

NBCCEDP BREAST CANCER EXPERT PANEL
WHITE PAPER ON TECHNOLOGIES
FOR THE EARLY DETECTION OF BREAST CANCER

INTRODUCTION

The National Breast and Cervical Cancer Early Detection Program (NBCCEDP), administered by the Centers for Disease Control and Prevention (CDC) helps low-income, uninsured, and underserved women gain access to lifesaving screening services for the early detection of breast and cervical cancers. The NBCCEDP is implemented in all 50 states, 4 U.S. territories, the District of Columbia, and 13 American Indian/Alaska Native organizations. Through these grantees, the program implements a wide range of activities, including a) public education to raise awareness of the benefits of screening and the availability of subsidized screening services; b) outreach to recruit high-risk women; c) provision of breast and cervical cancer screening exams and diagnostic testing; d) case management to facilitate access to care and assure completion of recommended follow-up testing; and e) professional education and quality assurance to ensure the highest standard of care for women in the program. Although the program has screened 1.9 million women and provided 4.6 million screening examinations since it was established in 1991, it reaches fewer than 20 percent of eligible women annually, primarily due to limited Congressional appropriations.

Fiscal management of the multifaceted NBCCEDP poses many challenges; one in particular is the determination of which screening tests should be paid by the program. Appropriate stewardship of federal funds requires that decisions be evidence-based, yet there are market factors that influence the daily realities of the program. Since the program's inception, research and scientific advances have resulted in both changing recommendations regarding the timing and subjects of screening, but also the introduction of new technologies. Determinations about whether the NBCCEDP should pay for newer screening tests and procedures are complicated. The program must balance a wide range of factors, including, for example, standards of care for women in the program, the public health mandate to serve as many women as possible, limited program funds, varying local health services infrastructures, and the impact of changes in program policies on program operating procedures and partners.

With regard to breast imaging, currently the NBCCEDP provides reimbursement for film mammography only. Digital mammography, magnetic resonance imaging (MRI), and ultrasound are not reimbursed as screening tests. Computer aided detection (CAD) of digital mammograms or of digitized films also is not reimbursed. These reimbursement policies are consistent with the U.S. Preventive Services Task Force (USPSTF) 2002 recommendations.¹ The USPSTF report reviewed studies of film mammography and clinical breast examination (CBE) screening, but did not explicitly address digital mammography, CAD, or ultrasound.

Recognizing the complexity of the task of reviewing NBCCEDP reimbursement policies and their considerable impact on individual BCCEDP programs, CDC initially sought to gather information about programs' experiences with current reimbursement policies. Key informant

interviews with NBCCEDP Program Directors representing eight state programs and two CDC program staff were conducted to identify the range of issues that should be considered in CDC's reconsideration of reimbursement policies. The report of these interview findings is presented in Appendix A. Additionally, CDC identified key scientific references to provide general background about current and newer technologies. Evidence overviews and discussions with experts revealed a lack of scientific evidence in many relevant areas, particularly direct comparisons of test performance characteristics, such as sensitivity and specificity, and in utilization patterns among the technologies. Also evident from these sources was the lack of a clear and consistent definition of 'high risk' for breast cancer. One reason for this inconsistency is that definitions of risk used in studies and public health shift as new scientific evidence emerges. Most studies assessing new screening technologies for use among women at high risk define high risk as either those with BRCA 1/2 or a family history of breast cancer. In the context of this paper, discussions of the use of new technologies directed to women at high risk relies on the various definitions used in current studies. The panel does recommend further work, however, to more clearly define concepts of risk within the NBCCEDP.

Because of the complexity of program issues and the paucity of directly relevant scientific evidence, the CDC sought to implement a review process relying primarily on expert opinion to guide its decision-making. An expert panel was established including researchers, clinicians, public health practitioners and NBCCEDP Program Directors. A list of expert panel members is included in Appendix B. This expert panel was charged with a) identifying minimum criteria for establishing new reimbursement policies, b) identifying a framework of issues to be considered in policy review, c) providing specific recommendations for reimbursement policies, and d) providing guidance concerning procedures for future reviews of reimbursement policies.

Members of the expert panel on breast cancer reimbursement policies conferred in subgroups and as a full committee through a series of conference calls and a face-to-face meeting held in Atlanta on March 29 and 30, 2005. This report provides the background for and final recommendations of this expert panel. The first two sections of this paper provide general information about the epidemiology of breast cancer and the women served by the NBCCEDP. The next two sections provide context for assessing individual technologies by defining the minimum criteria that must be met in order to recommend reimbursement and the specific test characteristics and public health factors that must be assessed in making reimbursement policy decisions. The final two sections specifically review the test characteristics and public health factors for each technology under consideration and present the expert panel's recommendations for reimbursement policies as well as recommendations for additional research and surveillance to provide a firmer foundation for future assessments of program policies.

BREAST CANCER

Breast cancer is the most frequently diagnosed cancer among women in the United States and the second leading cause of cancer death. An estimated 211,240 women will learn they have breast cancer and an additional 40,410 will die from breast cancer in the United States in 2005. A woman's chances of being diagnosed with breast cancer sometime during her life are about 1 in 7 and her chances of dying from breast cancer are about 1 in 33. Currently, just over 2 million

women in the United States have been diagnosed with and treated for breast cancer. Although the disease is most prevalent among women, 1,690 men also will be diagnosed with breast cancer and 460 men will die from the disease in 2005.

In addition to the new cases of invasive breast cancer that will be identified in 2005, 58,490 new cases of in situ breast cancer will be identified as well. Almost 85 percent of these will be ductal carcinoma in situ (DCIS).² In situ cancers are an early stage of cancer, when the disease is still confined to its site of origin. Increases in the detection of these early stage cancers are a direct result of screening with mammography. DCIS is estimated to account for as much as 20 percent of all cancers diagnosed by mammography, about 1 in every 1,300 mammograms.

Mammography has been shown to be better at detecting DCIS than invasive cancers, in one study finding 86 percent of DCIS cases and 75 percent of invasive breast cancers.³

While the use of mammography to find early stage breast cancers before physical signs of disease are evident is attributed with overall reductions in breast cancer mortality over the past decade, mammography does have limitations.² Mammography is estimated to miss as many as 25 percent of cancers and about 10 percent of findings require additional testing in women who later are found not to have breast cancer.^{4,5} However, routine mammography screening among asymptomatic, age-appropriate women to detect early stage breast cancers remains the best public health defense against breast cancer. Despite the identification of several risk factors for breast cancer, such as increasing age, family history of breast cancer, high breast tissue density, and inherited genetic mutations, more than 50 percent of breast cancers occur in women with no known risk factors.²

SCREENING FOR BREAST CANCER IN THE NBCCEDP

The NBCCEDP serves low-income, uninsured women. When the NBCCEDP began in 1991, CDC followed recommendations for breast cancer screening that emphasized the value of screening mammography both for women 40–49 years of age and for women 50 years of age and older. CDC funded programs were permitted to establish their own age guidelines within these parameters. In 1994, however, the NBCCEDP established a more stringent policy for funding breast cancer screening that was consistent with the best use of very limited resources. The new NBCCEDP policy required that at least 75 percent of mammograms paid with NBCCEDP funds be provided to women 50 years of age or older. In addition, in 1998, when Medicare began to pay for screening mammography, NBCCEDP policy changed to exclude women 65 years of age with Medicare Part B coverage. Over time, these changes have produced an age shift in women screened in the program. Although about 48 percent of mammograms were provided to women ages 50-64 in the first 5 years of NBCCEDP screening, this proportion has increased to 72 percent in the most recent 5 years (2000-2004).

Looking at aggregate data from 1991-2002, approximately 50 percent of the women screened in the program are white. Increasing focus on recruiting foreign-born women and those least likely to be previously screened, however, lowered this proportion to 43 percent from 2001-02, with corresponding increases among minority women, particularly Hispanic women and Asian/Native Hawaiian, and Pacific Islanders.⁶

A study of re-screening in four NBCCEDP programs found that 72 percent of women in these programs were re-screened within 18 months and 82 percent within 30 months, which is similar to the proportion of women in the general population that have been re-screened. Hispanics, women with a history of breast cancer before their initial program mammogram, and women who had used hormone replacement therapy before their initial program mammogram, were more likely to have been re-screened at 30 months.⁷

Approximately 11 percent of first round screening mammograms performed by the program between 1991 and 2002 were abnormal. This proportion decreased to about 7 percent for second round mammograms. The percentage of women reporting symptoms also was greater in the first screening round than in subsequent rounds (11 vs. 7 percent, respectively). The proportion of abnormal screening mammograms decreased with increasing age during this same time period (12 percent in women 40 to 49 years of age vs. 7 percent in women 65 years of age and older).⁶

Between 1991 and 2002, 9,956 women were diagnosed with invasive breast cancer through the NBCCEDP. Seventy four percent of these cancers were identified at an early stage (stage I or II). Overall and adjusted for age, about 9.4 cases of in situ or invasive breast cancer are diagnosed per 1,000 mammograms in the NBCCEDP. This rate is higher in white women, but lower in all other racial and ethnic groups. Regardless of age, race, or ethnicity, the detection rates for carcinoma in situ and invasive cancer were substantially lower in subsequent screening rounds compared to the initial program screening.⁶

REIMBURSEMENT DECISION CRITERIA

Review of NBCCEDP reimbursement for new screening technologies must consider the overall advantages and disadvantages of the new technology relative to the mission of the NBCCEDP and current screening approaches. Because screening is performed on healthy, asymptomatic women, each new technology must clearly demonstrate its ability to perform equally to or better than current technologies. Overall the technology must meet certain minimum criteria. These include:

- Reduce Breast Cancer Morbidity and Mortality – The technology must contribute to reductions in morbidity and mortality across the population of program eligible women. For breast cancer screening, reductions in morbidity and mortality come from identifying and treating early stage cancers including in situ carcinomas.
- Sustain or Enhance Overall Public Health Benefit – Use of the technology should sustain or enhance the number of program eligible women served by the NBCCEDP, for example by maintaining or increasing access to services or maintaining or increasing dollars available to pay for services.
- Sustain or Enhance Overall Quality of Care – Use of the technology should sustain or enhance the quality of services provided by the NBCCEDP, for example by maintaining or enhancing effectiveness, reducing false positive findings, or improving test acceptability and patient adherence.
- Sustain or Enhance Overall Program Operations – Use of the technology should sustain or enhance program operations across NBCCEDP sites, for example by

streamlining administrative procedures, maintaining or increasing provider enrollment, or enhancing clinical efficiency.

