with different prevalences of disease.

We also analyzed the “underestimation rate”. Lesions diagnosed as ductal carcinoma in situ (DCIS; a non-invasive early stage of breast cancer) by core-needle biopsy that were found to be invasive by the reference standard were counted as DCIS underestimates. Similarly, lesions diagnosed as benign atypical hyperplasia (ADH) by core-needle biopsy that were found to instead be invasive by the reference standard were counted as ADH underestimates. The underestimation rate was then calculated as the number of underestimates per number of DCIS (or ADH) diagnoses.

The goal of biopsy procedures is to provide an accurate diagnosis while minimizing complications, disfigurement, cost, and other adverse effects to the patient. Minimizing adverse effects is particularly important to consider because the majority of lesions being biopsied will be benign. The primary outcome measure is the false negative rate (missed cancers). Secondary outcomes include patient preference and satisfaction; pain; scarring; disfigurement; complications such as infection, hematoma formation, and wound healing problems; scheduling and waiting time; time off from work or family responsibilities; length of stay in a facility; additional procedures required, and costs.

Conclusions

Key Question 1. In women with a palpable or non-palpable breast abnormality what is the accuracy of different types of large core breast biopsy compared with open biopsy for diagnosis?

Our literature searches identified 104 studies of 54,393 breast lesions that met the inclusion criteria. All of the studies were diagnostic cohort studies that enrolled a population of women found to have suspicious breast abnormalities on routine screening. The women were sent for various types of breast biopsies, and the accuracy of the breast biopsy was determined by comparing the results of the breast biopsy to the results of a combination of open surgery and patient followup. The quality of the studies was evaluated with a standardized checklist; the evidence was rated as being of uniformly low quality due primarily to poor reporting of study and patient details. Our conclusions for Key Question 1 are summarized in Table 1 and Figure 1 through Figure 4. Factors affecting the accuracy are summarized in Table 3. Our key conclusions are stated below.

Stereotactically-guided vacuum-assisted core-needle biopsies have a sensitivity of 99.2% (95% CI: 97.9 to 99.7%). Strength of evidence: Low

Stereotactically-guided automated gun core-needle biopsies have a sensitivity of 97.8% (95% CI: 95.8 to 98.9%). Strength of evidence: Low

Ultrasound-guided vacuum-assisted core-needle biopsies have a sensitivity of 96.5% (95% CI: 81.2 to 99.4%). Strength of evidence: Low
Ultrasound-guided automated gun core-needle biopsies have a sensitivity of 97.6% (95% CI: 97.0 to 98.1%). Strength of evidence: Low

Freehand automated gun core-needle biopsies have a sensitivity of 85.8% (95% CI: 75.8 to 92.1%). Strength of evidence: Low

There was insufficient evidence to estimate the accuracy of MRI-guided core-needle biopsies. The included studies assumed that open surgical biopsy was 100% accurate. We obtained information about the actual accuracy of open surgical biopsy from recent review articles and therefore the strength of the evidence was not rated for conclusions about open surgical biopsy.

Key Question 2. In women with a palpable or non-palpable breast abnormality what are the harms associated with different types of large core breast biopsy compared with open biopsy for diagnosis?

We recorded the complications and harms reported by the 104 studies that met the inclusion criteria for Key Question 1. Our results are summarized in Table 2, and factors found to affect complication rates are summarized in Table 3. Severe complications following core-needle biopsy of any type are very rare, affecting fewer than 1% of procedures. Vacuum-assisted procedures may be associated with slightly more severe bleeding events than automated gun core-needle biopsies. Strength of evidence: Low. Information about harms of open surgical biopsy in the included studies was scanty and we supplemented it with information from recent review articles and therefore the strength of the evidence was not rated for conclusions about open surgical biopsy.

Key Question 3. How do open biopsy and various large-core 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?

Due to the nature of Key Question 3, we did not use formal inclusion criteria; we collected information relevant to the topic from many sources, including interviews with experts. Core-needle biopsy was generally agreed to cost less than open surgical biopsy, to consume fewer resources, and to be preferred by patients. Women were generally satisfied with the cosmetic results of core-needle procedures. Women who underwent a core-needle biopsy as their first invasive test to diagnosis a breast cancer had, on average, fewer surgical procedures than women who underwent an open biopsy procedure as their first invasive test. There was insufficient information available to evaluate the impact of equipment or pathologist availability.

Remaining issues

Well-reported retrospective chart reviews, retrospective database analyses, or prospective studies are needed to address the as-yet-unanswered questions as to what factors affect the accuracy and harms of core-needle breast biopsy. Answers to such questions are important for both patients and clinicians when faced with the decision of what type of breast biopsy is best for each individual patient.
Table 1. Summary of key accuracy findings (Key Question 1)

<table>
<thead>
<tr>
<th>Type of biopsy</th>
<th>Number of missed cancers expected for every 1,000 biopsiesa</th>
<th>Risk of malignancy following a “benign” test resultb</th>
<th>Number of malignancies expected per 1,000 biopsy diagnoses of “high risk” lesion</th>
<th>Number of invasive cancers expected per 1,000 biopsy diagnoses of DCIS</th>
<th>Strength of evidence supporting the conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open surgical</td>
<td>3 to 6</td>
<td>0 to 1%</td>
<td>0</td>
<td>0</td>
<td>Not rated</td>
</tr>
<tr>
<td>Freehand automated gun</td>
<td>24 to 73</td>
<td>3.4 to 10%</td>
<td>Insufficient data to estimate</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>US guidance automated gun</td>
<td>6 to 9</td>
<td>1 to 2%</td>
<td>202 to 364</td>
<td>245 to 493</td>
<td>Low</td>
</tr>
<tr>
<td>Stereotactic guidance automated gun</td>
<td>3 to 13</td>
<td>0.5 to 2%</td>
<td>357 to 517</td>
<td>180 to 321</td>
<td>Low</td>
</tr>
<tr>
<td>MRI guidance automated gun</td>
<td></td>
<td>Insufficient data to estimate</td>
<td></td>
<td></td>
<td>Inconclusive</td>
</tr>
<tr>
<td>US guidance vacuum-assisted</td>
<td>2 to 56</td>
<td>0.3 to 8%</td>
<td>Insufficient data to estimate</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Stereotactic guidance vacuum-assisted</td>
<td>1 to 6</td>
<td>0.1 to 1%</td>
<td>178 to 266</td>
<td>111 to 151</td>
<td>Low</td>
</tr>
</tbody>
</table>

a. For a population of women with a prevalence of malignancy of 30%, assuming a 100% specificity (no false-positives)
b. For a woman with a BIRADS 4 score following mammography expected to have an approximate pre-biopsy risk of malignancy of 30%. Note that an individual woman’s risk may be different from these estimates, depending on her own individual characteristics.
c. Primarily ADH lesions

Table 2. Summary of key harms findings (Key Question 2)

<table>
<thead>
<tr>
<th>Type of biopsy</th>
<th>Number of deaths expected for every 1,000 biopsies</th>
<th>Number of cases of severe bleeding expected for every 1,000 biopsies</th>
<th>Number of cases of hematomas requiring treatment expected for every 1,000 biopsies</th>
<th>Number of infections expected for every 1,000 biopsies</th>
<th>Strength of evidence supporting the conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open surgical</td>
<td>0</td>
<td>Insufficient data to estimate</td>
<td>20 to 100</td>
<td>38 to 63</td>
<td>Not rated</td>
</tr>
<tr>
<td>Automated gun core needle</td>
<td>0</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>Low</td>
</tr>
<tr>
<td>Vacuum-assisted core needle</td>
<td>0</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>Low</td>
</tr>
</tbody>
</table>
### Table 3. Summary of impact of various factors on accuracy and harms

<table>
<thead>
<tr>
<th>Category</th>
<th>Factor</th>
<th>Impact on Accuracy</th>
<th>Impact on Harms</th>
<th>Strength of Evidence Supporting the Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient characteristics</td>
<td>Insufficient data for any patient characteristics</td>
<td></td>
<td></td>
<td>Inconclusive</td>
</tr>
<tr>
<td>Lesion characteristics</td>
<td>Insufficient data for any lesion characteristics</td>
<td></td>
<td></td>
<td>Inconclusive</td>
</tr>
<tr>
<td>Biopsy methods</td>
<td>Patient position</td>
<td>Insufficient data</td>
<td>Vasovagal reactions occur more often in patients seated upright</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Needle gauge</td>
<td>Does not affect accuracy</td>
<td>Insufficient data</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Insufficient data for any other factor related to biopsy methods</td>
<td></td>
<td></td>
<td>Inconclusive</td>
</tr>
<tr>
<td>Clinician characteristics</td>
<td>Operator experience</td>
<td>Accuracy improves with experience</td>
<td>Insufficient data</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Insufficient data for any other factor related to clinician characteristics</td>
<td></td>
<td></td>
<td>Inconclusive</td>
</tr>
<tr>
<td>Facility type</td>
<td>Type of facility</td>
<td>Does not affect accuracy</td>
<td>Insufficient data</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>Geographic location of facility</td>
<td>Does not affect accuracy</td>
<td>Insufficient data</td>
<td>Low</td>
</tr>
</tbody>
</table>
Figure 1. Sensitivity of different types of biopsy

- Freehand automated gun
- US vacuum-assisted
- US automated gun
- Stereotactic automated gun
- Stereotactic vacuum-assisted
- Open surgical

Figure 2. Negative likelihood ratios of different types of biopsy

- Freehand automated gun
- US vacuum-assisted
- US automated gun
- Stereotactic automated gun
- Stereotactic vacuum-assisted
- Open surgical
Figure 3. DCIS underestimation rates of different types of biopsy

Figure 4. ADH underestimation rates of different types of biopsy
Chapter 1. Introduction

Background

Breast Cancer

Breast cancer is the second most common malignancy of women. The American Cancer Society estimates that in the U.S. in 2008, 67,770 women will be diagnosed with new cases of in situ cancer, 182,460 women will be newly diagnosed as having invasive breast cancer, and there will be 40,480 deaths due to this disease. In the general population, the cumulative risk of being diagnosed with breast cancer by age 70 is estimated to be 6% (lifetime risk of 13%).

Ductal carcinoma, including ductal carcinoma in situ (DCIS), is the most common malignancy of the breast. It arises within the ducts of the breast. DCIS is early breast cancer confined to the inside of the ductal system, and invasive (also called infiltrating) ductal carcinoma is a later stage that has broken through the walls of the ducts and invaded nearby tissues. Lobular carcinoma is similar to ductal carcinoma, first arising in the terminal ducts of the lobules and then invading through the walls of the ducts and into nearby tissues. Atypical lobular hyperplasia (ALH) and lobular carcinoma in situ (LCIS) are caused by abnormal cellular proliferation within the terminal ducts of the lobules. The two conditions are distinguished primarily by the degree to which the ducts are filled by cells, and some pathologists have suggested the use of the term lobular neoplasia to describe a continuum of disease from ALH to LCIS. There are no clinical or mammographic findings seen with LCIS; it is, however, often detected as an incidental finding at the time a breast biopsy is performed for other reasons. Women diagnosed with ALH or LCIS are at elevated risk of developing an invasive carcinoma in future.

Other types of breast abnormalities that have been linked to elevated risk of invasive carcinoma or a finding of associated invasive carcinoma upon excision are atypical ductal hyperplasia (ADH), papillary lesions, and radial scars.

Breast Biopsy

Initial detection of breast cancer is usually the result of lumps noticed upon physical examination or areas of abnormal density identified by x-ray screening mammography. Survival rates depend on the stage of disease at diagnosis. At stage 0 (carcinoma in situ) the five-year survival rate is 100%. The five-year survival rate for women with stage IV (cancer has spread beyond the breast) is only 27%. These observations suggest that breast cancer mortality rates can be significantly reduced by identifying cancers at earlier stages. Because early breast cancer is asymptomatic, the only way to detect it is through population-wide screening. Mammography is a widely accepted method for breast cancer screening.

Mammography uses x-rays to examine the breast for calcifications, masses, or other abnormal structures. Currently most professional organizations recommend that all women older than fifty years of age receive a yearly mammogram. Some professional organizations recommend that routine breast cancer screening begin earlier, though x-ray mammography screening is less effective in younger women. Most experts believe that annual x-ray mammographic screening of all women who are between the age of 50 and 70 can reduce mortality from breast cancer.
The American College of Radiology has created a standardized system for reporting the results of mammography, the Breast Imaging Reporting and Data System (BIRADS). There are seven categories of assessment and recommendation:

0  Assessment is incomplete, and additional imaging evaluation is needed.
1  Negative. There is no appreciable abnormality to report.
2  Benign finding. Benign finding such as benign calcifications, intramammary lymph nodes and calcified fibroadenomas.
3  Probably benign finding. An abnormality that has a high probability of being benign.
4  Suspicious abnormality. Biopsy should be considered.
5  Highly suggestive of malignancy. Biopsy is very strongly recommended.
6  Confirmed diagnosis of malignancy.

After identification of an abnormality (e.g., BIRADS 3 or higher) on screening mammography, women typically undergo additional imaging studies (e.g., diagnostic mammography, ultrasound, magnetic resonance imaging (MRI)) and a physical examination. If these studies suggest the abnormality may be malignant, a biopsy of the suspicious area may be recommended. Biopsy material may be obtained by fine-needle aspiration, core-needle biopsy, or open surgical procedures. The combination of physical examination, imaging studies, and needle biopsy is sometimes referred to as the “triple assessment”.

Open surgical biopsy involves removing a sample of tissue from the suspicious area through an open incision. To aid in location of a non-palpable lesion, it may be marked with a wire, dye, or carbon particles using an imaging method (mammography, ultrasound, MRI) to guide placement of the marker. The procedure may be performed under general anesthesia, sedation plus local anesthesia, or local anesthesia only. The surgeon may attempt to remove the entire lesion during the biopsy procedure (excisional biopsy) if the lesion is fairly small. After removing the tissue sample, the incision is closed with sutures.

Open surgical biopsy is the “gold standard” method of evaluating a suspicious breast lesion. However, it is a surgical procedure that, like all surgeries, places the patient at risk of experiencing morbidities and, in rare cases, mortality. The majority of women who undergo breast biopsy procedures do not have cancer. Lacquement et al. examined a series of 668 women who underwent biopsy, and reported that only 23% of these women were diagnosed with breast cancer after biopsy. Exposing large numbers of women who do not have cancer to invasive surgical procedures may be considered to be an undesirable medical practice. A less invasive method for evaluation of suspicious breast lesions would be preferable if it were sufficiently accurate.

A core-needle biopsy is a procedure that involves removing small samples of breast tissue through a hollow core needle inserted through the skin. Basic core-needle biopsy uses a special 11-, 14-, or 16-gauge needle (the smaller the gauge the larger the diameter of the needle). The suspicious lesion may be located by palpation or by imaging (stereotactic mammography, ultrasound, MRI). The procedure is usually performed under local anesthesia. Multiple core-needle samples may be taken from the suspicious area.

A variant on core-needle biopsy is vacuum-assisted biopsy. After locating the suspicious area by stereotactic mammography or ultrasound, the probe of the device is inserted into the suspicious
area. The device uses vacuum suction to help remove tissue samples. Multiple samples may be taken from the suspicious area, but, unlike with traditional core-needle biopsy, the device need only be inserted through the skin once. The Mammotome (manufactured by Johnson & Johnson Ethicon Endo-Surgery) is a commercially available vacuum-assisted biopsy device.

Another variant on core-needle biopsy is large core breast biopsy. Large core breast biopsy is intended to be a minimally invasive method of removing a fairly large sample of breast tissue, or even to remove an entire small lesion. After locating the suspicious area by stereotactic mammography a wire is inserted to mark the location. The device then removes a large core of breast tissue through a cannula. Sutures are required to close the skin at the entry site. There are no large-core biopsy devices commercially available in the United States at the time this report was prepared.

**Prognostic and Predictive Factors**

Pathological prognostic and predictive factors are used in clinical practice to guide treatment planning. One of the major concerns about core-needle biopsy techniques is under-sampling of important areas of the lesion. If important areas are missed, the pathology report may be misleading. Categories of prognostic and predictive factors include tumor type, histological grade, and immunophenotype of the tumor. These categories are briefly discussed below.

Tumor typing is evaluation of the basic type of the tumor, e.g., DCIS, infiltrating ductal carcinoma, medullary carcinoma, infiltrating lobular carcinoma, tubular carcinoma, mucinous carcinoma, or inflammatory breast cancer. Tumor typing of mixed-type tumors by core-needle biopsy may be incorrect due to the inability of needle biopsy to sample all parts of the tumor.

Histological grade is only used to describe invasive tumors. The grade is based on how closely cells in the sample tissue resemble normal breast tissue. Different grading systems are in use, but in general the higher the grade, the more abnormal the tissue structure and cells. Interpretation of grade from core-needle biopsy material has been reported to commonly under-estimate the grade by one level as compared to surgical specimens. Rakha and Ellis have suggested that the discrepancy is often due to the fact that core-needle samples are generally taken from the interior of the tumor and surgical specimens for grading are usually taken from the periphery of the tumor, where the most active growth is occurring.

