Hypothesis/Commentary
See related minireview by Paci et al., p. 1159

The Value of Modern Mammography Screening in the Control of Breast Cancer: Understanding the Underpinnings of the Current Debates

Robert A. Smith

Abstract

Since the introduction of mammography screening, debates about the value of screening have endured and been contentious. Recent reviews of the randomized controlled trials reach different conclusions about the absolute benefit of screening, as do evaluations of population trends in breast cancer mortality and the evaluations of service screening. Conclusions about the value of screening commonly are expressed in terms of the balance of benefits and harms, which can differ greatly even when derived seemingly from the same data. It can be shown when different estimates are adjusted to a common screening and follow-up scenario, differences in balance sheet estimates diminish substantially. The strong evidence of benefit associated with exposure to modern mammography screening suggests that it is time to move beyond the randomized controlled trial estimates of benefit and consider policy decisions on the basis of benefits and harms estimated from the evaluation of current screening programs. *Cancer Epidemiol Biomarkers Prev; 23(7); 1139–46. ©2014 AACR.*

Introduction

Debates over the value of mammography screening have been ongoing since the earliest published evidence demonstrated reduced breast cancer mortality associated with an invitation to screening (1–4). While the accumulation of experimental and observational evidence (5–7) provided sufficient confidence for many health systems to introduce regular mammography screening, growth has been accompanied by enduring and typically unwavering debates over the balance of benefits and harms in specific age groups and overall (8–11).

The Nature of Disputes about Mammography

Some screening debates have reflected different interpretations and judgments about the same data. Early examples are the continuing debates about inviting women ages 40 to 49 years to screening (10, 12, 13), whereas a more contemporary example is whether the age to begin screening and the screening interval should be tailored to individual risk (14). Today, the former still is more of a judgment-based debate, whereas the latter is regarded as a worthy goal, but opinions differ about the potential and practicality for risk-based screening (15). Some debates endure by shifting to new issues as earlier ones are resolved. The debate over screening women in their 40s was first defined by the absence of evidence from randomized controlled trial (RCT) data showing a statistically significant reduction in mortality. That debate might have been resolved with greater understanding of the poorer performance of biennial screening in women ages 40 to 49 years (16) and once meta-analysis of the RCT data demonstrated a statistically significant mortality reduction (17) due to longer follow-up and the inclusion of results from 2 second-generation RCTs (18, 19). Instead of the issue being considered settled, the debate then shifted from efficacy to cost-effectiveness (20), then from relative to absolute benefit (21), and eventually to the balance of benefits and harms, where benefits are defined as a reduction in the risk of breast cancer death, and harms mostly are defined by the false-positive rate (22, 23). Unlike the early debate, the modern debate is complex and very much defined by the data that are chosen to represent the magnitude of benefits and harms, that is, the balance sheet (24).

Supporting evidence for benefits and harms may rely on different methodologies applied to the same data or different data, and thus estimates may differ considerably. Modern examples include estimates of the relative risk of dying from breast cancer associated with invitation or exposure to screening, the absolute benefit of screening measured by the number needed to be invited (NNI) or screened (NNS) to save one life, the age to begin and end screening, and how to judge the downsides of screening, such as radiation exposure, false-positive examinations, and overdiagnosis. While conclusions about whether or not mammography screening can be recommended on the basis of a balance sheet of estimated benefits and harms

Author's Affiliation: American Cancer Society, Atlanta, Georgia

Corresponding Author: Robert A. Smith, American Cancer Society, 250 Williams St, Suite 600, Atlanta, GA 30303. Phone: 404-329-7610; Fax: 404-327-6404; E-mail: robert.smith@cancer.org

©2014 American Association for Cancer Research.

may seem to be uncomplicated, the underlying data supporting the estimates may not be evident or well understood. Differences in the mortality reduction associated with screening can differ more than 4-fold (25, 26), the number needed to invite/screen to save one life differ more than 20-fold (27, 28), and estimates of overdiagnosis range from 0% to greater than 50% (29). At the core of the extreme views against the value of mammography are challenges to the credibility of the experimental evidence (11), trend data of published incidence and mortality rates (30, 31), and claims of very high rates of overdiagnosis and overtreatment (32, 33). Reports advancing these findings and concluding that mammography screening either never was justified, or no longer is, often are published in leading journals and have generated considerable media attention. Taken together, the modern arguments against breast cancer screening emphasize an unfavorable balance of benefits and harms, stating that harms, principally false-positive findings and overdiagnosis, significantly exceed the small absolute benefit, which is further diminishing due to progress in therapy and increased breast awareness (22, 26, 34, 35). Others looking at the same data conclude that screening is beneficial and that benefits outweigh harms (22, 36–38). Indeed, beyond favorable interpretation of the RCT evidence, there is a steadily growing body of literature of the evaluation of service screening showing breast cancer mortality reductions as great as and commonly greater than those observed in the RCTs (25, 39–43). These reports conclude that modern mammography and improvements in therapy are contributing to a decline in the breast cancer mortality rate, and while both are important, early detection is likely to be the more significant factor (39, 44).