- Reduce Overall Health Disparities – Use of the technology should further NBCCEDP goals to reduce disparities in the delivery of services to and health outcomes of low-income, uninsured, and underserved women.

Beyond these minimum criteria for establishing reimbursement policies, consideration must be given to two additional factors. First, policies must accommodate differences across programs. NBCCEDP programs differ considerably in public health infrastructures as well as local health care capacities and systems. Reimbursement policies must be consistent across programs while still affording flexibility in how NBCCEDP programs implement these policies across local communities.

Second, as a federal government agency, the CDC must consider related policies established by other federal agencies, in particular the Food and Drug Administration (FDA) and the Centers for Medicare and Medicaid Services (CMS). Each federal agency establishes policies consistent with its unique mission. Unlike the CDC, FDA and CMS are regulatory agencies. The FDA provides market approval for new drugs and devices and CMS provides payment approval and establishes reimbursement rates for the delivery of medical services under mandated federal entitlement programs. The NBCCEDP relies on the rate structure established by CMS for reimbursement of early detection and diagnostic services in Medicare and it is statutorily mandated that NBCCEDP reimbursement not exceed these Medicare rates.

Reflective of the different missions of these agencies, the procedures each uses to establish policies differ. FDA seeks to establish whether a medical drug or device is safe and as effective as existing drugs or devices. FDA relies in part on input from industry and industry-sponsored studies in making this determination.⁸ CMS seeks to identify medical procedures for reimbursement under Medicare and Medicaid. Its determinations are based on whether a procedure, device, or technology is “reasonable and necessary” for the diagnosis and treatment of a medical condition.⁹ Like the FDA, CMS also invites industry collaboration and comment during their approval process. Importantly, however, neither CMS nor FDA approval of a new procedure, drug, device, or technology indicates that it is *more* effective than existing procedures, drugs, devices or technologies.

Some components of these approval procedures overlap across federal agencies. For example, CMS requires that drugs or devices be approved as safe and effective by the FDA before it will provide approval for reimbursement under Medicare or Medicaid. But it is also true that some components remain independent. For example, CMS provides approval for some procedures, such as counseling about preventive service, that do not fall within the authority of FDA’s mandate to establish safety and efficacy because it is not a drug or device.

Establishment of reimbursement policies under CDC’s NBCCEDP must first reflect the unique mission of the program, maximizing reductions in breast cancer morbidity and mortality in the eligible population of low-income, uninsured women. Procedures for establishing these policies rely primarily on scientific evidence, expert opinion, and program considerations. In this context it is not surprising that CDC policies in some cases will overlap with those of the FDA and CMS,

while in others they may not. For example, while CDC might require that all reimbursed technologies be approved by FDA as safe and effective for the same use, there may be program services for which FDA has no authority (e.g., preventive services counseling). Similarly, there may be circumstances where CMS has approved a technology or procedure and established associated reimbursement rates, but the benefits of the technology for the NBCCEDP are outweighed by disadvantages such as high costs, lack of clinical availability, or program inefficiencies.

For these reasons, absolute requirements for FDA and/or CMS approval for all NBCCEDP reimbursed technologies were considered overly restrictive. Further, any requirement that the NBCCEDP reimburse for all FDA and/or CMS approved technologies was considered inappropriate as this might result in limiting the program's ability to achieve its mission to extend services to as many eligible women as possible in order to maximize reductions in breast cancer morbidity and mortality. Thus, it is recommended that:

- for all technologies and procedures within FDA authority, the technology should be approved by the FDA for the use under consideration, and
- for all technologies and procedures within CMS authority, the technology should be approved by CMS and have established Medicare rates, but not all CMS approved technologies need to be reimbursed by the NBCCEDP.

BASIS FOR TECHNOLOGIES ASSESSMENT

The basis for decisions about whether the NBCCEDP should provide reimbursement for any new technology combines the full range of test characteristics as well as program factors. This section presents an overview of the components of this assessment. These issues combine uniquely for each technology. For example, some new technologies bring more favorable test characteristics, but at a test or program cost that on balance does not support the overall public health goals of the NBCCEDP. Other new technologies might bring only comparable test performance characteristics, but add program efficiencies or reduce test costs that potentially allow more women to be screened by the program.

Test Characteristics

Test characteristics include a combination of five performance and cost characteristics that will be unique for each technology. Comparison of technologies across these characteristics provides the basis for assessing test-specific advantages and disadvantages. These characteristics include:

Accuracy – test accuracy in identifying early stage breast cancers is reflected in several measures, including sensitivity, specificity, positive predictive value, negative predictive value and level of test uncertainty. Sensitivity and specificity are related measures. Sensitivity refers to the proportion of all true cancers detected by a test within a specified timeframe, usually one year. Specificity refers to the proportion of true negative results (e.g., no cancer present) for which a negative test result is obtained within a specified timeframe, usually one year. High sensitivity increases the probability that cancers will not be missed while high specificity reduces the probability that women will undergo unnecessary follow-up procedures, such as repeat mammograms, adjunctive imaging (ultrasound or MRI), fine needle aspiration, and biopsies.

While the negative consequences of missing cancers are high, the adverse physical and emotional consequences of unnecessary medical procedures also are high. For any single test, specificity generally decreases as sensitivity increases.

From a public health perspective, the trade-offs between different levels of test sensitivity and specificity is substantial. For example, in a population of 100,000 women for which a true prevalence of cancer is 5 percent, 95,000 women would be normal (95 percent) and 5,000 would have cancer. A test having a sensitivity of 80 percent would find 4,000 cancers, but would miss 1,000 cancers. An increase in test sensitivity of 10 percent, to a sensitivity of 90 percent, would result in half as many missed cancers, or 500 fewer missed cancers. More dramatically, however, if test specificity is 90 percent, 10 percent of the 95,000 women without cancer would receive a false-positive result. In this scenario, 9,500 women would incorrectly receive a positive test result. A 5 percent absolute decrease in specificity to 85 percent translates into an additional 4,750 women receiving a false-positive test result. Decreases in test specificity which often accompany improvements in sensitivity can yield substantial increases in follow-up tests such as image guided needle biopsies that do not result in a diagnosis of malignancy and the costs associated with unnecessary follow-up tests. In the example given, an additional detection of 100 cancers came at a cost of additional work up of 4,950 normal women. The critical issue for any test is the extent to which both sensitivity and specificity can be balanced to yield an optimal public health outcome.

Two additional related measures, positive and negative predictive value, also provide valuable information about test performance. These measures assess the diagnostic value of a test. Positive predictive value reflects the proportion of times a positive test finding leads to diagnosis of disease, while negative predictive value reflects the proportion of times a negative test finding is obtained among women who do not have cancer. Similar to the scenarios described above for test sensitivity and specificity, the consequences of low positive predictive value (PPV) are realized in missed cancers and the consequences of low negative predictive value (NPV) are realized in unnecessary follow-up tests and patient anxiety.

One final indicator of test accuracy is the level of uncertainty about test results. Uncertainty can result for example, from ambiguity in a test image or lack of clarity about interpretation of specific image characteristics.

Reproducibility - Test reproducibility refers to the consistency of the image or sample produced by the test as well as the consistency of interpretation of the image or sample. Reproducibility is particularly relevant for an examination in which subsequent images are compared to a baseline image, such as with mammography. Poor reproducibility can result in repeat screening examinations to enhance overall test precision.

Population Characteristics – Some tests perform better among women with certain characteristics, particularly for imaging technologies. For example, image capture or display characteristics might accentuate identification of abnormalities in dense breasts or testing procedures might reduce patient discomfort and potentially increase compliance. Test characteristics that maximized test performance among subpopulations may introduce important

new benefits, but also can introduce challenges and potentially additional costs associated with outreach, communications and monitoring in NBCCEDP programs.

Interval – Screening interval refers to the recommended time to repeat routine screening following a normal test. Frequent screening can lead to increased costs because more tests are performed. But particularly long screening intervals reduce the lead time gained from more frequent screening and can introduce compliance problems, particularly if the interval differs from normal health routines.