Immunophenotype of the tumor refers to determining the status of certain biomarkers. The presence of estrogen receptors, progesterone receptors, and HER-2 overexpression are important features of tumor biology that need to be incorporated into treatment decisions. For example, estrogen receptor positive tumors may be effectively treated with hormone-blocking medications such as tamoxifen, and tumors that over-express HER-2 may be treated with trastuzumab (Herceptin) or lapatinib (Tykerb). Core-needle specimens can be utilized in tests to determine the immunophenotype of the tumor.

**Staging**

Final treatment decisions are based on the stage of the tumor. Breast cancer is most commonly staged with the American Joint Committee on Cancer (AJCC) TNM system. The T stands for tumor, and is assigned a number from 0 to 4 to describe the size and local spread of the primary tumor, determined by imaging studies such as mammography and CT scanning. The N stands for lymph nodes, and is assigned a number from 0 to 3 to indicate whether the cancer has spread to the lymph nodes and to how many lymph nodes, determined by sentinel lymph node biopsy or
axillary lymph node dissection. The M stands for metastasis, and is assigned either 0 or 1 to
indicate whether the cancer has spread to distant locations, determined by imaging studies such
as CT scanning and bone scintigraphy. Breast cancer stage may also be expressed as a number
from 0 to IV, where stage 0 is ductal carcinoma in situ (DCIS) and stage IV is metastatic
cancer.

**Negative Surgical Excision after Needle Core Biopsy**

Sometimes a core-needle biopsy specimen suggests that a tumor is present, and thus surgery is
performed, only to find no tumor present. Many experts suggest that in these cases the core-
needle biopsy procedure removed the entire tumor. This may be the case. It is also possible that
the pathology report for either procedure was incorrect, or that the open procedure missed the
lesion.

**Choice of Biopsy Method**

Medical indications may direct the preference of one type of procedure over another such as size
and location of the lesion, imaging characteristics of the lesion, and likelihood of eventual
surgical excision. However, other factors such as patient preferences, access, and practice and
referral patterns also influence decisions about which procedure should be performed.

The large number of possible methods of performing breast biopsy can be bewildering to patients
and healthcare providers alike. Which method to choose? Is a particular method clearly superior,
or does the method of choice depend upon individual patient characteristics? We have performed
a systematic review intended to evaluate the accuracy of different methods of performing breast
biopsy, and to explore what factor(s) may impact the accuracy and possible harms of different
methods of performing breast biopsy.

The goal of biopsy procedures is to provide an accurate diagnosis while minimizing
complications, disfigurement, cost, and other adverse effects to the patient. Minimizing adverse
effects is particularly important to consider because the majority of lesions being biopsied will be
benign. The primary outcome measure is the false negative rate (missed cancers). Secondary
outcomes include patient preference and satisfaction; pain; scarring; disfigurement;
complications such as infection, hematoma formation, and wound healing problems; scheduling
and waiting time; time off from work or family responsibilities; length of stay in a facility;
additional procedures required, and costs.

**Conceptual Framework**

The analytical framework (Figure 5) demonstrates the links between patients, tests, interventions,
and outcomes. The numbers on the diagram refer to the Key Questions (see next section) and
their placement on the diagram exhibits the many links separating the Key Questions from the
patient-oriented outcomes. Fryback and Thornbury have proposed a six-level model of assessing
diagnostic efficacy. Demonstration of efficacy at each lower level is logically necessary, but
not sufficient, to assure efficacy at higher levels. This systematic review is primarily concerned
with Level 2, the diagnostic accuracy of various methods of performing breast biopsies.
Scope and Key Questions

This systematic review was commissioned by the Agency for Healthcare Research and Quality (AHRQ) to address the following key questions:

1. In women with a palpable or non-palpable breast abnormality what is the accuracy of different types of large core breast biopsy compared with open biopsy for diagnosis? (The primary outcomes for determination of accuracy are missed cancers [the false negative rate] and the false positive rate.)

   1a. What factors associated with the patient and her breast abnormality influence the accuracy of different types of large core breast biopsy compared with open biopsy for diagnosis of a breast abnormality?

      Patient and lesion-associated factors include, but may not be limited to:

      Age, characteristics of lesion on mammography or other imaging, breast density, tissue type(s) and architecture of breast lesion, location of breast lesion, or other patient clinical health issues that may affect biopsy (i.e., clotting disorder).

   1b. What factors associated with the procedure itself influence the accuracy of different types of large core breast biopsy compared with open biopsy for diagnosis of a breast abnormality?

      Procedure-related factors include, but may not be limited to:

      Equipment used, gauge of large core needle used, # of cores, area/amount of specimen obtained, use of vacuum, specific device used, and use of imaging guidance (e.g., MRI, US, stereotactic techniques).

   1c. What clinician and facility factors influence the accuracy of large core breast biopsy compared with open biopsy for diagnosis of a breast abnormality?

      Clinician and facility factors include, but may not be limited to:

      Training and experience of clinicians performing the diagnostic procedure and interpreting breast specimen (e.g., specialized breast team, pathologist), annual volume of each procedure performed at facility, geographic location (where in country/world), practice setting (e.g., group, solo), facility setting (e.g., office, ambulatory surgical center, hospital).

2. In women with a palpable or non-palpable breast abnormality what are the harms associated with large core breast biopsy compared to the open biopsy technique in the diagnosis of breast cancer? (The primary outcomes for determination of harms are inconclusive findings and the re-biopsy rate, dissemination of cancerous cells along needle track, complications, patient centered outcomes including satisfaction, quality of life metrics, time to recovery, use of pain medications and subsequent false positive and false negative rate on mammography.)

   2a. What factors associated with the patient and her breast abnormality influence the harms of large core breast biopsy compared with the open biopsy technique in the diagnosis of a breast abnormality?
Patient and lesion-associated factors include, but may not be limited to:

Age, characteristics of lesion on mammography or other imaging, breast density, tissue type(s) and architecture of breast lesion, location of breast lesion, or other patient clinical health issues that may affect biopsy (i.e., clotting disorder).

2b. What factors associated with the procedure itself influence the harms of large core breast biopsy compared with the open biopsy technique in the diagnosis of a breast abnormality?

Procedure-related factors include, but may not be limited to:

Equipment used, Gauge of large core needle used, # of cores, area/amount of specimen obtained, use of vacuum, specific device used, and use of imaging guidance (e.g., MRI, US, stereotactic techniques).

2c. What clinician and facility factors influence the harms of large core breast biopsy compared with the open biopsy technique for diagnosis of a breast abnormality?

Clinician and facility factors include, but may not be limited to:

Training and experience of clinicians performing the diagnostic procedure and interpreting breast specimen (e.g., specialized breast team, pathologist), annual volume of each procedure performed at facility, geographic location (where in country/world), practice setting (e.g., group, solo), facility setting (e.g., office, ambulatory surgical center, hospital)

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

This report focuses on the use of core-needle biopsies to evaluate suspected cancer confined to the breast. Fine-needle aspiration is outside the scope of this report. Other uses of biopsy for diagnosing and managing breast cancer, or any other issue not mentioned in the Key Questions, are outside the scope of this report.
Figure 5. Analytical framework

1. Women referred for biopsy after detection of a breast abnormality

2. Adverse events related to biopsy procedure

3. Core-needle or open biopsy?

Decision-making

- Treat or followup or return to routine screening?

- Change in clinical decisions

- Treatment, further testing

- Clear surgical margins
- Response to treatment
- Cosmetic results

Intermediate outcomes

Patient management

Patient-oriented outcomes

Survival Quality of life
Chapter 2. Methods

In response to Section 1013 of the Medicare Modernization Act, AHRQ requested an evidence report to synthesize the evidence on the comparative effectiveness of core needle and open surgical biopsy for diagnosis of breast cancer. ECRI Institute established a team and a work plan to develop the evidence report. The process consisted of recruiting experts, formulating and refining the key questions, performing a comprehensive literature search, abstracting data, constructing evidence tables, synthesizing the data, and submitting the report for peer review.

Topic Development

The topic for this report was nominated in a public process. At the beginning of the project, AHRQ recruited a panel of technical experts to give input on key steps including the selection and refinement of the questions to be examined. The expert panel membership is provided in Appendix A.

The technical experts and representatives of AHRQ developed the Key Questions that are presented in the Scope and Key Questions section of the Introduction. Draft Key Questions were posted on a public Web site for additional feedback.

Search Strategy

The medical literature was searched from December 1990 through May 1, 2008. The full strategy is provided in Appendix B. In brief, we searched 14 external and internal databases, including PubMed and EMBASE, for clinical trials addressing the Key Questions. To supplement the electronic searches, we also examined the bibliographies/reference lists of included studies, recent narrative reviews, and scanned the content of new issues of selected journals and selected relevant gray literature sources.

Study Selection

We selected the studies that we consider in this report using a priori inclusion criteria. Arriving at these criteria before beginning the analysis is one way of reducing bias. Some of the criteria we employed are geared towards ensuring that we used only the most reliable evidence. Therefore, some of our criteria are based on study design. For similar reasons, we developed other criteria to ensure that the evidence is not derived from unusual patients or interventions, and/or outmoded technologies.

Studies of diagnostic test performance compare results of the experimental test to a reference test. The reference test is intended to measure the “true” disease status of each patient. It is important that the results of the reference test be very close to the truth, or the performance of the experimental test will be poorly estimated. For the diagnosis of breast cancer, the “gold standard” reference test is open surgical biopsy. However, a difficulty with the use of the “gold standard” in large cohort studies of screening-detected breast abnormalities is the questionable ethics of subjecting women with probably benign lesions to open surgical biopsy. Restricting the evidence base to studies that used open surgery as the reference standard for all enrolled subjects would eliminate the majority of the evidence. Therefore, we have chosen to use a combination of follow-up and open surgical biopsy as the reference standard for our analysis.

For Key Question 1 we used the following formal criteria to determine which studies would be included in our analysis. Many of our inclusion criteria for Key Question 1 were intended to
reduce the potential for spectrum bias. Spectrum bias refers to the fact that diagnostic test performance is not constant across populations with different spectrums of disease. For example, patients presenting with severe symptoms of disease may be easier to diagnose than asymptomatic patients in a screening population; and a diagnostic test that performs well to diagnose the former population may perform poorly to diagnose the latter population. The results of our analysis are intended to apply to a general population of women at average risk of breast cancer participating in routine breast cancer screening programs, and therefore many of our inclusion criteria are intended to eliminate studies that enrolled “other” types of populations.

1. The study must have directly compared large core-needle biopsy to open surgical biopsy or patient follow-up for six months or longer in the same group of patients. 
   
   Although it is possible to estimate diagnostic accuracy from a two-group trial, the results of such indirect comparisons must be viewed with great caution. Diagnostic cohort studies, wherein each patient acts as her own control, are the preferred study design for evaluating the accuracy of a diagnostic test.\(^\text{18}\) Retrospective case-control studies and case reports were excluded. Retrospective case-control studies have been shown to overestimate the accuracy of diagnostic tests, and case reports often report unusual situations or individuals that are unlikely to yield results that are applicable to general practice.\(^\text{18}\) Retrospective chart reviews that selected cases for study on the basis of the type of lesion diagnosed by core-needle biopsy were excluded because the data such studies report cannot be used to calculate the overall diagnostic accuracy of core-needle biopsy.

2. The study enrolled female human subjects. 
   Animal studies or studies of “imaging phantoms” are outside the scope of the report. 
   Studies of breast cancer in men are outside the scope of the report.

3. The study must have enrolled patients referred for biopsy for the purpose of primary diagnosis of a breast abnormality. 
   Studies that enrolled women who were referred for biopsy after discovery of a possible breast abnormality by screening mammography or routine physical examination were included. Studies that enrolled subjects that were undergoing biopsy for any of these purposes were excluded as being out of scope of the report: breast cancer staging, evaluation for a possible recurrence of breast cancer, monitoring response to treatment, evaluation of the axillary lymph nodes, evaluation of metastatic or suspected metastatic disease, or diagnosis of types of cancer other than primary breast cancer. Studies that enrolled patients from high-risk populations such as BRCA1/2 mutation carriers are also out of scope. If a study enrolled a mixed patient population and did not report data separately, it was excluded if more than 15% of the subjects did not fall into the “primary diagnosis of women at average risk presenting with an abnormality detected on routine screening” category.

4. Fifty percent or more of the subjects must have completed the study. 
   Studies with extremely high rates of attrition are prone to bias and were excluded.
5. Study must be published in English. Moher et al. and Holenstein et al. have demonstrated that exclusion of non-English language studies from meta-analyses has little impact on the conclusions drawn.\textsuperscript{19,20} Although we recognize the possibility that requiring studies to be published in English could lead to bias, it is insufficiently likely that we cannot justify the time and cost of translations.

6. Study must be published as a peer-reviewed full article. Meeting abstracts were not included. Published meeting abstracts have not been peer-reviewed and often do not include sufficient details about experimental methods to permit one to verify that the study was well designed.\textsuperscript{21,22} In addition, it is not uncommon for abstracts that are published as part of conference proceedings to have inconsistencies when compared to the final publication of the study, or to describe studies that are never published as full articles.\textsuperscript{23-27}

7. The study must have enrolled 10 or more individuals per arm. The results of very small studies are unlikely to be applicable to general clinical practice. Small studies are unable to detect sufficient numbers of events for meaningful analyses to be performed, and are at risk of enrolling unique individuals.

8. When several sequential reports from the same patients/study are available, only outcome data from the most recent report were included. However, we used relevant data from earlier and smaller reports if the report presented pertinent data not presented in the more recent report.

9. Studies of biopsy instrumentation that are no longer commercially available were excluded. The ABBI device, the MIBB device, and SiteSelect have been discontinued by their manufacturers. Studies of the accuracy and harms related to the use of these devices are no longer clinically relevant.

10. Only studies that had a score of 5.0 or greater on the quality assessment instrument were included for data analysis. Studies with scores of 4.9 or less may be biased and cannot be considered to be reliable sources of information.

To address Question 2, we recorded any harms information reported in the studies included to address Question 1. In addition, we collected any articles, regardless of design, that addressed part of Question 2, namely the dissemination of cancer cells by the biopsy procedure. To address Question 3, we consulted a variety of information sources, including published literature, cost-effectiveness analyses, evidence-based clinical practice guidelines, published expert panel consensus statements, and consultations with experts. We did not use formal inclusion criteria for Question 3 due to the nature of the question.

The abstracts of articles identified by the literature searches were screened in duplicate for possible relevance by three research assistants. The first fifty abstracts screened by each research assistant were also screened in duplicate by the lead research analyst, and all exclusions at the
abstract level were approved by the lead research analyst. The full-length articles of studies that appeared relevant at the abstract level were then obtained and three research assistants examined the articles in duplicate to see if they met the inclusion criteria. All conflicts were resolved by the lead research analyst. The excluded articles and primary reason for exclusion are shown in Appendix C.

Data Abstraction

Standardized data abstraction forms were created and managed with the software system SRS (see Appendix D). Three research assistants abstracted the data. The first fifty articles were abstracted in duplicate. All conflicts were resolved by the lead research analyst.

Quality Assessment

Quality of the included studies was assessed by an instrument developed by ECRI Institute. This instrument is based on the QUADAS instrument. The instrument is shown in Appendix D. To estimate the quality of an individual study, we computed a normalized score on a scale from 0 to 10. A study with a score higher than 8.4 was considered to be a high-quality study; a study with a score greater than 6.7 was considered to be a moderate-quality study; and a study with a score less than 5.0 was considered to be of unacceptably low quality. To evaluate the overall strength of the evidence base for each conclusion we computed the median quality score of the studies contributing to that conclusion. If fewer than three studies provided data to address a key question or subpart of a key question we refrained from drawing a formal evidence-based conclusion and rated the strength of the evidence as “Inconclusive”.

Applicability

The issue of applicability was chiefly addressed by excluding studies that enrolled patient populations that were not a general population of asymptomatic women participating in routine breast cancer screening programs.

Data Analysis and Synthesis

Several key assumptions were made: 1) the “reference standard”, a combination of open surgery and follow-up for at least six months, was 100% accurate; 2) the pathologists diagnosing the biopsy results were 100% accurate in diagnosing the material submitted to them; and 3) core-needle diagnoses of malignancy (invasive or in situ) that could not be confirmed by an open surgical procedure were assumed to have been correct diagnoses where the lesion had been completely removed by the core-needle biopsy procedure. In addition, the majority of studies reported data on a per-lesion rather than a per-patient basis, and therefore we analyzed the data on a per-lesion basis assuming that statistical assumptions of data independence were not being violated.