Recent Reviews of Mammography Screening

In the last year, there have been 3 evaluations of the effectiveness of mammography: (i) an update of the Nordic Cochrane Institute's review (26); the independent review of the UK's National Health Service's breast cancer screening program (27); and the evaluation of the European breast cancer screening programs by the EURO-SCREEN working group (45), a summary of which is published in this issue of the journal (46). It is useful to examine these 3 recent evaluations of breast cancer screening because they represent different approaches to the data and reach a range of conclusions about the benefits and harms of mammography screening and whether or not the balance between them is favorable. Of particular note is whether the comparison of absolute benefit is based on the NNI or the NNS or both. Results from RCTs are always first reported on the basis of an intention-to-treat analysis and estimates of the absolute benefit of screening based on the NNI retain fidelity to the intention-to-treat analysis. However, the NNI is a rather nebulous concept for estimating the effectiveness of mammography because variable nonadherence to the invitation to screening and contamination in the control group

diminish the measurement of the effectiveness of screening due to deaths that occur among nonparticipants and deaths avoided due to cross-over. Here, estimating the NNS is preferable, especially when adjusted for selection bias, as it measures the effectiveness of screening in those who accept the invitation to screening. RCTs and observational studies can report both NNI and NNS, the latter with a correction for selection bias. Observational studies also may report the NNI to estimate the benefit of screening at the population level and to compare more modern results with those of the RCTs.

Nordic Cochrane Institute

The Nordic Cochrane Institute's report is the most recent update of their assessment of the effectiveness of mammography first reported in 2001 (26, 47). Then and now, the conclusion of the report is that the evidence does not support mammography screening. The early report challenged the experimental evidence, principally in terms of the randomization process, and concluded that the only RCTs with adequate randomization methods, that is, the 2 Canadian trials and the Malmö trial, showed no benefit from screening. The inclusion/exclusion criteria for their meta-analysis have been criticized as misguided (48), overly restrictive (49), and self-serving (50), as studies that showed a benefit from screening were excluded. Moreover, no other systematic evidence review has followed their restrictive methodology, with the exception that most will exclude the Edinburgh trial (and did so before the first Cochrane report) due to an imbalance in the socioeconomic status of the invited and control groups (36, 51, 52). The conclusions of the Cochrane authors have evolved over time, with the most recent report acknowledging a benefit from mammography screening but arguing that: (i) the absolute benefit of mammography screening today is smaller due to advances in treatment and women's increased awareness and (ii) that the harms of screening, principally overdiagnosis, exceed the benefit (26). The most current Cochrane estimate of the mortality reduction from the RCTs associated with an invitation to screening is 19%, but then the author's downgrade it to a 15% reduction in mortality based on their judgment (not calculation) that suboptimal randomization biases the estimate upward (26). Even allowing for opinion and judgment in evidence-based medicine, an arbitrary, downward adjustment of the point estimate is extraordinary. For the balance sheet, the authors estimate that the rate of overdiagnosis is 30% and that 2,000 women must be screened for 10 years to save one life, at a cost of 10 overdiagnosed cases of breast cancer.

Independent UK Panel on breast cancer screening

The commission of an Independent UK Panel on breast cancer screening to review the UK National Health Service's (NHS) breast cancer screening program resulted partly from an open letter to Professor Sir Mike Richards (National Cancer Director, England) from Dr. Susan

Bewley challenging him to acknowledge the findings of the Nordic Cochrane Institute, conduct a truly independent review of the evidence, and based on the findings "adjust screening policy appropriately" (53). In response, the leadership of the NHS and Cancer Research UK asked Professor Sir Michael Marmot to assemble and chair a small group of independent exerts (with no prior published work on breast cancer screening) to review the evidence related to the benefits and harms of breast cancer screening in the United Kingdom. The panel concentrated their analysis on the RCTs of breast cancer screening and applied their findings to the UK breast cancer screening program. The panel's meta-analysis observed 20% fewer breast cancer deaths comparing invited versus control women. For the balance sheet, the panel based their estimates on 10,000 women ages 50 years invited to screening every 3 years for a 20-year period (ages 50–69 years), with the cumulative risk of breast cancer mortality measured from ages 55 to 79 years. On the basis of that scenario, the panel estimated that one breast cancer death would be prevented for every 235 women invited to screening or for every 180 women who attended screening (27). The panel based their estimate of overdiagnosis (19%) on the excess proportion of cancers diagnosed among women invited to screening compared with the control group in the Malmö and Canadian trials. Thus, according to the scenario above, 1 in 77 women invited to screening would be diagnosed with an overdiagnosed cancer or 3 cancers overdiagnosed for every one life saved. In the summary of their report published in *The Lancet*, the panel acknowledged that mammography screening was associated with both benefits and harms, but concluded, "the UK breast screening programmes confer significant benefit and should continue" (27).