Test Cost – All procedures reimbursed by the NBCCEDP are reimbursed at current Medicare rates. As reflected in Table 1 for the technologies being reviewed in this white paper, these rates vary across regions and technologies. These Medicare test reimbursement rates reflect lab and test costs and do not include the professional component. Generally, new technologies cost more initially on a per-test basis than existing technologies, although costs of new technologies tend to fall as adoption rises. The primary issue when comparing costs across technologies is the incremental cost difference between the new compared with the older technology.

CPT Code	Procedure	Low	High	Average	Median
76092	Screening Mammogram, Conventional, Bilateral	\$66.53	\$143.03	\$86.82	\$84.58
76090	Diagnostic Mammogram, Conventional, Unilateral	\$61.62	\$131.01	\$79.53	\$77.51
76091	Diagnostic Mammogram, Conventional, Bilateral	\$76.54	\$162.65	\$98.75	\$96.24
76082	CAD, w/ 76090, 76091, G0206, or G0204	\$14.23	\$32.91	\$19.97	\$19.23
76083	CAD, w/ 76092 or G0202	\$14.23	\$32.91	\$19.97	\$19.23
G0202	Screening Mammogram, Digital, Bilateral	\$101.53	\$225.94	\$137.24	\$132.14
G0204	Diagnostic Mammogram, Digital, Bilateral	\$108.33	\$237.97	\$144.55	\$139.37
G0206	Diagnostic Mammogram, Digital, Unilateral	\$87.53	\$192.40	\$116.87	\$112.68
76093	MRI, Breast, Unilateral	\$556.49	\$1,314.51	\$797.14	\$769.31
76094	MRI, Breast, Bilateral	\$727.71	\$1,732.85	\$1,050.75	\$1,013.17
76645	Breast Ultrasound, Unilateral/Bilateral (single rate)	\$54.12	\$117.08	\$71.07	\$69.16

Public Health Factors

Public health factors include a combination of clinical, patient, and program factors. Similar to test characteristics, each of these factors can have a large influence on the ability of the NBCCEDP programs to realize reductions in breast cancer morbidity and mortality. Unlike test characteristics, however, public health factors vary considerably across NBCCEDP programs. This variability is not systematically monitored and can be difficult to assess. Key informant interviews were conducted with select NBCCEDP sites to expand available information about the range of issues encountered by programs. But information about the prevalence of these issues across programs is not generally available. For this reason, recommendations are also

presented in this white paper for research and surveillance initiatives that might enhance public health information for future policy reviews.

Clinical Factors – Three types of clinical factors are considered in assessing test reimbursement, including practice patterns, clinical efficiency, and patient education requirements. Practice patterns refer to differences in adoption of new technologies across program localities. In localities where providers primarily utilize a technology that is not approved for reimbursement, the program provides reimbursement at the rate of the approved technology. But newer technologies often are more expensive, and the added cost difference must either be absorbed by providers or reimbursed through alternative funds, placing added strain on providers as well as on alternate funding sources. Further, as providers transition to newer technologies they perform older tests at lower frequencies, potentially reducing their proficiency. These situations also can reduce the efficiency of clinical practice. Finally, many new technologies require additional patient education. The clinical time associated with educating patients about the appropriate use of new technologies and interpretation of findings is an additional factor for consideration. The media, industry, and providers can add to this pressure by marketing new technologies directly to women, creating demand for services that are not reimbursable under the program.

Patient Factors – Patient factors relevant to the overall benefit of providing a particular technology through the NBCCEDP include the acceptability of the technology, compliance, the burden of disease and screening history among those appropriately screened by the technology, as well as quality of life impact. Acceptability by patients is influenced by a variety of factors, including the level of discomfort associated with the test as well as perceived disease risk and test benefits. Women's perceptions of their personal risk of getting breast cancer are considerably higher than their actual risk and they overestimate the benefits of breast cancer screening.¹¹ Thus women generally accept some test discomfort to ensure that a diagnosis of cancer is not missed. However, it is also true that poor test acceptability can cause delays in initial or routine screening. Further, controversial tests receiving media attention can stimulate confusion that dissuades women from receiving any test at all.¹¹

Patient characteristics, such as age, risk, and prior screening history, significantly influence the likelihood of finding breast cancer, and as a result change the cost/benefit estimate of screening. These are important considerations when assessing program benefits of reimbursement for technologies whose test performance varies across these patient characteristics. Finally, patient quality of life related to test characteristics is an important consideration. Despite women's willingness to accept additional procedures or discomfort to reduce their personal risk of dying from breast cancer, the consequences of these procedures and associated non-medical patient costs, such as time lost from work or child care expenses, are not trivial.¹²

Program Factors – Program factors play an important role in assessing the overall advantages and disadvantages of providing reimbursement for new technologies. Introduction of new technologies can influence program efficiency, provider enrollment, and women's access to program services. Program efficiencies can be either enhanced or reduced by changes in requirements for provider communication, patient outreach and education, and administrative procedures. When new technologies are accepted for reimbursement by the program, considerable program staff time is required to educate providers about new policies and

procedures, make modifications to reimbursement systems, modify data and reporting systems, and amend contracts with clinical providers. Providers similarly need time to implement new office procedures. But the converse is also true when providers are using technologies that are not reimbursed by the program. Providers need to find alternate funding to cover cost differentials. This takes time and resources, not only to find separate sources of funds, but also to establish systems that account for these separate funding sources.

Providers are essential to the NBCCEDP. Reimbursement policies can, in some rare circumstances, cause providers to drop out of the program altogether. This reduces the number of providers delivering services for the program and thereby reduces program access for women. Reduced provider capacity can both limit the programs' ability to meet demand for early detection services and cause delays in providing needed services. Callbacks introduce another barrier to program access when women must travel back to a facility to be retested.

Key informant interviews also revealed the potential for some reimbursement policies to adversely affect program credibility. Failure to reimburse technologies that have become common can convey an image of the program as 'out of step' with current practices or leave an impression that women in the program receive 'less than optimal' care. Educating patients and providers about the basis for reimbursement policies and the advantages and disadvantages of new technologies is an important program activity, which in these cases may require additional staff time and skill.

TECHNOLOGIES OVERVIEW

This section provides an overview of the relevant test characteristics and public health factors for breast cancer screening technologies currently reimbursed by the NBCCEDP and those being considered for reimbursement as screening tests. These tests include film mammography, digital mammography, computer assisted detection (CAD), magnetic resonance imaging (MRI), and ultrasound. Mammography and CAD are currently approved by the FDA and CMS for breast cancer screening, while MRI and ultrasound are approved only as diagnostic tests.

Film Mammography

Test Characteristics – The overall accuracy of film mammography is high. Film mammography yields significant reductions in breast cancer mortality, ranging from 21-30 percent^{13,14,15,16} and has resulted in an overall shift toward detection of small, low-grade tumors that have a better, long-term prognoses.^{17,18} Film mammography sensitivity varies as a function of breast density, achieving levels as high as 98 percent in fatty breasts and 84 percent in dense breasts.^{16,19} One recent study found sensitivities ranging from 63 percent in dense breasts to 87 percent in fatty breasts.⁴

The reproducibility of film mammography images and interpretation also are generally high. The technology has been used in clinical practice for more than 30 years and reporting and quality assurance systems are well established. The BI-RADS[®] system for film interpretation has undergone four revisions since its inception in 1992.

All major U.S. medical organizations recommend screening film mammography, with or without CBE, for women 40 years of age and older.^{1,2,12,20} The NBCCEDP emphasizes use of screening mammography in women 50 years of age and older by requiring that 75 percent of program mammograms be provided to this group. As reflected in Table 1, film mammography is one of the least expensive breast cancer screening tests currently available.

Public Health Factors – Film mammography is widely available²¹ and systems for quality assurance and uniform reporting are well established. Film mammography is a completely portable system, offering women the ability to take films from one center to another as needed. However, different procedures for reading film images, such as batch interpretation, can influence repeat testing not associated with an abnormal finding. Facilities that rely on batch interpretation without immediate review require that a woman return for a diagnostic mammogram for problems identified on the screening exam. Facilities that use batch interpretation, however, tend to have lower recall rates than facilities that perform online interpretation of mammograms.

Screening with film mammography has considerable market penetration. A recent study found that 60 percent of women had had their first mammogram by the end of their 40th year and almost 90 percent had begun screening by 50 years of age.²² Even among subpopulations having large barriers to routine medical care, high rates of mammography screening are evident. Women without private health insurance began screening at a median age of 46.6 years.²² Women who did not speak English began screening at a median age of 49.3 years.²² And even among women with no private health insurance and who don't speak English mammography screening was initiated at a median age of 55.3 years.²² While these rates of initial screening are encouraging, rates of routine screening are lower and vary considerably by region. Among women 50 years of age and older in the United States, 20 percent reported not having received a mammogram within the past 2 years. This rate varied from 12 to 31 percent across states.²³

Digital Mammography

Test Characteristics – The accuracy of digital mammography appears to be comparable to that of film mammography.²⁰ Three prospective screening trials, two with the women acting as their own control and one randomized trial comparing film mammography to digital mammography, demonstrated no statistically significant difference in sensitivity.²⁴⁻²⁷ One trial demonstrated that digital mammography had a statistically significant lower recall rate than film mammography,^{24, 25} while one showed no difference,²⁶ and the other showed a statistically significant higher recall rate.²⁷ Most differences between screen and digital mammography are thought to be due to technique rather than modality.^{24, 25} Additional data about the relative diagnostic accuracy and cost-effectiveness of digital compared to film mammography are expected within the next year from the Digital Mammographic Imaging Screening Trial (DMIST), a multi-center trial sponsored by the National Cancer Institute and coordinated by the American College of Radiology Imaging Network.