We performed two primary types of analyses - a standard diagnostic accuracy analysis and an analysis of underestimation rates. For the diagnostic accuracy analysis, true negatives were defined as lesions diagnosed as benign on core-needle biopsy that were found to be benign by the reference standard; false negatives were defined as lesions diagnosed as benign on core-needle biopsy that were found to be malignant (invasive or in situ) by the reference standard; true positives were defined as lesions diagnosed as malignant (invasive or in situ) on core-needle biopsy (refer to Key Assumption #3 above), and “high risk” lesions that were found to be malignant (invasive or in situ) on the reference standard were also counted as true positives; and
false positives were defined as lesions diagnosed as “high risk” (most commonly ADH lesions) on core-needle biopsy that were found to not be malignant (invasive or in situ) by the reference standard. We meta-analyzed the data reported by the studies using a bivariate mixed-effects binomial regression model as described by Harbord et al. All such analyses were computed by the STATA statistical software package using midas. The summary likelihood ratios and Bayes theorem were used to calculate the post-test probability of having a benign or malignant lesion. In cases where the data were too heterogeneous to fit a bivariate binomial regression model, we meta-analyzed the data using a random-effects model and the software package Meta-Disc. Meta-regressions were also performed with the Meta-Disc software package.

Diagnostic tests all have a trade-off between minimizing false-negative and minimizing false-positive errors. False-positive errors that occur on core-needle biopsy are not considered to be very clinically relevant. Women who experience a false-positive error will be sent for an additional biopsy procedure, and at worst will suffer minor temporary complications. However, women who experience a false-negative error may die from cancer. In addition, because all “positive” diagnoses of malignancy on core-needle biopsy are assumed to be correct, the “true” false positive rate is artificially reduced towards 0%. Thus false-positive errors, and diagnostic test characteristics that evaluate the impact of false-positive errors (specificity, positive predictive value, positive likelihood ratio), are not relevant for evaluating this technology.

We focused on measures that evaluate the extent of false-negative errors: sensitivity, negative likelihood ratio, and negative predictive value. A biopsy method with a very high sensitivity misses very few cancers. Negative likelihood ratios can be used along with Bayes’ theorem to directly compute an individual woman’s risk of having a malignancy following a “benign” diagnosis on core-needle biopsy. In general, the smaller the negative likelihood ratio the more accurate the diagnostic test is in predicting the absence of disease. However, each individual woman’s post-test risk varies by her pre-test risk of malignancy. Simple nonograms are available for in-office use that allow clinicians to directly read individual patients’ post-test risk off a graph without having to go through the tedium of calculations. Negative predictive value applies to specific populations of women and can be used to predict how many women in that particular population do not have a malignancy following a “benign” diagnosis on core-needle biopsy. Negative predictive values vary by the prevalence of disease in each specific population and should not be applied to other populations with different prevalences of disease.

The second type of analysis we performed was an analysis of underestimation rates. Lesions diagnosed as DCIS by core-needle biopsy that were found to be invasive by the reference standard were counted as underestimates. Similarly, “high risk” (most commonly ADH lesions) that were found to be malignant (in situ or invasive) by the reference standard were counted as underestimates. The underestimation rate was then calculated as the number of underestimates per number of DCIS (or “high risk”) diagnoses. We meta-analyzed the underestimation rates with a random-effects model using the CMA software package.

We did not assess the possibility of publication bias because statistical methods developed to assess the possibility of publication bias in treatment studies have not been validated for use with studies of diagnostic accuracy.
Chapter 3. Results

Question 1. In women with a palpable or non-palpable breast abnormality what is the accuracy of different types of large-core breast biopsy compared with open biopsy for diagnosis?

Evidence Base

Our literature searches identified 1,108 potentially relevant articles. After review of the abstracts, the full-length articles of 558 of these studies were obtained and examined in full. Of these, 104 studies met the inclusion criteria. The excluded studies and primary reason for exclusion are shown in Appendix C. The studies are briefly described in Table 4. Full details about the included studies, the enrolled patients, the biopsy methods, and the characteristics of the breast lesions are shown in the evidence tables in Appendix E.

Thirty-four of the 104 studies were prospective in design. Forty-eight were conducted in the United States. Ninety were carried out in general hospitals. A total of 54,393 breast lesions were enrolled in the 104 studies. The quality of all of the studies, and of the entire evidence base, was rated as low (median score 6.1, range 5.4 to 6.8).

Accuracy of Core-Needle Biopsy

We attempted to fit a bivariate binomial regression model to the data reported by all 104 studies but the data were too heterogeneous to allow a valid model to be fitted. Due to obvious differences across studies of biopsy methods and enrolled patient populations, we did not perform further analyses on the full set of data. In the following analyses we have grouped the studies by the type of core-needle biopsy used in the study. The analyses are summarized in Table 9. Full details of the analyses and reported data are provided in Appendix F.

Freehand Core-Needle Biopsies

Five studies reported data on the accuracy of non-guided, i.e., freehand, core-needle biopsies performed with automated biopsy gun devices. We fitted a bivariate binomial model. There was very little heterogeneity in the data ($I^2 = 6.95\%$). The summary sensitivity was 85.8% (95% CI: 75.8 to 92.1%) and the summary negative likelihood ratio was 0.143 (95% CI: 0.082 to 0.250). None of the studies reported underestimation rates. Because there were only five studies we did not perform any sub-group or meta-regression analyses.

Cusick et al. noted that small lesions were more likely to be misdiagnosed.34 In contrast, Hamed et al. commented that neither tumor size nor patient age affected the accuracy of the procedure; however, tumors located in the right breast were much more likely to receive false-negative diagnoses, perhaps due to the fact that the persons performing the biopsy procedures were right handed.35 Hamed et al. also noted that operator inexperience was a key factor in misdiagnoses.35

Ultrasound Guided Automated Gun Core-Needle Biopsies

Fifteen studies of 5,686 biopsies used ultrasound guidance and an automated biopsy gun. We could not fit a bivariate binomial model due to heterogeneity. The random-effects model found a
summary sensitivity of 97.6% (95% CI: 97.0 to 98.1%) and a summary negative likelihood ratio of 0.031 (95% CI: 0.024 to 0.040). Nine of the fifteen studies reported data on atypia underestimation rates; the summary atypia underestimation rate was 0.276 (0.202 to 0.364). Ten studies reported data on DCIS underestimation rates; the summary DCIS underestimation rate was 0.360 (0.245 to 0.493).

We then proceeded to explore factors that might affect the accuracy of the biopsies by performing meta-regressions. We only performed meta-regressions if all of the studies reported information about the factor being analyzed and at least three studies were different from the rest of the studies for that factor.

**Patient and breast lesion factors**

The studies reported insufficient information about characteristics of the lesions or the patients to explore the impact of these factors on the accuracy of the biopsies.

**Biopsy procedure factors**

Only seven of the studies reported information about patient position during the procedure, and six of these seven reported the patients were supine while the seventh reported the patients were seated. All but two of the studies reported using a 14G needle.

Three of the fifteen studies verified all core-needle findings with surgery (the rest used a combination of surgery and patient followup), and six of the studies did not followup all patients for at least two years. Meta-regression did not find a statistically significant impact of methods of verification of biopsy on the accuracy of the biopsies.

One study, de Lucen et al., evaluated the impact of number of cores taken on the accuracy of the procedure. The authors of the study reported that taking more than 2 cores did not improve the accuracy of the procedure. However, Fishman et al. reported that taking more than 2 cores did improve the accuracy of the biopsy, with 4 cores being the optimal number.

**Clinician and facility factors**

All but one of the studies were performed in general hospitals. The studies were set all over the world; meta-regression did not find a statistically significant effect of geographic location on the accuracy of the biopsies. Most of the studies did not report data about the training or experience of the persons performing the biopsies.

**Stereotactic-guided Automated Gun Core-needle Biopsies**

Thirty-three studies of 7153 biopsies used stereotactic guidance and an automated biopsy gun. We were able to fit a bivariate binomial model. The summary sensitivity was 97.8% (95% CI: 95.8 to 98.9%) and the summary negative likelihood ratio was 0.022 (95% CI: 0.012 to 0.043). Twenty-three of the 33 studies reported data on atypia underestimation rates and 14 reported data on DCIS underestimation rates. The atypia underestimation rate was 0.435 (95% CI: 0.357 to 0.517) and the DCIS underestimation rate was 0.244 (95% CI: 0.180 to 0.321).

We then proceeded to explore factors that might affect the accuracy of the biopsies by performing meta-regression. We only performed meta-regressions if all of the studies reported information about the factor being analyzed and at least three studies were different from the rest of the studies for that factor.
Patient and breast lesion factors

Koskela et al. reported one false-negative out of 96 procedures performed on lesions detected as masses on mammography but zero false-negatives out of 106 procedures performed on lesions with microcalcifications. However, Walker et al. reported that the sensitivity of core-needle biopsy was much lower for microcalcifications than for any other type of lesion.

The majority of the studies appeared to have enrolled patients with only non-palpable lesions but many of the studies did not report on the palpability of the lesions. The studies reported insufficient information about other characteristics of the lesions or the patients to explore the impact of these factors on the accuracy of the biopsies.

Biopsy procedure factors

All but three of the studies used 14G needles, and meta-regression did not find a statistically significant impact of needle size on biopsy accuracy. Twenty of the studies reported the patients were prone, three reported the patients were seated, one reported the patients were in the decubitus position, one reported patients were either prone or seated, but six did not report information about patient positioning.

Nine of the studies verified all core-needle findings with surgery (the rest used a combination of surgery and patient followup), and 22 of the studies did not followup all patients for at least two years. Meta-regression did not find a statistically significant impact of methods of verification of biopsy on the accuracy of the biopsies.

Koskela et al. reported that more than three cores need to be taken from lesions before an accurate diagnosis can be made.

Clinician and facility factors

Twenty-nine of the studies were conducted at a single center. Twenty-seven of the studies were conducted in general hospitals, four were conducted in free-standing dedicated cancer centers, one was conducted in a breast cancer screening clinic, and one was conducted in multiple centers of different types. Eighteen of the studies were conducted within the United States and the rest were scattered worldwide. Meta-regressions did not find that any of these factors had a statistically significant impact on biopsy accuracy.

The majority of studies reported that radiologists performed the biopsies, but many studies did not report information about the training of the operators. Very few of the studies reported the degree of experience of the operators or their caseloads.

Ultrasound-guided Vacuum-assisted Core-needle Biopsies

Seven studies of 507 biopsies used ultrasound guidance and a vacuum-assisted device to perform breast biopsies. There was no significant heterogeneity in the data ($I^2 = 0.0\%$). We fitted a bivariate binomial model to the data. The summary sensitivity was 96.5% (95% CI: 81.2 to 99.4%) and the summary negative likelihood ratio was 0.036 (95% CI: 0.006 to 0.212). The studies reported no cases of atypia underestimation and only a single case of DCIS underestimation.

Due to the lack of heterogeneity in the data, we did not perform any meta-regressions to explore the impact of factors on accuracy. The following differences between studies do not appear to affect accuracy.
Patient and breast lesion factors

The studies reported very little information about the patients or lesions.

Biopsy procedure factors

All of the studies verified core-biopsy results by a combination of open surgery and patient followup. Only one of the studies followed up all patients for at least two years.

Five of the studies used the MammoStat device with an 11G needle; one study used a VACORA device with a 10G needle, and one study did not report information about the device or needle gauge. Four of the studies reported the patients were supine and the others did not report details of patient positioning.

Clinician and facility factors

Two of the studies were conducted in free-standing cancer centers and the others were performed in general hospitals. The studies were conducted in many different countries worldwide. The studies generally did not report information on operator training or experience.

Stereotactic-guided Vaccum-assisted Core-needle Biopsies

Twenty studies of 6,255 biopsies used stereotactic guidance and a vacuum-assisted device to perform core-needle biopsies. We were able to fit a bivariate binomial model. The summary sensitivity was 99.2% (95% CI: 97.9 to 99.7%) and the summary negative likelihood ratio was 0.009 (95% CI: 0.003 to 0.023). Eighteen studies reported information about atypia underestimation rates and 19 studies reported information about DCIS underestimation rates. The summary atypia underestimation rate was 0.219 (95% CI: 0.178 to 0.266) and the summary DCIS underestimation rate was 0.130 (95% CI: 0.111 to 0.151).

We then proceeded to explore factors that might affect the accuracy of the biopsies by performing meta-regressions. We only performed meta-regressions if all of the studies reported information about the factor being analyzed and at least three studies were different from the rest of the studies for that factor.

Patient and breast lesion factors

Two studies reported that core-needle biopsy was equally accurate for lesions with microcalcifications and lesions detected as masses on mammography.\(^\text{40,41}\)

Nine of the studies reported that all of the lesions were non-palpable but the other studies reported no information on palpability of enrolled lesions. The studies reported insufficient information about characteristics of the lesions or the patients to explore the impact of these factors on the accuracy of the biopsies.

Biopsy procedure factors

All 20 studies used the MammoStat device either exclusively or in part. Seventeen of the studies used an 11G needle; two used a 14G needle; one used either a 14G or an 11G needle; and one did not report the size of the needle. All but one of the studies used a combination of open surgery and patient followup to verify the results of the biopsies. Only three studies followed up all patients for at least 24 months. Meta-regression found that method of biopsy verification did not statistically significantly affect the accuracy of the biopsies.
The majority of the studies reported that patients were prone; two reported that patients were seated, and one did not report information about patient positioning.

Lomoschitz et al. reported that 12 cores were necessary for accurate diagnosis and taking more than 12 cores did not improve accuracy.\textsuperscript{40}

\textbf{Clinic\textup{\textendash}and facility factors}

Only two of the studies were multi-center studies. Four of the studies were conducted in free-standing dedicated cancer centers, one was conducted in an ambulatory surgical center, and the rest were conducted in general hospitals. Six of the studies were conducted in the USA and 12 were conducted in Europe. Meta-regression did not find that the type or location of facility affected the accuracy of the biopsies.

Very few of the studies reported any information about the training or experience of the persons performing the biopsies. Pfarl et al. noted that for six of the seven false-negatives that occurred in the study, the biopsy procedure had been performed by an operator who had previously performed fewer than 15 stereotactic-guided biopsies.\textsuperscript{41}

\textbf{MRI Guided Core-Needle Biopsies}

Only one study reported data on the accuracy of MRI-guided biopsies performed with automated biopsy guns.

\textbf{Perforated Compression Grid Guided Core-Needle Biopsies}

Only one study reported data on the accuracy of biopsies performed with automated biopsy guns guided by a perforated compression grid.

\textbf{Multiple Methods}

There were an additional 24 studies that used multiple biopsy methods in their studies and did not report the data for different biopsy methods separately. Some of these studies reported information relevant to this topic as discussed below.

\textbf{Patient and breast lesion factors}

Abdasaleh et al. reported that technical failures were more likely to occur with women with very dense breast tissue.\textsuperscript{42}

The authors of Ciatto et al., who used multiple methods of performing core-needle biopsy, reported the percentage of procedures that gave false-negative results by lesion type: 2.7\% palpable lesions, 2.2\% nonpalpable lesions, 2.3\% masses on mammography, 1.4\% distortions on mammography, and 2.5\% of microcalcifications.\textsuperscript{43} Cipolla et al. reported that correspondence between core-needle biopsy and surgical biopsy results was 100\% for palpable lesions but only 88\% for nonpalpable lesions.\textsuperscript{44} Fajardo reported that the sensitivity of core-needle biopsies for nonpalpable lesions and lesions with microcalcifications was 90.7\%, much lower than the 97.4\% sensitivity of core-needle biopsy for masses detected on mammography.\textsuperscript{45}

\textbf{Biopsy procedure factors}

Abdasaleh et al. reported that taking two cores instead of one increased the accuracy of the procedure.\textsuperscript{42}
Helbich et al. randomly assigned patients to be biopsied in different positions—seated upright, supine, or prone. The accuracy data were not reported separately for each group, but the authors did comment that patient position did not affect the biopsy procedure.46

**Clinician and facility factors**

Ciatto et al. reported that sensitivity of core-needle biopsies improved as the operators gained experience, from 88% in the first year of the study to 96% in the last year of the study.43
### Table 4. Studies addressing Key Questions 1 and 2