The EUROSCREEN working group

The EUROSCREEN working group consists of 30 experts involved in the planning and evaluation of cancer screening programs in the European Screening Network (54, 55). Unlike the 2 reports described above, the EUROSCREEN Working Group's report, published in a supplement to the *Journal of Medical* Screening (45), and summarized here (46), was both an evaluation of the European breast cancer screening programs and a review of the methodologic issues involved in the evaluation of observational data of mammography screening (25, 45, 56, 57), including trend studies, case–control studies, and incidence-based mortality (IBM) studies (25, 56, 57). In addition, the EUROSCREEN working group produced a comprehensive evaluation of the European literature on overdiagnosis (29) and false-positive outcomes (58).

The EUROSCREEN group began with a literature search focused on the effect of mammography screening in Europe, and after applying inclusion criteria, including (i) mortality as an endpoint, (ii) age groups ≥ 50, and (iii) nonrandomized controlled trial study design, they identified 17 trend studies, 20 IBM studies, and 8 case–control studies for evaluation. The countries represented in these studies included Denmark, Finland, Iceland, Italy, Netherlands, Norway, Spain, Sweden, and the United Kingdom. All national programs initiated screening from 1987 to 2007, and literature evaluated in the EURO-SCREEN analyses was published from 2001 to 2012. Trend and IBM studies compared breast cancer mortality before and after the introduction of screening, whereas the case–control studies compared breast cancer deaths (cases) with matched controls from the same population on the basis of screening history before the diagnosis date of the breast cancer case. All studies were evaluated for heterogeneity.

The trend studies typically were based on an evaluation of registry data, and either described breast cancer mortality trends before and after the introduction of screening ($n = 5$) or attempted to quantify the effect of mammography screening on breast cancer mortality ($n = 12$). Studies with adequate follow-up (3 single-country studies) after the full implementation of screening (≥10 years) observed mortality reductions ranging from 28% to 36%. Given the limitations of trend analyses described by Moss and colleagues (56), and the varied methodology of these studies, no attempt was made to produce pooled estimates of the effect of screening on breast cancer mortality. IBM studies offer advantages over trend studies because the analysis is restricted to deaths arising from cases that were diagnosed after the first invitation to screening. Among the several strengths of this approach is avoiding the sizeable contamination in the post-screening period by breast cancer deaths among women diagnosed before the introduction of screening. On the basis of 7 incidence-based mortality studies that had the strongest designs, the EUROSCREEN Working Group estimated a 25% breast cancer mortality reduction associated with an invitation to screening and a 38% mortality reduction associated with exposure to screening. Among the case–control studies, the mortality reductions were 31% and 48%, respectively. Among all pooled analyses, only the combined mortality reduction for invitation to screening in the case–control studies showed significant heterogeneity (25).

Overdiagnosis was a special focus of the EURO-SCREEN group's evaluation. Overdiagnosis refers to the diagnosis of breast cancer by screening that never would have presented clinically in the woman's lifetime. To the extent that overdiagnosis exists, it represents a significant harm as women will undergo treatment unnecessarily. Overdiagnosis must be understood as a statistical concept, as by definition, there are no defining histologic features to distinguish an overdiagnosed (i.e., nonprogressive) cancer from one that is progressive. The ideal method for estimating overdiagnosis is to compare the cumulative incidence in the invited and control groups in an RCT including several years after the end of the study and before the control group has begun screening. Opportunities are limited to do this, and there has been an

uncertainty with the existing RCT data over continued screening in the invited group or post-trial contamination in the control group. Accurately estimating the duration of the sojourn time and whether or not ductal carcinoma *in situ* (DCIS) is included also pose challenges. Given the limited opportunity to address this question with RCT data, Puliti and colleagues sought to evaluate the growing number of estimates of overdiagnosis in the service screening setting (29). To avoid bias, estimates of overdiagnosis must compare incidence rates over time in screened and unscreened populations that are equivalent with respect to the underlying risk of disease and the effects of lead time. The EUROSCREEN Working Group divided 13 studies based on whether or not overdiagnosis estimates were adjusted for lead time associated with screening and contemporaneous trends in incidence. Failure to adjust for these influences on incidence rates largely accounts for the wide range in the estimates of overdiagnosis (0%–57%), in particular extraordinary large estimates (59). Among the reliable estimates of overdiagnosis, that is, those that did adjust for lead time and changing risk over time, the investigators estimated overdiagnosis of invasive and *in situ* cancers at 6.5%. From the above, the group estimated that among 1,000 women ages 50 years screened every 2 years until age 69 years (and followed to age 79) 7 to 9 lives are saved (of 30 deaths expected in the absence of screening) and 4 women are overdiagnosed (of 67 incident cases expected in the absence of screening; ref. 45). Unlike the balance sheets described above, the EUROSCREEN Working Group estimates that more lives are saved because of screening compared with cancers overdiagnosed.

