DOI:10.1093/jnci/djs289
Advance Access publication on July 17, 2012.

Screening Mammography: What Good Is It and How Can We Know If It Works?

Michael W. Vannier

Correspondence to: Michael W. Vannier, MD, Department of Radiology, University of Chicago Medical Center, 5841 S. Maryland Ave, MC 2026, Chicago, IL 60637 (e-mail: mvannier@radiology.bsd.uchicago.edu).

Nobody said it would be easy. Cancer in general and breast cancer specifically are targets of the War on Cancer that began in the 1970s with substantial investment in research and fundamental changes in clinical practice (1). The prevailing wisdom has been that if you find cancer early, you can save lives and the outcome will be better. Many advances in technology and therapy for breast cancer emerged and were adopted, so the tools available today have replaced the methods adopted in the 1970s. Breast cancer remains a major public health problem; there is heightened awareness and interest in the disease and strong demand for measures to control or eradicate it.

The advocacy for screening to detect breast cancer early is well organized, resourceful, and highly motivated, with support from all sectors of society. The advocates play a central role in adoption and expansion of screening programs, the growth of research, and advances in therapy. Mammography was among the first imaging procedures used to detect cancers early, followed by other imaging techniques for the lung, colon, prostate, and cervix.

But how well does screening work? And did it deliver on its promise to lengthen life? Conventional wisdom has been to use mortality as the end point for screening-program evaluation, despite the fact that diagnosis, staging, treatment, and retreatment for recurrence take place before the end of life (2).

How much can screening be expected to prolong life given all of the other events that intercede between disease detection and death? And when every step in this process changes year by year with new methods, technologies, and drugs, how can we separate the effect of screening? Should we use total mortality rather than cancer-specific mortality to judge cancer screening programs? The answer to this controversial question is divided between some researchers who believe that all-cause mortality is a more reliable measure of the effectiveness of screening (3) and others who think that it is too stringent (4).

A decade ago, JNCI published a report on all-cause mortality in randomized trials of cancer screening (5) that was accompanied by an editorial (6) that observed that screening trials are even more difficult than we thought. Black et al. (5) concluded that the benefits of screening can be overestimated using all-cause mortality records. We now recognize that it is difficult to estimate a screening benefit using mortality reduction and that distortions may arise because of the natural history of the disease, the frequency of screening, and the duration of follow-up, all of which contribute to the time patterns in the mortality reductions observed in trials. So, without appropriate analyses, results from cancer screening trials will be distorted (7).

Recently, epidemiologic studies have been completed on national screening programs, which have revealed that mortality reduction due to breast cancer screening is questionable. Previous reports on the national screening programs in Norway (8) and the Netherlands (9) were complemented by a recent study on the program in Sweden (10). The Norwegian report was accompanied by a pessimistic editorial observing that screening mammography may be a long run for a short slide (11). The Dutch report was very positive regarding the mortality benefit of breast screening, but like all of these studies, it has been controversial (12).

Two new reports on mammography screening and all-cause county-specific mortality in Sweden (13,14) show negative results for a program that was started in 1974, with nationwide implementation by 1997. But again, do the mortality records measure what we want or need to know regarding screening?

Mortality from breast cancer results from the limitations of treatment to a greater extent than from the shortcomings of screening. Because of the widespread availability of mortality records, it has been common practice to use them in estimating the value of medical procedures. Despite the remoteness of screening from death, the mortality difference in screened and unscreened populations has been used as the principal metric for public health benefit from a screening program. However, even if screening were 100% successful, if diagnosis and treatment were delayed or ineffective, the number of deaths might not be affected.

No, it is not simple to understand what benefits screening provides, and we should regard any generalizations with skepticism. If mortality reduction is not the principal benefit of screening, then what is? Certainly our patients are not comfortable with uncertainty regarding a cancer diagnosis—and, if cancer is found, they want it to be treated. Limitations in diagnosis and treatment are present for all cancers, so they are understood and accepted by patients. Screening and overtreatment are both fueled by conditioning of the public and the physicians who care for them to detect cancer early and treat it aggressively, presumably providing real benefit. Breast cancer remains a heterogeneous set of diseases that, when treated as a single entity, behaves less predictably. Our assessments are clouded by heterogeneity in disease effects and natural history.

We now know that isolating screening as an evaluable entity using death records fails to reveal major benefits. Whether this

result can be reproduced in the future is unclear, but as treatment improves and more importantly, as we gain the ability to discriminate risk of progression from biomarkers linked to targeted therapies, the picture can change.

