Use of GeneSearch Breast Lymph Node Assay to Detect Sentinel Node Metastases in Early Stage Breast Cancer

Assessment Program
Volume 22, No. 8
February 2008

Executive Summary

Background

The spread of breast cancer to the axillary lymph nodes is a key prognostic indicator and affects treatment choices. For women with early stage breast cancer and no palpable lymph nodes, sentinel lymph node biopsy (SLNB) is common. Currently, two approaches are used for evaluating sentinel nodes. In the first, the sentinel lymph node is removed, sent to a pathologist, and the surgery is completed. Nodal histology is evaluated, with or without immunohistochemical stains, which takes 1–2 days. A positive result usually leads to a second surgery, where axillary lymph node dissection (ALND) is performed and additional nodes are removed for examination. In the second approach, the sentinel lymph node is removed and sent for evaluation during the surgery. The two most common intraoperative techniques are frozen section histology or imprint cytology, both of which are less accurate than postoperative histology of permanently fixed tissue. Patients with positive lymph nodes generally receive ALND during the same surgery.

The GeneSearch™ breast lymph node (BLN) assay (hereafter, referred to as "GeneSearch" or "the assay") is proposed as an alternative intraoperative test. Developed and commercialized by Veridex, LLC (Warren, NJ), GeneSearch uses real-time polymerase chain reaction (RT-PCR) to qualitatively evaluate nodal sections for the presence of two gene expression markers, mammaglobin and cytokeratin 19. These genes are expressed at higher levels in breast tissue, but not in normal nodal tissue. The assay is automated and performed in a homogeneous, one-step, fully contained reaction; test results are available in about 35 to 40 minutes. The assay cutoff for positivity is designed to allow detection of metastases that are larger than 0.2 mm in size; however, the assay does not discriminate between micrometastases (i.e., between 0.2 mm and 2 mm) and macrometastases (i.e., larger than 2 mm). The results would allow an intraoperative decision regarding the need for ALND.

Objective

To evaluate the use of the intraoperative GeneSearch™ BLN assay in detecting early stage breast cancer metastases in the sentinel lymph nodes compared to

1. postoperative histology or
2. other commonly used intraoperative evaluation methods, namely, intraoperative frozen section histology or imprint cytology.

Search Strategy

A MEDLINE® search was conducted via PubMed, covering references entered into the database from the year 2000 through December 2007. Text word searching was performed for the terms "GeneSearch" and Veridex. The search was limited to articles in English that discussed treatment of human patients. In addition, clinical data were sought from the U.S. Food and Drug Administration (FDA) and Veridex websites.

BlueCross BlueShield Association

An Association of Independent Blue Cross and Blue Shield Plans

Selection Criteria

Articles in peer-reviewed journals or FDA submissions reporting specifically on the GeneSearch assay were selected.

Main Results

FDA-approved tests are extensively reviewed for analytic validity. The GeneSearch BLN assay includes positive and negative controls to ensure reagent quality and instrument performance in each run; control failure invalidates the run. An internal control, consisting of mRNA detection from a constitutively expressed gene in lymph node tissue, monitors the quality of the sample preparation and RT-PCR reaction. A study to ensure replicability of results was also performed.

There is only one published study that provides evidence on the performance of the assay; additional information includes the manufacturer's submissions to the FDA, materials from the meeting of the Immunological Devices Panel meeting on November 16, 2006, and several abstracts presented at professional society meetings. The published study reports on 416 subjects from 11 sites in the U.S. Patients were those with a diagnosis of breast cancer and who were 18 years or older and who were scheduled for sentinel lymph node biopsies. The GeneSearch assay was performed alongside other tests commonly used at each facility; the GeneSearch results were not used for clinical management. The assay was compared to a reference standard, postoperative histology (primary study outcome), and to other intraoperative techniques where routinely conducted by the participating institution, namely, frozen section histology and imprint cytology. Among the 416 patients with postoperative histology results, 29.1% had positive lymph nodes for metastases, with a range across sites of 14.3% to 45.5%.

Compared to the reference standard, the sensitivity of the GeneSearch assay was 87.6% (95% CI: 80.4–92.9%); the specificity was 94.2% (95% CI: 90.9–96.6%); the positive predictive value was 86.2% (78.8–91.7%); and the negative predictive value was 94.9% (91.7–97.1%). The sensitivity was higher (97.9%; 95% CI: 92.5–99.7%) for patients with macrometastases than for those with micrometastases (56.5%; 95% CI: 34.5–76.8%). The authors assert that the false-positive rate may be lower than reported, because these results may be based on examination of parts of the lymph node that contain metastases, while the postoperative histology slides may come from a different slice of the node where there were no metastases, i.e., differences may be due to sampling distinct parts of the lymph node. While the researchers provide suggestive evidence, further research is needed to confirm or deny this hypothesis.

Among the 319 patients with results for both the assay and frozen section histology, the test performance of each technique compared to the reference standard was as follows:

Test	n	Sensitivity (95% CI)	Specificity (95% CI)	PPV	NPV
GeneSearch	319	95.6 (89.0–98.8)	94.3 (90.5–96.9)	86.9 (78.6–92.8)	98.2 (95.4–99.5)
Macrometastases	76	100			
Micrometastases	11	81.8			
Frozen section	319	85.6 (76.6–92.1)	97.8 (95.0–99.3)	93.9 (86.3–98.0)	94.5 (90.8–97.0)
Macrometastases	76	90.8			
Micrometastases	11	54.5			

The FDA analyzed the differences in test characteristics between these approaches using the same data. They concluded that the difference in sensitivity between GeneSearch and frozen section histology was statistically significant (difference=10; 95% CI: 2.5–17.7); as is the difference in the

negative predictive value (difference=3.7; 95% CI: 0.8–6.4). The FDA characterized the difference in specificity as borderline significant (difference=-3.5; 95% CI: -7.4–0.0) and the difference in the positive predictive value as insignificant (difference=-7.0; 95% CI: -14.8–1.3). However, this comparison of GeneSearch and the alternative intraoperative techniques was not part of the original protocol. Rather, it was a post-hoc analysis that did not include all participating sites; nor did sites performing alternative tests follow a uniform protocol for conducting the tests and for adjudication of results (as was used for the preplanned analysis). While the results are suggestive, they need additional confirmation. Additional analyses by the FDA suggest that both tests may operate along the same receiver operating characteristic (ROC) curve, in other words, that one test is not better overall than the other one. Data are inadequate to compare the performance of the GeneSearch assay to imprint cytology.

Taking the sensitivity and specificity of the assay in all 416 subjects included in the overall analysis, one can estimate the number of second surgeries avoided if a positive intraoperative assay result lead to immediate ALND and, conversely, the number of unnecessary ALNDs that would be performed because of false-positive assay results. Assuming a population of 1,000 patients, use of the GeneSearch assay to guide intraoperative decision-making regarding ALND would prevent 255 second surgeries and incur 41 unnecessary ALNDs, assuming the study average prevalence of 29.1% patients with positive lymph nodes. The number of second surgeries avoided would fall to 88 if the prevalence were 10%, with 52 unnecessary ALNDs. The number of second surgeries avoided would rise to 438 if the prevalence were 50%, with 29 unnecessary ALNDs. In reviewing the GeneSearch application, the FDA estimated that roughly 11% of patients undergoing ALND develop lymphedema and about the same proportion develop impaired shoulder movement (12–15% for patients undergoing radiotherapy; 7–8% for those not having radiotherapy).

There are several operational issues that add difficulty to the use of the GeneSearch assay, including the need for fresh specimens (rather than putting them in formalin for permanent fixation), the learning curve involved in reducing both the percentage of invalid results (from about 15% initially to 4–8% for more experienced technicians) and the time to perform the test (from 35–55 minutes to 30–35 minutes), and the potentially longer time required to perform the test compared to alternative intraoperative techniques (which take less than 15 minutes). Although not as much of an experienced pathologist's time is required to evaluate the test results, skill is needed in preparing the specimen.

Author's Conclusions and Comments

Question 1: What are the outcomes (i.e., surgeries avoided vs. unnecessary ALNDs) of using the GeneSearch BLN assay intraoperatively compared to waiting for postoperative histology results before deciding on whom to perform ALNDs?

Does the evidence permit conclusions? The GeneSearch BLN assay has potentially wide use, since a majority of the 180,000 (http://www.cancer.gov/cancertopics/types/breast) new patients per year with breast cancer undergo sentinel lymph node biopsy. A solid evidence base is especially important in this context, because the use of this assay could potentially lead to a major change in practice. There are a number of strengths to the study presented to the FDA, including the clear distinction between the training and test sets and double-reading of the permanent histology slides to increase the reliability of the reference standard. However, there are a number of weaknesses as well:

1. Patient recruitment is not described. For example, were these consecutive SLNB patients? Without this information, it is not possible to assess the potential for spectrum bias in the study results. The wide range of prevalence of positive nodes across sites (14.3% to 43.5%) adds weight to the importance of understanding how the subjects were selected.
2. There was substantial variation in the performance of the assay across sites. Sensitivity varied from 50% to 100%, and specificity ranged from 83.3% to 100%. Therefore, the Immunological Devices Panel questioned whether or not pooling into summary statistics was valid.

3. There are no corroborating studies. The FDA has requested two postapproval studies of the assay, which will substantially increase the evidence base. Other research is also underway, based on presentations at ASCO and other specialty society meetings.
4. GeneSearch does not distinguish between micro- and macrometastases. Although ASCO guidelines currently recommend that ALND be performed if either is detected, they also state that the prognostic significance of micrometastases is unknown, as is the importance of ALND in these patients. If GeneSearch has higher sensitivity because it detects more micrometastases (which is not clear), then the clinical significance of finding these patients may not balance the harms of ALND.
5. There is a learning curve for those conducting this assay. How best to train new staff to perform this test is unknown, as is the impact on patient care if and when the use of this assay is disseminated. The initial failure rate (i.e., invalid test result) was about 15%; it decreased to about 4–8% after personnel gained experience. The failure rate outside a research study is not known. The failure rate was higher with small nodes than with larger ones. GeneSearch proponents suggest that small nodes are less likely to contain metastases. The assay could have been rerun with the leftover homogenate or mRNA for the invalid tests, but this was not routinely done. Doing so would provide a better idea of the impact of these invalid results.