Estimating the Absolute Benefits and Harms of Mammography

In meta-analyses of the RCTs conducted to date, the relative risks of dying from breast cancer have been consistent, varying only on the basis of which RCT data are included in the analysis. This also is the case in the 2 most recent meta-analyses conducted as part of the systematic reviews described above. In contrast, absolute risk estimates derived from the various systematic reviews vary by orders of magnitude. How do we explain that the UK Panel and the Nordic Cochrane Institute each estimated a 20% and 19%, respectively, reduction in breast cancer deaths associated with an invitation to screening (26, 27) but then differed nearly 10-fold in their estimate of the number of women who needed to be invited (NNI) to screening to save one live (1/250 vs. 1/2,000, respectively)? The United States Preventive Services Task Force estimates of the NNI to screening to save one life for women ages 50 to 59 and 60 to 69 years (respectively, 1 in 1,339 and 1 in 377) also are considerably different than the UK Panel and the Nordic Cochrane Institute estimates, although their estimates are for narrower age groups (60). Each of these, in turn, is different from the estimates from the EUROSCREEN

Working Group, which estimated both NNI and NNS from observational data. Closer examination reveals that each measure of absolute risk differs in terms of the reference population, mortality benefit (originally a 19% mortality reduction became 15% in the Nordic Cochrane Institute's estimate), duration of screening and follow-up, and whether invitation or exposure to screening is being compared. In particular, the most extreme estimate of the NNI to save one life is from the Nordic Cochrane Institute, which is influenced by the shortest observation period (10 years of screening with no follow-up beyond that) and an absolute benefit estimated from a subset of the RCTs that are dominated by women in their 40s (28). In an effort to understand the disparity between estimates in absolute benefit, given the similarity of the common relative risk estimates derived from the RCTs, Duffy and colleagues standardized the Nordic Cochrane Institute, USPSTF, and EUROSCREEN estimate of absolute benefit to a common scenario, that is, the recent UK Independent Panel estimate of the effect of screening UK every 3 years in women ages 50 to 69 years on breast cancer mortality from ages 55 to 79 years. This is fairly straightforward, that is, and involves applying the Nordic Cochrane estimate of benefit to the estimated cumulative mortality in the UK screening program. For example, as Duffy and colleagues have shown, in the UK the cumulative breast cancer mortality among women ages 55–79 years is 17 per 1,000. From the perspective of the Nordic Cochrane review's claim of a 15% mortality reduction associated with an invitation to screening, the breast cancer mortality rate in the absence of screening would be 20 per 1,000 (17/0.85), with the effect of an invitation to screening being 3 deaths prevented per 1,000 women invited. As the attendance rate to screening in the UK is 77%, the effect of being screened would be 3.89 breast cancer deaths avoided per 1,000 women invited (3/0.77) or 257 women needed to screen to prevent one breast cancer death (1,000/3.89; ref. 28).

When the 4 leading estimates (Nordic Cochrane Institute, UK Independent Review Panel, USPSTF, and EUROSCREEN Working Group) of absolute benefit are adjusted to the screening scenario used by the UK Independent Review Panel (the effect of screening every 3 years for a 20-year period beginning at the age of 50 years on breast cancer mortality between ages 55–79 years), the difference in the estimates of absolute benefit are substantially reduced from 20- to 4-fold (Table 1) and thus are not as disparate as the original estimates, especially when considered over a 25-year period (28). Duffy and colleagues also converted the NNI estimates to the NNS as the NNI is artificially inflated by breast cancer deaths in the invited group among women who were nonadherent to the invitation to screening, which is nontrivial across all rounds of screening and all RCTs. The NNI is not generalizable and has no meaning to a woman contemplating screening or having already attended screening. It should not be used in a balance sheet to communicate benefits and harms.

