In this paper, the historic and recent evidence supporting the value of breast cancer screening will be described, along with the underpinnings of the current debate over the relative and absolute benefit of regular mammography screening.

**Introduction**

In the world, breast cancer is the most common cancer diagnosed in women and the most common cause of death from cancer. Globally, an estimated 1.4 million new cases of breast cancer were diagnosed in 2008.[1] In the United States, breast cancer is the most common cancer diagnosed among women, and the second leading cause of death from cancer.[2]

Breast cancer is a progressive disease that becomes systemic as the size of the tumor increases. Tumor size, nodal involvement, and histologic grade all are linked to prognosis, and tumor size is strongly associated with nodal involvement and advanced histologic grade, each of which portends worse survival.[3] When breast cancer is diagnosed while still localized to the breast, 5-year relative survival is 98.6%, compared with 83.8% for regional disease, and 23.3% for distant disease.[4] The effects of these prognostic indicators also are evident in longer-term follow-up, that is, of 20 years or more.[3]

In the early 1970s, experimental evidence from a prospective randomized controlled trial (RCT) demonstrated conclusively that screening with a combination of clinical breast examination (CBE) and mammography using general purpose x-ray equipment could reduce the risk of dying from breast cancer.[5] Later, the Swedish Two-County Trial, utilizing dedicated mammography units, demonstrated that mammography alone resulted in a significant reduction in breast cancer deaths.[6] Since then, eight additional RCTs have been conducted, which together provide firm evidence of the efficacy of screening in reducing breast cancer mortality.[7,8] Moreover, recent evidence from a 29-year follow-up of the Swedish Two-County Trial demonstrates the substantial long-term benefit of invitation to mammography screening in relative and absolute terms.[9]

**Modern Breast Cancer Screening**

**Physical examinations of the breast**

Before the introduction of mammography, breast cancer was detected principally by women themselves, either by finding signs of a tumor incidentally or during breast self-examination (BSE), or by a health professional at a CBE. BSE is a formal method of structured self-examination of each breast (usually performed monthly), and is distinct from self-detection of symptoms incidentally during daily activities. The examination protocol for CBE is similar to that for BSE, but CBE is performed by a health professional. The potential contribution of self-detection and CBE to clinical outcomes depends on whether or not mammography screening programs are available. In settings in which women have access to regular mammography screening, self-detection and CBE continue to play important roles in detecting interval cancers; mammography will not detect all breast cancers, because of the heterogeneous nature of the disease and the large variations that exist in breast density. In the Two-County Trial of breast cancer screening, women in the group invited to screening who were diagnosed with an interval cancer had better long-term survival than the control group, which was likely the result of counseling that patients received after the mammography exam, to be alert to breast changes.[3] Among those who are not within the age range of women offered mammography screening, incidental self-detection, BSE, and CBE are the only methods that can lead to an earlier detection of a palpable mass.

**Mammography**
Mammography is a low-dose x-ray examination of the breasts performed with dedicated imaging equipment to detect abnormalities that may be breast cancer. The image receptor in modern mammography equipment is either screen-film or digital. Full-field digital mammography (FFDM) has largely replaced screen-film units over the past decade, and according to the US Food and Drug Administration (FDA), 88% of all accredited mammography units in the US are FFDM units.[10] Dedicated mammography equipment is specifically designed to produce high-quality images of the breast at a minimum x-ray dose (approximately 3 to 4 mGy per view).[11] In the US, image quality and interpretive skills have also been improved through early efforts by the American College of Radiology’s Mammography Accreditation Program (ACRMAP), and subsequently as a result of the passage of the Mammography Quality Standards Act (MQSA) of 1992, which requires a facility to meet a broad range of technical and personnel standards in order to be certified by the US Food & Drug Administration (FDA).[12] The incremental improvement in image quality is evident in examples of mammograms derived from different mammography imaging technologies from the 1940s to today (Figure 1).