What Is the Net Benefit of the Assay? The key factor in assessing this assay is the trade-off between avoiding a second surgery to remove axillary lymph nodes if the sentinel node is positive versus risking unnecessary ALNDs if the assay produces a false positive result. The reason SLNB is the standard of care in early breast cancer, despite the fact that trials on its impact are still incomplete, is the desire to avoid ALND where possible, given the potential for significant, long-term morbidity. While the researchers conducting the GeneSearch study suggest that many of the false-positive assay results detect real metastases in the additional sections of the node used for the assay, the evidence presented to support this argument is insufficient. Given that GeneSearch was not used for patient treatment decisions in this study, longer-term follow-up of patients with false-positive test results to see if and when axillary metastases develop might be informative, although the utility of this approach may be compromised by the small number of patients involved.

As for the trade-off between avoiding a second surgery and undergoing an unnecessary ALND, patient preferences should play an important role in this decision. The FDA is requiring informed consent for participants in one of the postmarketing studies. Although the number of unnecessary ALNDs is substantially smaller than the number of second surgeries avoided, the "harms" from the former are more common and generally longer lasting than the harms associated with needing a second surgery. The sequelae of ALND can last for years or even a lifetime. The inconvenience of a second surgery, while difficult for patients coping with a serious diagnosis and substantial course of treatment, is likely to be short-lived. For these women, the harms of ALND are unavoidable.

Question 2: How does the sensitivity and specificity of the GeneSearch BLN assay performed intraoperatively during sentinel lymph node biopsy compare to the sensitivity and specificity of alternative intraoperative tests, namely imprint cytology and frozen section histology?

Does the evidence permit conclusions? The test set data provided to the FDA is useful in that it provides a direct comparison of the GeneSearch assay versus alternative intraoperative tests on the same patients. Several concerns, on the other hand, are outlined below:

1. The comparison of GeneSearch to alternative, intraoperative tests was not planned; not all study sites performed such tests.
2. The sample size of 29 for the imprint cytology comparison is too small to be meaningful.
3. The alternative tests were not conducted with the rigorous double-reading, review by central laboratory pathologists, and reconciliation of discrepancies applied to the reference standard.
4. The statistical approach to evaluating the difference in the test characteristics between GeneSearch and frozen section histology is not well described.

What Is the Net Benefit of the Assay? The focus here is on which technique—the assay, frozen section histology, or imprint cytology—performs best at avoiding second surgeries versus undergoing unnecessary ALNDs. The data on imprint cytology is inadequate and will not be discussed further. There is also inadequate evidence on whether the GeneSearch assay outperforms frozen section histology. However, an ROC analysis conducted by the FDA suggests that the GeneSearch assay and frozen section histology operate at different points on the same or very similar ROC curves. If this is correct (the analysis is not described in sufficient detail to assess this), then the question is not whether one technique is better than the other overall, but rather what the optimal tradeoff (and therefore assay cutoff) is between false-positive and false-negative results. Given the longer term and potentially more serious sequelae from ALND than from a second surgery, increasing the sensitivity of the test while sacrificing some specificity, as the GeneSearch assay does, may not be optimal.

The GeneSearch assay also provides less information for staging than other intraoperative procedures, since it cannot distinguish between micro- and macrometases. Nor can it indicate the location of the metastasis (inside or outside of the node). Postoperative histology is therefore required in all cases. It is less crucial when frozen section histology is performed, since pathologists can judge the size of the metastasis and its location from this test, although distortion is possible.

To summarize, the data available is inadequate to assess the clinical utility of the GeneSearch assay compared to either postoperative histology alone or to alternative intraoperative tests such as imprint cytology and frozen section histology. In addition, the balance of benefits versus harms may require higher specificity to avoid unnecessary ALNDs and their sequelae, whereas the GeneSearch design emphasizes sensitivity. Patient preferences should also play a key role in this calculation.

Based on the available evidence, the Blue Cross and Blue Shield Association Medical Advisory Panel made the following judgments about whether the use of GeneSearch breast lymph node assay to detect sentinel node metastases in early stage breast cancer meets the Blue Cross and Blue Shield Association Technology Evaluation Center (TEC) criteria.

1. **The technology must have final approval from the appropriate governmental regulatory bodies.**

The GeneSearch™ BLN Test Kit was cleared for marketing by the FDA on July 16, 2007.

2. **The scientific evidence must permit conclusions concerning the effect of the technology on health outcomes.**

There is only one published study on the performance of the GeneSearch assay, which reports on the same study used by the FDA to determine approval. This study has a number of limitations, including the lack of a description of patient recruitment, inadequate descriptions of several analyses performed, substantial variations in test performance across sites, and an ad hoc comparison of the assay to other intraoperative techniques. Additional validation studies are needed that address these issues. The FDA has mandated two postmarketing studies that will also help fill in the gaps. It would also be helpful to have information on patient preferences between avoiding second surgeries versus running the risk of unnecessary axillary lymph node dissection. Therefore, the data available on the GeneSearch assay are insufficient to permit conclusions regarding the effect of this technology on health outcomes.

3. **The technology must improve the net health outcome.**

The available data are insufficient to determine whether or not the technology improves the net health outcome.

4. The technology must be as beneficial as any established alternatives.

The available data are insufficient to determine whether or not the technology is as beneficial as the established alternatives.

5. The improvement must be attainable outside the investigational settings.

The available data are insufficient to determine whether or not the technology improves the net health outcome and therefore, whether any improvement is attainable outside investigational settings. There appears to have been a learning curve for using this technology, which was addressed in the study through additional training.

Based on the above, the use of GeneSearch breast lymph node assay to detect sentinel node metastases in early stage breast cancer does not meet the TEC criteria.

Contents

Published in cooperation with Kaiser Foundation Health Plan and Southern California Permanente Medical Group.

TEC Staff Contributors

Author—Barbara M. Rothenberg, Ph.D.; **TEC Executive Director**—Naomi Aronson, Ph.D.; **Managing Scientific Editor**—Kathleen M. Ziegler, Pharm.D.; **Research/Editorial Staff**—Claudia J. Bonnell, B.S.N., M.L.S.; Maxine A. Gere, M.S.

Blue Cross and Blue Shield Association Medical Advisory Panel

Allan M. Korn, M.D., F.A.C.P.—Chairman, *Senior Vice President, Clinical Affairs/Medical Director, Blue Cross and Blue Shield Association;* **Alan M. Garber, M.D., Ph.D.**—Scientific Advisor, *Staff Physician, U.S. Department of Veterans Affairs; Henry J. Kaiser, Jr., Professor, and Professor of Medicine, Economics, and Health Research and Policy, Stanford University*; **Steven N. Goodman, M.D., M.H.S., Ph.D.**—Scientific Advisor, *Associate Professor, Johns Hopkins School of Medicine, Department of Oncology, Division of Biostatistics (joint appointments in Epidemiology, Biostatistics, and Pediatrics)*—American Academy of Pediatrics Appointee. ■ **Panel Members Peter C. Albertsen, M.D.,** *Professor, Chief of Urology, and Residency Program Director, University of Connecticut Health Center;* **Sarah T. Corley, M.D.,** *Physician Consultant, NexGen Healthcare Information Systems, Inc.*—American College of Physicians Appointee; **Helen Darling, M.A.** *President, National Business Group on Health;* **Josef E. Fischer, M.D., F.A.C.S.,** *William V. McDermott Professor of Surgery, Harvard Medical School and Chair, Department of Surgery, Beth Israel Deaconess Medical Center*—American College of Surgeons Appointee; **Willard K. Harms, M.D., M.M.,** *Medical Director, Medical Policy and Adjudication, Blue Cross Blue Shield of Illinois;* **I. Craig Henderson, M.D.,** *Adjunct Professor of Medicine, University of California, San Francisco;* **Mark A. Hlatky, M.D.,** *Professor of Health Research and Policy and of Medicine (Cardiovascular Medicine), Stanford University School of Medicine;* **Walter A. Hollinger, M.D., M.M., M.H.P.E., F.A.C.P.,** *Senior Medical Director, Care Management, Blue Cross and Blue Shield of Florida;* **Bernard Lo, M.D.,** *Professor ofMedicine and Director, Program in Medical Ethics, University of California, San Francisco;* **Barbara J. McNeil, M.D., Ph.D.,** *Ridley Watts Professor and Head of Health Care Policy, Harvard Medical School, Professor of Radiology, Brigham and Women's Hospital;* **Joel Owerbach, Pharm.D.,** *Vice President and Chief Pharmacy Officer, Excellus Health Plans;* **William R. Phillips, M.D., M.P.H.,** *Clinical Professor of Family Medicine, University of Washington*—American Academy of Family Physicians' Appointee; **Maren T. Scheuner, M.D., M.P.H.,** *Natural Scientist in the Division of Behavioral and Social Sciences, RAND Corporation; Adjunct Associate Professor, UCLA School of Public Health*—American College of Medical Genetics Appointee; **J. Sanford Schwartz, M.D.,** *Professor of Medicine, Department of Medicine, University of Pennsylvania School of Medicine and Professor, Health Care Systems, Health Management & Economics, The Wharton School;* **Earl P. Steinberg, M.D., M.P.P.,** *President and CEO, Resolution Health, Inc.;* **A. Eugene Washington, M.D., M.Sc.,** *Executive Vice Chancellor and Provost, University of California, San Francisco;* **Jed Weissberg, M.D.,** *Associate Executive Director for Quality and Performance Improvement, The Permanente Federation.*

Assessment Objective

To evaluate the use of the intraoperative GeneSearch™ breast lymph node assay (hereafter referred to as "GeneSearch" or "the assay") in detecting early stage breast cancer metastases in the sentinel lymph nodes compared to

1. postoperative histology or
2. other commonly used intraoperative evaluation methods, namely, intraoperative frozen section histology or imprint cytology.