Table 1. Original and adjusted estimate of the absolute benefit of breast cancer screening from four systematic reviews

Systematic review group	Age group, y	Intervention	Screening period/ years of follow-up	Mortality reduction	Original NNS/NNI[b]	Adjusted NNS[c]
UK Independent Panel[a]	50–69	Screening	20 y, ages 50–69/ 25 y, ages 55–79	20%	180	180
USPSTF[d]	50–59	Invitation	Avg. 7 y, ages 50–59/ Avg. 14, ages 50–59	14%	1,339	
	60–69	Invitation	Avg. 7 y, ages 60–69/ Avg. 14, ages 60–69	32%	337	
	50–69		Combined	19%		193
EUROSCREEN Working Group	50–69	Screening	20 y, ages 50–69/ 30 y, ages 50–79	38%–48%	111–143	64–96
Nordic Cochrane Institute	40–74	Invitation	10 y, ages 40–74/ 10 y, ages 40–74	15%	2,000	257

[a]Reference population and protocol.
[b]NNS: The UK Independent Panel and EUROSCREEN estimates of absolute benefit are based on the number needed to screen.
[c]Derived from Duffy and colleagues. Original estimates are adjusted to the same scenario used in the UK Independent Review, that is, the impact of screening UK women ages 50–69 years every 3 years for 20 years on mortality in women ages 55–79 years.
[d]The estimate for the mortality reduction for women ages 50–69 years is based on taking an inverse variance-weighted average of the 2 relative risks in the logarithmic scale [relative risk (RR) = 0.81 (95% CI, 0.72–0.92)].

The Enduring but Limited Value of the RCTs of Mammography Screening

Relying on the RCTs to measure the effectiveness of screening is conservative and based principally on the hierarchy of evidence. This is the logic advanced by the USPSTF, the UK Independent Review Panel, and the Nordic Cochrane Center for only considering RCT evidence in their systematic reviews. However, with the efficacy of breast cancer screening established more than 4 decades ago, does it make sense to rely on these data to estimate the effectiveness of modern mammography? The RCTs varied considerably on important protocol factors (i.e., 1- vs. 2-view mammography, the screening interval, etc.) number of screening rounds, adherence with the invitation to screening, and control group contamination. Duration of follow-up also is highly variable, and lack of long-term follow-up reduces the ability to measure the absolute benefit of screening, as was just shown in the example above, and also recently in the 29-year follow-up of the Swedish Two County Trial (61). In the Two County Trial, based on the local endpoint committee assessment, an invitation to screening was associated with a 31% reduction in breast cancer deaths over the 29-year period. However, the NNS to save one life became steadily more favorable over the duration of the follow-up period. According to the trial protocol, the NNS for 7 years (∼2–3 rounds) to save one life with 10 years of follow-up was 922. At 29 years of follow-up, the NNS to screen to save one life had dropped to 414. The investigators estimated that if screening had continued for 10 years, the

NNS to prevent one breast cancer death would be 300. The follow-up data also reveal that fewer than half of the deaths prevented were observed in the first 10 years. These observations reinforce that 10 to 15 years of follow-up should be regarded as the minimum needed to measure the benefit of screening mammography and 20+ years is preferable. Indeed, as shown in Table 1, failing to allow for longer follow-up accounts for a great deal of the difference between estimates of absolute risk, as does measuring absolute risk in terms of NNI versus NNS. Understanding the underpinnings of these estimates is important as they contribute to summary judgments about the balance of benefits and harms and screening recommendations.

Meta-analysis of the RCTs further underestimates the true effectiveness of screening. As would be expected, there are differing opinions over which among the RCTs best reveals the true association between an invitation to screening and breast cancer death, and these judgments figure heavily into meta-analysis inclusion and exclusion decisions. What is often overlooked among the RCT findings is that the relative risk of being diagnosed with an advanced breast cancer is strongly associated with the relative risk of dying from breast cancer (62). Thus, to argue that it is inappropriate to pick winners and losers among the RCTs and rely instead on meta-analysis of all trials as a more credible estimate of benefit ignores this important and consistent pattern, which more than any other observation explains why some trials demonstrated reduced breast cancer mortality and some did

not. This pattern of more favorable tumor characteristics associated with exposure to screening is also evident in the evaluation of service screening (63), lending further support for moving beyond the RCTs to estimate the effectiveness of modern mammography.

The EUROSCREEN Working Group has made a great contribution by reviewing, summarizing, and critiquing the European literature on the evaluation of modern mammography screening. Many of the key methodologic imperatives for the proper evaluation of mammography screening programs have been understood for some time, but they are commonly neglected, leading to erroneous conclusions about the value of breast cancer screening. For example, trend analyses of breast cancer mortality or incidence contemporaneous with the introduction of screening are deceptively attractive due to their simplicity, but these analyses are not well-suited to measuring the effectiveness of screening (56). For example, in the first decade after the introduction of screening, usually more than half of the breast cancer deaths are attributable to diagnoses that occurred before screening was introduced, highlighting the need to censor deaths in the post-screening era attributable to diagnoses in the pre-screening era (64). Deaths from breast cancer among women not invited to screening and among those who are invited but do not attend screening can also account for a significant fraction of deaths after screening is introduced. There also is a need to appreciate that it takes time to launch a screening program and invite and screen the target population, a process that can take several years, and long follow-up also is needed. Trends in the death rate also can be influenced by cohort effects on incidence over time and by improvements in therapy. For these reasons, the most valid designs for the evaluation of service screening are those in which individual longitudinal data linking screening history and cause of death are evaluated with either an IBM or a case–control approach (25). With respect to overdiagnosis, Puliti and colleagues have revealed the basic methodologic flaws in analyses that report alarmingly high rates of overdiagnosis, mainly a failure to account for lead time and trends in increasing breast cancer incidence, but they also have shown that credible attempts to estimate overdiagnosis can be made if the underlying risk of breast cancer is similar, if there is proper adjustment for lead time, and there is careful attention to other methodologic issues (29).

