ARTICLES

Randomized Trial of Breast Self-Examination in Shanghai: Final Results

David B. Thomas, Dao Li Gao, Roberta M. Ray, Wen Wan Wang, Charlene J. Allison, Fan Liang Chen, Peggy Porter, Yong Wei Hu, Guan Lin Zhao, Lei Da Pan, Wenjin Li, Chunyuan Wu, Zakia Coriaty, Ilonka Evans, Ming Gang Lin, Helge Stalsberg, Steven G. Self

Background: **Among women who practice breast self-examination (BSE), breast cancers may be detected when they are at an earlier stage and are smaller than in women who do not practice BSE. However, the efficacy of breast self-examination for decreasing breast cancer mortality is unproven. This study was conducted to determine whether an intensive program of BSE instruction will reduce the number of women dying of breast cancer.** ***Methods:*** **From October 1989 through October 1991, 266 064 women associated with 519 factories in Shanghai were randomly assigned to a BSE instruction group (132 979 women) or a control group (133 085 women). Initial instruction in BSE was followed by reinforcement sessions 1 and 3 years later, by BSE practice under medical supervision at least every 6 months for 5 years, and by ongoing reminders to practice BSE monthly. The women were followed through December 2000 for mortality from breast cancer. Cumulative risk ratios of dying from breast cancer were estimated using Cox proportional hazards models. All statistical tests were two-sided.** ***Results:*** **There were 135 (0.10%) breast cancer deaths in the instruction group and 131 (0.10%) in the control group. The cumulative breast cancer mortality rates through 10 to 11 years of follow-up were similar (cumulative risk ratio for women in the instruction group relative to that in the control group = 1.04, 95% confidence interval = 0.82 to 1.33; *P* = .72). However, more benign breast lesions were diagnosed in the instruction group than in the control group.** ***Conclusions:*** **Intensive instruction in BSE did not reduce mortality from breast cancer. Programs to encourage BSE in the absence of mammography would be unlikely to reduce mortality from breast cancer. Women who choose to practice BSE should be informed that its efficacy is unproven and that it may increase their chances of having a benign breast biopsy. [J Natl Cancer Inst 2002;94:1445–57]**

Whether practicing breast self-examination (BSE) ultimately reduces the number of women who die from breast cancer is unclear. Breast cancers detected while practicing breast self-examination tend to be diagnosed at an earlier stage *(1–10)* and to be smaller *(2,5,6,9,10)* than cancers diagnosed in the absence of any screening. Women who report practicing BSE tend to have their tumors diagnosed at an earlier stage than women who do not report practicing BSE *(8,10–12).* Tumor size has been inversely associated with the frequency of practicing BSE *(10,11,13–18),* and women who regularly and competently practice BSE are more likely to find their tumor themselves than women who practice BSE less diligently *(3,15,18).* Improved survival has also been associated with BSE practice in some studies *(16,19,20)* but not in others *(8,21,22).* One study *(10)* showed better survival in women who detected their tumor while practicing BSE than in unscreened women who did not detect their tumors while practicing BSE, but two others *(7,9)* did not.

Although no apparent breast cancer mortality benefit of BSE was observed in one prospective study *(23),* no information was available on the frequency or competency of BSE. Two prospective studies showed reduced breast cancer mortality in women who received detailed BSE instruction *(12,24).* However, one study *(12)* also showed a reduced risk of all-cause mortality, suggesting uncontrolled confounding, and results from the other study *(24)* may have been influenced by differences in treatment received by women in the BSE and comparison groups.

In two case–control studies *(25,26),* increasing trends in the risk of late stage or fatal breast cancer with frequency of BSE practice were observed, possibly because women in the case group reported self-examinations in response to symptoms of their disease. However, a small decrease in the risk of advanced or fatal breast cancer in women who practiced BSE with high proficiency was observed in one of these investigations *(26).* In two additional case–control studies nested within nonrandomized *(20)* and randomized *(27)* trials in which BSE practice was ascertained before any of the women developed breast cancer, one study showed a reduced risk of dying from breast cancer associated with attendance at BSE classes *(20),* and the other showed decreased trends in risk of advanced or fatal breast cancer in relation to the frequency and level of proficiency of BSE *(27).*