As a newer technology, systems for quality assurance and standardization of digital mammography are less well established than those for film mammography. Unlike film

mammography, the image capture and display components of digital mammography are separated and there is considerable variability for each of these elements across different digital systems. Programming differences in image capture not only affect the characteristics of the image and thus reproducibility across systems, but also the ability to transfer images to other systems as a woman moves through the health care system. Differences in display characteristics, such as pixel size and contrast, also affect the reproducibility of image interpretation.

Recommended screening intervals using digital mammography are the same as those for film mammography. As reflected in Table 1, however, digital mammography costs much more than film mammography, approximately \$55 more per screening examination. These increased costs and the costs associated with potentially higher recall rates could substantially reduce the overall number of screening examinations that could be provided through the NBCCEDP within existing appropriation levels.

Public Health Factors – As of 2003, only 413 full field digital mammography units were accredited under the Mammography Quality Standards Act (MQSA) in the United States.²⁸ It is estimated that 6.8 percent of all mammography facilities use digital mammography, although these are generally high-volume facilities (Personal Communications, Pamela A. Wilcox). While market penetration of digital mammography is generally low at this time, it is anticipated that this may change due to direct marketing of the technology. Digital manufacturers have launched extensive market campaigns to both medical centers and the public. Facilities, having made substantial investments in the technology, also have marketed digital mammography to the public extensively as they seek to recover their capital expenditures. These factors have inflated the perceived value of the technology, at least based on current evidence, and have stimulated public demand. Further increases in the adoption of digital mammography may depend greatly on the results of the DMIST trial.

Because few facilities use digital mammography, few NBCCEDP programs have noted problems with provider enrollment or program access due to the lack of reimbursement for this technology. However, because high volume facilities appear more likely to be using digital mammography, the lack of reimbursement for digital mammography may have a disproportionate impact of failure to reimburse for digital mammography in some areas.

From the perspective of the patient, the acceptability of film and digital mammography are comparable. The tests are virtually indistinguishable at the point of image capture. As a result of this and equivalent screening intervals, compliance appears equal across film and digital mammography.

CAD

Test Characteristics – CAD is not a screening technology but a detection aid and it is unclear whether CAD improves the accuracy of screening mammography. Evidence suggests that cancer detection rates may be slightly enhanced by using CAD,^{25, 29} particularly among less experienced radiologists.³⁰ One large prospective community-based study comparing breast cancer detection with and without CAD demonstrated a cancer detection rate of 3.2 cancers/1000

women screened without CAD and 3.8 cancers/1000 women screened with CAD, a 19.5 percent increase. However, these higher detection rates appear to come at the expense of increased recall rates. Recall rates in this same study increased from 6.5 to 7.7 percent.^{25,29} Using a non-commercial CAD system in a screening situation, Helvie et al.³¹ detected 10/11 malignancies for a 91 percent sensitivity, which was identical to the radiologists' sensitivity. The missed cancer was different for each modality. Due to CAD results, recall increased 9.7 percent, from 14.4 to 15.8 percent. Interestingly, in a 1-year follow-up, five patients developed cancer, two of whom were marked by CAD the preceding year. In a recent article by Gur, et al.,³⁰ the recall rate for 24 radiologists interpreting 115,751 screening mammograms (59,139 with CAD and 56,432 without CAD), demonstrated a similar recall rate with and without CAD (11.39 versus 11.4 percent, respectively) and similar breast cancer detection rates with and without CAD (3.49 versus 3.55/1000, respectively). These data, however, were not adjusted for possible differences in the characteristics of the women screened and whether the examination was the woman's first or subsequent exam.

Different algorithms are used in different CAD systems and no evidence is available about differences across these systems or the reproducibility of interpretation results. Algorithms have been refined over time and these refinements have proceeded even for systems within clinical trials. Further, procedures for how CAD is used to complement radiologists' review of digital images are not uniform. CAD adds approximately \$20 to the cost of a screening mammogram, and CAD has been shown to substantially increase the amount of time needed to interpret each mammogram.

Public Health Factors – CAD is widely available and is rapidly achieving substantial market penetration. CAD introduces an additional step in the interpretation process. Following initial review and interpretation of mammography images, CAD results are reviewed and the mammography images may then be re-reviewed to assess specific CAD findings. Thus, use of CAD would not be expected to increase clinical efficiency. Among potential concerns are that CAD may be reviewed before initial interpretation and that CAD may alter radiologists' normal search and decision-making process. Over reliance on CAD prompts could limit search in some areas of the digital image.^{32,33} And while CAD may provide an objective source of information in litigation, there also is evidence of misuse of the technology by litigators to generate independent interpretations of digital images without radiologists' involvement.

While CAD is intended to be used after the initial interpretation of the mammogram to assure that results do not bias the radiologists' interpretation, there are numerous anecdotal reports that CAD results are reviewed while mammograms are being interpreted. The studies that have assessed CAD have carefully limited its use as an adjunct after the initial interpretation of a mammogram. It is possible that the results of these CAD studies are not generalizable to community practice. As a result, community recall rates from CAD may be even higher than those found in studies.

MRI

Test Characteristics – MRI is not a primary screening test for women at average risk for breast cancer. MRI has been used to detect malignancies in women who have problematic diagnostic

mammograms or unknown primary malignancies, to detect recurrences in women who have been treated conservatively for breast cancer, and/or to search for additional occult foci in women with a known malignancy. Studies of MRI have primarily assessed MRI as a screening test for breast cancer in women high-risk for the disease (e.g., BRCA1/2 carriers).

Studies of MRI among women at high risk for breast cancer demonstrate substantially higher sensitivity than mammography in detecting cancer. Warner, et al.,³⁴ reported sensitivities among women at high risk for breast cancer of approximately 36 percent for mammography compared to 77 percent for MRI, using BI-RADS[®] 1 to 3B as negative findings. Using similar criteria, Kriege et al.³⁵ reported sensitivities of 24 percent and 47 percent for mammography and MRI, respectively. When Kriege et al.³⁵ included BI-RADS[®] 3 as abnormal, sensitivities for mammography and MRI were 40 percent and 71 percent, respectively. However, these higher sensitivities also come with lower specificity.²⁰ Approximately 10 to 25 percent of high risk women screened with MRI received a false-positive result.^{34, 35} MRI has not been shown to decrease morbidity or mortality in any group of women. Further, the unique combination of consequences from increased false positive findings and the challenge of accurately conveying patient risk for breast cancer among women at high-risk for breast cancer for whom the test might be appropriate increases the likelihood of errors in therapeutic decision making.

An important limitation of the test is the general lack of capacity to perform MRI-guided biopsy to verify occult findings. This limitation is particularly noteworthy given the high false positive rates associated with the test. When abnormal and suspicious findings are identified, there is no way to confirm that the finding is benign without surgical resection or short interval re-evaluation. Further, protocols for performing breast MRI are not standardized and there are few expert readers for breast MRI. Like mammography, a BI-RADS[®] lexicon system has been established to guide the interpretation of MRI findings. But unlike mammography, the BI-RADS[®] lexicon for MRI is less well developed or tested. There are no accreditation programs for breast MRI interpreters and understanding of MRI BI-RADS[®] reports are generally low in clinical practice. The reproducibility of MRI is not known, but given these factors is likely lower than mammography. Some centers have begun providing breast MRI without a dedicated breast coil.

MRI as a screening test among at women high risk for breast cancer would be an adjunct to, not a substitute for, a screening mammogram. MRI would not be necessary following an abnormal mammogram. MRI is an expensive procedure, more than 10 times the cost of film mammography. CMS only reimburses for MRI as a diagnostic procedure in women at high risk for breast cancer.

Public Health Factors – While MRI is generally available in most major clinical centers, breast MRI requires a breast coil for accuracy, and breast MRI using a breast coil is not widely available. Financial and marketplace incentives exist for increased use of MRI. MRI centers are profit sources for hospitals and are marketed to women as cutting edge technology with distinct advantages over mammography.

Patient acceptability of breast MRI is questionable. MRI is an invasive examination, requiring injection with a contrast agent. Further, patients must lie in an imaging cylinder for 30 to 60

minutes. Many find the conditions claustrophobic and are bothered by the noises associated with the procedure, in some cases requiring sedation and increasing the complexity of the procedure. While women at higher-risk for breast cancer may be more motivated to comply with screening recommendations than average risk women, patient acceptability of breast MRI may be substantially lower than for other imaging modalities such as mammography.

Directing a screening exam to a subpopulation of NBCCEEDP eligible women at higher risk for breast cancer would have considerable impact on program operations. Standard reporting categories and criteria would need to be established for characterizing women as eligible for MRI based on some minimum genetic or breast density criteria. New testing procedures for assessing genetic risk would need to be implemented, confidentiality protected, and associated genetic counseling provided. Data and financial systems would need to be changed to accommodate the collection and reporting of risk criteria. It is likely that case management demands would rise to meet the needs of women receiving non-standard testing and/or to address new patient issues.