The weight of the conclusions of the recent systematic reviews provide strong support for the importance of mammography screening in the control of breast cancer and sweep aside alarmist claims that the benefits of mammography in the era of adjuvant therapy are few, and the harms, specifically high rates of overdiagnosis, unacceptably high. While continued analysis of the RCT databases will remain fruitful, especially where contemporary question can be addressed with the benefit of decades of follow-up, it is time to move beyond the RCTs and use modern service screening data evaluated with appropriate methodology to inform screening policy, a sentiment echoed by others (54). The EUROSCREEN Working Group has provided us with a comprehensive blueprint of current state-of-the art methods for the evaluation of mammography service screening. We can expect these methods to evolve, but it also is important to support the archiving of data that will improve the evaluation of screening programs, especially in the United States, which lacks nation-wide population-based longitudinal data linking screening and outcomes. Indeed, insofar as measuring the balance of benefits and harms should be ongoing and regularly assessed, contemporary data on harms should be compared with contemporary data on benefits. This is preferable to populating the balance sheet with old data on benefits from meta-analysis of RCT data and contrasting it with modern data from all women exposed to mammography to measure harms. Finally, the outcomes that populate the balance sheet may improve as a result of new technology or interventions focused on improving accuracy. Guidelines developers should be clear about why the balance of benefits and harms does or does not support a recommendation for screening. When the balance is judged to be unfavorable, they also should state clearly what change in the level of benefits, harms, or both would be sufficient to endorse screening.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Shapiro S, Strax P, Venet L. Periodic breast cancer screening in reducing mortality from breast cancer. JAMA 1971;215:1777–85.
2. Miller AB. Early detection and mass screening for cancer. Can Fam Physician 1972;18:76–9.
3. Morgan RW, Vakil DV. Etiology of breast cancer. III. Opportunities for prevention. Can Med Assoc J 1974;111:1105–7.
4. Bailar JC 3rd. Mammography: a contrary view. Ann Intern Med 1976;84:77–84.
5. Baker LH. Breast Cancer Detection Demonstration Project: five-year summary report. CA Cancer J Clin 1982;32:194–225.
6. Tabar L, Fagerberg CJ, Gad A, Baldetorp L, Holmberg LH, Grontoft O, et al. Reduction in mortality from breast cancer after mass screening with mammography. Randomised trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. Lancet 1985;1:829–32.
7. Wald N, Chamberlain J, Hackshaw A. Report of the European Society of Mastology Breast Cancer Screening Evaluation Committee. Tumori 1993;79:371–9.
8. Breast-cancer screening with mammography in women aged 40–49 years. Swedish Cancer Society and the Swedish National Board of Health and Welfare. Int J Cancer 1996;68:693–9.
9. Fletcher SW, Black W, Harris R, Rimer BK, Shapiro S. Report of the international workshop on screening for breast cancer. J Natl Cancer Inst 1993;85:1644–56.