<table>
<thead>
<tr>
<th>Study</th>
<th>Type(s) Core Biopsy</th>
<th>Quality Score</th>
<th>Type of Study</th>
<th>Number of Centers</th>
<th>Care Setting</th>
<th>Country Conducted in</th>
<th>Funded by</th>
<th>Number of Lesions Enrolled</th>
<th>Followup</th>
<th>% Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peters et al. 2008</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>5.9</td>
<td>Retrospective</td>
<td>4</td>
<td>General hospital</td>
<td>Netherlands</td>
<td>NR</td>
<td>948</td>
<td>2 years</td>
<td>5%</td>
</tr>
<tr>
<td>Tonegutti and Girardi 2008</td>
<td>Stereotactic guidance vacuum-assisted 11G</td>
<td>5.9</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>Italy</td>
<td>NR</td>
<td>268</td>
<td>2 years</td>
<td>0%</td>
</tr>
<tr>
<td>Youk et al. 2008</td>
<td>US guidance automated gun 14G</td>
<td>5.9</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>South Korea</td>
<td>NR</td>
<td>4,359</td>
<td>2 years</td>
<td>44%</td>
</tr>
<tr>
<td>Ciatto et al. 2007</td>
<td>Multiple methods</td>
<td>6.1</td>
<td>Retrospective</td>
<td>1</td>
<td>Dedicated breast cancer center</td>
<td>Italy</td>
<td>Funded in part by a National Health and Medical Research Council (NHMRC) grant</td>
<td>4,035</td>
<td>1 year</td>
<td>26%</td>
</tr>
<tr>
<td>de Lucena et al. 2007</td>
<td>US guidance automated gun 14G</td>
<td>5.9</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>Brazil</td>
<td>NR</td>
<td>150</td>
<td>Immediate surgery</td>
<td>0%</td>
</tr>
<tr>
<td>Uematsu et al. 2007</td>
<td>Stereotactic guidance vacuum-assisted 11G</td>
<td>6.4</td>
<td>Prospective</td>
<td>1</td>
<td>General cancer center</td>
<td>Japan</td>
<td>NR</td>
<td>100</td>
<td>Mean: 26 months Range: 5 to 44 months</td>
<td>0%</td>
</tr>
<tr>
<td>Vag et al. 2007</td>
<td>US guidance vacuum-assisted 10G</td>
<td>5.7</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>Germany</td>
<td>NR</td>
<td>70</td>
<td>2 years</td>
<td>0%</td>
</tr>
<tr>
<td>Chapellier et al. 2006</td>
<td>Stereotactic guidance vacuum-assisted 11G</td>
<td>5.9</td>
<td>Prospective</td>
<td>1</td>
<td>General cancer center</td>
<td>France</td>
<td>NR</td>
<td>318</td>
<td>Range: 4 to 16 months</td>
<td>0%</td>
</tr>
<tr>
<td>Cipolla et al. 2006</td>
<td>Multiple methods</td>
<td>6.1</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>Italy</td>
<td>NR</td>
<td>426</td>
<td>1 year</td>
<td>0%</td>
</tr>
<tr>
<td>Dhillon et al. 2006</td>
<td>Stereotactic guidance vacuum-assisted 11G</td>
<td>6.4</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>UK</td>
<td>NR</td>
<td>150</td>
<td>Median: 48 months</td>
<td>0%</td>
</tr>
<tr>
<td>Bolivar et al. 2005</td>
<td>US guidance automated gun 14G</td>
<td>6.3</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>Spain</td>
<td>NR</td>
<td>214</td>
<td>2 years</td>
<td>5%</td>
</tr>
<tr>
<td>Crystal et al. 2005</td>
<td>US guidance automated gun 14G</td>
<td>6.1</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>Israel</td>
<td>NR</td>
<td>715</td>
<td>Median: 39 months Range: 27 to 60 months</td>
<td>0%</td>
</tr>
</tbody>
</table>
Table 4. Studies addressing key questions 1 and 2 (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Type(s) Core Biopsy</th>
<th>Quality Score</th>
<th>Type of Study</th>
<th>Number of Centers</th>
<th>Care Setting</th>
<th>Country Conducted in</th>
<th>Funded by</th>
<th>Number of Lesions Enrolled</th>
<th>Followup</th>
<th>% Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dillon et al. 2005&lt;sup&gt;56&lt;/sup&gt;</td>
<td>Multiple methods</td>
<td>5.9</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>Ireland</td>
<td>NR</td>
<td>2,427</td>
<td>Median: 24 months Range: 3 to 67 months</td>
<td>19%</td>
</tr>
<tr>
<td>Koskela et al. 2005&lt;sup&gt;38&lt;/sup&gt;</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>6.4</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>Finland</td>
<td>Kuopio University Hospital (the center it was conducted in)</td>
<td>213</td>
<td>Mean: 24 months Range: 6 to 39 months</td>
<td>4%</td>
</tr>
<tr>
<td>Sauer et al. 2005&lt;sup&gt;57&lt;/sup&gt;</td>
<td>US guidance automated gun 14G</td>
<td>6.1</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>Germany</td>
<td>NR</td>
<td>962</td>
<td>Mean: 22.2 months Median: 21 months Range: 8 to 36 months</td>
<td>13%</td>
</tr>
<tr>
<td>Weber et al. 2005&lt;sup&gt;58&lt;/sup&gt;</td>
<td>Stereotactic guidance vacuum-assisted 11G</td>
<td>6.6</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>Switzerland</td>
<td>NR</td>
<td>225</td>
<td>Median: 2.1 years Range: 0.5 to 4.4 years</td>
<td>15%</td>
</tr>
<tr>
<td>Wu et al. 2005&lt;sup&gt;59&lt;/sup&gt;</td>
<td>US guidance vacuum-assisted 11G</td>
<td>6.1</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>Taiwan</td>
<td>NR</td>
<td>113</td>
<td>1 year</td>
<td>0%</td>
</tr>
<tr>
<td>Alonso-Bartolome et al. 2004&lt;sup&gt;40&lt;/sup&gt;</td>
<td>US guidance vacuum-assisted 11G</td>
<td>6.1</td>
<td>Prospective</td>
<td>2</td>
<td>General hospital</td>
<td>Spain</td>
<td>NR</td>
<td>102</td>
<td>6 to 12 months</td>
<td>0%</td>
</tr>
<tr>
<td>Delle and Terinde 2004&lt;sup&gt;41&lt;/sup&gt;</td>
<td>US guidance automated gun 14G</td>
<td>6.1</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>Germany</td>
<td>NR</td>
<td>169</td>
<td>2 years</td>
<td>0%</td>
</tr>
<tr>
<td>Fajardo et al. 2004&lt;sup&gt;45&lt;/sup&gt;</td>
<td>Multiple methods</td>
<td>6.8</td>
<td>Prospective</td>
<td>22</td>
<td>Academic and community practice clinical sites</td>
<td>USA</td>
<td>National Cancer Institute</td>
<td>2,403</td>
<td>2 years</td>
<td>30%</td>
</tr>
<tr>
<td>Kettritz et al. 2004&lt;sup&gt;42&lt;/sup&gt;</td>
<td>Stereotactic guidance vacuum-assisted 11G</td>
<td>5.5</td>
<td>Prospective</td>
<td>5</td>
<td>General hospital</td>
<td>Germany</td>
<td>NR</td>
<td>2,893</td>
<td>Mean: 25 months Range: 6 to 67 months</td>
<td>22%</td>
</tr>
<tr>
<td>Study</td>
<td>Type(s) Core Biopsy</td>
<td>Quality Score</td>
<td>Type of Study</td>
<td>Number of Centers</td>
<td>Care Setting</td>
<td>Country Conducted in</td>
<td>Funded by</td>
<td>Number of Lesions Enrolled</td>
<td>Followup</td>
<td>% Attrition</td>
</tr>
<tr>
<td>------------------------------</td>
<td>---------------------------------------------</td>
<td>---------------</td>
<td>---------------</td>
<td>-------------------</td>
<td>--------------------</td>
<td>----------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>---------------------------</td>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>Lomoschitz et al. 2004(^{58})</td>
<td>Stereotactic guidance vacuum-assisted 11G</td>
<td>6.1</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>Austria</td>
<td>One author partially supported by both Ethicon Edonsurgery and Biopsys Medical</td>
<td>100</td>
<td>2 years</td>
<td>0%</td>
</tr>
<tr>
<td>Abdsaleh et al. 2003(^{52})</td>
<td>Multiple methods</td>
<td>6.3</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>Sweden</td>
<td>NR</td>
<td>180</td>
<td>1 year</td>
<td>21%</td>
</tr>
<tr>
<td>Ambrogetti et al. 2003(^{63})</td>
<td>Stereotactic guidance vacuum-assisted 11G</td>
<td>5.9</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>France</td>
<td>NR</td>
<td>364</td>
<td>Mean: 15.8 months Range: 6 to 36 months</td>
<td>35%</td>
</tr>
<tr>
<td>Fishman et al. 2003(^{27})</td>
<td>US guidance automated gun 14G</td>
<td>6.1</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>73</td>
<td>Mammo-graphic and US followup Median: 21 months Range: 4 to 30 months</td>
<td>33%</td>
</tr>
<tr>
<td>Han et al. 2003(^{64})</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>5.7</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>Korea</td>
<td>NR</td>
<td>271</td>
<td>At least 6 months</td>
<td>27%</td>
</tr>
<tr>
<td>Kirshenbaum et al. 2003(^{59})</td>
<td>Multiple methods</td>
<td>6.1</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>506</td>
<td>Mean: 2.1 years Range: 3 months to five years.</td>
<td>23%</td>
</tr>
<tr>
<td>March et al. 2003(^{60})</td>
<td>US guidance vacuum-assisted 11G</td>
<td>6.3</td>
<td>Prospective</td>
<td>2</td>
<td>Dedicated breast cancer center</td>
<td>USA</td>
<td>RSNA Seed Grant and the Rays of Hope charitable fund</td>
<td>34</td>
<td>6 months</td>
<td>9%</td>
</tr>
<tr>
<td>Pfleiderer et al. 2003(^{27})</td>
<td>MRI guidance automated gun 14G</td>
<td>5.9</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>Germany</td>
<td>NR</td>
<td>14</td>
<td>2 years</td>
<td>0%</td>
</tr>
<tr>
<td>Study</td>
<td>Type(s) Core Biopsy</td>
<td>Quality Score</td>
<td>Type of Study</td>
<td>Number of Centers</td>
<td>Care Setting</td>
<td>Country Conducted in</td>
<td>Funded by</td>
<td>Number of Lesions Enrolled</td>
<td>Followup</td>
<td>% Attrition</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------------------------------------------------</td>
<td>---------------</td>
<td>---------------</td>
<td>-------------------</td>
<td>--------------------</td>
<td>----------------------</td>
<td>-------------------------------</td>
<td>------------------------------</td>
<td>--------------------------------</td>
<td>--------------</td>
</tr>
</tbody>
</table>
| Philpotts et al. 2003<sup>6a</sup> | Multiple methods                                        | 5.9           | Retrospective | 1                 | General hospital   | USA                  | NR                           | 281                          | Mean: 19 months
Range: 3 to 53 months for 14G
Mean: 13 months
Range: 1 to 24 for 11G | 24%                                                        |              |
| Wong and Hisham 2003<sup>5b</sup> | Freehand automated gun 14 or 16G                        | 6.3           | Prospective   | 1                 | General hospital   | Malaysia             | NR                           | 150                          | Range: 6 to 13 months | 0%          |
| Apesteguia et al. 2002<sup>2b</sup> | Stereotactic guidance vacuum-assisted 11G                | 6.3           | Prospective   | 1                 | General hospital   | Spain                | NR                           | 132                          | 1 year                          | 0%          |
| Georgian-Smith et al. 2002<sup>2</sup> | Stereotactic guidance vacuum-assisted 11G                | 6.3           | Retrospective | 4                 | General hospital   | USA                  | NR                           | 185                          | Range: 6 to 12 months | 21%         |
| Jackman and Lamm 2002<sup>7</sup> | Multiple methods                                        | 6.3           | Retrospective | 1                 | General hospital   | USA                  | Funded in part by Biopsy Medical | 31                           | At least 6 months                  | 0%          |
| Johnson et al. 2002<sup>73</sup> | US guidance vacuum-assisted 11 or 8G                     | 6.1           | NR            | 1                 | General hospital   | USA                  | Fashion Footwear of NY       | 101                          | Mean: 9.5 months                  | 24%         |
| Liberman et al. 2002<sup>74</sup> | Stereotactic guidance vacuum-assisted 11G                | 5.9           | Retrospective | 1                 | General cancer center | USA                 | NR                           | 800                          | At least 1 year                  | 29%         |
| Meloni et al. 2002<sup>75</sup>  | Stereotactic guidance vacuum-assisted                    | 6.3           | Retrospective | 1                 | General hospital   | Italy                | NR                           | 129                          | Mean: 18.7 months Range: 14 to 26 months | 0%          |
| Morris et al. 2002<sup>76</sup>  | Stereotactic guidance vacuum-assisted 14G                | 6.1           | Prospective   | 1                 | Dedicated breast cancer center | USA              | NR                           | 21                           | Median: 46 months Range: 40-54 months | 10%         |
| Pfarl et al. 2002<sup>41</sup>   | Stereotactic guidance vacuum-assisted 11G                | 6.3           | Retrospective | 1                 | General hospital   | Austria              | NR                           | 332                          | Immediate surgery               | 4%          |
Table 4. Studies addressing key questions 1 and 2 (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Type(s) Core Biopsy</th>
<th>Quality Score</th>
<th>Type of Study</th>
<th>Number of Centers</th>
<th>Care Setting</th>
<th>Country Conducted in</th>
<th>Funded by</th>
<th>Number of Lesions Enrolled</th>
<th>Followup</th>
<th>% Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verkooijen et al. COBRA 2002</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>6.6</td>
<td>Prospective</td>
<td>5</td>
<td>General hospital</td>
<td>the Netherlands</td>
<td>Dutch National Health Insurance Fund Council</td>
<td>984</td>
<td>Immediate surgery</td>
<td>11%</td>
</tr>
<tr>
<td>Becker et al. 2001</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>6.1</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>Canada</td>
<td>NR</td>
<td>232</td>
<td>Range: 6 to 12 months</td>
<td>27%</td>
</tr>
<tr>
<td>Brenner et al. 2001</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>5.7</td>
<td>Prospective</td>
<td>7</td>
<td>Cancer centers and hospitals</td>
<td>USA</td>
<td>NR</td>
<td>1,003</td>
<td>Mean: 19.3 months Range: 0 to 36 months</td>
<td>1%</td>
</tr>
<tr>
<td>Cangiarella et al. 2001a</td>
<td>Stereotactic guidance vacuum-assisted 11G</td>
<td>5.9</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>160</td>
<td>Mean: 20.5 months Range: 6 to 35 months</td>
<td>38%</td>
</tr>
<tr>
<td>Dahlstrom and Jain 2001</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>6.1</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>Australia</td>
<td>NR</td>
<td>301</td>
<td>Range: 2.4 to 7.5 years</td>
<td>0%</td>
</tr>
<tr>
<td>Lai et al. 2001</td>
<td>Stereotactic guidance vacuum-assisted 11G</td>
<td>6.1</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>Canada</td>
<td>NR</td>
<td>673</td>
<td>Mean: 6.7 months Range: 6 to 24 months</td>
<td>29%</td>
</tr>
<tr>
<td>Levin et al. 2001</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>6.3</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>Canada</td>
<td>Physician's Services Incorporated Foundation</td>
<td>70</td>
<td>Immediate surgery</td>
<td>0%</td>
</tr>
<tr>
<td>Margolin et al. 2001</td>
<td>Multiple methods</td>
<td>6.1</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>1,333</td>
<td>Mean: 14 months Range: 6 to 24 months; missing data was collected from SEER database; at the time of accession of SEER data followup ranged from 15 to 75 months</td>
<td>3%</td>
</tr>
</tbody>
</table>
Table 4. Studies addressing key questions 1 and 2 (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Type(s) Core Biopsy</th>
<th>Quality Score</th>
<th>Type of Study</th>
<th>Number of Centers</th>
<th>Care Setting</th>
<th>Country Conducted in</th>
<th>Funded by</th>
<th>Number of Lesions Enrolled</th>
<th>Followup</th>
<th>% Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perez-Fuentes et al. 2001&lt;sup&gt;56&lt;/sup&gt;</td>
<td>US guidance vacuum-assisted 11G</td>
<td>5.4 NR</td>
<td>1</td>
<td>Dedicated breast cancer center</td>
<td>Venezuela</td>
<td>NR</td>
<td>88</td>
<td>Median: 11.1 months Range: 4 to 24 months.</td>
<td>33%</td>
<td></td>
</tr>
<tr>
<td>Smith et al. 2001&lt;sup&gt;56&lt;/sup&gt;</td>
<td>US guidance automated gun 14G</td>
<td>5.9 NR</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>500</td>
<td>Mean: 22 months Median: 14 months Range: 12 to 60 months</td>
<td>21%</td>
<td></td>
</tr>
<tr>
<td>White et al. 2001&lt;sup&gt;57&lt;/sup&gt;</td>
<td>Multiple methods</td>
<td>6.1 Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>1,042</td>
<td>Median: 29 months, at least 1 year</td>
<td>29%</td>
<td></td>
</tr>
<tr>
<td>Wunderbaldinger et al. 2001&lt;sup&gt;59&lt;/sup&gt;</td>
<td>US guidance automated gun 14G</td>
<td>6.1 Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>Austria</td>
<td>author supported by Erwin Schroedinger Auslandsstipenium of the Austrian Science Fund</td>
<td>45</td>
<td>Immediate surgery</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>Yeow et al. 2001&lt;sup&gt;59&lt;/sup&gt;</td>
<td>US guidance automated gun 14 or 16G</td>
<td>6.3 Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>China</td>
<td>NR</td>
<td>98</td>
<td>Mean: 4 years Range: 3 to 5 years</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>Beck et al. 2000&lt;sup&gt;50&lt;/sup&gt;</td>
<td>Stereotactic guidance vacuum-assisted 11G</td>
<td>6.1 NR</td>
<td>1</td>
<td>General hospital</td>
<td>Germany</td>
<td>NR</td>
<td>594</td>
<td>1 year</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>Kirwan et al. 2000&lt;sup&gt;91&lt;/sup&gt;</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>5.7 Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>UK</td>
<td>NR</td>
<td>72</td>
<td>Immediate surgery</td>
<td>13%</td>
<td></td>
</tr>
<tr>
<td>Latojsinsky et al. 2000&lt;sup&gt;79&lt;/sup&gt;</td>
<td>Multiple methods</td>
<td>6.1 Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NIH grant</td>
<td>692</td>
<td>Median: 17.2 months Range: 2.8 to 43 months</td>
<td>42%</td>
<td></td>
</tr>
<tr>
<td>Liberman et al. 2000&lt;sup&gt;93&lt;/sup&gt;</td>
<td>Multiple methods</td>
<td>5.9 Retrospective</td>
<td>1</td>
<td>General cancer center</td>
<td>USA</td>
<td>NR</td>
<td>155</td>
<td>Median: 53 months Range: 24 to 69 months</td>
<td>32%</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Studies addressing key questions 1 and 2 (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Type(s) Core Biopsy</th>
<th>Quality Score</th>
<th>Type of Study</th>
<th>Number of Centers</th>
<th>Care Setting</th>
<th>Country Conducted in</th>
<th>Funded by</th>
<th>Number of Lesions Enrolled</th>
<th>Followup</th>
<th>% Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Makoske et al. 2000</td>
<td>Multiple methods</td>
<td>6.1</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>817</td>
<td>Mean: 1.7 years</td>
<td>30%</td>
</tr>
<tr>
<td>Ward et al. 2000</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>6.3</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>Canada</td>
<td>NR</td>
<td>121</td>
<td>Mean: 16 months, Range: 4 to 36 months</td>
<td>7%</td>
</tr>
<tr>
<td>Welle et al. 2000</td>
<td>Multiple methods</td>
<td>5.4</td>
<td>Retrospective</td>
<td>3</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>225</td>
<td>Range: 6 to 24 months</td>
<td>20%</td>
</tr>
<tr>
<td>Helbich et al. 1999</td>
<td>Multiple methods</td>
<td>6.1</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>Vienna</td>
<td>Ludwig-Boltzmann Institute for Radiologic Tumor Research; one author was supported by a grant from the Max Kade Foundation</td>
<td>44</td>
<td>Immediate surgery</td>
<td>0%</td>
</tr>
<tr>
<td>Jackman et al. 1999</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>6.1</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>483</td>
<td>Median: 55 months</td>
<td>1%</td>
</tr>
<tr>
<td>Meyer et al. 1999</td>
<td>Multiple methods</td>
<td>6.1</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>1,836</td>
<td>At least 1 year</td>
<td>25%</td>
</tr>
<tr>
<td>Puglisi et al. 1999</td>
<td>Perforated compression grid automated gun 14G</td>
<td>6.3</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>Italy</td>
<td>NR</td>
<td>106</td>
<td>At least 6 months</td>
<td>1%</td>
</tr>
<tr>
<td>Soo et al. 1999</td>
<td>Multiple methods</td>
<td>6.1</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>116</td>
<td>Mean: 16 months, Range: 5 to 31 months</td>
<td>19%</td>
</tr>
<tr>
<td>Caruso et al. 1998</td>
<td>Multiple methods</td>
<td>6.3</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>Italy</td>
<td>NR</td>
<td>92</td>
<td>Immediate surgery</td>
<td>13%</td>
</tr>
<tr>
<td>Doyle et al. 1998</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>6.1</td>
<td>Retrospective</td>
<td>1</td>
<td>Dedicated breast cancer center</td>
<td>New Zealand</td>
<td>NR</td>
<td>151</td>
<td>Range: 6 to 36 months</td>
<td>11%</td>
</tr>
<tr>
<td>Fuhrman et al. 1998</td>
<td>Multiple methods</td>
<td>6.1</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>1,440</td>
<td>At least 6 months</td>
<td>18%</td>
</tr>
</tbody>
</table>
Table 4. Studies addressing key questions 1 and 2 (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Type(s) Core Biopsy</th>
<th>Quality Score</th>
<th>Type of Study</th>
<th>Number of Centers</th>
<th>Care Setting</th>
<th>Country Conducted in</th>
<th>Funded by</th>
<th>Number of Lesions Enrolled</th>
<th>Followup</th>
<th>% Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heywang-Kobrunner et al. 1998</td>
<td>Stereotactic guidance vacuum-assisted 11 or 14G</td>
<td>5.9</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>Germany</td>
<td>NR</td>
<td>261</td>
<td>6 months</td>
<td>31%</td>
</tr>
<tr>
<td>Ioffe et al. 1998</td>
<td>Multiple methods</td>
<td>5.7</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>224</td>
<td>Range: 6 to 12 months</td>
<td>14%</td>
</tr>
<tr>
<td>Liberman et al. 1998</td>
<td>US guidance automated gun 14G</td>
<td>6.1</td>
<td>NR</td>
<td>1</td>
<td>General cancer center</td>
<td>USA</td>
<td>NR</td>
<td>151</td>
<td>Median: 20 months Range: 6 to 48 months</td>
<td>23%</td>
</tr>
<tr>
<td>Schulz-Wendtland et al. 1998</td>
<td>US guidance automated gun 14G</td>
<td>6.1</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>Germany</td>
<td>NR</td>
<td>307</td>
<td>2 years</td>
<td>0%</td>
</tr>
<tr>
<td>Vega-Bolivar et al. 1998</td>
<td>Stereotactic guidance Surecut 15G</td>
<td>5.9</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>Spain</td>
<td>NR</td>
<td>182</td>
<td>Mean: 27 months Range: 6 to 47 months</td>
<td>6%</td>
</tr>
<tr>
<td>Whitman et al. 1998</td>
<td>Stereotactic guidance automated gun 16G</td>
<td>5.5</td>
<td>Retrospective</td>
<td>2</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>12</td>
<td>Immediate surgery</td>
<td>0%</td>
</tr>
<tr>
<td>Zannis and AliaNo 1998</td>
<td>Multiple methods</td>
<td>6.3</td>
<td>Retrospective</td>
<td>1</td>
<td>Ambulatory surgical center</td>
<td>USA</td>
<td>NR</td>
<td>424</td>
<td>At least 6 months</td>
<td>31%</td>
</tr>
<tr>
<td>Bauer et al. 1997</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>6.3</td>
<td>Retrospective</td>
<td>NR</td>
<td>NR</td>
<td>USA</td>
<td>NR</td>
<td>799</td>
<td>Mean: 9 months</td>
<td>0%</td>
</tr>
<tr>
<td>Britton et al. 1997</td>
<td>Multiple methods</td>
<td>6.1</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>UK</td>
<td>NR</td>
<td>202</td>
<td>Mean: 20.1 months Range: 5.3 to 30.8 months</td>
<td>2%</td>
</tr>
<tr>
<td>Helbich et al. 1997</td>
<td>Multiple methods</td>
<td>6.4</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>Vienna</td>
<td>NR</td>
<td>210</td>
<td>Immediate surgery</td>
<td>0%</td>
</tr>
<tr>
<td>Khattar et al. 1997</td>
<td>US guidance automated gun</td>
<td>5.9</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>Denmark</td>
<td>NR</td>
<td>106</td>
<td>Immediate surgery</td>
<td>43%</td>
</tr>
<tr>
<td>Liberman et al. 1997</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>5.7</td>
<td>Retrospective</td>
<td>1</td>
<td>General cancer center</td>
<td>USA</td>
<td>NR</td>
<td>442</td>
<td>Median: 18 months Range: 6 to 46 months</td>
<td>34%</td>
</tr>
</tbody>
</table>
Table 4. Studies addressing key questions 1 and 2 (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Type(s) Core Biopsy</th>
<th>Quality Score</th>
<th>Type of Study</th>
<th>Number of Centers</th>
<th>Care Setting</th>
<th>Country Conducted in</th>
<th>Funded by</th>
<th>Number of Lesions Enrolled</th>
<th>Followup</th>
<th>% Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitre et al. 1997</td>
<td>Stereotactic guidance automated gun</td>
<td>6.1</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>128</td>
<td>1 year</td>
<td>8%</td>
</tr>
<tr>
<td>Stolier et al. 1997</td>
<td>Multiple methods</td>
<td>6.1</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>244</td>
<td>Mean: 12.8 months Range: 6 to 39 months</td>
<td>NR</td>
</tr>
<tr>
<td>Sutton, et al. 1997</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>5.5</td>
<td>Retrospective</td>
<td>1</td>
<td>Screening clinic</td>
<td>Australia</td>
<td>NR</td>
<td>206</td>
<td>1 year</td>
<td>32%</td>
</tr>
<tr>
<td>Walker et al. 1997</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>6.1</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>UK</td>
<td>NR</td>
<td>200</td>
<td>Range: 6 to 36 months</td>
<td>10%</td>
</tr>
<tr>
<td>Frazee et al. 1996</td>
<td>Stereotactic guidance automated gun</td>
<td>6.3</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>103</td>
<td>At least 6 months</td>
<td>0%</td>
</tr>
<tr>
<td>Fuhrman et al. 1996</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>5.5</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>451</td>
<td>1 year</td>
<td>22%</td>
</tr>
<tr>
<td>Head and Haynes 1996</td>
<td>Stereotactic guidance automated gun 18G</td>
<td>5.9</td>
<td>Prospective</td>
<td>1</td>
<td>Dedicated breast cancer center</td>
<td>USA</td>
<td>NR</td>
<td>115</td>
<td>2 years</td>
<td>8%</td>
</tr>
<tr>
<td>Mainiero et al. 1996</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>5.9</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>138</td>
<td>At least 6 months</td>
<td>14%</td>
</tr>
<tr>
<td>Meyer et al. 1996</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>5.5</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>388</td>
<td>1 year</td>
<td>30%</td>
</tr>
<tr>
<td>Nguyen et al. 1996</td>
<td>Multiple methods</td>
<td>5.9</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>American Cancer Society, UCLA Jonsson Comprehensive Cancer Center, and the Stein-Oppenheim Foundation</td>
<td>431</td>
<td>At least 6 months</td>
<td>10%</td>
</tr>
<tr>
<td>Pettine et al. 1996</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>5.7</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>25</td>
<td>6 month repeat mammography for benign</td>
<td>0%</td>
</tr>
<tr>
<td>Study</td>
<td>Type(s) Core Biopsy</td>
<td>Quality Score</td>
<td>Type of Study</td>
<td>Number of Centers</td>
<td>Care Setting</td>
<td>Country Conducted in</td>
<td>Funded by</td>
<td>Number of Lesions Enrolled</td>
<td>Followup</td>
<td>% Attrition</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------------------------------------</td>
<td>---------------</td>
<td>---------------</td>
<td>-------------------</td>
<td>--------------</td>
<td>----------------------</td>
<td>-----------</td>
<td>--------------------------</td>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>Rosenblatt et al. 1996</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>5.9</td>
<td>Retrospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>25</td>
<td>1 year</td>
<td>16%</td>
</tr>
<tr>
<td>Scopa et al. 1996</td>
<td>Freehand TruCut</td>
<td>5.9</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>Greece</td>
<td>NR</td>
<td>120</td>
<td>Immediate surgery</td>
<td>0%</td>
</tr>
<tr>
<td>Cross et al. 1995</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>5.7</td>
<td>NR</td>
<td>1</td>
<td>Dedicated breast cancer center</td>
<td>USA</td>
<td>NR</td>
<td>250</td>
<td>1 year</td>
<td>12%</td>
</tr>
<tr>
<td>Doyle et al. 1995</td>
<td>Multiple methods</td>
<td>6.1</td>
<td>Prospective</td>
<td>1</td>
<td>General Hospital</td>
<td>USA</td>
<td>NR</td>
<td>150</td>
<td>Range: 6 to 24 months</td>
<td>3%</td>
</tr>
<tr>
<td>Hamed et al. 1995</td>
<td>Freehand Biopsy-cut</td>
<td>5.9</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>United Kingdom</td>
<td>NR</td>
<td>122</td>
<td>Immediate surgery</td>
<td>0%</td>
</tr>
<tr>
<td>Burbank et al. 1994</td>
<td>Multiple methods</td>
<td>5.5</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>105</td>
<td>At least 6 months</td>
<td>0%</td>
</tr>
<tr>
<td>Gisvold et al. 1994</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>6.1</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>160</td>
<td>Immediate surgery</td>
<td>0%</td>
</tr>
<tr>
<td>Parker et al. 1994</td>
<td>Multiple methods</td>
<td>5.4</td>
<td>Retrospective</td>
<td>20</td>
<td>Various hospitals, breast care centers, clinics</td>
<td>USA</td>
<td>NR</td>
<td>6,152</td>
<td>At least 6 months</td>
<td>39%</td>
</tr>
<tr>
<td>Smyth and Cederbom 1994</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>5.5</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>58</td>
<td>Immediate surgery</td>
<td>0%</td>
</tr>
<tr>
<td>Elvecrog et al. 1993</td>
<td>Stereotactic guidance automated gun 14G</td>
<td>6.4</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>100</td>
<td>Immediate surgery</td>
<td>0%</td>
</tr>
<tr>
<td>Parker et al. 1993</td>
<td>US guidance automated gun 14G</td>
<td>6.1</td>
<td>NR</td>
<td>1</td>
<td>Specialized imaging center</td>
<td>USA</td>
<td>NR</td>
<td>181</td>
<td>Range: 12 to 36 months</td>
<td>0%</td>
</tr>
<tr>
<td>McMahon et al. 1992</td>
<td>Multiple methods</td>
<td>6.3</td>
<td>Prospective</td>
<td>1</td>
<td>General hospital</td>
<td>UK</td>
<td>NR</td>
<td>151</td>
<td>Median: 11 months</td>
<td>0%</td>
</tr>
<tr>
<td>Hamed et al. 1991</td>
<td>Freehand automated gun 18G</td>
<td>5.9</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>UK</td>
<td>NR</td>
<td>107</td>
<td>Immediate surgery</td>
<td>0%</td>
</tr>
</tbody>
</table>
Table 4. Studies addressing key questions 1 and 2 (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Type(s) Core Biopsy</th>
<th>Quality Score</th>
<th>Type of Study</th>
<th>Number of Centers</th>
<th>Care Setting</th>
<th>Country Conducted in</th>
<th>Funded by</th>
<th>Number of Lesions Enrolled</th>
<th>Followup</th>
<th>% Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cusick et al. 1990(^\text{14})</td>
<td>Freehand</td>
<td>5.7</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>96</td>
<td>Immediate surgery</td>
<td>0%</td>
</tr>
<tr>
<td>Parker et al. 1990(^\text{13a})</td>
<td>Stereotactic guidance automated gun</td>
<td>5.7</td>
<td>NR</td>
<td>1</td>
<td>General hospital</td>
<td>USA</td>
<td>NR</td>
<td>103</td>
<td>Immediate surgery</td>
<td>0%</td>
</tr>
</tbody>
</table>