Ultrasound

Test Characteristics – Ultrasound is not a primary screening test for women at average risk for breast cancer. Ultrasound has been used as a diagnostic test in women who have suspicious abnormalities based on physical examinations or screening mammography. Studies of screening ultrasound primarily assess the test as an adjunctive screening exam for breast cancer in women for whom mammography is less effective (e.g., women with dense breasts).

Ultrasound is widely used as a diagnostic test to further evaluate masses found on physical examination or mammography. Ultrasound discriminates well between solid lesions that require biopsy and cystic lesions that do not require follow-up. Twenty-five to 50 percent of breast masses are benign cysts. Thus, the role of ultrasound in the evaluation of suspected breast masses is important and well established. A large number of publications have reported that ultrasound can be used effectively to characterize solid breast masses and to estimate the risk of cancer.³⁶

Ultrasound has been studied in several small observational and uncontrolled studies for its ability to detect breast cancer among women who have dense breasts. When used as an adjunctive screening test for women with dense breasts, ultrasound resulted in high false-positive rates leading to large numbers of additional diagnostic procedures with only a small gain in the number of cancers detected.³⁷⁻⁴² However, the American College of Radiology Imaging Network (ACRIN) is conducting a large screening ultrasound trial, which may provide important new information in the near future.

Ultrasound is highly operator dependent. Further, ultrasound is a real time examination and diagnostic value is lost if not interpreted in real time. Despite reduced diagnostic value of static images, failure to capture these images precludes re-review and requires repeating the entire procedure if re-review is needed. A bilateral screening examination can take from 15 to 60 minutes. While most facilities have ultrasound equipment, few providers are trained specifically for whole breast screening examination. Protocols for performing breast ultrasound are not standardized and are not implemented uniformly. Similar to mammography, a BI-RADS®

lexicon system has been established to guide the interpretation of breast ultrasound. But unlike mammography, the BI-RADS® lexicon for ultrasound is less well developed or tested. There is an accreditation program for breast ultrasound but very few sites have applied for accreditation and understanding of ultrasound BI-RADS® reports is generally low in clinical practice. The reproducibility of ultrasound and its interpretation are unclear, but appear lower than mammography.

Because screening with ultrasound may be appropriate only for women with dense breasts and breast ultrasound is used primarily as a diagnostic exam to distinguish between solid lesions that require biopsy and cystic lesions that do not require follow-up, the distinction between a screening and diagnostic ultrasound and associated determination of a woman's routine screening cycle could become confused. CMS reimburses for ultrasound as a diagnostic procedure. The addition of ultrasound as a screening exam to mammography among women with dense breast tissue would double the cost of screening.

Public Health Factors – Ultrasound equipment is available in nearly all facilities that perform breast imaging, but many facilities use ultrasound systems that are old and equipment variability is high. The time requirements of the examination reduce its feasibility as a screening exam. Further, high false positive rates would require increased time for patient education.

Directing a screening exam to a higher risk subpopulation of NBCCEDP eligible women would have considerable impact on program operations. Standard reporting categories and criteria would need to be established for characterizing women as eligible for ultrasound based on some minimum breast density criteria. It is likely that case management demands would rise to meet the needs of women receiving non-standard testing and/or to address new patient issues. The proportion of eligible women that might be classified as having dense breasts and thus eligible for ultrasound screening is unknown, but could be as high as 20 to 25 percent of program eligible women.

Provider education would be required to address issues related to distinctions between screening and diagnostic ultrasound and determinations of women's screening cycles for program eligibility. Education also would be required about program criteria for defining breast density and consequent eligibility for screening ultrasound.

RECOMMENDATIONS

Reimbursement Policies

Following careful review of the test characteristics and public health factors associated with each technology, the NBCCEDP Expert Panel on Breast Cancer Reimbursement Policies discussed potential reimbursement policies and the supporting rationale for each option. Panel members reached consensus on specific recommendations for reimbursement policies and identified the key factors providing the rationale for their recommendation. These recommendations and the key rationale points for each are presented below.

Digital Mammography

Recommendation:

Digital mammography should be reimbursed only at the conventional rate for film mammography. This recommendation should be reassessed following release of DMIST study findings.

Rationale:

- ***Cost*** – The per-test cost of digital mammography would substantially increase screening costs and consequently reduce the total number of women who could be screened by the program.
- ***Access*** – The current limited market penetration of digital mammography suggests that access to the NBCCEDP program will not be substantially affected by the lack of reimbursement for the technology.
- ***Accuracy*** – There is insufficient evidence that digital mammography would contribute to reductions in morbidity/mortality over that achieved by film mammography. This lack of evidence is particularly problematic given the large cost differential between the two technologies.
- ***Reproducibility*** – Lack of standardization and current levels of image and interpretation reproducibility limit the overall accuracy of the exam.

CAD

Recommendation:

CAD should not be reimbursed at this time.

Rationale:

- ***Cost*** – The costs associated with the addition of CAD to current interpretation procedures and the increase in the number of needed follow-up tests for increased false positive findings based on CAD would substantially increase program costs and consequently reduce the total number of women who could be screened by the program. The added cost of 3 CAD procedures would eliminate program funds for one film mammogram
- ***Accuracy*** – There is insufficient evidence that CAD would contribute to reductions in morbidity/mortality over that achieved by film mammography. Further, increased rates of false positive findings would result in unnecessary follow-up procedures and anxiety for women.

MRI

Recommendation:

MRI should not be reimbursed as a screening examination for either (BRCA 1/2) women at high-risk or average risk for breast cancer at this time. This recommendation should be reassessed following release of ACRIN study findings and formal, clear definition of “high risk”.

Rationale:

- ***Program Operations***– Development and implementation of program systems and procedures to direct MRI screening to a subpopulation of women at high risk and to provide necessary case management and genetic counseling support are overly prohibitive for the relatively small potential public health gain.
- ***Accuracy*** – While sensitivity may be increased among women at high risk, false positive rates are unacceptably high, resulting in unnecessary tests and anxiety for women.
- ***Reproducibility*** – Lack of standardization of breast MRI imaging and interpretation limit the overall reproducibility of the exam across settings.
- ***Access*** – Staff time and program resources to implement directed screening could limit resources to provide screening across the population of eligible women.

Ultrasound***Recommendation:***

Ultrasound should not be reimbursed as a screening examination for either normal or high risk women at this time. Reimbursement should continue for ultrasound as a diagnostic procedure for all women after an abnormal breast examination finding and/or mammogram.

Rationale:

- ***Accuracy*** – Test sensitivity is lower than that achieved by mammography and false positive rates among women with dense breasts are higher, resulting in unnecessary test procedures and anxiety for women.
- ***Access*** –Time requirements and the increased costs of the exam, could limit program access to services and disproportionately divert provider time away from other program services.
- ***Reproducibility*** – Lack of standardization of the technology, appropriate credentialing and expertise for operators, as well as equipment variability limits the reproducibility of the exam.
- ***Population characteristics*** – Because younger women are more likely to have denser breast tissue and the risk of breast cancer is substantially lower in these younger age groups, the proportional number of cancers identified from use of the exam directed to this subpopulation would be extremely low.

Research and Surveillance

In addition to specific reimbursement policy recommendations, the panel developed recommendations to address the general paucity of data to inform policy determinations. These recommendations include:

- Fund pilot studies in a subset of NBCCEDP programs to assess current levels of use of CAD.
- Consider pilot assessments of specific reimbursement policy changes on technology practice patterns and the effects of such changes on program operations.

- Initiate planning efforts to more clearly and practically define criteria for high risk.

Future Reimbursement Policy Reviews

The panel recommended that the CDC assess on an annual basis whether new technologies and/or data have emerged that could change existing reimbursement policies. In the presence of new technologies and/or data, an expert panel review of policies should be undertaken. A full policy review should be undertaken at least every 5 years. USPSTF evidence reviews should be utilized to prevent duplication of effort.