10. National Institutes of Health Consensus Development Panel. National Institutes of Health Consensus Development Conference Statement: Breast Cancer Screening for Women Ages 40–49, January 21–23, 1997. National Institutes of Health Consensus Development Panel. J Natl Cancer Inst 1997;89:1015–26.
11. Olsen O, Gotzsche PC. Screening for breast cancer with mammography (Cochrane Review). Cochrane Database Syst Rev 2001;4: CD001877.
12. Leitch AM, Dodd GD, Costanza M, Linver M, Pressman P, McGinnis L, et al. American Cancer Society guidelines for the early detection of breast cancer: update 1997. CA Cancer J Clin 1997;47:150–3.
13. U.S. Preventive Services Task Force. Screening for breast cancer: recommendations and rationale. Ann Intern Med 2002;137:344–6.
14. Schousboe JT, Kerlikowske K, Loh A, Cummings SR. Personalizing mammography by breast density and other risk factors for breast cancer: analysis of health benefits and cost-effectiveness. Ann Intern Med 2011;155:10–20.
15. Mandelblatt JS, Stout N, Trentham-Dietz A. To screen or not to screen women in their 40s for breast cancer: is personalized risk-based screening the answer? Ann Intern Med 2011;155:58–60.
16. Tabar L, Chen HH, Fagerberg G, Duffy SW, Smith TC. Recent results from the Swedish Two-County Trial: the effects of age, histologic type, and mode of detection on the efficacy of breast cancer screening. J Natl Cancer Inst Monogr 1997;22:43–7.
17. Hendrick RE, Smith RA, Rutledge JH 3rd, Smart CR. Benefit of screening mammography in women aged 40–49: a new meta-analysis of randomized controlled trials. J Natl Cancer Inst Monogr 1997;22:87–92.
18. Andersson I, Janzon L. Reduced breast cancer mortality in women under age 50: updated results from the Malmo Mammographic Screening Program. J Natl Cancer Inst Monogr 1997;22:63–7.
19. Bjurstam N, Bjorneld L, Duffy SW, Smith TC, Cahlin E, Eriksson O, et al. The Gothenburg breast screening trial: first results on mortality, incidence, and mode of detection for women ages 39–49 years at randomization. Cancer 1997;80:2091–9.
20. Salzmann P, Kerlikowske K, Phillips K. Cost-effectiveness of extending screening mammography guidelines to include women 40 to 49 years of age. Ann Intern Med 1997;127:955–65. Erratum in Ann Intern Med 1998;128:878
21. Berry DA. Benefits and risks of screening mammography for women in their forties: a statistical appraisal. J Natl Cancer Inst 1998;90:1431–9.
22. US Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2009;151:716–26.
23. Mandelblatt JS, Cronin KA, Bailey S, Berry DA, de Koning HJ, Draisma G, et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. Ann Intern Med 2009;151:738–47.
24. Harris R, Sawaya GF, Moyer VA, Calonge N. Reconsidering the criteria for evaluating proposed screening programs: reflections from 4 current and former members of the U.S. Preventive services task force. Epidemiol Rev 2011;33:20–35.
25. Broeders M, Moss S, Nystrom L, Njor S, Jonsson H, Paap E, et al. The impact of mammographic screening on breast cancer mortality in Europe: a review of observational studies. J Med Screen 2012;19 Suppl 1:14–25.
26. Gotzsche PC, Jorgensen KJ. Screening for breast cancer with mammography. Cochrane Database Syst Rev 2013;6:CD001877.
27. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. Lancet 2012;380:1778–86.
28. Duffy SW, Chen TH, Smith RA, Yen AM, Tabar L. Real and artificial controversies in breast cancer screening: a perspective article. Breast Cancer Manage 2013;2:519–28.
29. Puliti D, Duffey SW, Miccinesi G, de Koning H, Lynge E, Zappa M, et al. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. J Med Screen 2012;19:42–56.
30. 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 mortality database. BMJ 2011;343:d4411.
31. Esserman L, Shieh Y, Thompson I. Rethinking screening for breast cancer and prostate cancer. JAMA 2009;302:1685–92.
32. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. N Engl J Med 2012;367: 1998–2005.
33. Zahl PH, Gotzsche PC, Maehlen J. Natural history of breast cancers detected in the Swedish mammography screening programme: a cohort study. Lancet Oncol 2011;12:1118–24.
34. Baum M. Should routine screening by mammography be replaced by a more selective service of risk assessment/risk management? Womens Health (Lond Engl) 2010;6:71–6.
35. Welch HG. Screening mammography–a long run for a short slide? N Engl J Med 2010;363:1276–8.
36. IARC Working Group on the Evaluation of Cancer-Preventive Strategies. Breast cancer screening. Lyon, France: IARC Press; 2002.
37. Smith RA, Saslow D, Sawyer KA, Burke W, Costanza ME, Evans WP III, et al. American Cancer Society guidelines for breast cancer screening: update 2003. CA Cancer J Clin 2003;53:141–69.
38. American College of Obstetricians-Gynecologists. Practice bulletin no. 122: breast cancer screening. Obstet Gynecol 2011;118:372–82.
39. Tabar L, Vitak B, Tony HH, Yen MF, Duffy SW, Smith RA. Beyond randomized controlled trials: organized mammographic screening substantially reduces breast carcinoma mortality. Cancer 2001;91: 1724–31.
40. Paci E, Duffy SW, Giorgi D, Zappa M, Crocetti E, Vezzosi V, et al. Quantification of the effect of mammographic screening on fatal breast cancers: The Florence Programme 1990–96. Br J Cancer 2002;87:65–9.
41. Olsen AH, Njor SH, Vejborg I, Schwartz W, Dalgaard P, Jensen MB, et al. Breast cancer mortality in Copenhagen after introduction of mammography screening: cohort study. BMJ 2005;330:220.
42. Coldman A, Phillips N, Warren L, Kan L. Breast cancer mortality after screening mammography in British Columbia women. Int J Cancer 2007;120:1076–80.
43. Duffy SW, Tabar L, Olsen AH, Vitak B, Allgood PC, Chen TH, et al. Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. J Med Screen 2010;17:25–30.
44. Otto SJ, Fracheboud J, Looman CW, Broeders MJ, Boer R, Hendriks JH, et al. Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. Lancet 2003;361:1411–7.
45. Euroscreen Working Group. Summary of the evidence of breast cancer service screening outcomes in Europe and first estimate of the benefit and harm balance sheet. J Med Screen 2012;19 Suppl 1:5–13.
46. Paci E, Broeders M, Hofvind S, Puliti D, Duffey S, Euroscreen Working Group. European breast cancer service screening outcomes: a first balance sheet of the benefits and harms. Cancer Epidemiol Biomarkers Prev 2014;23:1159–63.
47. Olsen O, Gotzsche PC. Cochrane review on screening for breast cancer with mammography. Lancet 2001;358:1340–2.
48. de Koning HJ. Assessment of nationwide cancer-screening programmes. Lancet 2000;355:80–1.
49. Health Council of the Netherlands. The benefit of population screening for breast cancer with mammography. Hague, the Netherlands: Health Council of the Netherlands; 2002. Report No.: 2002/03E.
50. Freedman DA, Petitti DB, Robins JM. On the efficacy of screening for breast cancer. Int J Epidemiol 2004;33:43–55.
51. Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 2002;137:347–60.
52. Nelson HD, Fu R, Griffin JC, Nygren P, Smith ME, Humphrey L. Systematic review: comparative effectiveness of medications to reduce risk for primary breast cancer. Ann Intern Med 2009;151: 703–15, W-226–35.
53. Bewley S. The NHS breast screening programme needs independent review. BMJ 2011;343:d6894.