Background

Axillary Lymph Node Metastases

Whether breast cancer has spread to the axillary lymph nodes is the most significant prognostic factor for patients with early stages of the disease (Lyman et al. 2005). Staging evaluation of the axillary lymph nodes is used to define prognosis and to determine appropriate therapy (see National Comprehensive Cancer Care Guidelines at http://www.nccn.org/professionals/physician_gls/PDF/breast.pdf). Although breast cancer treatment varies in accordance with a number of other factors as well, such as tumor size, hormone receptor status, and HER2 status, in general, the greater the number of positive nodes, the more aggressive the treatment. Patients are categorized as node negative, 1–3 positive nodes, 4–9 positive nodes, or 10 or more positive nodes.

Removing the lymph nodes and staging them histologically is still the most accurate method for determining the presence or absence of disease. Removing axillary lymph nodes (axillary lymph node dissection or ALND), may also be therapeutic, if removing tumor-involved nodes improves local control. Results of the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-04 trial did not find that ALND improves survival, although the data were insufficient to rule out the possibility of a small, but clinically meaningful, benefit (Rockette et al. 1982). A Bayesian meta-analysis of 6 other randomized controlled trials showed a survival advantage collectively of 5.4% (95% CI: 2.7–8.0%) (Orr 1999); however, these patients did not receive adjuvant chemotherapy. A survival advantage has been suggested in analyses of large cancer data bases (Harlow et al. 2005; Luini et al. 2005). Thus, the true impact of ALND on survival in patients receiving recommended chemotherapy remains uncertain.

Sentinel Lymph Node Biopsy

Sentinel lymph node biopsy (SLNB) is a less-extensive surgery that has largely supplanted ALND for staging certain groups of patients, even in the absence supporting evidence from long-term clinical trials. Several trials comparing SLNB and ALND are currently underway, including the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-32. This trial has accrued 5,611 subjects and examines the impact of SLNB on survival and morbidity (Lyman et al. 2005).

According to guidelines from the American Society for Clinical Oncology (ASCO), indications for SLNB include T1 or T2 tumors, multicentric tumors, and ductal carcinoma in situ (DCIS) with mastectomy. Contraindications include T3 or T4 tumors, inflammatory breast cancer, pregnancy, prior axillary surgery, prior non-oncologic breast surgery, prior systemic therapy, and clinically palpable lymph nodes (Lyman et al. 2005).

SLNB uses a tracer (usually blue dye and a radiolabeled colloid) injected into the breast tissue around the primary tumor, which drains through the lymphatics toward the axilla. The earliest lymph nodes identified by the tracer are designated the "sentinel" nodes. These nodes are surgically removed and directly analyzed for presence of tumor. Surgeons become more adept at identifying sentinel lymph nodes with experience. The ASCO guideline reported that the American Society of Breast Surgeons recommends a rate of sentinel lymph node identification of at least 85% with a false-negative rate of no more than 5% (Lyman et al. 2005). That this is achievable is supported by studies that have successfully identified sentinel lymph nodes in 94–96% of patients with a low false-negative rate (Ozmen and Cabioglu 2006).

An unresolved question regarding SLNB is the clinical significance of the size of the metastasis. Metastases are categorized as macrometastases, with a size greater than 2 mm in widest diameter, or micrometastases, with a size between 0.2 mm and 2 mm. Smaller lesions (<0.2 mm) or even individual cells may be detected, but these are currently considered negative findings. Contradictory information is available on the prognostic value of micrometastases, and clinical trials are currently under way to gauge their long-term impact. In its SLNB guidelines, ASCO notes the lack of

definitive information on the impact of micrometastases and isolated tumor cells.[1] Citing McCready et al. (2004), however, the guideline notes that metastases are found in the non-sentinel lymph nodes of about 10% of patients with isolated tumor cells in the sentinel lymph node and in 20–35% in patients with micrometastases. The guidelines therefore currently recommend ALND following the detection of micrometastases in sentinel lymph nodes (Lyman et al. 2005).

Morbidity Associated with Axillary Lymph Node Dissection and Sentinel Lymph Node Biopsy

While removal of axillary lymph nodes with metastatic disease clearly improves cancer prognosis, the potential morbidity of ALND is also well established. The morbidity of axillary lymph node dissection depends on the extent of removal of lymph nodes, categorized into three levels: I (most superficially located), II, and III (most deeply located). Most staging axillary lymph node dissections currently involve removal of nodes from levels I and II, with Level III removed only if gross disease is visible in level II nodes.

Adverse effects from ALND include seroma formation, impairment of shoulder movement, neuropathy, and arm lymphedema; in some cases, these effects last for years. Wide ranges of incidence have been reported for most of these: 4% to 52% for seroma formation; up to 70–80% for numbness and paresthesia after division of intercostobrachial nerve; and 6–30% for lymphedema (Purushotham et al. 2005). In a review article on lymphedema, Sakorafas et al. (2006) state that the incidence of lymphedema has fallen over time as surgeries become less radical and that it is now about 10–20% among women treated for breast cancer.

In a study of 298 patients with early breast cancer (tumor size less than 3 cm, no neoadjuvant chemotherapy, no clinically involved lymph nodes, no multifocal breast disease), 155 received ALND (control group) and 143 underwent SLNB (Purushotham et al. 2005). In the SLNB arm, 57 subsequently had ALND either because the sentinel lymph nodes could not be found (9) or the SLN biopsy was positive (48). Thirty-four percent of the patients assigned to SLNB were node positive, as were 26% of those assigned to ALND; this difference was not statistically significant. Of particular interest is the comparison of patients with negative nodes who were assigned to ALND versus those assigned to SLNB, who did not have ALND. Patients undergoing SLNB alone had a significantly smaller change in arm volume (corrected for overall weight changes as reflected in the contralateral arm) than those having ALND during the first 6 months after surgery (mean difference at 1 month=61.0 mL, $p<0.001$; at 6 months=43.5 mL, $p=0.03$); the difference was not statistically significant at 12 months. In both groups of patients, the volume of the ipsilateral arm increased over this period and was still larger at 12 months.

The odds ratio for subjectively identified lymphedema (by physician or patient) for SLNB patients vs. ALND patients ranged from 0.18 to 0.24 during the year after surgery (p value ranged from 0.007 to 0.03 and was 0.007 at 1 year). The percentage of patients with lymphedema was not reported. There was no statistically significant difference in the development of lymphedema among node-positive patients in each arm of the trial. The incidence of seroma was also higher in node-negative patients undergoing ALND (24% vs. 11% for SLNB patients, $p=0.01$). For any paresthesia, node-negative ALND patients were more likely to have any paresthesia than node-negative SLNB patients (48% vs. 23%, OR=0.31, $p<0.001$).

Postoperative sequelae, including impaired arm functioning, numbness in the ipsilateral arm, and poor range of movement, have also been reported for patients in the large ALMANAC trial. However, the comparison between the arms of the trial is complicated by the fact that (1) the ALND arm includes Level I-III dissection and four-node sampling; (2) positive results in the SLNB arm are followed up by either ALND or radiotherapy, and (3) results are not reported by nodal status. This makes the comparison of sequelae from ALND versus SLNB less clear, because SLNB patients with positive nodes will also undergo ALND and because the sequelae from ALND itself can vary by nodal status, as Purushotham et al. (2005) show.

[1] According to the 6th edition (2006) of the AJCC Cancer Staging Manual, "ITCs are defined as single cells or small clusters of cells not greater than 0.2 mm in largest dimension, usually with no histologic evidence of malignant activity (e.g., proliferation or stromal reaction)" (http://caonline.amcancersoc.org/cgi/content/full/56/1/37#TBL2).

In reviewing the GeneSearch application, the U.S. Food and Drug Administration (FDA) asserted that roughly 11% of patients undergoing ALND develop lymphedema and about the same proportion develop impaired shoulder movement (12–15% for patients undergoing radiotherapy; 7–8% for those not having radiotherapy) (FDA 2006d).

Patients who test positive for metastases during the postoperative histology evaluation and who do not have axillary dissection during the SLNB surgery (because no intraoperative assessment was performed or the result was a false negative) will undergo a second surgery to remove additional axillary lymph nodes. In addition to the obvious inconvenience and discomfort for the patient of having to undergo repeated surgery, there is a small risk from the administration of anesthesia and there are varying opinions on whether the operating through the scar tissue from the SLNB makes it more difficult (FDA 2006a). Although the risk from anesthesia varies with the type of surgery and the patients' overall physical condition, the American Society of Anesthesiologists estimates that the risk of death from anesthesia is 1 in 250,000 (wwww.asahq.org).

Laboratory Examination of Sentinel Lymph Nodes

Resected sentinel lymph nodes are embedded in paraffin and sectioned for microscopy, e.g., in 2-mm increments (Cody 1999). Slides are stained with hematoxylin and eosin (H&E) to enhance histologic features and microscopically examined by a pathologist for evidence of metastatic malignant cells. Nodes for which no visual evidence of metastasis is found can be further evaluated by immunohistochemistry (IHC) using antibodies that are reactive with pancytokeratin, a sensitive and specific marker for epithelial cells. Because of the time and labor involved, the histologic examination is conducted postoperatively.