BIBLIOGRAPHY

1. U.S. Preventive Services Task Force. Screening for Breast Cancer: Recommendations and Rationale. *Ann Intern Med.* 2002;137((5 Part 1)):344-346.
2. American Cancer Society. *Cancer Facts & Figures.* Atlanta, GA: American Cancer Society; 2005.
3. Ernster VL, Ballard-Barbash R, Barlow WE, et al. Detection of ductal carcinoma in situ in women undergoing screening mammography. *J Natl Cancer Inst.* 2002;94(20):1546-1554.
4. Carney P, Miglioretti D, Yankaskas B, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Ann Intern Med.* 2003;138(3):168-175.
5. Smith-Bindman R, Chu P, Miglioretti D, et al. Physician Predictors of Mammographic Accuracy. *J Natl Cancer Inst.* 2005;97:358-367.
6. Ryerson A, Benard V, Major A. 1991-2002 National Report: Summarizing the first 12 years of partnerships and progress against breast and cervical cancer. *Centers for Disease Control and Prevention.* Available at: <http://www.cdc.gov/cancer/nbccedp/Reports/NationalReport/index.htm>. Accessed April 6, 2005, 2005.
7. Bobo J, Shapiro D, Schulman J, et al. On-schedule mammography rescreening in the National Breast and Cervical Cancer Early Detection Program. *Cancer Epidemiol Biomarkers Prev.* 2004;13(4):620-630.
8. Food and Drug Administration. *21CFR 814.2: Food and Drug Administration;* 2004.
9. Federal Register Notice 03-24361. *Medicare Program: Revised Process for Making Medicare National Coverage Determinations:* Department of Health and Human Services, Centers for Medicare and Medicaid Services; 2003.
10. Center for Medicare and Medicaid Services. Medicare Physician Fee Schedule. *Center for Medicare and Medicaid Services.* Available at: <http://www.cms.hhs.gov/physicians/mpfsapp/step1.asp>. Accessed March 17, 2005.
11. Joy J, Penhoet E, Petitti D, et al., eds. *Saving Women's Lives: Strategies for improving breast cancer detection and diagnosis.* Washington, DC: Institute of Medicine and National Research Council; 2005.
12. Fletcher SW, Elmore JG. Mammographic screening for breast cancer. *N Engl J Med.* 2003;348:1672-1680.
13. Smart C, Byrne C, Smith R, et al. Twenty-year follow-up of the breast cancer diagnosed during the Breast Cancer Detection Demonstration Project. *CA Cancer J Clin.* 1997;47:134-149.
14. Tabar L, Vitak B, Tong H, et al. Beyond randomized controlled trials: organized mammographic screening substantially reduces breast carcinoma mortality. *Cancer.* 2001;91:1724-1731.
15. Nystrom L, Andersson I, Bjurstam N, et al. Long-term effects of mammography screening: updated overview of the Swedish randomized trials. *Lancet.* 2002;359(9310):909-919.
16. Kerlikowski K, Grady D, Barclay J, et al. Effect of age, breast density and family history on the sensitivity of first screening mammography. *JAMA.* 1996;276:33-35.

17. Tabar L, Vitak B, Chen H. The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up. *Radiol Clin North Am.* 2000;38:625-651.
18. Carlson K, Helvil M, Roubedoux M, et al. Relationship between mammographic screening intervals and size and histology of ductal carcinoma in situ. *AJR Am J Roentgenol.* 1999;172:313-317.
19. Kerlikowski K, Carney P, Geller B, et al. Performance of screening mammography among women with and without a first-degree relative with breast cancer. *Ann Intern Med.* 2000;133:855-863.
20. Elmore JG, Armstrong K, Lehman CD, et al. Screening for Breast Cancer. *JAMA.* 2005;293(10):1245-1256.
21. United States General Accounting Office. *Mammography Capacity Generally Exists to Deliver Services.* Washington, D.C.: US General Accounting Office; April 2002. GAO-02-532.
22. Colbert JA, Kaine EM, Bigby J, et al. The age at which women begin mammographic screening. *Cancer.* 2004;101:1850-1859.
23. National Center for Chronic Disease Prevention and Helath Promotion. Risk Factors and Use of Preventive Services: United States. *Centers for Disease Control and Prevention.* Available at: <http://www.cdc.gov/nccdphp/burdenbook2004/Section03/mammog.htm>. Accessed April, 26, 2005.
24. Lewin JM, D'Orsi CJ, Hendrick RE, et al. Clinical comparison of full-field digital mammography and screen-film mammography for detection of breast cancer. *Am J Roentgenol.* 2002;179:671-677.
25. Lewin JM, Hendrick RE, D'Orsi CJ, et al. Comparison of full-field digital mammography with screen-film mammography for cancer detection: Results of 4,945 paired examinations. *Radiology.* 2001;218(3):873-880.
26. Skaane P, Young K, Skjennald A. Population-based mammography screening: Comparison of screening film and full-field digital mammography with soft-copy reading - The Oslo I Study. *Radiology.* 2003;229:877-880.
27. Skaane P, Skhennald A. Screen-film mammography versus full-field digital mammography with soft copy reading: Randomized trial in a population-based screening program - The Oslo II Study. *Radiology.* 2004;232:197-204.
28. Food and Drug Administration. *Office of Health and Industry Programs Fiscal Year 2003 Annual Report: Radiologic Health Programs.* Washington DC: Food and Drug Administration; 2003.
29. Freer TW, Uliddey MJ. Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center. *Radiology.* 2001;220(3):781-786.
30. Gur D, Sumkin JH, Rockette HE, et al. Changes in breast cancer detection and mammography recall rates after the introduction of a computer-aided detection system. *J Natl Cancer Inst.* 2004;96(3):185-190.
31. Helvil MA, Hodjuski L, Makariou Eea. Sensitivity of non-commercial computer-aided detection system for mammographic breast cancer detection: pilot clinical trial. *Radiology.* 2004;231:208-214.

32. Krupinski EA. An eye-movement study on the use of a CAD information during mammographic search. Paper presented at: Seventh For West Imaging Perception Conference; October 16-18, 1997; Tucson, Az.
33. Krupinski EA. Computer-aided detection in clinical environment. Benefits and challenges for radiologists. *Radiology*. 2004;231:7-9.
34. Warner E, Plewes DB, Hill KA, et al. Surveillance of BRCA1 and BRCA2 mutation carriers with magnetic resonance imaging, ultrasound, mammography, and clinical breast examination. *JAMA*. 2004;292(11):1317-1325.
35. Kriege M, Brekelmans C, Boetes C, et al. Efficacy of MRI and mammography for breast cancer screening in women with a familial or genetic predisposition. *N Engl J Med*. 2004;351:427-437.
36. Stavros AT, Thickman D, Rapp CL, et al. Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. *Radiology*. 1995;196(1):123-134.
37. Irwig L, Houssami N, van Vliet C. New technologies in screening for breast cancer: a systematic review of their accuracy. *Br J Cancer*. 2004;90 (11):2118-2122.
38. Leconte I, Feger C, Galant C, et al. Mammography and subsequent whole-breast sonography of nonpalpable breast cancers: the importance of radiologic breast density. *Am J Roentgenol*. 2003;180(6):1675-1679.
39. Gordon PB, Goldenberg SL. Malignant breast masses detected only by ultrasound. A retrospective review. *Cancer*. 1995;76(4):626-630.
40. Kolb TM, Lichy J, Newhouse JH. Occult cancer in women with dense breasts: detection with screening US -- diagnostic yield and tumor characteristics. *Radiology*. 1998;207(1):191-199.
41. Buchberger W, Niehoff A, Obrist P, et al. Clinically and mammographically occult breast lesions: detection and classification with high-resolution soography. *Semin Ultrasound CT MR*. 2000;21(4):325-336.
42. Buchberger W, DeKoekkoek-Doll P, Springer P, et al. Incidental findins on sonography of the breast: clinical significance and diagnostic workup. *AJR Am J Roentgenol*. 1999;173(4):921-927.

APPENDIX A: KEY INFORMANT INTERVIEWS

EVALUATION OF NBCCEDP REIMBURSEMENT POLICIES FOR NEW BREAST AND CERVICAL CANCER SCREENING TECHNOLOGIES

INTRODUCTION

The National Breast and Cervical Cancer Early Detection Program (NBCCEDP), administered by the Centers for Disease Control and Prevention (CDC), helps low income, uninsured, and underserved women gain access to lifesaving screening programs for the early detection of breast and cervical cancers. The program implements a wide range of activities, including a) public education and outreach to increase access to services; b) administration of breast and cervical cancer screening exams and diagnostic testing; c) case management to facilitate access to care and utilization of best practices; and d) professional education and quality assurance to ensure the highest standard of care for women in the program. The NBCCEDP is implemented in all 50 states, 4 U.S. territories, the District of Columbia, and 13 American Indian/Alaska Native organizations. While the program has screened 1.9 million women and provided 4.6 million screening examinations since its inception in 1991, it reaches fewer than 20 percent of eligible women, primarily due to financial limitations.

While the size and complexity of the NBCCEDP poses many challenges, one challenge has been the determination of which screening and diagnostic tests should be paid for by the program. Since the program's inception, scientific advances have resulted not only in improvements to existing screening and diagnostic tests and implementation procedures, but also in the introduction of new technologies. Determinations about whether the NBCCEDP should pay for use of newer screening and diagnostic tests and procedures are complicated. The program must balance a wide range of factors, including for example, standard of care for women in the program, the public health mandate to serve as many women as possible, limited program funds, varying local health services infrastructures, and the impact of changes in program policies on program operating procedures and partners.

The CDC is reviewing the NBCCEDP reimbursement policies for breast and cervical cancer screening and diagnostic services. For breast cancer, the NBCCEDP currently provides reimbursement for film mammography only. Digital mammography, magnetic resonance imaging (MRI), and ultrasound are not reimbursed as screening tests. Computer aided detection (CAD) of digital mammograms is not reimbursed. For cervical cancer, the NBCCEDP provides reimbursement for conventional pap tests, but not for liquid-based pap tests. HPV/DNA testing is reimbursed only for women with ASC-US findings on pap.