54. Hackshaw A. The benefits and harms of mammographic screening for breast cancer: building the evidence base using service screening programmes. J Med Screen 2012;19 Suppl 1:1–2.

55. Zappa M, Federici A. The potential for systematic early detection and treatment of breast cancer to reduce the burden of disease is widely recognized in the European Union (EU). Introduction. J Med Screen 2012;19 Suppl 1:3–4.

56. Moss SM, Nystrom L, Jonsson H, Paci E, Lynge E, Njor S, et al. The impact of mammographic screening on breast cancer mortality in Europe: a review of trend studies. J Med Screen 2012;19 Suppl 1:26–32.

57. Njor S, Nystrom L, Moss S, Paci E, Broeders M, Segnan N, et al. Breast cancer mortality in mammographic screening in Europe: a review of incidence-based mortality studies. J Med Screen 2012; 19:33–41.

58. Hofvind S, Ponti A, Patnick J, Ascunce N, Njor S, Broeders M, et al. False-positive results in mammographic screening for breast cancer in Europe: a literature review and survey of service screening pro-grammes. J Med Screen 2012;19 Suppl 1:57–66.

59. Puliti D, Miccinesi G, Zappa M, Manneschi G, Crocetti E, Paci E. Balancing harms and benefits of service mammography screening programs: a cohort study. Breast Cancer Res 2012;14:R9.

60. Nelson HD, Tyne K, Naik A, Bougatsos C, Chan BK, Humphrey L. Screening for breast cancer: an update for the U.S. Preventive Services Task Force. Ann Intern Med 2009;151:727–37, W237-42.

61. Tabar L, Vitak B, Chen TH, Yen AM, Cohen A, Tot T, et al. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. Radiology 2011;260:658–63.

62. Smith RA, Duffy SW, Gabe R, Tabar L, Yen AM, Chen TH. The randomized trials of breast cancer screening: what have we learned? Radiol Clin North Am 2004;42:793–806, v.

63. Swedish Organized Service Screening Evaluation Group. Effect of mammographic service screening on stage at presentation of breast cancers in Sweden. Cancer 2007;109:2205–12.

64. Duffy SW, Tabar L, Chen HH, Holmqvist M, Yen MF, Abdsalah S, et al. The impact of organized mammography service screening on breast carcinoma mortality in seven Swedish counties. Cancer 2002;95:458–69.

The Value of Modern Mammography Screening in the Control of Breast Cancer: Understanding the Underpinnings of the Current Debates

Robert A. Smith

Updated version	Access the most recent version of this article at: http://cebp.aacrjournals.org/content/23/7/1139

Cited articles	This article cites 62 articles, 4 of which you can access for free at: http://cebp.aacrjournals.org/content/23/7/1139.full#ref-list-1
Citing articles	This article has been cited by 2 HighWire-hosted articles. Access the articles at: http://cebp.aacrjournals.org/content/23/7/1139.full#related-urls

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions	To request permission to re-use all or part of this article, use this link http://cebp.aacrjournals.org/content/23/7/1139. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.