Intraoperative evaluation of sentinel nodes potentially allows determination of nodal status during surgery and ALND for patients with evidence of metastases during the same operation. The most commonly used intra-operative methods are touch imprint cytology and frozen section histology. Both can be completed within a short enough time frame to allow an intraoperative decision regarding the need for ALND. An ideal intraoperative test would, when compared to postoperative histology, detect all positive nodes and avoid all second surgeries; and would deliver no false-positive results, as these would result in unnecessary ALNDs and possible sequelae.

Touch imprints are made by bisecting the lymph node and pressing each cut half onto a glass slide. The slides are fixed, stained and examined for malignant cells. Slide preparation takes no more than a few minutes. This procedure is most effective when metastases are large. Tew et al. (2005) conducted a meta-analysis of 31 studies of intraoperative touch imprint cytology versus postoperative histology. The pooled sensitivity for metastases was 63% (95% CI: 57–69) and pooled specificity was 99% (95% CI: 98–99). The pooled sensitivity for macrometastases (>2 mm) was 81% but for micrometastases (0.2 mm to 2 mm) was only 22%.

A portion of the node (e.g., alternating sections) may be used for frozen section. In this procedure, tissue is frozen quickly to about -20 degrees C, cut into thin sections, placed onto slides, then fixed and stained with H&E. The slides are examined for malignant cells; the morphology of frozen sectioned tissue is adequate, although as the potential for distortion of the tissue architecture is greater than with paraffin-embedded and sectioned material. The slide preparation takes approximately 10 minutes. Tew et al. (2005) reported that frozen section had better sensitivity than imprint cytology in 3 of 4 direct comparison studies included in the meta-analysis of imprint cytology. This is in keeping with several studies that report frozen section sensitivities of approximately 75% to 85%, and specificities of 99 to 100% in patient populations numbering 70 to more than 2,000 (Al-Shibli et al. 2005; Leung et al. 2007; Tanis et al. 2001; Celebioglu et al. 2006; Schrenk et al. 2005). Although frozen section is typically compared to postoperative histology as the reference standard, frozen section may detect small metastases that are not detected by post-operative histology (e.g., if metastases are small and presumably present only in the sections examined by frozen section) and thus positive results by frozen section or by postoperative histology may be accepted (Arora et al. 2008).

Some centers perform both imprint cytology and frozen section intraoperatively (Regan Fulton, MD, personal communication) whereas others believe that imprint cytology adds no useful additional information (Celebioglu et al. 2006; Al-Shibli et al. 2005). More recently, rapid

pancytokeratin IHC methods for intraoperative frozen section have been reported. In one study, frozen section-IHC improved sensitivity (Johnston et al. 2006) compared to frozen section alone, but in another study there was no difference (Celebioglu et al. 2006).

Another method of detecting nodal metastases is by using reverse transcriptase polymerase chain reaction (RT-PCR) to detect one or a few gene expression markers of epithelial cells (reviewed in Cserni 2007). In this type of assay, tissue architecture is not preserved. Thus, sampling the lymph node at different levels for RT-PCR plus using the remaining sample to conduct histology appears to be preferred.

The GeneSearch™ BLN assay, developed and commercialized by Veridex, LLC (Warren, NJ), uses real-time RT-PCR to qualitatively evaluate nodal sections for the presence of the gene expression markers, mammaglobin and cytokeratin 19. These genes are expressed at higher levels in breast tissue, but not in normal nodal tissue. The assay is automated and performed in a homogeneous, one-step, fully contained reaction; test results are available in 35 to 40 minutes. The assay cutoff for positivity is designed to allow detection of metastases that are greater than 0.2 mm in size (i.e., includes detection of micrometastases); however, the assay does not discriminate between micro- and macrometastases.

FDA Status. The premarket approval application for the GeneSearch™ BLN Test Kit was approved by the U.S. Food and Drug Administration (FDA) on July 16, 2007. According to the GeneSearch™ kit insert (available at http://www.veridex.com/pdf/MKG-1545_BLN_EnglishOnly_IFU_Reprint._Rev1.pdf), the intended use of the assay is "to guide the intra-operative or post-operative decision to remove additional lymph nodes." Furthermore, "Post-operative histological evaluation of permanent sections of the tissue specimen, in accordance with usual diagnostic practice and using the Veridex lymph node cutting scheme [alternating sections of ≤3 mm], is required." Lymph nodes that are too small to provide samples sufficient for both the GeneSearch™ BLN Test and to examine histology should be used to examine the histology alone. The FDA's consumer information sheet for this test (http://www.fda.gov/cdrh/mda/docs/P060017.html) also states "The GeneSearch™ Breast Lymph Node (BLN) test may be used with sentinel lymph node biopsy for patients who have been informed about the risks, benefits and limitations of this test." The GeneSearch test must be performed in CLIA-licensed laboratories.

The FDA also has required that Veridex, LLC, conduct two postapproval studies on the GeneSearch assay (see approval decision at http://www.fda.gov/cdrh/pdf6/p060017a.pdf):

1. A "positive-predictive-concordance" (defined as the proportion of BLN assay positive subjects with histologically detectable breast cancer metastases) study between the assay and routine H&E histology of both sentinel lymph nodes and, when available, non-sentinel axillary lymph nodes. The secondary objectives include estimating the PPV against more extensive H&E sectioning, analysis based on size of metastases, evaluation of proportion of invalid assay results, assessment of site-to-site variability of assay performance, assessment of "negative predictive concordance" against sentinel lymph node routine histology only, and assessment of positive and negative likelihood ratios. The study will enroll at least 1,000 patients, with a minimum of 246 lymph node positive patients. Patients with lymphoma or prior chemotherapy or hormonal therapy will be excluded. All subjects must be counseled on the use of the assay "as part of physician-patient discussion of lymph node dissection procedures and complications." Persons performing the assay must be trained before assay results can be used for clinical decision-making. The study will be completed by July 2010.
2. A study to assess assay timing, specifically, how long it takes and whether the result was received in time to make an intraoperative decision. Data on other surgical procedures performed at the same time will be collected, to determine whether use of the assay lengthened the total surgical time. The sample size will be at least 320 patients and will be completed by July 2008.

Methods

Search Methods

A MEDLINE® search was conducted via PubMed, covering references entered into the database between 2000 and December 2007. Text word searching was performed for the terms "GeneSearch" and Veridex. The search

was limited to articles in English that discussed treatment of human patients. In addition, we searched for information on the GeneSearch™ BLN assay on the FDA website and looked for references to additional studies on the Veridex website.

Study Selection

Articles in peer-reviewed journals or FDA submissions reporting specifically on the GeneSearch assay were selected. One article on GeneSearch was found in peer-reviewed journals. Additional information include the manufacturer's submissions to the FDA; materials from the meeting of the FDA Center for Devices and Radiological Health's Immunology Devices Panel meeting on November 16, 2006 (available at http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfAdvisory/details.cfm?mtg=671); and several abstracts presented at professional meetings. Additional sources included the GeneSearch™ kit insert (available at http://www.veridex.com/pdf/MKG-1545_BLN_EnglishOnly_IFU_Reprint._Rev1.pdf). GeneSearch BLN is the only FDA-approved, commercially available assay that evaluates sentinel lymph nodes for metastases. Therefore, articles on molecular testing of sentinel lymph node tissue were not included even if they reported on the expression of the same genes as used in the GeneSearch assay.

Medical Advisory Panel Review

This Assessment was reviewed by the Blue Cross and Blue Shield Association Medical Advisory Panel (MAP) on October 17, 2007. To maintain the timeliness of the Assessment's scientific information, literature search updates were performed subsequent to the Panel's review (see "Search Methods"). If the search updates identified any additional studies that met the criteria for detailed review, the results of these studies were included in the tables and text where appropriate. There were no studies that would change the conclusions of this Assessment.

Formulation of the Assessment

Patient Indications

The GeneSearch™ BLN assay is intended for patients diagnosed with invasive breast cancer who are scheduled to undergo a sentinel lymph node biopsy and who have been informed about the benefits, risks, and limitations of this assay. The assay is marketed for intraoperative use on freshly excised sentinel lymph node tissue; the results are used to decide whether to proceed with axillary lymph node dissection during the same surgery.

Technologies to be Compared

The GeneSearch™ BLN assay is compared to 1) the reference standard, which is final, postoperative histology of sentinel lymph node tissue, and 2) other intraoperative tests, namely, frozen section histology and imprint cytology, where direct comparison data are available.

Health Outcomes

Benefits. Patients with true positive results benefit by having a necessary axillary lymph node dissection during the same surgery as the sentinel node biopsy and avoiding a second surgery.

Harms. Patients with false-positive results are harmed by having an unnecessary axillary lymph node dissection and in some cases by experiencing sequelae, including lymphedema.

Patients with true- or false-negative results, or whose intraoperative result is not available due to an assay failure, experience the same outcomes as when no intraoperative assay is used.

Specific Assessment Questions

1. What are the outcomes (i.e., surgeries avoided vs. unnecessary ALNDs) of using the GeneSearch BLN assay intraoperatively compared to waiting for postoperative histology results before deciding on whom to perform ALNDs?
2. How do the sensitivity and specificity of the GeneSearch BLN assay performed intraoperatively compare to alternative intraoperative tests, namely, frozen section histology and imprint cytology?

Review of Evidence

Analytic Validity

The analytic validity of a genetic test defines its ability to accurately and reliably measure the genotype of interest. Components of analytic validity include analytic sensitivity and specificity, laboratory quality control, and assay robustness. Although analytic validity will ultimately be reflected in the evidence gathered to support clinical validity and clinical utility, a separate review allows estimation of its likely contribution to any uncertainty regarding

clinical validity and utility as well as information on likely routine laboratory performance.

