EDITORIALS

Screening Mammography — A Long Run for a Short Slide?

H. Gilbert Welch, M.D., M.P.H.

No screening test has ever been more carefully studied than screening mammography. In the past 50 years, more than 600,000 women have participated in 10 randomized trials, each involving approximately 10 years of follow-up. Given this extraordinary research effort, it is ironic that screening mammography continues to be one of the most contentious issues within the medical community.

The juxtaposition of such a charged medical debate in the face of such an exhaustive scientific investigation is in itself instructive. For context, one trial involving fewer than 150 men who were followed for less than 2 years was sufficient to convince physicians of the value of treating severe hypertension.[1] That physicians are still debating the relative merits of screening mammography despite the wealth of data suggests that the test is surely a close call, a delicate balance between modest benefit and modest harm.

In this issue of the *Journal*, Kalager et al.[2] provide additional data that the benefit of mammography is modest. Making use of the opportunity provided by the staggered implementation of a national screening program in Norway, the investigators were able to isolate the benefit of the screening program from other factors that may have changed over time, including increased breast-cancer awareness and improvements in treatment. They report that the benefit of the Norwegian screening program was disappointingly small: a 10% reduction in breast-cancer mortality among women between the ages of 50 and 69 years.

Moreover, this reduction in mortality reflected the combined effect of the two interventions that make up the Norwegian screening program: screening mammography and multidisciplinary teams instituted to better treat breast cancer. Kalager et al. provide data that the latter may be the more important of the two factors, since women over the age of 70 years, who were exposed to the program's multidisciplinary teams but were not invited to undergo mammography, had an 8% reduction in breast-cancer mortality. Thus, the relative reduction in mortality due to screening mammography alone could be as low as 2%.

Clinicians who follow the mammography debate will reasonably wonder why the benefit estimated by Kalager et al. is so much smaller than the reduction in mortality of 15 to 23% estimated by the U.S. Preventive Services Task Force.[3] The easiest explanation would be that the Kalager estimate is wrong. Although the task force uses data from randomized trials, the Norwegian data are observational — and as with all observational data, the primary threat to validity is the comparability of the comparison groups.

But the staggered cohort design that was used by Kalager et al. mitigates the concern that the women in the four study groups are somehow different, since many of the women in the study actually contributed data to each group at different points in their life. Contamination is a more relevant concern. If the women in the nonscreening groups were exposed to opportunistic mammography screening or began to benefit from the multidisciplinary teams, which had to be in place before the screening program was initiated, then the background effect of time may have been overestimated. This would have led to an underestimation of the benefit of the screening program. Furthermore, the follow-up period may be too short to fully capture the benefits of screening. The authors argue that these effects are small.

So another explanation must be considered:

the estimates of both the task force and Kalager et al. are correct. But where the randomized trials reflect the world before 1990, the observational data reflect the world after 1990. It is quite plausible that screening mammography was more effective in the past than it is now. If women with new breast lumps now present earlier for evaluation, the benefit of screening will be less. If treatment of clinically detected breast cancer (i.e., tumors that are detected by means other than screening) has now improved, the benefit of screening will be less. Thus, the increased awareness about the importance of promptly seeking care for overt breast abnormalities (there is no debate about diagnostic mammography) and the widespread use of adjuvant therapy have probably combined to make screening now less important.[4,5]

Nevertheless, the public widely perceives screening mammography to be one of the most important services provided by modern medicine. The perception is largely the product of well-crafted public health messaging, such as the American Cancer Society's print campaign in the 1980s that featured the headline "If you haven't had a mammogram, you need more than your breasts examined." Given current data, such messaging must become more balanced.

If we assume that mammography screening is associated with a 10% reduction in the rate of death from breast cancer (making the optimistic assumption that all the benefit comes from screening mammograms), the 10-year risk of breast-cancer death for a 50-year-old woman in the United States is now about 4 per 1000 women.[6] If we assume that this risk already incorporates the benefit of screening mammography, the risk estimate without mammography would be about 4.4 per 1000 women.

Because we are all subject to framing effects, it is important to consider the reverse frame. The number of women who will not die from breast cancer rises from 995.6 to 996 per 1000 women with the addition of screening mammography. Although readers may each respond differently to these frames, both reflect the same absolute benefit: 0.4 per 1000 women. In other words, 2500 women would need to be screened over a 10-year period for 1 to avoid death from breast cancer (Table 1).

What happens to the other 2499 women who had to undergo screening to achieve this benefit is also relevant. Estimates of harm vary considerably. In the United States, more than 1000 women would be expected to have at least one false positive result,[7] a number that would be considerably lower in Europe.[8] Less frequent but more worrisome is the problem of overdiagnosis. Somewhere between 5 and 15 women would be expected to be needlessly treated for a condition that was never going to bother them, with all the accompanying harms.[9,10]

Table 1. Estimated Benefits and Harms Associated with a 10-Year Course of Screening Mammography for 2500 Women Who Are 50 Years of Age.*

Benefit	Harm
One woman will avoid dying from breast cancer.	Up to 1000 women will have at least one "false alarm," about half of whom will undergo biopsy.
	Breast cancer will be overdiagnosed in 5 to 15 women, who will be treated needlessly with surgery, radiation, chemotherapy, or a combination.