It has recently been suggested that BSE has been shown not to be efficacious *(28).* Although this conclusion seems unwarranted, the efficacy of BSE in reducing deaths from breast cancer is uncertain. The current position of the U.S. Preventive Health Services Task Force is that there is insufficient evidence

Affiliations of authors: D. B. Thomas, R. M. Ray, C. J. Allison, P. Porter, W. Li, C. Wu, Z. Coriaty, I. Evans, M. G. Lin, S. G. Self, Fred Hutchinson Cancer Research Center, Seattle, WA; D. L. Gao, W. W. Wang, F. L. Chen, Y. W. Hu, G. L. Zhao, L. D. Pan, Station for Prevention and Treatment of Cancer, Shanghai Textile Industry Bureau, Shanghai, China; H. Stalsberg, Institute of Medical Biology, University of Tromsø, Tromsø, Norway.

Correspondence to: David B. Thomas, M.D., Dr.P.H., Program in Epidemiology, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. North, Seattle, WA 98109-1024 (e-mail: dbthomas@fhcrc.org).

See "Notes" following "References."

to recommend for or against the teaching of BSE *(29)*. The need for randomized trials of BSE was recognized in 1983 by a World Health Organization consultation *(30)*, and in 1989 at a workshop of the International Union Against Cancer *(31)*. The first randomized trial of BSE was initiated in 1985 in Leningrad (now St. Petersburg) and Moscow. Results have been published only from the St. Petersburg portion of the trial *(32–34)*. After approximately 10 years of follow-up, nearly equal numbers of women in the BSE instruction and control groups had died from breast cancer, and the breast cancers were not diagnosed at a smaller size or at a less advanced stage in the women in the BSE instruction group. Possible reasons for these findings include the relatively low level of compliance with the BSE instructions and absence of a beneficial effect of BSE in women whose breasts were also examined clinically.

In this article, we report the final results from the only randomized trial of BSE, to our knowledge, that provides additional information on the efficacy of BSE instruction in reducing mortality from breast cancer. It was conducted in women currently or previously employed in the Shanghai textile industry.

Subjects and Methods

Prior Report

A detailed description of the methods used in this study was published with the preliminary results in 1997 *(35)*. The number of study subjects included in the analyses presented in this report differs from those in the first publication as a result of subsequent data cleanup, that is, the removal of duplicate records and ineligible study subjects. In addition, we removed data for women from one factory of unknown size that was erroneously included in the initial report as a control factory but had actually been excluded from the study before randomization. No baseline or subsequent trial activities were conducted in that factory.

Study Setting

This trial was conducted on women employed by the Shanghai Textile Industry Bureau (STIB), which included more than 520 factories when the study began in 1988. From the time the socialist system in China was established after 1949 until economic reforms began in about 1994, women entering the work force were assigned to a factory where they typically remained until they retired. They lived in modest housing units close to their place of employment and were provided shower facilities in their factory. They also received their primary medical care in clinics located in their factory. On retirement, women continued to receive their primary medical care and housing, as well as their pensions, through their factory, and Retired Workers Committees in each factory maintained regular contact with the retired women and maintained records of their vital status. Women requiring medical care beyond that obtained from the primary care facility were referred either to one of three hospitals operated by the STIB or to other hospitals having health care contracts with individual factories. No mammographic screening was available for women in the STIB. Periodic clinical breast examination had been performed by factory medical workers in some of the factories before the initiation of the trial. Although the medical workers were asked to substitute trial activities for these examinations, it is possible that medical workers in a small number of factories continued these examinations.

Participant Consent and Study Approval

Informed consent was obtained from all study participants. The study was approved by the Institutional Review Boards of the Fred Hutchinson Cancer Research Center and the Station for Prevention and Treatment of Cancer of the Shanghai Textile Industry Bureau, in accordance with an assurance filed with the Office for the Protection from Research Risks (OPPR) of the National Institutes of Health.

Recruitment and Baseline Data Collection

A team of 34 specially trained former factory medical workers, hereafter referred to as BSE workers, and approximately 5000 factory medical workers conducted the field operations. All field procedures were developed jointly by American and Chinese scientists. Study instruments and protocols were developed in English, translated into Chinese by native Shanghainese-speaking persons, pilot tested in one to seven pilot study factories, and modified as necessary before use in the trial.