FDA-approved tests are extensively reviewed for analytic validity; results are summarized in the GeneSearch™ BLN kit insert and in the FDA Advisory Committee materials (see Methods, Study Selection). Briefly, the GeneSearch BLN assay includes positive and negative controls to ensure reagent quality and instrument performance in each run; control failure invalidates the run. An internal control, consisting of mRNA detection from a constitutively expressed gene in lymph node tissue, monitors the quality of the sample preparation and RT-PCR reaction. Potential interfering substances include excess fat tissue, lymphoma cells, and contaminating breast or breast tumor tissue. The impact of the use of neoadjuvant (presurgical) chemotherapy on the accuracy of the GeneSearch assay is not known.

Two operators from each of three sites measured four different lymph node samples (72 total replicates per sample) across kit lot numbers and across days. Percent CVs[2] for all marker results were ≤6.82% for intra-run, inter-run, inter-site, inter-operator and inter-lot analyses, indicating a high degree of precision and robustness. Although the assay result is qualitative, the assay met the criteria for a linear response between the lowest level detected (about 10^2–10^3) and highest level tested (10^8) copies per µL for all markers tested. In a test set of 274 samples (see Clinical Validity, training set, following), measured by GeneSearch and compared to known histology (i.e., known presence or absence of markers), the analytic sensitivity was 91.1% (95% CI: 82.5–96.4%) and the analytic specificity was 95.9% (95% CI: 92.1–98.2%). The kit insert includes a detailed procedure for sample preparation and analysis that should allow for similar performance in other laboratories.

Clinical Validity

The clinical validity of a genetic test defines its ability to detect or predict the associated disorder, phenotype, or outcome. In the case of GeneSearch™ BLN, the test is intended to detect the presence or absence of clinically relevant metastases (defined as greater than 0.2 mm in size).

The primary evidence on the performance of the GeneSearch™ BLN assay comes from documents submitted by the manufacturer, Veridex, to the FDA, as well as the published article on this study (Blumencranz et al. 2007a). Two major studies were reported: in one study, the assay cutoff value for a positive result was determined (training set); in the other, clinical validity of the assay was evaluated (test set; designated 'Pivotal Study' by the manufacturer). Both studies were multicenter (all located in the U.S.) and enrolled patients 18 years or older diagnosed with invasive breast cancer.

Freshly excised lymph nodes were bisected along the short axis into sections of 1.5–3 mm in thickness. After touch preparation intraoperative cytology (if part of site's procedure), alternate sections were designated for GeneSearch BLN or histology, the latter including both intraoperative frozen section (if part of site's routine procedure) and final postoperative histology (routine hematoxylin and eosin [H&E] stained slide evaluation and immunohistochemistry if H&E negative). Histology was reported by on-site pathologists and also by at least 2 pathologists at a central trial facility (disagreements resolved by a third pathologist). Site slides reported positive for nodal metastases but centrally reported negative (on separate slides) were sent to the central facility and re-evaluated; they were counted as positive if the central facility pathologists agreed with that designation.

Training Set (determination of cutoff). In this study 306 samples were tested by GeneSearch™ BLN Assay and postoperative histology. Data from 30 subjects with invalid GeneSearch™ BLN Assay results (9.8%) and 2 with indeterminate postoperative histology results were excluded, since they could not contribute to the cutoff determination. Cutoffs for the internal control and for the mammaglobin (MG) and cytokeratin 19 (CK19) gene expression markers were chosen to achieve a minimum specificity of 95% compared with the histological result.

Test Set (comparison of GeneSearch to postoperative histology). In this study 416 patients were tested by GeneSearch BLN assay and postoperative histology (Blumencranz et al. 2007a,b; FDA 2006a-c, 2007a,b). The patients

[2] Variability across multiple measurements of the positive and negative controls, and of positive and negative nodal tissue samples, was measured and reported as percent coefficient of variance (%CV = standard deviation/mean × 100).

were recruited from 11 sites in the U.S. with previous research experience in SLNB studies through the American College of Surgeon's Oncology Group or the National Surgical Adjuvant Bowel and Breast Project. The patient selection process is not described in detail, e.g., whether they were consecutive patients scheduled for SLNB. Characteristics of the study participants are summarized in Table 1.

The test characteristics for the GeneSearch assay are shown in Table 2. There was substantial variation in the number of subjects and the sensitivity and specificity across sites (data not shown). The sample size per site ranged from 7 to 124; the sensitivity ranged from 50% to 100%, and the specificity ranged from 83.3% to 100%. A Breslow-Day test for heterogeneity across sites yielded a p value of 0.066, which the FDA characterized as borderline significant. The data presented to the FDA relied on a fixed-effects model. The FDA statistician suggested that a random-effects model might account for variation across sites, which sometimes is not detected by low-powered tests for heterogeneity. Random-effects models generally produce larger confidence intervals. Little detailed information was found on these analyses, which were only alluded to in the transcript of the Immunology Devices Panel meeting.

The assay is significantly more sensitive in detecting macrometastases than micrometastases. According to the researchers, there is no evidence of different performance based on tumor histology, size, or stage, although the sample size in some of these subgroups is small; no specific statistical tests were reported.

A patient was considered positive for metastases if at least 1 node was positive. According to the postoperative histology, 29.1% of the patients had nodal metastases, with a range across sites of 14.3% to 43.5%. Table 3 reports the positive and negative predictive values for a similar range of prevalence of positive results.

When the prevalence of positive nodes is lower, the positive predictive value is also lower and the negative predictive value is higher. When the prevalence of positive nodes increases, the positive predictive value improves, but the negative predictive value decreases, i.e., there are more false-negative findings.

The results of the reference test, postoperative histology, varied across readers and between the test site and the central laboratory. The study design addressed this by using two central readers (93.8% agreement re: node positive or negative), with a third tie breaker where needed, and by attempting to reconcile differences between site and central laboratory readings. While differences across readers of the same slide or image occur in a number of applications, evaluating sentinel lymph nodes is complicated by the fact that different readers may be looking at different sections of the node. Metastases, especially micrometastases, may not be evident in all sections of the node.

Table 1. Characteristics of Participants in Test Set

Parameter	Results
Age	Mean: 60 years; range: 27 to 92 years
Gender	418 females, 5 males
Stage	I: 62.3% II: 32.0% III: 5.3% IV: 0.5%
Type of cancer	Invasive ductal alone or with other types: 80.4% Invasive lobular with no invasive ductal: 13.9% Other than ductal or lobular: 5.7%
Estrogen receptor positive	79.2%
Progesterone receptor positive	67.8%
HER2 negative	74.2%
Patients receiving chemotherapy or radiation therapy	9, chemotherapy, 1, radiotherapy

Table 2. Test Characteristics of GeneSearch™ BLN Assay

	Sensitivity (95% CI)	Specificity (95% CI)	PPV	NPV	Patients with Positive Nodes*
All patients	87.6% (80.4–92.9%)	94.2% (90.9–96.6%)	86.2% (78.8–91.7%)	94.9% (91.7–97.1%)	29.1%
Patients with macrometastases by histology	97.2% (92.5–99.7%)				
Patients with micrometastases by histology	56.5% (34.5–76.8%)				
Patients with Stage I disease (n=248)	82.9% (66.4–93.4%)	95.3% (91.5–97.7%)			14.2%
Patients with Stage II disease (n=128)	85.9% (75–93.4%)	88.5% (77.8–95.3%)			51.2%
Patients with Stage III disease (n=21)	100% (73.5–100%)	100% (63.1–100%)			60.0%
Patients with invasive ductal cancer (n=338)	88.8% (80.8–94.3%)	94.1% (90.3–96.7%)	86.1%	95.3%	29.2%
Patients with invasive lobular cancer (n=59)	80.0% (56.3–94.3%)	91.9% (78.1–98.3%)	84.2%	89.5%	35.1%
Patients with other invasive cancer (n=24)	100% (29.2–100%)	100% (83.9–100%)	100%	100%	13.0%

*The sample size for the prevalence sometimes varies slightly from the reported sample size used in calculating test characteristics.

Table 3. Variation in Gene Search Test Characteristics for Varying Prevalence of Nodal Metastases

% Prevalence of a Positive Node	% Sensitivity	% Specificity	% Positive Predictive Value	% Negative Predictive Value	# Second Surgeries Avoided* (Assume n=1,000)	# Unneeded ALNDs* (Assume n=1,000)
29.1 (Test set)	87.6	94.2	86.2	94.9	255	41
10	"	"	62.8	98.6	88	52
20	"	"	79.2	96.8	175	46
30	"	"	86.5	94.7	263	41
40	"	"	90.9	91.9	350	35
50	"	"	93.8	88.4	438	29

*Compared to no intraoperative nodal testing and therefore no axillary node dissection during surgery for sentinel lymph node biopsy.

Therefore, a discrepancy between site and central readers may be due to the fact that metastasis is present in the section of the node examined in one location but not in the section examined at the other.

To illustrate this issue, the researchers compared site readings against central laboratory readings, assuming that the latter were correct. The site pathologists had a sensitivity of 94.1% and a specificity of 95.8%. Central laboratory pathologists obtained the site slides for 21 nodes reported as positive by site pathologists and negative by central pathologists. Upon reviewing the site slides, the central lab pathologists changed their decision and agreed that 13 of the 21 were positive; 4 were still judged to be negative (4 were missing central pathologist evaluation). These revised numbers were used in calculating the sensitivity and specificity of the assay. Therefore, it is possible that some of the nodes found positive by GeneSearch and negative by postoperative histology are, in fact, positive. However, because the GeneSearch process homogenizes the issue, it is not possible to verify this after the fact.

