Reflections on Breast Self-examination

Cornelia J. Baines*

The potential for breast self-examination (BSE) to aid women in the control of breast cancer has been recognized for more than 60 years (1). Nevertheless, in the mid-1990s, the efficacy of BSE has yet to be established firmly. Thus, the implementation of a large, randomized, controlled trial of BSE in Shanghai by Thomas et al. (2), the design of which is described in meticulous detail in this issue of the Journal, is to be applauded.

Evidence on BSE began to appear in the late 1970s and was derived largely from correlational studies seeking an association between reported BSE practice and the diagnosis of early stage breast cancer (3,4). These studies and subsequent ones [reviewed in (5)] were followed by enthusiastic endorsements by women and by the American Cancer Society (6). Thus, it was disappointing when the U.K. trial of breast screening, using a nonrandomized design comparing communities, yielded contradictory results on BSE (7). An increased relative risk for breast cancer mortality was reported for one community, Nottingham (1.13; 95% confidence interval [CI] = 0.95-1.35), and a decreased risk was reported for another, Huddersfield (0.78; 95% CI = 0.61-1.00). More encouraging to BSE proponents were the findings from a nested case-control study in the same trial (8). Attendance at BSE classes was associated with a decreased odds ratio for death from breast cancer: 0.70 (95% CI = 0.50-0.97). Furthermore, substantial encouragement could be derived from a meta-analysis of BSE by Hill et al. (9) that led to the conclusion that there were good grounds to encourage women to practice BSE regularly.

In a Finnish cohort study (10), breast cancer mortality among 29,000 women who participated in a BSE program was compared with the expected mortality in the Finnish population. The observed-to-expected ratio was 0.71 (95% CI = 0.57-0.87). Of great interest was the fact that decreased mortality from breast cancer occurred in all age groups, with those 50 years of age and younger showing greater benefit than those older than 50. Furthermore, half of the deficit in breast cancer deaths occurred within 3-6 years of enrollment, in stark contrast to what has been observed with mammography screening in the same age group. Such results raised an obvious question: How could BSE reduce breast cancer mortality in young women when the benefit from mammography screening has been so difficult to establish (11)? Whenever an outcome in a cohort of volunteers is compared with the outcome for the general population, someone will inevitably dispute the validity of the conclusions regardless of the analytic methodology employed. Such skepticism is bolstered by the observed all-cause mortality in the Finnish cohort, which, being less than expected, suggests a "healthy volunteer effect."

In addition, there is the indisputable possibility that the Finnish population may differ in many ways from the North American population.

Newcomb et al. (12) provided another slant on BSE. They found that women who reported more thorough BSE had a significantly decreased occurrence of advanced breast cancer than women who performed BSE less competently or not at all. These observations are consistent with the results of a nested case-control study of BSE in the Canadian National Breast Screening Study (Harvey BJ, Miller AB, Baines CJ, Corey PN: manuscript submitted for publication). However, BSE skeptics can continue to point out the incongruity of BSE apparently competing with mammography and also emphasize the logistic difficulties in achieving widespread competence in BSE.

Finally, the gold standard for evaluation of screening remains the randomized, controlled trial. Other than the Shanghai trial, only one other trial is known to me. The World Health Organization trial of BSE being conducted in Russia (13) has yet to publish overall results, although early and preliminary results for Leningrad alone were not encouraging. Now we have early results from the Shanghai trial. More importantly, we have a detailed description of the study protocol—the problems that have been encountered and the complexity of the issues that arise when a trial is conducted with human beings, with wills of their own, and not laboratory animals, taking place in cities and not clinics, and involving large, not small, populations.

More than a quarter of a million women 31-64 years of age were randomly assigned in clusters on the basis of 520 work sites (major textile factories) to receive either intensive BSE instruction (BSE group) or sessions on the prevention of low back pain (control group). A major advantage of the trial locale was that screening mammography was not available. Compliance was good in the BSE group, and demonstrated BSE proficiency was higher in the BSE group than in the control group. A less happy finding is that breast cancer detection has been equal in both groups (331 in the BSE group versus 322 in the control group). One expects that, in any comparison of breast screening with no breast screening, more breast cancers will be detected in the screened group. Furthermore, there has been no substantial stage shift in the screened group, although differences in tumor mortality were not statistically significant (14).

*Correspondence to: Cornelia J. Baines, M.D., Department of Preventive Medicine and Biostatistics, University of Toronto, 12 Queen’s Park Crescent W., Toronto, ON, Canada, M5R 1G6.

Journal of the National Cancer Institute, Vol. 89, No. 5, March 5, 1997
size and stage were both “in the expected direction only for incident cases in women aged 50 years or older, which is the same group for whom mammography has been shown to be efficacious.”

It seems that the main consequence of screening has been to more than double the number of false-positive detections, if indeed it is fair to designate the detection of benign breast lesions as a false-positive phenomenon. Perhaps the increased awareness that such a finding indicates may yet yield some advantage to those performing BSE.

Unpredicted problems always arise during the course of randomized, controlled trials. In Shanghai, economic reform has disrupted the expected stability of the factory populations, resulting in some mingling of BSE and control groups and rendering follow-up more difficult. Also, dealing with language and cultural differences, as well as with geographic distances, has been no mean feat.

Chalmers (14) has acutely and wittily described the perils that follow when reports of randomized, controlled trials that have yielded unpopular results are published. The fact that no BSE effect has yet been observed in the Shanghai trial may encourage BSE skeptics and discourage BSE proponents, but neither should leap to conclusions, given the slow-paced manifestation of the benefits from mammography screening (11). If, ultimately, the Shanghai trial demonstrates a reduction in breast cancer mortality, the BSE skeptics may argue that, because the setting in China is so vastly different from that in North America, the results are irrelevant. If no mortality reduction is observed, BSE advocates may regret that women in their thirties were included; they even may be tempted, with no basis in fact, to argue that the trial (the design, the implementation, the investigators, or the analysis) was flawed.

Anyone committed to finding the best available version of the truth will not be daunted by these prospects. Thomas et al. (2) deserve praise (as do the funding agencies) for addressing such an important question, for dealing with such complex issues, and for soldiering on.

References