After excluding the pilot study factories, the remaining 519 factories were stratified by total numbers of workers (five strata) and hospital affiliation (one stratum for each of the three STIB hospitals and one for all others). Factories in each stratum were randomly allocated to the BSE instruction or control group. All eligible women in each factory were assigned to the study arm of their factory. There were 260 factories in the instruction group and 259 in the control group.

All women born between 1925 and 1958 who were permanent residents of Shanghai and either current or retired employees of the STIB were eligible for the trial. In 1988, all eligible women were identified from factory records, and their identifying data were recorded in a notebook for each factory and entered into a computer database. Between October 1989 and October 1991, the medical workers in each factory attempted to administer a four-page optically scannable questionnaire to all eligible women in their factory. Additional eligible women not included in the database were added to the study at that time and also interviewed by the medical workers. Information was collected on the major recognized and suspected risk factors for breast cancer, use of tobacco and alcohol, contraceptive practices, prior breast cancer, previous clinical breast examinations, and previous breast self-examinations.

Nearly 290 000 women were originally identified for the trial, of whom 17 005 were subsequently considered ineligible because of changes in their status before the baseline questionnaire was administered in their factory (Table 1): 3349 had transferred out of the STIB, 3799 had moved out of Shanghai, 8293 could not be located and had presumably terminated their association with the STIB, and 1467 had died. An additional 97 women were subsequently found to be ineligible because of their date of birth.

Table 1. Numbers of recruited women in the instruction and control groups in the randomized trial of breast self-examination in Shanghai

	Instruction group	Control group	Total
Originally identified	146 437	142 955	289 392
Not eligible	9099	7906	17 005
Total eligible	137 338	135 049	272 387
No baseline questionnaire	3656	1331	4987
Total with questionnaire	133 682	133 718	267 400
Prior breast cancer	703	633	1336
Total in analyses	132 979	133 085	266 064

The numbers of women excluded for each reason were similar in the two arms of the study (data not shown). Of the 272 387 eligible women, 4987 (1.8%) were excluded from the analyses because they did not complete a baseline questionnaire. They also did not participate in any subsequent trial activities. More women in the instruction group (2.6%) than in the control group (1.0%) did not complete a baseline questionnaire for each reason recorded: 1485 in the instruction group and 227 in the control group refused, 492 and 289 were judged by the medical worker to be mentally unable to participate, 1150 and 380 were judged to be physically unable, and 529 and 435 had transferred to another STIB factory before the questionnaire was administered in the woman's original factory. Of the 1485 women in the instruction group who refused, 1177 were from a single factory and, with the exception of this factory, the numbers of women who refused were similar in the two groups (308 and 227 in the instruction and control groups, respectively). An additional 703 women in the instruction group and 633 women in the control group who gave a history of breast cancer on the baseline questionnaire were excluded from the analyses, leaving approximately 133 000 women in each arm of the trial included in the analyses.

BSE Instruction, Reinforcement, and Compliance Monitoring

At the time the baseline questionnaire was administered to women in the instruction group factories, BSE instruction was given by the BSE workers to groups of about 10 women. The BSE workers used a variety of visual aids and provided information on normal breast anatomy, breast cancer, and correct BSE technique. A three-step BSE technique was taught that included inspection in front of a mirror for evidence of asymmetry and dimpling and palpation in both standing and lying positions with the ipsilateral arm above the head. Palpation instruction emphasized using a circular motion with the pads of three middle fingers while pressing firmly, systematic coverage of the entire breast using a circular search pattern, palpation of the axilla, and squeezing the nipple to detect any discharge. The sessions also included a group discussion of perceived barriers to regular BSE practice and ended with individual instruction and practice by each woman on silicone breast models and then on themselves. These baseline activities began in October 1989 and were completed by October 1991. The baseline activities took no more than 6 months to complete in any single factory and were not conducted in all factories at the same time. Subsequent reinforcement sessions, at about 1 and 3 years after initial BSE instruction, were conducted in 1990 through 1992 and 1993 through 1995, respectively.