NR = Not Reported
USA = United States of America
UK = United Kingdom
Question 2. In women with a palpable or non-palpable breast abnormality what are the harms associated with large core breast biopsy compared to the open biopsy technique in the diagnosis of breast cancer?

Forty-eight of the 104 included studies did not report any harms (see Appendix F); whether this was because no harms occurred is unclear. Five studies only reported that no severe complications or harms occurred.

Tonegutti and Giradi reported that complications only occurred during the first year of performing stereotactically-guided vacuum-assisted biopsies.48

Very few of the included studies reported information about complications occurring in association with open surgical biopsy procedures. We consulted a narrative review published in 2007 to obtain further information about complications of open surgical biopsy procedures. In this review, Vitug and Newman report that 2 to 10% of breast surgeries are complicated by hematoma formation, and that 3.8% are complicated by infections.139 Rissanen et al. reviewed a series of 425 wire-localized open biopsy procedures and reported that 10.2% were complicated by vasovagal reactions.140

Use of Pain Medications

Four studies reported information on the use of pain medications. These studies reported that 100% of patients were sent home with narcotics after an open biopsy procedure, and only one patient (0.17%) required narcotics after a core-needle procedure. Twenty (3.5%) patients were reported to have required acetaminophen after a core-needle procedure.

Bruising, Bleeding, and Hematomas

Twenty four studies of 17,585 core-needle biopsy procedures reported that only 0.085% were complicated by hematomas that required treatment. Fifty-three percent of the hematomas that required treatment occurred after a vacuum-assisted procedure.

Twenty four studies of 8,474 core-needle biopsy procedures reported that less than 1% were complicated by troublesome bleeding. More than 60% of the bleeding events were reported to have occurred during a vacuum-assisted procedure.

Nine studies reported that bruising occurred after core-needle biopsy procedures. Three of the nine reported that bruising was a common event; two reported that approximately 50% of patients had bruising; and four studies reported that 45 out of 976 patients (4.6%) had severe bruising. These nine studies used a variety of different core-needle procedures. The fact that they report that bruising is common suggests that the other 95 studies, instead of not experiencing patient bruising, chose to not report this adverse event.

Infections

March et al. reported that 2.1% of open biopsy procedures were complicated by the development of an abscess, but zero abscesses complicated 234 ultrasound-guided vacuum-assisted core-needle procedures.66 Tonegutti and Giradi reported that one abscess that required surgical
treatment occurred in a series of 268 stereotactically-guided vacuum-assisted procedures. None of the other studies reported the occurrence of abscesses.