Table 1: Performance Measures for 1,960,150 Screening Mammography Examinations performed in the US From 2002 to 2006 by Age, Based on Breast Cancer Surveillance Consortium Data as of 2009

A mammography screening examination is performed by specially trained radiology technologists, and involves two views of each breast: a craniocaudal (CC) view and a mediolateral oblique (MLO) view. The recall rate generally is higher at the first exam compared with subsequent exams, but overall it typically ranges from 7% to 10% and is slightly higher in younger women compared with older women.[13] In most instances, the use of a multimodality approach (detailed mammography examination, hand-held ultrasound exam, and magnetic resonance imaging [MRI] in selected cases, combined with the occasional use of interventional methods such as fine needle-aspiration biopsy or larger bore-needle biopsy) provides a preoperative microscopic diagnosis (Figure 2). The sensitivity and specificity of mammography vary somewhat by age, with sensitivity, specificity, and the positive predictive value (PPV) improving with increasing age (Table 1).[14] Historically, age-specific differences in the accuracy of mammography have been a central issue in the debate over the value of screening women under age 50, and generally mammography performance in women aged 40 to 49 has been compared with that in women 50 years of age and older. This comparison has led to the mistaken impression that performance measures are uniform in postmenopausal women, and measurably poorer in premenopausal women. As shown in Table 1, although there is improvement in sensitivity, specificity, and PPV with increasing age, mammography performance in adjacent decades of life is quite similar. Not shown in Table 1 is PPV as the percentage of biopsies that result in a diagnosis of breast cancer, of which the median percentage is approximately 32%.[15] While an intuitive interpretation of the PPV is that it mainly is influenced by the false-positive rate, in fact the PPV mostly is affected by the underlying prevalence of disease. Therefore, even if screening performance is exactly the same, PPV will be lower in the age group with lower disease prevalence.
The randomized controlled trials

FIGURE 2

The patient is a 47-year-old woman with known bilateral breast cysts

The efficacy of mammography screening has been well established through the accumulation of evidence from the RCTs and more recent evaluations of modern service screening. There have been eight population-based RCTs of breast cancer screening, and two RCTs that randomized volunteers (the Canadian National Breast Screening Studies [NBSS] 1 and 2). Among the RCTs, five were carried out in Sweden, three in North America, and two in the UK. The earliest RCT of breast cancer screening was the Health Insurance Plan of Greater New York Trial (HIP), which was initiated in 1963.[5] Nearly 30 years later (1991), the United Kingdom (UK) launched the Age Trial, which was designed specifically to measure the benefit of mammography screening among women in their 40s, without any age migration past 50 years.[7]

The RCT with the longest follow-up (29 years) is the Swedish Two-County Trial, which was the first trial to demonstrate a breast cancer mortality reduction associated with invitation to mammography screening without CBE. The results from the long-term follow-up are worth highlighting because they illustrate the importance of long-term follow-up of the RCTs to measure the full impact of screening. In the Two-County Trial, 133,065 women aged 40 to 74 and residing in two Swedish counties were randomized into a group invited to three to four rounds of mammographic screening over a 7-year period, and a control group receiving usual care. From the first evaluation 8 years after the study start throughout the 29 years of follow-up, there was a highly significant 31% fewer deaths in the group invited to screening than in the control group (relative risk [RR] = 0.69; 95% confidence interval [CI], 0.56–0.84; \( P < .0001 \)).[9] It is noteworthy that the statistically significant reduction in breast cancer mortality not only lasted over the duration of the follow-up period, but in addition, the absolute benefit of an invitation to screening improved over time and was still improving after 20 years. In fact, most of the breast cancer deaths prevented occurred 10 years or longer after the inception of screening. This observation illustrates the importance of very-long-term follow-up (ie, > 20 years) to approach measuring the full impact of mammography screening, in particular the absolute benefit. For example, the number of women needed to screen (NNS) every 2 to 3 years over a 7-year period to save one life was 922 at 10 years, 464 at 20 years, and 414 at 29 years. If screening had been carried out for 10 years and the same relative benefit had been achieved, the absolute benefit as measured by the NNS would have been greater, that is, an estimated 300 women needed to screen to save one life. The absolute benefit can be expressed in other ways as well: at 29 years of follow-up, one life was saved for every 1334 screening mammograms, or for every 1000 women from the ages of 40 to 69 screened every 2 years, between 8 and 11 breast cancer deaths would be prevented. FIGURE 3
Relative risks of breast cancer mortality

Figures 3a and 3b show the summary relative risk (RR) of breast cancer mortality in the groups invited to screening compared with the control groups for the 8 population-based RCTs (RR = 0.77; 95% CI, 0.73–0.86) and 10 population-based and non–population-based RCTs (RR = 0.79; 95% CI, 0.73–0.86), a combined estimate of 23% and 21% fewer deaths associated with an invitation to screening. Over the years, meta-analyses with various RCT inclusion criteria have been conducted to provide overall RRs and age-specific RRs, ranging from 14% to 23% fewer breast cancer deaths associated with an invitation to screening. As seen in Figure 3, however, five of the trials showed greater mortality reductions than the summary statistic.