There are several operational issues regarding the use of the GeneSearch assay that may affect its use but that are not relevant for assessing its clinical utility:

- The assay requires fresh tissue, which would require a change in procedure. Tissue currently is placed in formalin when it is transported from the operating room to the pathology laboratory.
- The assay requires protection against cross-contamination. Because the assay detects breast tissue in the node and not cancer specifically, any contamination by breast tissue that might, for example, be removed during concurrent breast-conserving surgery could contaminate the sample and produce a false-positive result.
- While much of the test itself is automated and does not require a pathologist as with the other intraoperative tests, a highly skilled technician or pathologist is needed to separate the node from surrounding fat and to slice the node. Precise slicing is not needed for the GeneSearch assay, but it is needed to conduct postoperative histology. The presence of too much fat in the tissue used for the assay can also contaminate the sample. Furthermore, the assay should be performed as quickly as possible after the tissue is removed from the patient, to minimize RNA degradation.
- It takes about 30–35 minutes for experienced technicians to perform the assay on 1 to 3 nodes, compared to about 35–55 minutes for those with less practice. This appears to be longer than required for the alternative, intraoperative tests (which take less than 15 minutes).

Clinical Utility

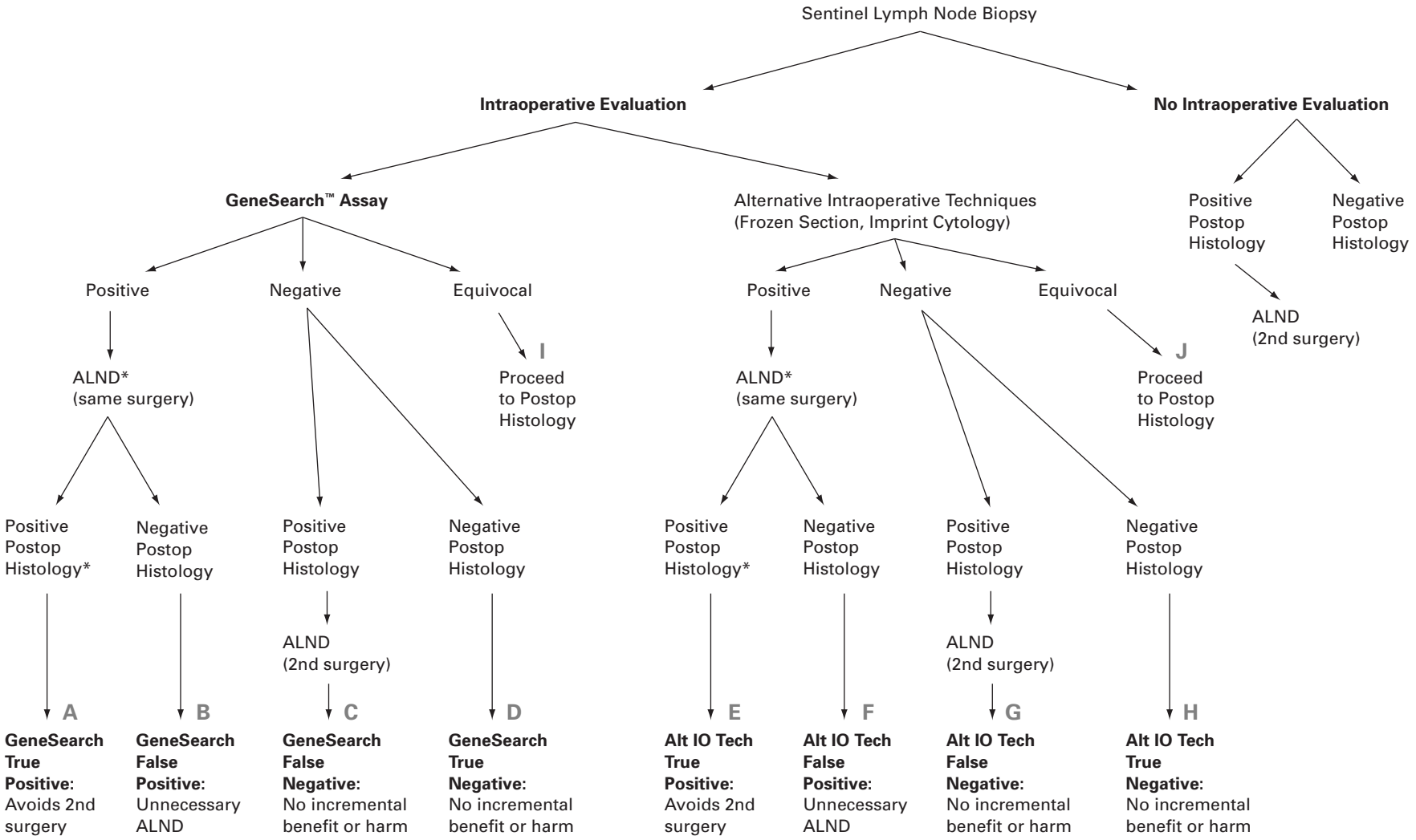
A genetic test has clinical utility if the use of the test result to influence management decisions improves patient outcomes compared to not using the test. When the genetic test replaces another test already in use, comparative effectiveness and incremental benefit on outcomes should also be evaluated.

The Figure provides a flow chart that shows the decision process for using the results from the GeneSearch assay versus alternative strategies. Surgeons performing SLNB select whether or not to perform intraoperative testing. Those who do not use intraoperative tests send the sentinel lymph nodes to the histology laboratory, where they are permanently fixed and stained; the results are usually available within 1–2 days. If metastases are found, patients usually undergo a second surgery, where ALND is performed and additional nodes are removed for examination. If the surgeon selects an intraoperative test, the lymph node is removed and immediately sent for evaluation during surgery. If the patient is having breast-conserving surgery, the sentinel lymph nodes are often removed first, and the lumpectomy is performed while the nodes are being tested.

If the nodes are positive, the surgeon usually immediately performs an axillary lymph node dissection; if the node is negative, the remaining axillary nodes are left intact. Postoperative histology is still performed, and patients with intraoperative false-negative results will generally have a second surgery to undergo axillary node dissection.

The documentation regarding GeneSearch is ambiguous about whether the assay is intended as a replacement for other intraoperative tests, as a complement to other intraoperative tests, or as a replacement for postoperative histology testing. The current kit insert specifies that the use of the assay must be followed by postoperative histology testing. The test set trial was

Figure. Decision Tree for GeneSearch vs. Alternative Strategies



*Postoperative Histology=Reference test; ALND=Axial Lymph Node Dissection; Alt IO Tech=Alternative intraoperative technique

designed to assess the use of GeneSearch as a substitute for other intraoperative testing, so the researchers did not evaluate whether and how the assays results might be combined with those of other assays.

The GeneSearch™ BLN assay results were not used to make patient management decisions in the Pivotal Study, so there is no direct evidence that using the test results in improved net outcome; thus, clinical utility is inferred by performance results. While indirect, inferences for clinical utility are probably quite reliable, because the algorithm for the use of the test is straightforward.

Benefits and Harms. The rationale for intraoperative testing is to identify metastases in time to perform axillary node dissection, when needed, during the same surgery and thereby avoid second surgery to remove the axillary nodes when the results of the postoperative histology become available. Identification of true-positive sentinel nodes during surgery avoids second surgery, with the associated inconvenience, pain, discomfort, and small risk from anesthesia. The harms are twofold: increased anesthesia duration during the first surgery (unless another necessary procedure, such as lumpectomy, is performed during the time needed to test the sentinel nodes) and the potential for false-positive results and, therefore, unnecessary ALNDs, with the accompanying sequelae. As the two right-hand columns of Table 2 show, given the sensitivity and specificity of GeneSearch and the prevalence of positive sentinel nodes reported from the test data, using GeneSearch would enable 255 patients per 1,000 patients to have the ALND performed during the SLNB, and thereby avoid a second surgery; while 41 patients per 1,000 would have apparently unnecessary ALNDs. Roughly 6 second surgeries would be avoided per unnecessary ALND. As the prevalence of metastatic cancer in a population rises, the number of second surgeries avoided per "unnecessary" ALND also increases, up to about 15 for a group of patients where half have metastases in the sentinel node(s). Although 50% prevalence is about double the rate typically found in large groups of patients undergoing SLNB, one site in the study did have a prevalence of 45%; the variation in prevalence across sites was large. Conversely, as the prevalence of metastatic cancer among SLNB patients declines, the number of second surgeries avoided per unnecessary ALND also decreases, to about 1.7 in a population with 10% prevalence.

As noted, it is possible that many of the apparently false-positive GeneSearch results are actually true-positive results in which the reference test produced a negative result because the samples used for each technique were different. An abstract was presented at the annual meeting of the American Society of Clinical Oncology in 2007 that attempts to address this issue (Kurosami et al. 2007), and presentation slides are available online. The authors examined 129 lymph nodes from 80 patients, using histological analysis of 10 µm sections at 0.2 mm intervals and GeneSearch assay on all of the tissue between these sections. The sensitivity of the assay was 100% and the specificity was 93.7%. These results are interesting, but further details on the study are needed. For example, the patients whose lymph nodes were used are not described. Also, the reference standard used was not postoperative histology using fixed section but rather frozen section histology.

Four other abstracts on the GeneSearch assay were presented at this meeting: two are by the authors of the studies submitted to the FDA (training and test sets), described in detail above (Blumencranz et al. 2007b; Whitworth et al. 2007); one is a validation study on the GeneSearch assay in Britain to assess its usefulness in the British healthcare system (n=48; Mansel et al. 2007)); one is on the use of the GeneSearch assay in patient management in Europe (n=50; Schobens et al. 2007). Insufficient information is available on these studies to assess their contributions, and the sample sizes are small. Abstracts on GeneSearch have also been presented at other meetings, including the 2006 San Antonio Breast Cancer Symposium. Two paralleled the abstracts by Blumencranz et al. (2007b) and Kurosumi et al. (2007); one looked at the presence of additional gene markers in nodes tested using the GeneSearch assay (Cao et al. 2006).