Recognizing the complexity of this task and the significant impact on individual NBCCEDP programs, the CDC sought to gather additional information about programs' experiences with reimbursement policies. Key informant interviews with NBCCEDP program directors representing eight state programs and two CDC program staff were conducted to gather information about the range of issues that should be considered in CDC's evaluation of its reimbursement policies. Specifically, interviews sought to provide information about:

- a) The type and magnitude of NBCCEDP challenges resulting from current reimbursement policies for screening technologies;

- b) NBCCEDP approaches for addressing challenges associated with current reimbursement policies;
- c) The range and nature of NBCCEDP modifications that would need to be made to adjust to potential modifications of current reimbursement policies for new screening technologies; and
- d) How appropriate balance might be achieved across scientific, infrastructure, programmatic, and public health impact factors in decision-making concerning NBCCEDP reimbursement policies.

METHODS

Interviews were conducted in December, 2004 with NBCCEDP program directors representing eight state programs and two CDC program staff. NBCCEDP program directors volunteered to participate in key informant interviews following an invitation from the NBCCEDP Science and Epidemiology Subcommittee. Program Directors could include other program staff in interviews at their discretion.

Email interview confirmations included an overview of the key informant assessment and a list of questions to be addressed in each interview. Four of the eight interviews with NBCCEDP program directors focused on breast cancer and the remaining four focused on cervical cancer. Interviewees were not restricted, however, from identifying issues beyond the specific cancer focus for their interview and most interviewees addressed reimbursement issues related to both cancers. Each interview was conducted by telephone by Dr. Marianne H. Alciati. Interviews lasted between 45 and 75 minutes. Handwritten interview notes were taken during each interview and a typed summary was prepared following each interview. These summaries were used as the primary information source for analysis. Interviews were tape recorded for verification purposes only and all tapes were destroyed at the end of the analysis.

Each interview summary was reviewed to identify themes and representative issues. Because the purpose of this assessment was to identify the range and nature of reimbursement challenges faced by the NBCCEDP and the sample size was so small, the specific numbers of mentions for each issue and the number of interviewees mentioning each issue was not calculated. However, general comments are presented reflecting whether a particular issue was identified by multiple sites.

LIMITATIONS

It is important to recognize that while the data from these interviews provides a valid picture of issues across the eight programs and from the perspectives of two CDC staff, it does not provide information about the pervasiveness of these issues across NBCCEDP sites and only generally provides perspective on the magnitude of each issue within NBCCEDP programs. While it is generally accurate that the eight programs combined with CDC staff perspectives are typical of NBCCEDP programs, the diversity across NBCCEDP programs and the method for selecting key informant interviewees suggests that the experiences of these programs may not be representative of all programs. It is possible and even likely, that some additional issues or examples exist within other programs. However, these interviews do provide a clear and accurate picture of the majority of issues resulting from current reimbursement policies and changes in policy.

RESULTS

NBCCEDP programs are complex local partnerships, involving extensive networks of providers and health care organizations who deliver screening and diagnostic examinations and help provide and coordinate follow-up care. Reimbursement for screening and diagnostic services is at the heart of the program, representing a significant driving force for how the NBCCEDP programs operate within local communities. Reimbursement policies influence not only what services these programs provide, but also how efficiently they provide those services and how the programs are perceived within their local communities and nationally.

Interviewees identified a broad range of issues associated with existing reimbursement policies as well as historic and current procedures for modifying these policies and communicating revisions. The vast majority of these issues were similar for both breast and cervical cancer reimbursement policies. For this reason, this presentation of results focuses on these issues and their common characteristics with illustrative examples from breast and cervical cancer. While most of the interview results focus on factors that influence demand for new technologies and the challenges posed by current reimbursement policies and review procedures, two significant overriding perspectives were emphasized by the majority of interviewees. First, the NBCCEDP provides a critical public health service and program participants are extremely committed to the NBCCEDP's success. Second, interviewees were extremely appreciative of the opportunity to provide input to the policy review process and of the CDC's commitment to and efforts on behalf of the NBCCEDP.

All NBCCEDP programs are required to reimburse at rates that do not exceed state Medicare rates. Although different state formulas may be used to establish these rates (e.g., urban vs. rural rates), they are quite low and in some cases below the actual cost of delivering the service. Several interviewees pointed out that some costs associated with providing diagnostic and follow-up procedures to this population are not reimbursable using CDC funds. These costs are often paid by state funds (not available in all states), grants, donations, or other sources; or absorbed by the facility or provider. But both of these options add pressure to the system of delivering NBCCEDP services. Newer technologies further exacerbate this pressure because they are often more expensive, although costs tend to decline over time. The consequence of higher costs for individual screening and diagnostic exams is a reduction in the programs' overall capacity to "achieve the greatest good for the greatest number." The reality that the program currently reaches only 20 percent of the eligible target population makes these trade-offs particularly difficult.

Program Consequences: But as revealed in these interviews, the issues go well beyond simple cost calculations. A broad range of consequences result from NBCCEDP reimbursement policies. These are presented below in five broad categories, including a) program performance, b) relationship with providers, c) practice patterns, d) standards of care, and e) program credibility.

Program Performance: Interviewees emphasized that the cost to individual programs of different reimbursement policy decisions have affects well beyond just the cost of individual examinations. In some areas, the failure to reimburse newer technologies has reduced the

number of providers who deliver services for the program causing program shortages. In other cases, providers have used their size or banded together with other providers to pressure the program to reimburse for newer technologies at the approved Medicare rates.

Reduced provider capacity can limit both the programs' ability to meet demand for early detection services as well as cause delays in providing needed services. Delays, in turn impact Minimal Data Element (MDE) reporting and a program's ability to achieve service delivery targets. Examples were noted in NBCCEDP's failure to reimburse for liquid-based pap (LBP) examinations. The paucity of providers performing conventional pap in some areas required women to travel for services, resulting in screening delays or failures to get screening.

Another impact of reimbursement policies on program performance relates to efficiency. In cases of an abnormal pap, use of conventional pap rather than LBP requires a second office visit and additional call-back efforts. This process was noted both to increase the likelihood that follow-up HPV testing would not be accomplished and to drain limited resources due to the need to find women and to pay for a second office visit. Other inefficiencies emerge as well. The need for alternate funding to cover costs for un-reimbursed services takes time and resources, not only to identify sources of funds, but to establish systems that account for separate sources of funding.

Beyond complications associated with existing policies, changes in reimbursement policies have extraordinary implications for program operations. Providers and their staff need to be made aware of new policies, corresponding CPT codes need to be identified and populated in reimbursement systems, data and reporting systems need to be modified, and contract requirements need to be adjusted. Ideally, program policy manuals also would be updated. Some programs indicated that listings of reimbursed procedures are not included in their program manuals because of the unpredictability of policy changes and, in at least one case, the reversal of a policy within a six month timeframe. Failures to include reimbursement information in policy manuals introduces another set of operational requirements, such as development of a separate listing of reimbursable services and increased communication to clarify reimbursement policies and procedures with providers and their staff.

Relationship with Providers: Many interviewees discussed the pressures on providers and their relationship with the program resulting not only from low reimbursement rates, but from a complex interplay of other factors. Providers historically have born much of the responsibility for ensuring follow-up and treatment for women diagnosed through the program. For breast cancer in particular, medical liability risks are high. Failure to diagnose breast cancer is the primary cause in the U.S. for malpractice claims and the second-leading reason for subsequent claimant payments. Providers also are challenged to keep pace with complex scientific evidence and medical advances. Media publicity further complicates this challenge as patients request and sometimes demand newer technologies that may not be reimbursable through the program. These factors are compounded when newer technologies become available in the market but are not reimbursed by the program and when the NBCCEDP changes what services can be reimbursed under the program.

Many interviewees commented on the extra financial burden to providers when they must absorb the additional cost difference between BCCEDP approved technologies and newer technologies.

While most interviewees commented on the high level of commitment of providers to the NBCCEDP, this added burden is perceived to strain that commitment. In some areas providers have left the program, but more often interviewees indicated that under current policies, providers remain with the program in hopes of upcoming policy changes.

Other consequences for providers were noted, particularly in the ethical dilemma of delivering what, in some cases providers believe to be less than the best care available. In this way, reimbursement policies are viewed as driving the practice of medicine, changing the role of the provider, and changing the patient/provider relationship. Providers in these situations are “pressured” to offer only covered services. In this role, as one interviewee commented, the program is not a “legitimate partner.” Further, many women will not get services until they are assured that they will not be billed. This tension is compounded when patients learn about new technologies through the media, advocacy organizations, or other sources and question the care they receive through the program. Differential treatment as noted by some interviewees fuels distrust between patients and providers.

Reimbursement policies that do not include newer technologies, particularly when they are available within a provider’s health care setting, also increase liability risks. Failure to provide a test or procedure in situations where a cancer is later identified increases the providers’ vulnerability to litigation, particularly if the decision appears based on cost.

All these factors combine to define the relationship between the programs and providers. All interviewees commented on the importance of building and maintaining strong relationships with the providers in their program. Several noted that reimbursement issues have created tension, most notably reflected in ‘uncomfortable’ dialogues in which program staff find themselves ‘arguing with providers’ about interpretations of scientific evidence, or countering a provider’s direct experience with a technology (e.g., LBP is easier to read). Interviewees noted that they expend a lot of time and effort communicating with their providers about the science and rationale behind current reimbursement policies. Some position these policy communications as the program staff and providers on one side and CDC on the other. Often program staff appears to be ‘stretching’ the commitment of providers until policies change in time.