Twenty studies of 16,407 core-needle procedures reported that only 0.15% of the procedures were complicated by infections. Zannis and Aliano reported that 6.3% of open surgical biopsies were complicated by infections.\textsuperscript{111}

**Pain**

Three vacuum-assisted biopsy procedures were reported to have been terminated after patients complained of severe pain. No other types of biopsy procedures were reported to have terminated due to patient complaints of pain. Seventeen studies of a wide variety of biopsy methods reported information about patient pain during the procedure, and overall only 1.7% of patients were reported to have experienced severe pain.

Frazee et al. reported the mean pain score (10-point VAS scale) was 2.5 for open biopsy procedures and 2.8 for stereotactically-guided automated gun core-needle biopsies (the difference was not statistically significant).\textsuperscript{119}

Wong and Hisham reported no difference in the amount of pain experienced by patients undergoing a 14G core-needle procedure vs. a 16G core-needle procedure.\textsuperscript{69} McMahon et al. reported that patients undergoing 18G core-needle procedures had significantly less pain than patients undergoing 14G core-needle procedures, but there was no significant difference in pain between 14G and 16G procedures.\textsuperscript{37}

**Vasovagal Reactions**

Twenty-one studies of 7,526 core-needle procedures reported that 1% were complicated by vasovagal reactions (fainting). More than 40% of the vasovagal reactions occurred in patients who were positioned upright sitting for the biopsy procedure.

Kirshenbaum et al. commented that the majority of vasovagal reactions occurred when inexperienced operators performed the biopsy procedures.\textsuperscript{65}

**Time to Recovery**

One study, Frazee et al., reported information about time to recovery. 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.\textsuperscript{119}

**Impact of Biopsy Procedure on Usual Activities**

One study, March et al., reported that ultrasound-guided vacuum-assisted procedures did not impact at all the usual activities of 47% of the women.\textsuperscript{66}

**Impact of Biopsy Procedure on Subsequent Mammographic Procedures**

Three studies reported information about the impact of core-needle biopsies on subsequent mammographic examinations. All three studies performed stereotactic-guided vacuum-assisted core-needle procedures. These three studies enrolled 3,748 patients of whom 3,345 (89.2%) were reported to have no mammographically visible scarring after the biopsy procedures. Only seven
of the patients (0.19%) were reported to have scars that were potentially diagnostically confusing on subsequent mammographic procedures.

**Miscellaneous Reported Harms**

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 method of performing the core-needle biopsies.

Two studies reported that one patient per study (out of 3,487 patients) 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 185 stereotactic-guided vacuum-assisted procedures reported that one patient vomited during the procedure.

**Dissemination of cancerous cells during the biopsy procedure**

To address this possible harm of a breast biopsy we did not use formal inclusion criteria; any clinical study that addressed the topic was included for discussion. Full details of the studies are shown in Appendix E. The results of the studies are summarized in Table 5.

We identified ten studies that used histopathology to demonstrate dissemination of cancerous cells by core-needle biopsy procedures. The percentage of needle tracks reported to contain displaced cancerous cells ranged from 0% to 65%. Diaz et al. demonstrated that the time elapsed between core-needle biopsy and examination of the needle track strongly influenced the findings, with fewer and fewer displaced cancerous cells observed the longer the interval, suggesting that the majority of displaced cancerous cells die off over time. However, we also identified six case reports of patients developing tumor recurrences at the site of prior core-needle biopsies, indicating that not all displaced cancerous cells are non-viable. Three of these six cases were reported to have not received radiation therapy for the primary tumor; for the other three cases it was not reported if they had or had not received radiation therapy.

The risk of tumor recurrence following biopsy was explored by four retrospective studies of 1,879 women. Three of these four studies reported that women who did not have a pre-operative needle biopsy had a higher rate of tumor recurrence than women who did receive a pre-operative needle biopsy; the fourth study reported the opposite. The majority of the women in these four studies were treated with breast-conserving surgery and radiation therapy.

The risk of seeding the lymph nodes with cancerous cells by biopsy procedures was examined in three retrospective studies of 3,103 patients. 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 pre-operative biopsy.
<table>
<thead>
<tr>
<th>Type of study</th>
<th>Number of studies</th>
<th>Number of patients</th>
<th>Summary of findings</th>
</tr>
</thead>
</table>
| Histopathological demonstration of dissemination of cells | 3 case reports\(^\text{142-144}\)  
1 retrospective study\(^\text{145}\)  
6 prospective studies\(^\text{141,146-150}\) | 786                | The percentage of needle tracks reported to contain displaced cancerous cells ranged from 0% to 65%.  
Factors reported to increase the risk of finding displaced cancerous cells include:  
duration of the biopsy procedure\(^\text{146}\), multiple passes of the needle\(^\text{147}\), and a short interval between core-needle procedure and surgical excision.\(^\text{141}\)  
Factors reported to decrease the risk of finding displaced cancerous cells include:  
diagnosis of invasive lobular carcinoma\(^\text{147}\), and use of vacuum-assisted core-needle biopsy.\(^\text{141}\) |
| Tumor recurrence at the biopsy site               | 3 case reports\(^\text{142,149,151}\) | 6                  | 6 cases of tumor recurrence at the biopsy site were presented. All were treated with skin-sparing mastectomy following core-needle biopsy, and three were reported to have not received radiation treatment.\(^\text{142,149}\) It was not reported whether the other 3 cases received radiation treatment.\(^\text{151}\) |
| Risk of tumor recurrence following biopsy        | 4 retrospective studies\(^\text{152-155}\) | 1879               | Three of the four studies reported that women treated with open excisional biopsies had a higher rate of tumor recurrence than women who received pre-operative core-needle biopsies\(^\text{152-154}\); the fourth study reported opposite findings.\(^\text{155}\) The majority of women in all four studies were treated with breast-conserving surgery and radiation therapy. |
| Risk of metastasis to the lymph nodes following biopsy | 3 retrospective studies\(^\text{156-158}\) | 3103               | Two studies reported that the method of biopsy did not correlate with the rate of metastases to the sentinel lymph nodes\(^\text{156,158}\); one study reported that the rate of metastases to the sentinel lymph nodes was higher in women who underwent some type of pre-operative needle biopsy than in women who underwent open excisional biopsy.\(^\text{157}\) |
Question 3. How do open biopsy and various large-core 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 did not use formal inclusion criteria to select literature that addressed Key Question 3 due to the nature of the question. Data addressing this question were collected and are shown in Appendix E. The data are summarized in Table 12, Table 13, and Table 14, and are discussed outcome-by-outcome below. Economic factors that may influence the choice of a particular technique are discussed first, followed by factors highly important to patients, followed by other factors such as availability of equipment.

Relative Costs

Articles identified by our searches that analyzed the costs of open and various large-core biopsy techniques in the U.S. health care system within the last five years are summarized in Table 7, with further details provided in Appendix E. The relative costs of open surgical biopsy and various large-core techniques have been evaluated by six studies. Some of the studies developed models, while others prospectively followed a patient population. When evaluating the costs of these techniques and procedures, the studies have reviewed factors such as the initial purchase price of the devices used, the costs of staffing, the costs of processing and analyzing the biopsy samples, the patient volume where the device will be utilized, if the device is used as a complimentary procedure, and what mammography results determine the use of a large core-needle biopsy technique.

According to the literature reviewed, the costs of open surgical biopsy are substantially higher than large-core techniques. A study by Hatmaker et al. in 2007 found that the average total cost of an open surgical biopsy performed in the operating room was $4,368 with a median cost of $3,479 and the average total cost of image-guided core-needle biopsy was $1,267 with a median cost of $1,239.159

The results of a mammographic screen help surgeons and radiologists decide which large-core technique, if any, would be beneficial and ultimately cost-effective for the patient and facility. Soo et al. used a decision analysis model to compare the costs of a 14-gauge core-needle biopsy to a 14-gauge and 11-gauge vacuum-assisted biopsy for noncalcified lesions. They found that the 14-gauge CNB is less costly for noncalcified lesions, which is not surprising since vacuum-assisted equipment is expensive.160 Golub et al. found that image-guided core-needle biopsy was favored for low suspicion lesions, calcifications, and masses.161

The cost to purchase a large core-needle biopsy device is another factor of interest to facilities. According to Kirshenbaum, in 2003 the average list price for a breast imaging unit biopsy ready (i.e add-on unit) was $90,000 and the average list price for a dedicated prone biopsy table was $226,000.65 Current quoted prices (not list prices) are about $170,000 for a dedicated table (which also requires a large dedicated room) and about $100,000 for an add-on unit.162 Unlike a dedicated prone biopsy table, a mammography unit with an add-on device can be used for general screening purposes when not being used for a biopsy
procedure. However, add-on units have limitations, including limited access angles, limited ability to restrict patient movement, and less patient comfort than dedicated units. Ultrasound-guided core-needle biopsies do not require special equipment and can be performed with a standard multi-purpose US device. Vacuum-assisted core-needle devices currently cost around $37,000 to purchase, and require $270 single-use probes. MRI-guidance is the most expensive method of performing core-needle biopsies, requiring expensive specialized equipment as well as access to an MRI facility.

**Spared Surgical Procedures**

We identified 30 studies that reported information on how the use of core-needle biopsy spares women additional surgical procedures (see Table 8; also see Appendix E for further details). Women who undergo open biopsy with positive findings usually also undergo a second surgical excision procedure to ensure the entire lesion has been removed. Women who undergo a core-needle biopsy procedure with positive findings may be able to undergo a single surgical procedure that simultaneously confirms the diagnosis and removes the entire lesion, thus being spared a second surgical procedure. Women who undergo a core-needle biopsy with negative findings may be able to avoid surgical procedures altogether. Liberman et al. reported that, before the introduction of core-needle biopsy, 29% of women diagnosed with cancer had only one surgical procedure, but after the introduction of core-needle biopsy that number rose to 84%. The studies consistently reported that approximately 75% of women who underwent a core-needle biopsy procedure were spared further procedures, with a mean of approximately 1.2 procedures per woman compared to 1.5 to 2.0 procedures per woman who went straight to open biopsy.

**Procedure Preference**

We identified 20 studies that reported data on patient preferences (see Table 8; also see Appendix E for further details). Ten of the 20 studied vacuum-assisted methods. The majority of the studies did not directly compare different biopsy procedures and instead reported information such as that the patients tolerated the procedure well or would recommend it to others in the future. One study reported that patients preferred the decubitus position to the prone position. Two studies reported that vacuum-assisted procedures were more comfortable than other types of core-needle biopsies. Two authors reported that patients lost less time to core-needle procedures than to open procedures. The majority of the studies concluded that core-needle biopsies were preferable to open biopsies, but one study reported that a survey of patients found that 90% were satisfied with their open surgical biopsy compared to only 80% satisfied with a vacuum-assisted core-needle biopsy.

**Cosmetic Results**

We identified ten studies that reported information on cosmetic results (see Table 8; also see Appendix E for further details). The studies all used vacuum-assisted core-needle biopsy methods. The authors of the studies reported information on how patients felt about the cosmetic results post-procedure. Overall, patients were reported to have been satisfied with the cosmetic results. Only one of the ten studies, Chun et al., compared a group of patients undergoing core-needle biopsy to a group of patients undergoing open biopsy. Chun et al. compared cosmetic results of patients undergoing wire-localized open biopsy to patients undergoing vacuum-assisted 11-gauge core-needle biopsy two years post-procedure. Ninety-five percent of the core-needle biopsy patients were satisfied with their cosmetic results.
biopsy group and only 25% of the open biopsy group were very satisfied with the appearance of their breast. None of the core-needle biopsy group said the cosmetic results were unacceptable compared to 20% of the open biopsy group who found the results unacceptable.

Although all of the studies reporting on cosmetic results used vacuum-assisted methods, it is likely the results apply to most forms of core-needle biopsy. Regardless of the needle gauge or method used, the actual incision cut in the skin is always approximately 1/4” long.162

**Physician Experience**

We identified ten studies that reported information concerning physician experience (see Table 8; for further details see Appendix E). Authors of some of the studies commented that certain devices were easier for inexperienced physicians to use. In general, however, the authors of the studies concluded that greater experience with particular devices improved the accuracy of the biopsy procedures, shortened procedure duration times, and led to a decrease in the number of open biopsies that were performed.

**Availability of a Qualified Pathologist**

We identified two studies that discussed pathologist qualifications and availability (see Table 8; for further details see Appendix E). One reported that whether a specimen was read by a local or central pathologist made little difference because concordance between readings was close to 100%.168 The authors of the other study speculated that lack of an experienced pathologist was the cause of the low accuracy of the core-needle biopsies performed during the course of their study.169

**Availability of Equipment**

We identified three studies that talked about the impact of equipment availability (see Table 8; for further details see Appendix E). One reported that vacuum-assisted devices were more commonly available in the U.S. than in Europe.170 One reported that wait times for access to core-needle procedures were significantly shorter than wait times for access to open surgical procedures.171 The authors of the third study reported that wait times for access to a dedicated prone biopsy table were longer than wait times for other types of core-needle biopsy.172

**Resource Usage**

We identified two studies that talked about resource usage (see Table 8; for further details see Appendix E). The authors of one study reported that vacuum-assisted procedures required more physician and room time than free-hand ultrasound-guided procedures.173 The other study reported that dedicated prone tables use four times as much space as non-prone units.174

**Procedure Duration Time**

We identified 40 studies that reported information about the duration of different biopsy procedures (see Table 8; for further details see Appendix E). The studies reported a wide range of times, from 10 minutes to 128 minutes. The wide range of times may be in part due to different definitions of when exactly the procedure was defined as starting and ending: for example, does the procedure start when the patient enters the room? When the incision is made? Does it end when the sample is collected or when the patient is released to go home? In general, study authors did not define what exactly they meant by procedure duration time.
The reported mean or median time to perform core-needle biopsies under ultrasound guidance ranged from 10 to 60 minutes; the mean or median time to perform core-needle biopsies under stereotactic guidance ranged from 19 to 70 minutes; and the mean or median time to perform core-needle biopsies under MRI guidance ranged from 31 to 70 minutes. Vacuum-assisted core-needle biopsies were reported to have a mean or median duration of 10 to 70 minutes. Open surgical biopsies were generally reported to have longer duration times than core-needle procedures, but only two studies reported estimated duration times of 40 to 45 minutes for open procedures.\textsuperscript{54,175}

**Wait Time for Test Results**

We identified two studies that reported mean or median times to get a diagnosis following a breast biopsy (see Table 8; for further details see Appendix E). The authors reported that wait times after a core-needle procedure were 7 to 10 days shorter than after an open excisional biopsy.\textsuperscript{169,171}
### Table 6. Economic considerations