So, what can we do with this information? One approach would be to discourage patients from participating in the screening regimen (15). However, in the absence of a better alternative and with inertia behind any national program, screening mammography is likely to continue. Recognizing that the cost and morbidity of this imperfect solution to breast cancer have raised the stakes when measuring the benefits of early screening has proven elusive, we have few options, and not all of them are practical. We could target screening to subpopulations according to risk, perhaps with the help of new biomarkers that improve specificity. Better diagnostic tools used to evaluate breast cancer candidates found on screening would compensate for the limitations of population-based imaging. Eventually, through better knowledge of breast cancer etiology and biology, we can address the concerns regarding overdiagnosis and overtreatment and see them minimized. As our tools improve, we can begin to fully realize the promise of breast cancer screening to arrest this dread disease at its earliest stage with the least morbidity and cost.

Nobody said it would be easy.

References

1. Dumaine B. Rethinking the war on cancer. *Fortune*. 2012;165(3):14, 16.
2. Peckham CS, Dezateux C. Issues underlying the evaluation of screening programmes. *Br Med Bull.* 1998;54(4):767–778.
3. Penston J. Should we use total mortality rather than cancer specific mortality to judge cancer screening programmes? Yes. *BMJ*. 2011;343:d6395.
4. Steele R, Brewster D. Should we use total mortality rather than cancer specific mortality to judge cancer screening programmes? No. *BMJ*. 2011;343:d6397.
5. Black WC, Haggstrom DA, Welch HG. All-cause mortality in randomized trials of cancer screening. *J Natl Cancer Inst*. 2002;94(3):167–173.
6. Juffs HG, Tannock IF. Screening trials are even more difficult than we thought they were. *J Natl Cancer Inst*. 2002;94(3):156–157.
7. Hanley JA. Measuring mortality reductions in cancer screening trials. *Epidemiol Rev*. 2011;33(1):36–45.
8. Kalager M, Zelen M, Langmark F, *Adami* H-O. Effect of screening mammography on breast-cancer mortality in Norway. *N Engl J Med.* 2010;363(13):1203–1210.
9. van Schoor G, Moss SM, Otten JD, et al. Increasingly strong reduction in breast cancer mortality due to screening. *Br J Cancer*. 2011;104(6):910–914.
10. Tabár L, Vitak B, Chen TH, et al. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology*. 2011;260(3):658–663.
11. Welch HG. Screening mammography—a long run for a short slide? *N Engl J Med.* 2010;363(13):1276–1278.
12. Jørgensen KJ. Flawed methods explain the effect of mammography screening in Nijmegen. *Br J Cancer*. 2011;105(4):592–593.
13. Autier P, Koechlin A, Smans M, Vatten L, Boniol M. Mammography screening and breast cancer mortality in Sweden. *J Natl Cancer Inst.* 2012;104(14):xxx.
14. Haukka J, Byrnes G, Boniol M, Autier P. Trends in breast cancer mortality in Sweden before and after implementation of mammography screening [published online ahead of print September 26, 2011]. *PLoS One*. 2011;6(9):e22422.
15. Nordic Cochrane Centre. *Screening for Breast Cancer with Mammography*. Copenhagen, Denmark: Nordic Cochrane Centre; 2012. http://www.screening.dk.

Affiliation of author: Department of Radiology, University of Chicago Medical Center, Chicago, IL.

DOI:10.1093/jnci/djs290
Advance Access publication on July 17, 2012.

Is the Breast Cancer Mortality Decrease in Sweden Due to Screening or Treatment? Not the Right Question

Nereo Segnan, Stefano Rosso, Antonio Ponti

Correspondence to: Nereo Segnan, MD, MSc Epi, CPO Piemonte, Unit of Cancer Epidemiology, ASO S Giovanni Battista University Hospital, Via S Francesco da Paola 31, Turin 10123, Italy (e-mail: nereo.segnan@cpo.it).

Autier et al. (1) attempt, in this issue of the Journal, to study the time trends of breast cancer mortality by county in Sweden from 1960 to 2009 by dividing the 21 counties into four groups depending on the year in which organized breast cancer screening started. The authors find a continuous decrease in breast cancer mortality during the period studied but claim that this decrease was not related to the introduction of screening. Their conclusion is based on a comparison between observed mortality trends and expected trends that they modeled assuming different levels of screening effectiveness. No effect on mortality was appreciated in two groups of Swedish counties, whereas the mortality trends in two other groups of Swedish counties declined by 5% and 8% more steeply than the mortality trends observed before screening started.

The analysis of time trends of breast cancer mortality rates following the introduction of screening is definitely not the most reliable method to assess its effectiveness. There are several limitations that produce a diluting effect, such as the inclusion of deaths from cancers diagnosed before screening started or before women reached screening age, the phased build-up of screening, or the presence of opportunistic screening that took place before organized screening started. Furthermore, it is well known that ecological studies suffer from important shortcomings (2); it is extremely difficult, if not impossible, to know and properly account for the multiple, diverse, and intertwined reasons for the observed trends, especially if time trends are compared across countries or regions. It is therefore paradoxical that descriptive analyses of this