The results of these meta-analyses should be regarded as conservative estimates of the effectiveness of mammography performed in the 1970s to 1980s, for several reasons: First, not all of the RCTs were equally successful in reducing the risk of being diagnosed with an advanced breast cancer, which is the principal purpose of mammography screening (Table 2). RCTs that succeeded in significantly reducing the risk of being diagnosed with an advanced breast cancer in the group invited to screening also eventually demonstrated a similar, significant reduction in the risk of dying from breast cancer. Second, the RCT results are based on intention-to-treat analyses, that is, the summary RRs are based on the difference in the breast cancer death rate in the invited (attended and not attended combined) and control groups regardless of individual exposure to mammography screening. Obviously, there is a lower mortality rate among those who attended screening regularly than among all invited women. Close examination of the individual studies, rather than blurring those differences in a meta-analysis, provides a more informative and evidence-oriented approach to evaluating the results of the RCTs. Third, breast imaging technology has improved considerably over the nearly 50 years since the first breast cancer RCT was launched. These improvements fall into two groups: 1) tailoring the screening interval to a woman’s age, based on research results demonstrating the different tumor growth rates according to age and histologic tumor types; and 2) technical improvements that include developments in screen-film systems and FFDM systems, the contribution of quality assurance algorithms to image quality, the appreciation of the importance of two-view mammography, double reading, the use of computer-aided detection (CAD), and the emerging use of automated ultrasound screening or tomosynthesis as an adjunct to two-view mammography only in women with dense breasts. When working up the screening findings, the addition of ultrasound and breast MRI to diagnostic evaluation, as well as the use of interventional methods to establish microscopic diagnosis preoperatively, play an important role in arriving at the final diagnosis and initiating proper patient management. Several decades of experience and significantly better imaging technology and performance have contributed to considerable improvements in breast cancer screening and diagnostic imaging.[16]

The evaluation of service screening

TABLE 2
The Randomized Trials of Breast Cancer Screening: Relative Risk (RR) of Diagnosis of a Node-Positive Breast Cancer and RR of Mortality

In the post RCT era, numerous evaluations of service screening have taken place, applying case-control and cohort study designs. Case-control studies tend to measure the actual effect of exposure to screening, whereas cohort studies may measure the effect of attending screening, an invitation to screening, or both. In the majority of studies in which mammography has been offered to the public for a significant duration of time, results show that mortality reductions associated with an invitation to screening usually are equal to or better than those observed in the RCTs.\[17\] In 2005, Gabe and Duffy summarized both the methodological challenges of evaluating screening in a nonexperimental setting, and results from 38 nonrandomized studies of breast cancer screening. The results indicated that breast cancer mortality reductions on the order of 30% to 40% were associated with screening. Since then, further observational studies have been published, the majority of which indicate a substantial and significant reduction in breast cancer mortality with screening.\[18-25\] In addition, in the evaluation of service screening, the association between the reduction in the risk of being diagnosed with an advanced breast cancer and the subsequent mortality reduction that was observed in the RCTs also are evident in the studies that have examined tumor characteristics in exposed and unexposed women. For example, in a large Swedish study (23,092 cancers and 10,177,113 person-years of observation) the rates of lymph node-positive cancers, of tumors with pathological size > 2 cm, and tumors of TNM stage II or worse were compared before and after the introduction of screening.\[26\] Rates were adjusted for changes in overall incidence during the period of study and stratified by age (40 to 49 and 50 to 69 years). In the period after screening was introduced, among women aged 40 to 49, there was a significant 45% reduction in tumors greater than 2 cm among women exposed to screening compared with the prescreening period, and a 33% reduction in the 50- to 69-year age group. Somewhat smaller but statistically significant reductions in lymph node-positive tumors and stage II tumors also were observed for all age groups in women exposed to screening, compared with the prescreening period.