<table>
<thead>
<tr>
<th>Reference</th>
<th>Source of Cost Data</th>
<th>Methods or Models of Analysis</th>
<th>Primary Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hatmaker et al. 2006(^{159})</td>
<td>The Massachusetts Utilization Multiprogramming System and the Decision Support System software packages were used to track costs of procedures, by Current Procedure Terminology (CPT) code and date of service.</td>
<td>Data were analyzed and described using the R statistical computing environment. Costs for all service related to each procedure were linked through billing procedures by the date of service and classified as related to radiology costs, to pathology or laboratory costs, or to procedural costs</td>
<td>The average total cost to evaluate a patient with a breast mass or mammographic abnormality through an OSB in the operating room was $4,368 (SD: $2,586) with a median cost of $3,479. The average total cost for a CNB was $1,267 (SD: $536) with a median cost of $1,239. For VA hospitals with available resources, the option of CNB is a cost-effective and more preferable alternative to OSB.</td>
</tr>
<tr>
<td>Orel et al. 2006(^{176})</td>
<td>NR</td>
<td>NR</td>
<td>The total Medicare allowance for one MR-guided vacuum-assisted CNB procedure is approximately $500. Additional investigation is needed to develop more cost-efficient systems. In addition, the cost of the needles will probably decrease as the use of them increases.</td>
</tr>
<tr>
<td>Shin et al. 2006(^{177})</td>
<td>NR</td>
<td>NR</td>
<td>If the surgeon chooses to perform a diagnostic core biopsy and then excise the lesion for definitive treatment, the overall cost would be between $12,000 and $15,000, depending on the initial modality used for biopsy. Extrapolating this to our small pilot study of 156 patients, the observation arm would cost $619,000 for ultrasound-guided CNB and $1,028,820 for stereotactic-guided CNB. OSB for diagnosis and treatment with routine screening follow-up would cost $1,454,544 at our institution.</td>
</tr>
<tr>
<td>Soo et al. 2005(^{160})</td>
<td>Cost &amp; probability variables were estimated from institution over a three year period. Ratios were used representing the relative dollar values of the estimated costs</td>
<td>Decision Analysis Model was used to compare costs of 14-gauge CNB to 14-gauge and 11-gauge vacuum-assisted CNB for stereotactic biopsy of noncalcified breast lesions</td>
<td>The 14-G vacuum-assisted CNB was 1.19 times as expensive as the multipass automated gun CNB method, and the 11-G vacuum-assisted CNB was 1.22 times as expensive as the multipass automated gun CNB. The 14-G CNB is less costly for stereotactic biopsy of non-calcified lesions over a wide range of cost estimates.</td>
</tr>
<tr>
<td>Reference</td>
<td>Source of Cost Data</td>
<td>Methods or Models of Analysis</td>
<td>Primary Conclusions</td>
</tr>
<tr>
<td>---------------------------</td>
<td>-----------------------------------------</td>
<td>----------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Golub et al. 2004</td>
<td>Patient billing records at the Lynn Sage Breast Center</td>
<td>A decision analytic model of the outcomes of all biopsy patients seen at the Lynn Sage Breast Center during a 2 year period was constructed. Costs were analyzed by considering only patients receiving breast-conserving surgery (lumpectomy alone), and subgroup based on degree of suspicion and on radiographic abnormality type. The sum of the mean costs determined from the patient billing records was used as the baseline outcome measures in the decision tree. Costs were measured from a societal perspective. Only direct costs related to inpatient care were considered, and they included CNB, OSB, lumpectomy with or without re-excision, lumpectomy with or without lymph node dissection, mastectomy with or without lymph node dissection, and lymph node dissection alone. Costs were derived by application of the institution’s cost-to-charge multiplier.</td>
<td>The total cost of diagnosis and surgical treatment was $1,849 for CNB versus $2,775 for OSB. When the probabilities were biased to favor OSB, the cost was $2,297 for CNB and $2,458 for OSB. CNB was favored for low suspicion lesions, calcifications, and masses. OSB was favored for high suspicion lesions and architectural distortion. Total costs were $926 less for the CNB group. CNB can be cost-saving compared with OSB, particularly when mammographic abnormality is classified as low suspicion or consists of calcifications or masses.</td>
</tr>
<tr>
<td>Kirshenbaum et al. 2003</td>
<td>NR</td>
<td>NR</td>
<td>A breast imaging center need spend only approximately $90,000 (average list price of add-on device) to make an existing mammography unit biopsy-ready. For a dedicated prone biopsy table, a center would need to spend $226,000 (average list price). If one includes the additional cost of purchasing a mammography machine (average $80,000) that might be required because the add-on unit is incompatible with the existing machine, the cost differential is substantially reduced. When not being used for biopsies, add-on units can be used for general screening and diagnostic work, whereas prone units can only be used for biopsies.</td>
</tr>
</tbody>
</table>
Table 7. Key Question 3: other outcomes

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Number of Studies</th>
<th>Number of Patients</th>
<th>Summary of Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cosmetic results</td>
<td>5 prospective</td>
<td>4,732</td>
<td>In eight of the ten studies, the authors reported how all included study patients felt about their scar appearance at some point in time from one week to six months post-procedure. Overall, patients were satisfied with the cosmetic outcome. In two of the ten studies, the authors made direct comparisons between two types of biopsy procedures. Weber et al. compared the cosmetic results of the Mammmotome with an 11-gauge needle to those of the ABBI. They found the ABBI group was less satisfied with the appearance of the biopsy site than those in the Mammmotome group. Chun et al. compared patients having either an ABBI or the Mammmotome with an 11-gauge needle to those undergoing a wire localized biopsy. These authors found that many patients in the wire localized group were unhappy with their cosmetic result, while all of the patients having Mammmotome or ABBI found the scar appearance to be acceptable or excellent.</td>
</tr>
<tr>
<td></td>
<td>studies and 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>retrospective</td>
<td></td>
<td>studies</td>
</tr>
<tr>
<td></td>
<td>66, 85, 178-180</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>58, 90, 167, 181, 182</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physician experience</td>
<td>5 prospective</td>
<td>23,332</td>
<td>Eight of the ten included studies described the study physicians’ level of experience and how that may have impacted the studies’ results. In two of these cases (Schneider et al. and Wunderbaldinger et al.), the study investigators were testing a new CNB device and concluded that the device is suitable for physician’s without a great deal of experience performing biopsies. The other two articles described how the availability of highly experienced biopsy operators has led to a decrease in the use of diagnostic excisional biopsies (Holloway et al. and Hoffman et al.).</td>
</tr>
<tr>
<td></td>
<td>studies and 5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>retrospective</td>
<td></td>
<td>studies</td>
</tr>
<tr>
<td></td>
<td>88, 137, 138, 164, 183</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>64-188</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedure time</td>
<td>23 prospective</td>
<td>6,121</td>
<td>A total of 40 studies reported procedure times for the various breast biopsy procedures. There was great variation in reported procedure times by study, with a range of between 10 and 128 minutes. Some studies indicated that changing from a conventional to an add-on unit and increased operator experience tended to decrease procedure times, while other studies suggested that cases in which benign epithelial cells were disseminated or where ABBI and wire localized procedures were used procedure times tended to be increased.</td>
</tr>
<tr>
<td></td>
<td>studies and 17</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>retrospective</td>
<td></td>
<td>studies</td>
</tr>
<tr>
<td></td>
<td>66, 167, 195-200</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>46, 52, 54, 60, 67, 70, 85, 88, 109, 135, 13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8, 146, 164, 165, 173, 180, 183, 189-194</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spared procedure rates</td>
<td>8 prospective</td>
<td>8,407</td>
<td>30 studies reported how diagnostic CNB spared patients a surgical procedure as compared with a diagnostic excisional biopsy. CNB appears to spare a majority of patients additional surgical procedures.</td>
</tr>
<tr>
<td></td>
<td>studies and 22</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>retrospective</td>
<td></td>
<td>studies</td>
</tr>
<tr>
<td></td>
<td>52, 54, 70, 85, 171, 172, 201, 202</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>74, 79, 93, 107, 118, 163, 200, 203-215</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcome</td>
<td>Number of Studies</td>
<td>Number of Patients</td>
<td>Summary of Findings</td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>-------------------</td>
<td>--------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Availability of a qualified pathologist</td>
<td>1 prospective</td>
<td>2,112</td>
<td>Two studies addressed the availability of a qualified pathologist for interpreting biopsy specimens. The first, Collins et al., found that whether a specimen was read by a local or central pathologist made very little difference. Agreement rates between the two were very high for both CNB and open biopsy, although agreement rates were somewhat lower for open biopsy specimens. The second study, Gukas et al., evaluated the accuracy of TruCut versus excisional biopsy in Nigeria. The pathologist used in their study did not have a lot of experience with the Trucut device, and the authors concluded that his lack of experience explains TruCut's poor performance compared with excisional biopsy.</td>
</tr>
<tr>
<td></td>
<td>study and</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 retrospective</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>study</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Availability of equipment</td>
<td>2 prospective</td>
<td>5,921</td>
<td>Three studies addressed the availability of various breast biopsy devices. One, Deurloo et al., explained that while vacuum-assisted CNB is on the rise in the United States, in Europe automated gun CNB is the preferred technique, suggesting that European women are much less likely to have access to a vacuum-assisted procedure than are women in this country. Verkooijen et al. report that median wait times, from initial physician referral to first diagnostic procedure, were shorter for patients having a CNB than those requiring an open biopsy (4 vs. 13 days, respectively), while Williams et al. found a longer wait list for prone CNB patients than for a historical cohort in the pre-prone table days.</td>
</tr>
<tr>
<td></td>
<td>studies and</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 retrospective</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>study</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resource usage</td>
<td>2 prospective</td>
<td>393</td>
<td>Two studies addressed how the various breast biopsy techniques impact resource usage. Mainiero et al. compared the amount of physician time and room time utilized by vacuum-assisted CNB compared to freehand ultrasound-guided CNB. They found the vacuum-assisted method required more physician and room time. Wunderbaldinger et al. reported that prone devices use four times the amount of hospital/office space as non-prone units.</td>
</tr>
<tr>
<td></td>
<td>studies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedure preference</td>
<td>12 prospective</td>
<td>5,001</td>
<td>Twenty studies collected data on patient preferences for breast biopsy procedures. Overall, these studies reported that patients tolerated the CNB procedure well and that a good percentage indicated they would recommend the procedure to others.</td>
</tr>
<tr>
<td></td>
<td>studies and</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8 retrospective</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>studies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wait time for test results</td>
<td>2 prospective</td>
<td>272</td>
<td>Two studies reported how long it may take patients to receive a diagnosis following either a CNB or open biopsy procedure. In both studies, wait times were shorter for the CNB (7.3 days less and 9 vs. 19 days, respectively).</td>
</tr>
</tbody>
</table>
Conclusions

Key Question 1. In women with a palpable or non-palpable breast abnormality what is the accuracy of different types of large core breast biopsy compared with open biopsy for diagnosis?

Our conclusions for Key Question 1 are summarized in Table 1, Figure 1 through Figure 4 (in the Executive Summary) and in Table 9. Factors affecting the accuracy are summarized in Table 3 (in the Executive Summary) and in Table 11. Our key conclusions are stated below.

Stereotactically-guided vacuum-assisted core-needle biopsies have a sensitivity of 99.2% (95% CI: 97.9 to 99.7%). Strength of evidence: Low

Stereotactically-guided automated gun core-needle biopsies have a sensitivity of 97.8% (95% CI: 95.8 to 98.9%). Strength of evidence: Low

Ultrasound-guided vacuum-assisted core-needle biopsies have a sensitivity of 96.5% (95% CI: 81.2 to 99.4%). Strength of evidence: Low

Ultrasound-guided automated gun core-needle biopsies have a sensitivity of 97.6% (95% CI: 97.0 to 98.1%). Strength of evidence: Low

Freehand automated gun core-needle biopsies have a sensitivity of 85.8% (95% CI: 75.8 to 92.1%). Strength of evidence: Low

A woman with a BIRADS 4 lesion (with an estimated pre-biopsy risk of having a malignancy of 30%\textsuperscript{219}) who has a benign diagnosis on core-needle biopsy would be expected to have a post-biopsy risk of malignancy of:

- Open surgical biopsy: 1% or less
- Stereotactically-guided vacuum-assisted core-needle biopsy: 1% or less
- Stereotactically-guided automated gun core-needle biopsy: 2% or less
- Ultrasound-guided automated gun core-needle biopsy: 2% or less
- Ultrasound-guided vacuum-assisted core-needle biopsy: 8% or less
- Freehand automated gun core-needle biopsy: 10% or less

Key Question 2. In women with a palpable or non-palpable breast abnormality what are the harms associated with different types of large core breast biopsy compared with open biopsy for diagnosis?

Severe complications following core-needle biopsy are very rare, affecting only 0.09 to 0.72% of procedures (summarized in Table 2 in the Executive Summary and in Table 10). Vacuum-assisted procedures may be associated with a slightly higher rate of severe bleeding events than automated gun core-needle biopsies. Two to 10% of open surgical biopsies may be affected by severe complications. Strength of evidence: Low
Other Conclusions

Conclusions about the impact of factors on accuracy and harm are summarized in Table 3 (in the Executive Summary) and in Table 11. Conclusions about Key Question 3 are summarized in Table 12, Table 13, and Table 14.
### Summary Tables

Table 8. Summary of accuracy by type of biopsy procedure

<table>
<thead>
<tr>
<th>Type of biopsy</th>
<th>N studies</th>
<th>N lesions</th>
<th>Prevalence of malignancy</th>
<th>Sensitivity (95% CI)</th>
<th>Negative likelihood ratio (95% CI)</th>
<th>Atypia underestimation rate</th>
<th>DCIS underestimation rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freehand automated gun</td>
<td>5</td>
<td>610</td>
<td>68.7%</td>
<td>85.8% (75.8 to 92.1%)</td>
<td>0.143 (0.082 to 0.250)</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>US guidance automated gun</td>
<td>15</td>
<td>5686</td>
<td>55.3%</td>
<td>97.6% (97.0% to 98.1%)</td>
<td>0.031 (0.024 to 0.040)</td>
<td>0.276 (0.202 to 0.364)</td>
<td>0.360 (0.245 to 0.493)</td>
</tr>
<tr>
<td>Stereotactic guidance, automated gun</td>
<td>33</td>
<td>7153</td>
<td>37.1%</td>
<td>97.8% (95.8% to 98.9%)</td>
<td>0.022 (0.012 to 0.043)</td>
<td>0.435 (0.357 to 0.517)</td>
<td>0.244 (0.180 to 0.321)</td>
</tr>
<tr>
<td>MRI guidance, automated gun</td>
<td>1</td>
<td>14</td>
<td>42.8%</td>
<td>83.3% (43.5% to 96.5%)</td>
<td>0.23 (0.05 to 0.95)</td>
<td>100% (1/1)</td>
<td>NR</td>
</tr>
<tr>
<td>Perforated compression grid automated gun</td>
<td>1</td>
<td>100</td>
<td>33%</td>
<td>91.4% (77.5% to 96.9%)</td>
<td>0.09 (0.03 to 0.26)</td>
<td>0.25 (1 out of 4)</td>
<td>0.286 (2 out of 7)</td>
</tr>
<tr>
<td>US guidance vacuum-assisted</td>
<td>7</td>
<td>507</td>
<td>15%</td>
<td>96.5% (81.2 to 99.4%)</td>
<td>0.036 (0.006 to 0.212)</td>
<td>None reported</td>
<td>Only one occurrence reported</td>
</tr>
<tr>
<td>Stereotactic guidance, vacuum-assisted</td>
<td>20</td>
<td>6255</td>
<td>32.8%</td>
<td>99.2% (97.9% to 99.7%)</td>
<td>0.009 (0.003 to 0.023)</td>
<td>0.219 (0.178 to 0.266)</td>
<td>0.130 (0.111 to 0.151)</td>
</tr>
</tbody>
</table>

NR = Not Reported

This DRAFT report is distributed solely for review purposes.
Table 9. Summary of harms complicating core-needle biopsies

<table>
<thead>
<tr>
<th>Harms</th>
<th>N Studies Reported</th>
<th>N Patients</th>
<th>N Occurrences</th>
<th>% Affected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did not report</td>
<td>48</td>
<td>25,562</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Reported no complications occurred</td>
<td>5</td>
<td>3,954</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Quality of life</td>
<td>0</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Patients dissatisfied with the procedure</td>
<td>2</td>
<td>328</td>
<td>2</td>
<td>0.61%</td>
</tr>
<tr>
<td>Hematomas requiring treatment</td>
<td>24</td>
<td>17,585</td>
<td>15</td>
<td>0.09%</td>
</tr>
<tr>
<td>Bleeding, severe</td>
<td>24</td>
<td>8,474</td>
<td>61</td>
<td>0.72%</td>
</tr>
<tr>
<td>Infections</td>
<td>20</td>
<td>16,407</td>
<td>24</td>
<td>0.15%</td>
</tr>
<tr>
<td>Pneumothorax</td>
<td>4</td>
<td>2,600</td>
<td>4</td>
<td>0.15%</td>
</tr>
<tr>
<td>Usual activities significantly affected by the biopsy procedure</td>
<td>1</td>
<td>34</td>
<td>4</td>
<td>11.80%</td>
</tr>
<tr>
<td>Time to recovery</td>
<td>1</td>
<td>103</td>
<td>1.5 days on average</td>
<td>NA</td>
</tr>
<tr>
<td>Bruising</td>
<td>9</td>
<td>3,256</td>
<td>Reported to be &quot;common&quot;</td>
<td>NR</td>
</tr>
<tr>
<td>Required pain medications</td>
<td>4</td>
<td>573</td>
<td>21</td>
<td>3.70%</td>
</tr>
<tr>
<td>Diagnostically confusing scars subsequent to the procedure</td>
<td>3</td>
<td>3,748</td>
<td>7</td>
<td>0.18%</td>
</tr>
<tr>
<td>Vasovagal reactions</td>
<td>21</td>
<td>7,526</td>
<td>75</td>
<td>1.00%</td>
</tr>
<tr>
<td>Severe pain during the biopsy procedure</td>
<td>17</td>
<td>3,128</td>
<td>52</td>
<td>1.70%</td>
</tr>
</tbody>
</table>

NR = Not Reported  
NA = Not Applicable
Table 10. Summary of the impact of factors on accuracy and harms