Comparative Effectiveness. The study did not include a formal test of GeneSearch versus alternative intraoperative methods, nor was such an analysis planned in advance. However, when sites performed these other tests as part of usual clinical practice, the results were collected and compared to the GeneSearch and postoperative histology results. The alternative tests were interpreted only by site pathologists

and were not subject to the same rigorous cross-checking by central laboratory pathologists as the postoperative histology. The results presented in Table 4 show the sensitivity and specificity of the alternative tests . GeneSearch may be more sensitive and less specific than frozen section histology.

The statistical tests of differences in test characteristics between GeneSearch and frozen section are provided in the FDA presentation by Gene Pennello, PhD (FDA 2006c). The FDA concluded that the difference in sensitivity between GeneSearch and frozen section histology is statistically significant (difference=10; 95% CI: 2.5–17.7); as is the difference in the negative predictive value (difference=3.7; 95% CI: 0.8–6.4). The FDA characterized the difference in specificity as borderline significant (difference=-3.5; 95% CI: -7.4–0.0) and the difference in the positive predictive value as insignificant (difference=-7.0; 95% CI: -14.8–1.3). No details are provided on the methods used, including whether samples were paired. The sample size from imprint cytology is too small to determine whether or not any differences between assays are statistically significant.

Dr. Pennello also presented the results of a receiver operator characteristic (ROC) analysis. He calculated ROC curves for the two genotype expression markers used in the GeneSearch assay separately. The area under the ROC curve is 0.94 (95% CI: 0.92–0.97) for cytokeratin (CK) 19 and 0.88 (95% CI: 0.84–0.92) for mammaglobulin. These are very high values, especially for CK 19, since a value of 1 indicates a perfect test. A third analysis was performed using the ROC curve for CK19 for the 319 patients with results by both frozen section histology and GeneSearch. The area under the ROC curve was 0.974. It appears that the operating points (i.e., sensitivity and [1-specificity]) for the frozen section histology and GeneSearch assay lie on the same or very similar ROC curves. If this is true, the GeneSearch assay does not represent a more accurate test but rather a different choice regarding the tradeoff between sensitivity and (1-specificity). It is possible, for example, that GeneSearch could achieve the same sensitivity and specificity as frozen section histology if the cutpoints used in the assay were altered. Similarly, it might be possible to increase the sensitivity and reduce the specificity of frozen section histology so that it performed the same as the assay by training pathologists to call more questionable lesions as positive.

Because the performance of these tests can vary with the type of cancer and the size of the metastases, a direct comparison is the best way to determine their relative accuracy. Therefore, indirect comparisons, using historical data for frozen section histology and imprint cytology, were not evaluated.

Discussion and Conclusions

Question 1: What are the outcomes (i.e., surgeries avoided vs. unnecessary ALNDs) of using the GeneSearch BLN assay intraoperatively compared to waiting for postoperative histology results before deciding on whom to perform ALNDs?

Does the evidence permit conclusions? The GeneSearch BLN assay has potentially wide use, since a majority of the 180,000 (http://www.cancer.gov/cancertopics/types/breast) new breast cancer cases each year undergo sentinel lymph node biopsy. A solid base of evidence is especially important in this context, because the use of this assay could lead to a major change in practice. There are a number of strengths to the study presented to the FDA, including the clear distinction between the training and test sets and double-reading of the permanent histology slides to increase the reliability of the reference standard. However, there are a number of weaknesses as well:

1. Patient recruitment is not described. For example, were these consecutive SLNB patients? Without this information, it is not possible to assess the potential for spectrum bias in the study results. The wide range of prevalence of positive nodes across sites (14.3% to 43.5%) adds weight to the importance of understanding how the subjects were selected.
2. There was substantial variation in the performance of the assay across sites. Sensitivity varied from 50% to 100%, and specificity ranged from 83.3% to 100%. Therefore, the Immunological Devices Panel questioned whether or not pooling into summary statistics was valid.
3. There are no corroborating studies. The FDA has requested two post-approval studies of

Table 4. Comparison of GeneSearch and Alternative Intraoperative Tests
(Reference standard=postoperative histology)

Test	n	Sensitivity (95% CI)	Diff (95% CI)	Specificity (95% CI)	Diff (95% CI)	PPV	Diff (95% CI)	NPV	Diff (95% CI)
Frozen section	319	85.6% (76.6–92.1)	10.0 (2.5, 17.7)	97.8% (95.0–99.3)	-3.5 (-7.4, 0)	93.9%	-7.0 (-14.8, -1.3)	94.5%	3.7 (0.8, 6.4)
Gene Search	319	95.6% (89.0–98.8)		94.3% (90.5–96.9)		86.9%		98.2%	
Imprint cytology	29	45.5% (16.7–76.6)	NR	100% (81.5–100)	NR	100%	NR	75.0%	NR
Gene Search	29	63.6%		100%		100%		81.8%	

NR=Not reported

the assay, which will substantially increase the evidence base. Other research is also underway, based on presentations at ASCO and other specialty society meetings.

4. GeneSearch does not distinguish between micro- and macrometastases. Although ASCO guidelines currently recommend that ALND be performed if either is detected, they also state that the prognostic significance of micrometastases is unknown, as is the importance of ALND in these patients. If GeneSearch has higher sensitivity because it detects more micrometastases (which is not clear), then the clinical significance of finding these patients may not balance the harms of ALND.
5. There is a learning curve for those conducting this assay. How best to train new staff to perform this test is unknown, as is the impact on patient care if and when the use of this assay is disseminated. The initial failure rate (i.e., invalid test result) was about 15%; it decreased to about 4-8% after personnel gained experience. The failure rate outside a research study is not known. The failure rate was higher with small nodes than with larger ones. GeneSearch proponents suggest that small nodes are less likely to contain metastases. The assay could have been rerun with the leftover homogenate or mRNA for the invalid tests, but this was not routinely done. Doing so would provide a better idea of the impact of these invalid results.

What Is the Net Benefit of the Assay? The key factor in assessing this assay is the trade-off between avoiding a second surgery to remove axillary lymph nodes if the sentinel node is positive versus risking unnecessary ALNDs if the assay produces a false positive result. The reason SLNB is the standard of care in early breast cancer, despite the fact that trials on its impact are still incomplete, is the desire to avoid ALND where possible, given the potential for significant, long-term morbidity. While the researchers conducting the GeneSearch study suggest that many of the false-positive assay results detect real metastases in the additional sections of the node used for the assay, the evidence presented to support this argument is insufficient. Given that GeneSearch was not used for patient treatment decisions in this study, longer term follow-up of "false-positive" patients to see if and when axillary metastases develop might be informative, although the utility of this approach may be compromised by the small number of patients involved.

As for the trade-off between avoiding a second surgery and undergoing an unnecessary ALND, patient preferences should play an important role in this decision. The FDA is requiring informed consent for participants in one of the postmarketing studies. Although the number of unnecessary ALNDs is substantially smaller than the number of second surgeries avoided, the "harms" from the former are more common and generally longer lasting than the harms associated with needing a second surgery. The sequelae of ALND can last for years or even a lifetime. The inconvenience of a second surgery, while difficult for patients coping with a serious diagnosis and substantial course of treatment, is likely to be short-lived. For these women, the harms of ALND are unavoidable.

Question 2: How does the sensitivity and specificity of the GeneSearch BLN assay performed intraoperatively during sentinel lymph node biopsy compare to the sensitivity and specificity of alternative intraoperative tests, namely imprint cytology and frozen section histology?

Does the evidence permit conclusions? The test set data provided to the FDA is useful in that it provides a direct comparison of the GeneSearch assay versus alternative intraoperative tests on the same patients. Several concerns, on the other hand, are outlined below:

1. The comparison of GeneSearch to alternative, intraoperative tests was not planned; not all study sites performed such tests.
2. The sample size of 29 for the imprint cytology comparison is too small to be meaningful.
3. The alternative tests were not conducted with the rigorous double-reading, review by central laboratory pathologists, and reconciliation of discrepancies applied to the reference standard.
4. The statistical approach to evaluating the difference in the test characteristics between GeneSearch and frozen section histology is not well described.

What Is the Net Benefit of the Assay? The focus here is on which technique—the assay, frozen section histology, or imprint cytology—performs best at avoiding second surgeries versus undergoing unnecessary ALNDs. The data on imprint cytology is inadequate and will not be discussed further. There is also inadequate evidence on whether the GeneSearch assay outperforms frozen section histology.

However, an ROC analysis conducted by the FDA suggests that the GeneSearch assay and frozen section histology operate at different points on the same or very similar ROC curves (FDA 2006c). If this is correct (the analysis is not described in sufficient detail to assess this), then the question is not whether one technique is better than the other overall, but rather what the optimal tradeoff (and therefore assay cutoff) is between false-positive and false-negative results. Given the longer term and potentially more serious sequelae from ALND than from a second surgery, increasing the sensitivity of the test while sacrificing some specificity, as the GeneSearch assay does, may not be optimal.

The GeneSearch assay also provides less information for staging than other intraoperative procedures, since it cannot distinguish between micro- and macrometases. Nor can it indicate the location of the metastasis (inside or outside of the node). Postoperative histology is therefore required in all cases. It is less crucial when frozen section histology is performed, since pathologists can judge the size of the metastasis and its location from this test, although distortion is possible.

To summarize, the data available is inadequate to assess the clinical utility of the GeneSearch assay compared to either postoperative histology alone or to alternative intraoperative tests such as imprint cytology and frozen section histology. In addition, the balance of benefits versus harms may require higher specificity to avoid unnecessary ALNDs and their sequelae, whereas the GeneSearch design emphasizes sensitivity. Patient preferences should also play a key role in this calculation.