An increasingly common expression of screening costs is the NNS or the number to needed to invite (NNI) to prevent one breast cancer death. Results from major organized service screening programs and long-term follow-up of the randomized trials indicate that screening 300 to 400 women for up to 10 years will result in prevention of one breast cancer death.\[9,23,27\] In contrast, the Nordic Cochrane review finds a very small benefit, of the order of 2000 women needed to invite to screening throughout a 10-year period to prevent one breast cancer death.\[28\] How do we reconcile these different estimates? One limitation of the NNI is that it is a poor proxy of exposure to screening, since it is inflated by the rate of nonparticipation in screening. This is especially the case, in this instance, when the NNI is derived from a meta-analysis, in which variable rates of nonattendance, number of screening rounds, and duration of follow-up in multiple RCTs further distort the estimate. Perhaps the greatest problem with the estimate from the Nordic Cochrane review is that it is not based on an observed outcome but rather on the authors’ estimate of the outcome based on their subjective judgment of the quality of the screening trials. In contrast, estimates that cluster between 300 and 400 women needed to screen to save one life are derived directly from empirically observed data in randomized trials and service screening programs.\[9,23,27\]

Potential harms associated with screening

Harms associated with screening derive mostly from adverse effects associated with further examination of the mammographic findings. These include added patient financial cost and inconvenience, anxiety associated with positive test results, biopsy for benign lesions, and overdiagnosis. One has to keep in mind that there are numerous hyperplastic breast changes that
mimic the mammographic appearance of breast cancer. Differentiating them from true malignancies necessitates the use of additional diagnostic tools, including needle biopsies, to arrive at a reliable microscopic diagnosis. While these procedures would constitute “harm” when findings are benign, they decrease the occurrence of unnecessary surgical biopsies, which constitute a greater harm. In contrast to the considerable benefits of screening, including a significant decrease in advanced cancers and decreased mortality, “harms” generally include experiences that range from hardly injurious at all to those that are somewhat injurious, such as invasive procedures to rule out the presence of a malignancy.

The potential harm that has received the greatest attention, and which receives the greatest emphasis from critics of screening, is overdiagnosis. Overdiagnosis usually is defined as the diagnosis of a breast cancer by screening that would never have been diagnosed in the patient’s lifetime if screening had not taken place. An “overdiagnosed” breast cancer is pathologically indistinguishable from a breast cancer that is progressive and potentially life-threatening, and thus, when treated, represents overtreatment. If molecular or histologic features could distinguish an indolent cancer from one that is progressive, overtreatment could be avoided. Since this is not the case, however, overdiagnosis must be understood as a statistical concept, measured by examining observed vs expected incidence rates associated with a screening program. Estimates of overdiagnosis vary widely, reflecting the different methodological approaches to its measurement.[23,29-36] There are two major challenges in estimating overdiagnosis of breast cancer due to screening: 1) taking account of pre-existing trends in breast cancer incidence when estimating the expected incidence in the absence of screening; and 2) distinguishing the excess incidence related to lead time from that resulting from overdiagnosis. Studies that fully take account of pre-existing trends and lead time tend to estimate overdiagnosis rates of 1% to 11%,[23,29,30,34-36] whereas those that do not take these factors into account estimate overdiagnosis rates of 30% or more.[31,32] Therefore, while it is likely that some overdiagnosis results from breast cancer screening, it appears to be a relatively minor phenomenon; moreover, the risk of overdiagnosis in a screened woman is less than the probability of a breast cancer death prevented.[23]

A more commonly encountered side effect of screening is the recall for further assessment of suspicious mammographic findings in women who do not have breast cancer, so-called “false positives.” In Europe, the cumulative risk of such a false-positive exam over 10 screening rounds is approximately 20%, with corresponding cumulative risks of invasive procedures ranging from 1% to 4%.[37] In the US, estimated cumulative risk of at least one false-positive examination over a 10-year period for women undergoing biennial screening was 41.6%, with a corresponding 4.8% cumulative risk of a recommendation for biopsy.[38] The cumulative risk of at least one false-positive is higher if a woman undergoes annual screening.