Practice Patterns: It became clear across interviews that different localities adopt newer technologies at different rates. For example, in some areas labs have gone exclusively to LBPs or CAD. In cases where only the newer technology is available, newer technologies are reimbursed at the rates of approved technologies. But newer technologies are often more expensive and the added cost difference must either be absorbed by providers or reimbursed through alternative funds, placing added strain on providers and alternate sources of funds. Several interviewees noted that procedures for providing and billing for new technologies at the rates of approved technologies preclude analysis of the frequency of this practice within the program.

Incompatibilities with existing local health care practices also can lead to inefficiencies and open the door for error. In some cases, the cost difference has been billed directly to women participating in the program. For example, a few interviewees conveyed stories of the cost difference between film mammography and digital mammography with CAD or between

conventional pap and LBP, estimated at about \$60 in each case, being billed directly to women. In some instances, these cases have gone into collections, placing extraordinary and unnecessary burden on women in the program. If an abnormality is identified, some providers back-bill this cost difference to Medicaid. While direct billing to women is disallowed by the program and the situations identified were ultimately resolved, they require considerable staff time and resource as each case must be addressed individually. These situations also extol a price in terms of women's negative experience with the program.

Another example provided by several interviewees of NBCCEDP reimbursed practices being out of step with local practices was the approval for cervical cancer testing using the Digene system. This process allows two samples to be captured during an initial patient visit, one for conventional pap and a second for HPV testing following an abnormal pap. But in most facilities, this procedure applied only to NBCCEDP clients and facilities did not have the capacity to properly store the second sample for potential follow-up. In many cases facilities were unfamiliar with the system altogether.

Another concern stemming from continued use of approved technologies for NBCCEDP women when facilities and providers have transitioned to newer technologies is perceived decline in proficiency by providers for technologies that they no longer perform with the same frequency. For example, one interviewee noted provider concerns about their proficiency interpreting pap slides due to declining frequency associated with increased use of LBP.

Standards of Care: As noted above, providers raise concerns about providing care through the NBCCEDP that is "less than optimal care." But these concerns appear to extend well beyond providers and in reality are fueled both by media coverage and public promotion of medical advances and pharmaceutical marketing efforts directed to providers that may oversell the science behind new technologies. Interviewees raised concerns about both the reality and perception that women in the NBCCEDP receive a different standard of care than those with the financial means to pay for health care. Several interviewees spoke of an emerging, two-tiered system of health care where the poor receive a lower level of care. This raised both public health and ethical concerns.

Program Credibility: Perceptions of a different standard of care for women in the NBCCEDP was viewed as one of several factors that undermine the credibility and reputation of the program. But several interviewees also noted that inefficiencies resulting from reimbursement policies that differ from common practice, as discussed above, also undermine the program's reputation. Resentment was reflected in one local program where providers 'banned together' to demand reimbursement at Medicare rates for LBP. Bad will is also generated when women are billed for differential costs, as in the cases noted above for LBP and CAD.

Perceptions that the NBCCEDP is 'out of step' with current technology has other ramifications as well. One program conveyed an interesting scenario in which their program was unable to participate in a collaborative research study with academia and the Indian Health Service to assess the impact of digital mammography on access to care for underserved, rural populations. The study was viewed as having great potential for expanding the program's reach, but the program's inability to participate because digital mammograms could not be reimbursed was

viewed as reducing program credibility. In this case and more broadly in the program provider relationship, some interviewees indicated these situations threatened the viability of the program as a credible partner in meeting the needs of underserved women.

Finally, several interviewees commented on discrepancies between the reimbursement policies of the NBCCEDP and policies of other federal programs, such as reimbursement by Medicare and approvals for use of new technologies by the Food and Drug Administration. These inconsistencies are confusing and increase the challenge and importance of program communications. Several interviewees also perceived these discrepancies as reducing NBCCEDP credibility

Review Procedures: The majority of interviewees commented on the historic and current process for revising reimbursement policies. Most expressed appreciation for the interview process and CDC's efforts to include their perspective in the current review of these policies. Continued involvement of multiple perspectives, and particularly NBCCEDP Program Directors was viewed very favorably. Many positive changes were noted in reimbursement policies over the past several years, in particular approvals for loop electrode excision and cold-knife conization of the cervix as diagnostic procedures and HPV testing as follow-up to ASC-US results on pap. Many also noted the improvements resulting from legislative action in 2001 to allow treatment reimbursement through Medicaid.

But the rare instances where policy changes had been made and reversed stood out. Reversals were perceived as program ineffectiveness and "taking something away." This situation required considerable staff time and resources to revise systems and communicate with program partners, and resulted in large credibility costs. In the context of policy revisions, interviewees again emphasized the large ripple effect of changes, requiring changes in recruitment and outreach, data and coding systems for reimbursement, provider education, and MDE reporting.

Several interviewees also commented specifically on the timing of policy revisions. These reviews are not conducted on a fixed schedule and announcements about revisions are not coordinated with impacted program cycles, such as contract renewal dates.

Systems for communicating policy revisions do not appear to be reaching all programs equally. Several interviewees emphasized the importance of enhancing communication about reimbursement policies as well as the process and rationale for policies, both between CDC and the programs, and between program staff and providers. Standardization of the process was often advocated, however, interviewees varied in their perspectives about how flexible final policies should be. Some saw value in flexibility, allowing the individual programs to adjust to local circumstances such as different practice patterns and rates of adoption of new technologies. Others advocated for "hard and fast rules" that they perceived to alleviate confusion shift the burden of unpopular reimbursement decisions to CDC rather than the local program. Some interviewees highlighted the importance of CDC support and assistance translating reimbursement policies into implementation procedures, such as aligning CPT codes to reimbursable procedures.

Finally, across interviewees a number of criteria for reimbursement policy determinations were identified. These included:

- Impact – ensure that policies extend the reach of the NBCCEDP.
- Scientific credibility – policies must be evidence-based, reflecting support for the most effective technologies.
- Cost-benefit – cost benefit analyses that account for all program costs – exam/procedure costs, implementation costs, and credibility costs – must support the overall benefit of new technologies.
- Current and future practice patterns – analysis of the rate of adoption of new technologies and the consequences of different program procedures must be considered.
- Consistency – policies should seek to minimize inconsistencies across national guidelines and federal programs that can adversely affect implementation.

CONCLUSIONS

The NBCCEDP is clearly a critical and valued public health program seeking to meet a need well beyond its resources. CDC, program staff, providers and many other key program partners demonstrate extraordinary commitment to the goals and implementation of the program. But the program is complex, with a broad array of factors influencing its capacity to maximize the delivery of services. Reimbursement policies for program services are at the apex of this web of influences. The key informant interviews conducted for this assessment identified and organized these influencing factors as a basis for more fully and systematically considering the impact of different reimbursement policies on the NBCCEDP. The primary factors identified include program performance, the program's relationship with providers, practice patterns, standards of care, and program credibility.

These interviews also identified strategies for improving the review and implementation process for reimbursement policy revisions, including a) involving multiple perspective, particularly at the program level, b) establishing a standardized process, and c) coordinating the timing of revisions with program cycles impacted by policy revisions. Clear criteria that consider program impact, scientific evidence, cost/benefit, practice patterns and continuity should be applied. And stronger systems must be established for communicating policy decisions and their rationale throughout the many partners of the NBCCEDP.

APPENDIX B: EXPERT PANEL PARTICIPANTS

Samantha Allison¹
SVP, Medical Group
GE Healthcare Financial Services

Helen Barr, MD
Director, Division of Mammography
Quality and Radiation Programs
Center for Devices and Radiological Health
Food and Drug Administration

Vernon R. Daly, MD
President
HEUREKA

Mary L. Dolan, MD,MPH
Division Chief, Women's Primary
Healthcare
University of North Carolina School of
Medicine

Carl D'Orsi, MD
Director of Breast Imaging
Emory University

Joann Elmore, MD, MPH
Associate Professor, Medicine
Adjunct Associate Professor, Epidemiology
Section Head, General Internal Medicine,
Harborview Medical Center
Associate Director, Robert Wood Johnson
Clinical Scholars Program
University of Washington

Janelle Guirguis-Blake, MD
Program Director, US Preventive Services
Task Force
Agency for Healthcare Research and Quality

Bradley Hutton, MPH²
Director, Cancer Services Program
New York State Department of Health

Jeffrey Marks, PhD
Associate Professor of Experimental
Surgery, Department of Surgery
Duke University Medical Center

Robert Smith, Ph.D
Director, Cancer Screening
American Cancer Society

Rebecca Smith-Bindman, MD
Associate Professor, Radiology,
Epidemiology and Biostatistics, Obstetrics
Gynecology and Reproductive Medicine
University of California, San Francisco

Barnarese Wheatley, MPH, MA, EDD¹
Health Care Program Administrator
Alameda County Medical Center

Lawrence D. Wagman, M.D., FACS
Professor and Chair, Division of Surgery
City of Hope National Medical Center

Pamela A. Wilcox, RN MBA
Assistant Executive Director
American College of Radiology

1) Member, Breast and Cervical Cancer Early Detection and Control Advisory Committee

2) Director, State Breast and Cervical Cancer Early Detection Program