Based on the available evidence, the Blue Cross and Blue Shield Association Medical Advisory Panel made the following judgments about whether the use of GeneSearch breast lymph node assay to detect sentinel node metastases in early stage breast cancer meets the Blue Cross and Blue Shield Association Technology Evaluation Center (TEC) criteria.

1. **The technology must have final approval from the appropriate governmental regulatory bodies.**

The GeneSearch™ BLN Test Kit was cleared for marketing by the FDA on July 16, 2007.

2. **The scientific evidence must permit conclusions concerning the effect of the technology on health outcomes.**

There is only one published study on the performance of the GeneSearch assay, which reports on the same study used by the FDA to determine approval. This study has a number of limitations, including the lack of a description of patient recruitment, inadequate descriptions of several analyses performed, substantial variations in test performance across sites, and an ad hoc comparison of the assay to other intraoperative techniques. Additional validation studies are needed that address these issues. The FDA has mandated two postmarketing studies that will also help fill in the gaps. It would also be helpful to have information on patient preferences between avoiding second surgeries versus running the risk of unnecessary axillary lymph node dissection. Therefore, the data available on the GeneSearch assay are insufficient to permit conclusions regarding the effect of this technology on health outcomes.

3. **The technology must improve the net health outcome.**

The available data are insufficient to determine whether or not the technology improves the net health outcome.

4. **The technology must be as beneficial as any established alternatives.**

The available data are insufficient to determine whether or not the technology is as beneficial as the established alternatives.

5. **The improvement must be attainable outside the investigational settings.**

The available data are insufficient to determine whether or not the technology improves the net health outcome and therefore, whether any improvement is attainable outside investigational settings. There appears to have been a learning curve for using this technology, which was addressed in the study through additional training.

Based on the above, the use of GeneSearch breast lymph node assay to detect sentinel node metastases in early stage breast cancer does not meet the TEC criteria.

References

Al-Shibli KI, Mohammed HA, Mikalsen KS. (2005). Sentinel lymph nodes and breast carcinoma: analysis of 70 cases by frozen section. *Ann Saudi Med*, 25(2):111-4.

Arora N, Martins D, Huston TL et al. (2008). Sentinel node positivity rates with and without frozen section for breast cancer. *Ann Surg Oncol*, 15(1):256-61.

Blumencranz P, Whitworth PW, Deck K et al. (2007a). Sentinel node staging for breast cancer: Intraoperative molecular pathology overcomes conventional histologic sampling errors. *Am J Surg*, 194:426-32.

Blumencranz P, Deck KB, Whitworth PW et al. (2007b). An investigational rapid RT-PCR assay for the detection of metastases in sentinel lymph nodes showed improved performance over frozen section H&E: Analysis by primary tumor characteristics. [Abstract] *J Clin Oncol*, 25 (18S):10561.

Cao W, Backus J, Vargo JM et al. (2006). Additional markers support results in a molecular breast sentinel lymph node diagnostic assay. [Abstract] Poster presented at the 2006 Annual San Antonio Breast Cancer Symposium.

Celebioglu F, Sylvan M, Perbeck L et al. (2006). Intraoperative sentinel lymph node examination by frozen section, immunohistochemistry and imprint cytology during breast surgery--a prospective study. *Eur J Cancer*, 42(5):617-20.

Cody HS. (1999). Sentinel lymph node mapping in breast cancer. *Oncology*, 13(1):25-34.

Cserni G. (2007). Pathological evaluation of sentinel lymph nodes. *Surg Oncol Clin N Am*, 16(1):17-34.

Harlow SP, Krag DN, Julian TB et al. (2005). Prerandomization surgical training for the National Adjuvant Breast and Bowel Project (NSABP) B-32 trial: A randomized phase III clinical trial to compare sentinel node resection to conventional axillary dissection in clinically node-negative breast cancer. *Ann Surg*, 241(1):48-54.

Johnston EI, Beach RA, Waldrop SM et al. (2006). Rapid intraoperative immunohistochemical evaluation of sentinel lymph nodes for metastatic breast carcinoma. *Appl Immunohistochem Mol Morphol*, 14(1):57-62.

Kurosumi M, Kobayashi Y, Takei H et al. (2007). The utility of an investigational real-time RT-PCR assay system for the detection of micrometastases in sentinel lymph nodes of breast cancer confirmed by 0.2 mm histological interval analysis. [Abstract] J Clin Oncol, 25 (18S):10560. Slides available at http://www.asco.org/portal/site/ASCO/menuitem.34d60f5624ba07fd506fe310ee37a01d/?vgnextoid=76f8201eb61a7010VgnVCM100000ed730ad1RCRD&vmview=abst_detail_view&confID=47&index=y&abstractID=33562.

Leung KM, Chan KW, Yeoh GP et al. (2007). Clinical relevance of intra-operative sentinel lymph node examination in breast cancer management. *Hong Kong Med J*, 13(1):8-11.

Luini A, Gatti G, Ballardini et al. (2005). Development of axillary surgery in breast cancer. *Ann Oncol*, 16:259-62.

Lyman GH, Guiliano AE, Somerfield MR et al. (2005). American Society of Clinical Oncology Guideline recommendations for sentinel lymph node biopsy in early stage breast cancer. *J Clin Oncol*, 23(30):7703-20.

Mansel RE, Goyal A, Douglas-Jones A et al. (2007). Intraoperative assessment of sentinel lymph nodes in breast cancer patients using real time RT-PCR: Results of the Cardiff validation study. Poster presented at the 2007 annual meeting of the American Society of Clinical Oncology.

McCready DR, Yong WS, Ng AK et al. (2004). Influence of the new AJCC breast cancer staging system on sentinel lymph node positivity and false-negative rates. *J Natl Cancer Inst*, 96:873-5.

Ozmen V, Cabioglu N. (2006). Sentinel lymph node biopsy for breast cancer: Current controversies. *Breast J*, 12(Suppl 2):S134-42.

Orr RK. (1999). The impact of prophylactic axillary node dissection on breast cancer survival: A Bayesian meta-analysis. *Ann Surg Oncol*, 6:109-16.

Purushotham AD, Upponi S, Klevesath MB et al. (2005). Morbidity after sentinel lymph node biopsy in primary breast cancer: Results from a randomized controlled trial. *J Clin Oncol*, 23(19):4312-21.

Rockette HE, Redmond CK, Fisher B. (1982). Impact of randomized clinical trials on therapy of primary breast cancer: the NSABP overview. *Control Clin Trials*, 3(3):209-25.

Sakorafas GH, Peros G, Cataliotti L et al. (2006). Lymphedema following axillary lymph node dissection for breast cancer. *Surg Oncol*, 15:153-65.

Schobens JC, Martinez DM, Veys I et al. (2007). Clinical experience with a new intraoperative molecular assay for sentinel lymph node metastases in breast cancer patients. Poster presented at the 2007 annual meeting of the American Society of Clinical Oncology.

Schrenk P, Konstantiniuk P, Wolfl S et al. (2005). Intraoperative frozen section examination of the sentinel lymph node in breast cancer. *Rozhl Chir*, 84(5):217-22.

Tanis PJ, Boom RP, Koops HS et al. (2001). Frozen section investigation of the sentinel node in malignant melanoma and breast cancer. *Ann Surg Oncol*, 8(3):222-6.

Tew K, Irwig L, Matthews A et al. (2005). Meta-analysis of sentinel node imprint cytology in breast cancer. *Br J Surg*, 92(9):1068-80.

U.S. Food and Drug Administration. (2006a). Transcript of the Immunology Devices Panel Meeting, Food and Drug Administration, Department of Health and Human Services, 16 November, 2006. Available online at: http://www.fda.gov/ohrms/dockets/ac/06/transcripts/2006-4249t1.pdf. Last accessed September 2007.

U.S. Food and Drug Administration. (2006b). Veridex GeneSearch breast lymph node assay: FDA panel questions. Powerpoint slides prepared by FDA scientific review team (Reeves JP, Pennello G, Robinowitz M, Horbowyi R). Available online at: http://www.fda.gov/ohrms/dockets/ac/06/slides/2006-4249s1-01-JPR-Introduction.pdf. Last accessed September 2007.

U.S. Food and Drug Administration. (2006c). Veridex BLN assay: Statistical design and analysis of pivotal study. Slides by Gene Pennello, PhD. Available online at: http://www.fda.gov/ohrms/dockets/ac/06/slides/2006-4249s1-04-pennello.pdf. Last accessed October 2007.

U.S. Food and Drug Administration. (2006d). Sentinel lymph node dissection in breast cancer. Slides by Roxolana Horbowyj, MD, FACS. Available online at: http://www.fda.gov/ohrms/dockets/ac/06/slides/2006-4249s1-02-horbowyj.pdf. Last accessed October 2007.

U.S. Food and Drug Administration. (2007a). Summary of safety and effectiveness for GeneSearch™ BLN assay. Available online at: http://www.fda.gov/ohrms/dockets/ac/06/briefing/2006-4249b1_01.pdf. Last accessed September 2007.

U.S. Food and Drug Administration. (2007b). Approval order for GeneSearch™ BLN assay, 16 July 2007. Available online at: http://www.fda.gov/cdrh/pdf6/p060017a.pdf. Last accessed October 2007.

Whitworth P, Deck KB, Rosenberg AL et al. (2007). Intraoperative assessment of sentinel lymph nodes using a rapid, reproducible RT-PCR based assay to detect clinically significant metastases. Poster presented at the 2007 annual meeting of the American Society of Clinical Oncology.

Technology
Evaluation
Center

Blue Cross and
Blue Shield Association
225 North Michigan Avenue
Chicago, Illinois 60601-7680
www.bcbs.com/tec