**Breast Cancer Screening Guidelines**

**TABLE 3**


Breast cancer screening guidelines for average-risk women, from the American Cancer Society (ACS)[39] and the United States Preventive Services Task Force (USPSTF),[40] are commonly cited by primary care physicians as very influential in guiding their recommendations to patients.[41] Both organizations endorse the importance of mammography screening, but they differ in their recommendations for the age to begin and end screening, and in the screening interval (Table 3). In 2002, the USPSTF had recommended mammography for women aged 40 years and older, with or
without CBE, every 1 to 2 years.[42] This recommendation was similar to the ACS guidelines, which recommend annual mammography and CBE for women beginning at 40 years of age.[39] Also, similar to the ACS, the USPSTF had stated that the evidence that mammography screening was beneficial in women aged 40 to 69 was generalizable to women 70 years of age and older “if their life expectancy is not compromised by comorbid disease”—and essentially concluding that as long as a woman was in good health, mammography screening could be expected to be beneficial. The ACS recommendation for annual screening was based on improved sensitivity of mammography with annual vs biennial screening, especially in women under age 50, whereas the USPSTF recommendation for screening every 1 to 2 years was based on RCT evidence showing a mortality reduction associated with annual and biennial testing, and the absence of an RCT comparing the two intervals. In 2009, the USPSTF modified its recommendations, withdrawing a recommendation that women aged 40 to 49 undergo routine mammography screening every 1 to 2 years in favor of individualized decisions, and changing the recommendation that women 50 to 74 years of age be screened with mammography every 1 to 2 years to a biennial screening interval.[40] In the updated recommendations, the USPSTF also concluded that there was insufficient evidence to assess the balance of benefits and harms of CBE, and also mammography screening for women 75 years of age and older.

Although evidence from age-specific meta-analyses of RCT data showed similar mortality reductions from women randomized in their 40s and 50s, the USPSTF concluded that since a woman’s absolute risk of breast cancer during her 40s was lower, and the harms of screening were judged to be similar or slightly higher compared with older women, there was “moderate certainty that the net benefit was small.”[40] The underpinnings of this decision can be challenged on several points; First, the USPSTF chose to rely only on meta-analysis of RCT data, which as described above results in an underestimate of the benefit of modern mammography. This is especially problematic because the first generation of RCTs screened women under age 50 with protocols (2-year interval, single-view mammography, etc) that were shown to be ineffective in younger women. Second, while the absolute risk to women in their 40s is less than that of women in their 50s and 60s, it is not trivial; death from breast cancer is the leading cause of cancer-related premature mortality in women, and a leading cause of premature death from all causes. Third, the USPSTF overestimated the impact of the harms relative to benefits in two ways: 1) screening benefits estimated by the RCTs were compared with harms based on modern rates of false-positive results from a mammography surveillance registry. Since the true benefit of screening is underestimated in the RCTs because some women in the invited group are not exposed to any or all screening rounds, comparing harms measured among all women attending community based screening overestimates harms relative to benefits; 2) evidence that showed women are aware of false-positive results and regard them as an acceptable tradeoff for the benefits of early detection was not included in the systematic review.[43]

Conclusion

There is strong evidence that mammography screening is an effective strategy to reduce the incidence of advanced breast cancer. Therefore, it is a means to reduce mortality by creating the opportunity to initiate breast cancer treatment earlier in the natural history of the disease. For screening programs to be successful, high rates of adherence to regular attendance must be coupled with high-quality screening and prompt delivery of state-of-the-art treatment for women diagnosed with breast cancer. Because screening is inherently imperfect, efforts must be equally focused on reducing the adverse effects of false-positive tests, especially the fraction that is avoidable. Organized screening programs tend to be more successful in delivering high-quality screening and thus are preferable to opportunistic screening.

Financial Disclosure: The authors have no significant financial interest or other relationship with the manufacturers of any products or providers of any service mentioned in this article.

References:


18. Autier P, Boniol M, Gavin A, Vatten LJ. Breast cancer mortality in neighbouring European countries with different levels of screening but similar access to treatment: trend analysis of WHO


Source URL:
http://www.diagnosticimaging.com/printpdf/breast-cancer-screening-evolving-evidence/page/0/1

Links:
[1] http://www.diagnosticimaging.com/review-article
[4] http://www.diagnosticimaging.com/authors/1%20Szil%C3%A1sz%20tab%C3%A1r-md