* The assumed benefit of screening mammography is a reduction of 10% in the rate of death from breast cancer, as reported by Kalagar et al.[2]

Screening mammography has become one of the most prominent measures of health care performance. Since the inception of health care report cards, such evaluations have focused on ensuring that all women undergo the test.[11] There were practical reasons for this: it was easily measured, easy to understand, and hard to argue against. But by highlighting that the mortality benefit is modest, Kalager et al. help confirm that the decision about whether to undergo screening mammography is, in fact, a close call. Many observers will argue that because it is a delicate decision — involving trade-offs among noncomparable outcomes — it must be left to informed individuals to decide. Others will argue that physicians should continue to persuade women to undergo screening and that the modest benefit is worth the associated harms.

But no one can argue that screening mammography is one of the most important services we provide in medicine. The time has come for it to stop being used as an indicator of the quality of our health care system.

Disclosure forms provided by the author are available with the full text of this article at NEJM.org.

From the Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth Medical School, Lebanon, NH.

1. Effects of treatment on morbidity in hypertension: results in patients with diastolic blood pressures averaging 115 through 129 mm Hg. JAMA 1967;202:1028-34.
2. Kalager M, Zelen M, Langmark F, Adami H-O. Effect of

screening mammography on breast-cancer mortality in Norway. N Engl J Med 2010;363:1203-10.
3. Mandelblatt JS, Cronin KA, Bailey S, et al. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. Ann Intern Med 2009;151:738-47.
4. Rostgaard K, Vaeth M, Rootzén H, Lynge E. Why did the breast cancer lymph node status distribution improve in Denmark in the pre-mammography screening period of 1978-1994? Acta Oncol 2010;49:313-21.
5. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005;365:1687-717.
6. The risk of developing/dying from cancer based on November 2009 SEER data submission. Bethesda, MD: National Cancer Institute. (Available at http://seer.cancer.gov/faststats/.)
7. Elmore JG, Barton MB, Moceri VM, Polk S, Arena PJ, Fletcher SW. Ten-year risk of false positive screening mammograms and clinical breast examinations. N Engl J Med 1998;338:1089-96.
8. Breast cancer screening. In: IARC handbooks of cancer prevention. Vol. 7. Lyon, France: International Agency for Research on Cancer, 2002:135-44.
9. Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP. Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. BMJ 2006;332:689-92.
10. Jørgensen KJ, Gøtzsche PC. Content of invitations for publicly funded screening mammography. BMJ 2006;332:538-41.
11. McIntyre D, Rogers L, Heier EJ. Overview, history, and objectives of performance measurement. Health Care Financ Rev 2001;22(3):7-21.

Superficial Phlebitis and Phase 3.5 Trials

Lee Goldman, M.D., and Jeffrey Ginsberg, M.D.

In this issue of the *Journal*, Decousus et al.[1] report on the efficacy and safety of fondaparinux for the treatment of superficial-vein thrombosis in the legs. The results of their carefully conducted, placebo-controlled trial show that treatment with fondaparinux, at a dose of 2.5 mg once daily for 45 days, as compared with placebo, reduced the probability that superficial-vein thrombosis in the legs would progress to deep-vein thrombosis or pulmonary embolism (1.3% with placebo vs. 0.2% with fondaparinux), without an increase in bleeding or other serious adverse events. The probability that patients would undergo surgery for superficial-vein thrombosis was reduced from 3.8% to 0.7%. Two patients in the fondaparinux group and one in the placebo group died, but none of the deaths were apparently the result of a pulmonary embolism. This study adds to previous work describing the natural history of superficial-vein thrombosis,[2-5] although it did not address which patients might be at an increased risk because of previously undiagnosed thrombophilia.[4,5]

To put the rates of deep-vein thrombosis and pulmonary embolism — the most important outcomes — into perspective, it is useful to consider the generally "acceptable" failure rates in strategies to diagnose venous thromboembolism. In the study by Decousus et al., the rate at which symptomatic deep-vein thrombosis or pulmonary embolism developed in untreated patients during follow-up (1.3%) was similar to the rate with widely accepted strategies for diagnosing deep-vein thrombosis and pulmonary embolism. For example, among patients who are evaluated for suspected deep-vein thrombosis but have normal results on a contrast venogram[6] or duplex ultrasonography,[7] about 1.3% and 0.6% of patients, respectively, will return with symptomatic deep-vein thrombosis or pulmonary embolism over the course of long-term follow-up. Similarly, among patients who have a suspected pulmonary embolism but then have normal results on a conventional pulmonary angiogram[8] or a computed tomographic pulmonary angiogram,[9] about 1.7% and 1.2%, respectively, will return with symptomatic deep-vein thrombosis or pulmonary embolism. These historical comparisons and the extremely low mortality among untreated patients with superficial-vein thrombosis support an initial "no anticoagulant treatment" approach, unless conservative measures fail to resolve symptoms or deep-vein thrombosis develops. It is also clear from the stringent inclusion and exclusion criteria in the study by Decousus et al. that treatment with fondaparinux for 45 days is clinically reasonable for patients with severe symptoms, thrombosis in the proximal saphenous vein, or recurrent disease.

Agents such as fondaparinux, low-molecular-weight heparins, and perhaps oral direct factor Xa inhibitors (apixaban, rivaroxaban) and thrombin inhibitors (dabigatran) have better risk profiles than do unfractionated heparin and warfarin, and the favorable risk-to-benefit ratio associated with them could lead to an extension