<table>
<thead>
<tr>
<th>Factors</th>
<th>N Studies Reported Data on the Impact of the Factor on Accuracy</th>
<th>Conclusion</th>
<th>N Studies Reported Data on the Impact of the Factor on Harms</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient Characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient age</td>
<td>1 Insufficient data</td>
<td>0 Insufficient data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast density</td>
<td>1 Insufficient data</td>
<td>0 Insufficient data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient co-morbidities</td>
<td>0 Insufficient data</td>
<td>0 Insufficient data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lesion characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Palpable vs. non-palpable</td>
<td>2 Insufficient data</td>
<td>0 Insufficient data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Microcalcifications vs. masses</td>
<td>4 Inconsistent findings</td>
<td>0 Insufficient data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distortions vs. masses</td>
<td>1 Insufficient data</td>
<td>0 Insufficient data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size of lesion</td>
<td>2 Insufficient data</td>
<td>0 Insufficient data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location of lesion</td>
<td>1 Insufficient data</td>
<td>0 Insufficient data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biopsy Methodology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of cores</td>
<td>3 Inconsistent findings</td>
<td>0 Insufficient data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient position</td>
<td>1 Insufficient data</td>
<td>21 Vasovagal reactions occur more often in patients seated upright</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference standard</td>
<td>68 Meta-regression found no impact</td>
<td>0 Insufficient data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of vacuum</td>
<td>78 Vacuum-assistance improved accuracy</td>
<td>24 Use of vacuum increased the percentage of procedures complicated by severe bleeding and hematoma formation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use of image guidance</td>
<td>78 Image guidance improved accuracy; stereotactic guidance was more accurate than US guidance</td>
<td>0 Insufficient data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Needle size</td>
<td>33 Meta-regression found no impact</td>
<td>1 Insufficient data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factors</td>
<td>N Studies Reported Data on the Impact of the Factor on Accuracy</td>
<td>Conclusion</td>
<td>N Studies Reported Data on the Impact of the Factor on Harms</td>
<td>Conclusion</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-----------------------------------------------------------------</td>
<td>------------------------------</td>
<td>----------------------------------------------------------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>Clinician and Facility Factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experience of operator</td>
<td>2</td>
<td>Insufficient data</td>
<td>0</td>
<td>Insufficient data</td>
</tr>
<tr>
<td>Training of operator</td>
<td>0</td>
<td>Insufficient data</td>
<td>0</td>
<td>Insufficient data</td>
</tr>
<tr>
<td>Facility location</td>
<td>68</td>
<td>Meta-regression found no impact</td>
<td>0</td>
<td>Insufficient data</td>
</tr>
<tr>
<td>Facility type</td>
<td>33</td>
<td>Meta-regression found no impact</td>
<td>0</td>
<td>Insufficient data</td>
</tr>
</tbody>
</table>
Table 11. Summary of economic aspects of core-needle biopsy

<table>
<thead>
<tr>
<th>Aspect</th>
<th>N Studies</th>
<th>Conclusions and Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative costs open biopsy vs. core-needle biopsy</td>
<td>8</td>
<td>All report that core-needle biopsy costs less than open biopsy procedures.</td>
</tr>
<tr>
<td>Relative costs of different types of core-needle biopsy</td>
<td>3</td>
<td>Insufficient data. All three studies reported information on different comparisons.</td>
</tr>
<tr>
<td>Resource usage</td>
<td>2</td>
<td>Insufficient data. Both studies reported information on different topics.</td>
</tr>
</tbody>
</table>

Table 12. Summary of patient perspectives on choice of biopsy method

<table>
<thead>
<tr>
<th>Aspect</th>
<th>N Studies</th>
<th>Conclusions and Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedure preference</td>
<td>20</td>
<td>The majority of the studies concluded that patients preferred core-needle procedures over open procedures</td>
</tr>
<tr>
<td>Spared surgical procedures</td>
<td>30</td>
<td>Approximately 75% of women who underwent a core-needle procedure were spared further procedures, with a mean of 1.2 procedures per woman compared to 1.5 to 2.0 procedures per woman who went straight to open biopsy.</td>
</tr>
<tr>
<td>Cosmetic results</td>
<td>10</td>
<td>Overall patients were satisfied with the cosmetic results of a vacuum-assisted core-needle procedure.</td>
</tr>
<tr>
<td>Procedure duration time</td>
<td>40</td>
<td>US-guided core-needle procedures took 10 to 60 minutes, stereotactically-guided core-needle procedures took 19 to 70 minutes, vacuum-assisted core-needle biopsies took 10 to 70 minutes. Open biopsy procedures were estimated to take 40 to 45 minutes.</td>
</tr>
<tr>
<td>Wait time for test result</td>
<td>2</td>
<td>Insufficient data</td>
</tr>
</tbody>
</table>

Table 13. Summary of clinician and facility factors related to core-needle biopsy

<table>
<thead>
<tr>
<th>Aspect</th>
<th>N Studies</th>
<th>Conclusions and Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinician experience</td>
<td>10</td>
<td>Greater experience with particular devices improved accuracy. Some types of devices were easier for inexperienced clinicians to use than others.</td>
</tr>
<tr>
<td>Availability of Equipment</td>
<td>3</td>
<td>Insufficient data</td>
</tr>
<tr>
<td>Availability of qualified pathologist</td>
<td>2</td>
<td>Insufficient data</td>
</tr>
</tbody>
</table>

This DRAFT report is distributed solely for review purposes.
Previously Published Systematic Reviews

Our searches identified two previously published systematic reviews. Verkooijen et al. reviewed the literature published prior to 1999 on core-needle biopsy of non-palpable lesions.\textsuperscript{220} Fahrbach et al. reviewed the literature published from 1996 to 2004 on core-needle biopsy of patients referred for biopsy after screening mammography.\textsuperscript{221}

We assessed the quality of each systematic review using the ‘assessment of multiple systematic reviews’ (AMSTAR) measurement tool.\textsuperscript{222} The AMSTAR consists of 11 items, which have been tested for face and content validity. The items assess whether or not a systematic review includes important elements, such as a comprehensive literature search, assessment of study quality, appropriate methods to combine study findings, and assessment of publication bias. Responses to each item are checked as ‘Yes’ if the review includes that item, ‘No’ if it does not, ‘Can’t tell’ if the item cannot be answered by the information provided in the review, or ‘Not applicable’ if the item is not applicable. The AMSTAR does not provide a method for rating the quality of a review. To rate the quality of the reviews, we applied the following criteria: a rating of ‘High’ if the review received mostly ‘yes’ responses (at least 8), a rating of ‘Low’ if the review received mostly ‘no’ responses (at least 8), and a rating of ‘Moderate’ if the review received mixed responses. Both systematic reviews were rated as Moderate quality. The reviews were not rated as High quality because neither systematic review stated conflicts of interest or incorporated ratings of the quality of the literature into their conclusions. See Appendix E for details about the quality rating.

Verkooijen et al. included only five cohort studies in their review. Their inclusion criteria were studies of non-palpable lesions, either surgical biopsy or at least two years of followup to verify the true diagnosis, and a minimum of five cores taken per lesion. All included studies happened to have used stereotactic guidance. The authors assumed core-needle biopsy had no false-positives (i.e., malignant diagnoses on core needle that were not found on open surgery were assumed to have been completely removed by the core-needle procedure). Their analyses found that the DCIS underestimation rate was 15\% (95\% CI: 8.0 to 26.0\%), the ADH underestimation rate was 40\% (95\% CI: 26.0 to 56.0\%), and the overall sensitivity of core-needle biopsy for non-palpable lesions was 97.0\% (95\% CI: 95.0\% to 99.0\%). Only two complications were reported, one hematoma and one case of infection.\textsuperscript{220}

Fahrbach et al. included 12 studies of stereotactically-guided vacuum-assisted core-needle biopsy and compared them to 25 studies of stereotactically-guided automated gun core-needle biopsy. One of their inclusion criterion was that the study must have been conducted in a western-style health care system (North America, Europe, Australia, or New Zealand). Their analyses found the false-negative rate of vacuum-assisted biopsy was 1.2\%, the DCIS underestimation rate was 13.7\%, and the ADH underestimation rate was 29.2\%. Automated gun core-needle biopsy had a false-negative rate of 2\%, a DCIS underestimation rate of 27.1\%, and an ADH underestimation rate of 47.4\%. Further, the authors performed analyses of possible factors that may have affected the results. Study location was a significant predictor of the false-negative rate, but type of reference standard and patient position had no significant impact on the results.
The authors of both systematic reviews concluded that core-needle biopsy rarely mis-diagnosed malignant lesions as benign. Fahrbach et al. concluded that vacuum-assisted biopsy may provide lower miss and underestimation rates than automated gun core-needle biopsy.\textsuperscript{221}
Chapter 4. Discussion

Conclusions

Open surgical biopsy is the “gold standard” method of evaluating a suspicious breast lesion. However, it is a surgical procedure that, like all surgeries, places the patient at risk of experiencing morbidities and, in rare cases, mortality. The majority of women who undergo breast biopsy procedures do not have cancer. Exposing large numbers of women to invasive surgical procedures when the majority of these women do not benefit from the procedure may be considered by most to be unacceptable. A less invasive method would be preferable if it were sufficiently accurate.

Open surgical biopsy has been reported to miss 1 to 2% of breast cancers. Our analysis found that stereotactically-guided vacuum-assisted core-needle biopsy is almost as accurate as open surgical biopsy with a much lower complication rate. US-guided core-needle biopsy may be almost as accurate as stereotactically-guided vacuum-assisted biopsy, and may have a slightly lower complication rate.

Diagnoses of “pure” DCIS determined on the basis of core-needle biopsy may be incorrect due to the inability of needle biopsy to sample all parts of the tumor. Rakha and Ellis reviewed the literature in 2007 and reported that 15 to 20% of cases diagnosed as “pure” DCIS by core-needle biopsy were subsequently found to contain associated invasive carcinoma upon excision. Our analyses found that DCIS underestimation rates ranged from 13% to 36%, justifying current clinical practice of referring all DCIS diagnoses for open surgery.

The management of “high risk” lesions such as ADH is somewhat controversial. Our analysis found that at least 20% of ADH diagnoses on core-needle biopsy are actually malignant, suggesting that patients diagnosed with atypia on core needle may benefit from open surgery as well.

Limitations of the Evidence Base

The evidence base is very large but of low quality. The majority of the available studies are poorly reported retrospective chart reviews. Poor reporting of biopsy methodology, patient characteristics, and details of lesions precluded answering the majority of the sub-questions about factors affecting the accuracy and harms of core-needle biopsy.

Applicability

We used inclusion criteria intended to restrict the evidence base to only those studies that included the population of interest: women of average risk undergoing breast biopsy after discovery of a suspicious lesion on routine screening. However, our analysis found that the prevalence of cancers in the study populations tended to be slightly higher than expected. The prevalence of cancers in the general population sent for breast biopsy (in the USA) has been reported to be around 23%. The studies in our analysis generally reported prevalences in the thirties to forties, up to 55% for freehand biopsies. This may be due to the fact that many of the studies were conducted in non-USA locations, where the prevalence of cancers in populations sent for biopsy has been reported to be 60 to 70%. It may also be an artifact caused by attrition. Many of the studies had fairly high rates of attrition, and most of the lost patients had been diagnosed as benign on core-needle biopsy. The lost patients were of necessity removed from the analysis, and this may have artificially elevated the prevalence of disease. Interestingly,
the studies of US-guided vacuum-assisted biopsy reported an overall prevalence of disease of only 15%, suggesting that lesions selected for this method may have a low probability of being malignant. Lesions selected for US-guided procedures generally do not contain microcalcifications and must be clearly visible on US.

**Possible Impact of Key Assumptions on the Conclusions**

Several key assumptions were made: 1) the “reference standard”, a combination of open surgery and follow-up for at least six months, was 100% accurate; 2) the pathologists examining the core-needle biopsy results were 100% accurate; and 3) core-needle diagnoses of malignancy (invasive or in situ) that could not be confirmed by open surgery were assumed to have been correct diagnoses where the lesion had been completely removed by the core-needle biopsy procedure. In addition, the majority of studies reported data on a per-lesion rather than a per-patient basis, and therefore we analyzed the data on a per-lesion basis.

Key assumption #1, that the reference standard was 100% accurate, is almost certainly not true. Open surgical biopsy has been reported to have a false-negative rate of 1 to 2%, and only six months of patient followup is unlikely to be long enough to detect all missed cancers.223 If a small percentage of the surgical biopsies were false-negatives then our estimates of the accuracy of core-needle biopsy are slightly lower than the actual “true” accuracy of core-needle biopsy. If a small percentage of the patients declared “benign” on six-month patient followup actually had cancers then our estimates of the accuracy of core-needle biopsy are higher than the actual “true” accuracy of core-needle biopsy. Logically one would expect short-term patient followup to be more prone to error than open surgical biopsy; thus it seems likely that our estimates of core-needle biopsy accuracy are slightly higher than the actual “true” accuracy. However, some of the studies did followup all patients for at least two years, and other studies did perform open biopsy on all patients. We performed meta-regressions and found no statistically significant impact of the type of reference standard used or length of followup on the reported accuracy of the core-needle biopsies.

Key assumptions #2 and #3 are inter-related and both depend on pathologists being 100% accurate in reading core-needle biopsy material. The literature reports pathology errors as being rare, affecting 0.08 to 1.2% of specimens examined.224 A 2006 review of medical malpractice suits filed against pathologists for breast biopsy misdiagnoses reported that about half the suits involved false-negative errors and about half involved false-positive errors.224 Even if a very small percentage of patients declared “true positive” in our analysis were actually false-positives and a very small percentage of patients declared “true negatives” were actually false-negatives it seems unlikely that our estimates of core-needle biopsy can be significantly different than the actual true accuracy.

Key assumption #4, that analyzing the data on a per-lesion rather than a per-patient basis would not violate statistical assumptions of independence, was unavoidable. Very few of the studies reported data on a per-patient basis. The percentage of patients with more than one lesion was, in most studies, quite low. Each lesion was subjected to an independent core-needle biopsy. A patient diagnosed with multiple benign lesions would have all lesions managed by followup, but a patient with one malignant lesion and a benign lesion may have had the benign lesion surgically biopsied at the same time as the malignant lesion was biopsied. Thus the independence of data at the per-lesion level is not quite complete. The impact of this minor lack of independence on the results of our analyses is most likely insignificant.
Correlation with Findings from Prior Systematic Reviews

As discussed previously, two prior systematic reviews of core-needle biopsy have been published.\textsuperscript{220,221} Both prior reviews and our review calculated very similar false-negative rates for stereotactically-guided automated gun core-needle biopsy: 2.2\%, 3.0\%, and 2.0\%. Both prior reviews and our review calculated very similar rates of ADH underestimation for stereotactically-guided automated gun core-needle biopsy: 40\%, 43.5\%, and 47.4\%. The DCIS underestimation rate reported by Verkooijen et al. for stereotactically-guided core-needle biopsy was much lower (only 15.0\%) than the DCIS underestimation rates (24.4\%, 27.1\%) reported by Fahrbach et al. and our review. This difference may be related to the fact that our review and Fahrbrach et al. included both palpable and non-palpable lesions in the analysis whereas Verkooijen et al. restricted their analysis to non-palpable lesions.

Verkooijen et al. did not study stereotactically-guided vacuum-assisted core-needle biopsy. Our review and Fahrbach et al. found very similar accuracy figures for stereotactically-guided vacuum-assisted core-needle biopsy: false negative rate, 1.2\% and 0.8\%; ADH underestimation rate, 29.2\% and 21.9\%; DCIS underestimation rate, 13.7\% and 13.0\%.

Fahrbach et al. found that study location was a significant predictor of the false-negative rate, but type of reference standard and patient position had no significant impact on the results. We also found that the type of reference standard had no impact on the results, but we found no impact of study location on the results. The reason for this apparent discrepancy may be that we included studies conducted worldwide, whereas Fahrbach et al. included only studies conducted in North America, Europe, Australia, or New Zealand.

Future Research Needed

The chain of evidence linking better patient outcomes to the use of open biopsy after detection of a breast abnormality is firmly established. There is no need to conduct randomized controlled trials to demonstrate that patients benefit from core-needle biopsy. Establishing that core-needle biopsy is safer than open biopsy and almost as accurate as open biopsy is sufficient to justify its use. However, well-reported retrospective chart reviews, retrospective database analyses, or prospective studies are needed to address the as-yet-unanswered questions as to what factors affect the accuracy and harms of core-needle breast biopsy. Answers to such questions are important for both patients and clinicians when faced with the decision of what type of breast biopsy is best for each individual patient.
References and Included Studies

(Note: There is a separate set of references at the end of Appendix F whose reference numbers are different from those in the text of the report.)


This DRAFT report is distributed solely for review purposes.


This DRAFT report is distributed solely for review purposes.


# List of Acronyms/Abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADH</td>
<td>Atypical Ductal Hyperplasia</td>
</tr>
<tr>
<td>AHRQ</td>
<td>Agency for Healthcare Research and Quality</td>
</tr>
<tr>
<td>ALH</td>
<td>Atypical Lobular Hyperplasia</td>
</tr>
<tr>
<td>BIRADS</td>
<td>Breast Imaging Reporting and Data System</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>DCIS</td>
<td>Ductal Carcinoma In Situ</td>
</tr>
<tr>
<td>FN</td>
<td>False Negative</td>
</tr>
<tr>
<td>FP</td>
<td>False Positive</td>
</tr>
<tr>
<td>G</td>
<td>Gauge</td>
</tr>
<tr>
<td>LCIS</td>
<td>Lobular Carcinoma In Situ</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>NA</td>
<td>Not Applicable</td>
</tr>
<tr>
<td>NR</td>
<td>Not Reported</td>
</tr>
<tr>
<td>TN</td>
<td>True Negative</td>
</tr>
<tr>
<td>TP</td>
<td>True Positive</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>US</td>
<td>Ultrasound</td>
</tr>
<tr>
<td>USA</td>
<td>United States of America</td>
</tr>
</tbody>
</table>