One year after the initial BSE instruction, the women in each BSE factory were brought together in groups of 10 to view a video developed by the study team titled "Protect Your Own Health with Your Own Hands," which emphasized the importance of BSE and reviewed proper BSE technique. After seeing the video, the women discussed the importance of BSE, using a reminder poster with the same name as the video as a point of focus. The reinforcement session concluded with each woman practicing BSE under the supervision of a BSE worker.

A second wave of reinforcement sessions began approximately 2 years after the first wave began. A video titled "BSE Right or Wrong" was shown to groups of about 10 women at a time. The video illustrated correct and incorrect BSE techniques. BSE workers showed a segment that included an incorrect BSE technique, stopped the tape, asked the women to identify the error, and conducted a discussion of the correct technique. This process was repeated until all incorrect procedures on the tape were seen and discussed. After viewing the video, the women again practiced BSE under the supervision of a BSE worker. Attendance at both reinforcement sessions was recorded on optically scannable forms.

During the first year following initial BSE instruction—at 1, 3, 6, and 9 months after the initial BSE instruction—the factory medical workers also scheduled current workers to come to the factory medical clinic to practice BSE under supervision. Retired workers were similarly scheduled for supervised BSE practice at 1 and 6 months after their initial BSE instruction. Beginning at month 12, supervised BSE practice was then scheduled every 6 months during years 2 through 5 after initial instruction. Attendance was recorded on optically scannable forms. The medical workers were instructed to observe the women practice BSE and to correct their technique, if necessary, but not to examine the women's breasts unless the women themselves reported a suspicious finding.

As detailed in our preliminary report *(35),* the medical workers in each factory were also encouraged to develop additional methods to remind women to practice BSE. These methods varied among the factories and included posters, reminders hung in shower rooms, personal contact in clinics and workshops, factory broadcasts, letters, home visits, and reminders at meetings held for other purposes or when retired workers visited their factory to receive their pension payments.

Women from factories in the control group received no information on breast cancer screening. However, concurrently with the second round of reinforcement sessions held in the BSE instruction factories, education sessions on prevention of low back pain were conducted in the control factories. These sessions included a video and discussion session. The purposes of this activity were to meet requests from control factory personnel for health information, thereby maintaining their interest and cooperation, and to provide a mechanism for conducting follow-up of women in the control group in a manner comparable to that in the instruction group.

Assessing BSE Proficiency

Random samples of 10 instruction and 10 control group factories, stratified on the number of women per factory, were selected before, immediately after, and 1 year after each of the two reinforcement sessions, with the exception that only five factories in each group were selected right after the first video. Twenty-five current and 25 retired workers (or all eligible women if fewer than 25 were associated with the factory) were randomly selected in each sampled factory. This resulted in the selection of 2404 women in the instruction group and 2463 women in the control group. If a selected woman could not be found or was not willing to participate, a replacement was recruited from lists of 20 additional women selected at random in each factory; 2374 women in the instruction group and 2426 in the control group were ultimately tested. All selected women were given 4 minutes to palpate three silicone breast models randomly chosen from a set of six models. The models, which contained varying numbers of lumps of varying sizes and consistencies at different locations and depths, were developed specifically for evaluation purposes *(36).* Women in the instruction

group were also asked about their practice of BSE and were asked to demonstrate their technique to a BSE worker who recorded the completeness of coverage and the inclusion of various components of correct BSE technique on a standardized form.

Follow-up of the Cohort

The BSE workers visited each factory every 1–2 months to inquire at the factory medical clinic, payroll office, and retirement committee office about deaths. Also, the medical workers in each factory were asked to report on standardized forms all retirements and transfers to other factories within the STIB or to jobs elsewhere. In addition, a tumor and death registry for the STIB received individual reports of deaths from the medical workers in each factory and annual summaries from each factory of all deaths during the preceding year. Attendance at the two BSE reinforcement sessions and the back pain prevention session provided additional opportunities to monitor continuation in the study. Data from all of these sources were used to update the trial database. Periodically during the trial, lists of women were generated for each factory with each woman's last known residential address, employment status, and vital status. The BSE workers checked these lists against factory records and, in some instances, other government sources. When necessary, home visits were also made to confirm and update the information on the study subjects. After 1995, when all BSE promotion activities ceased, follow-up activities were conducted in an identical manner in all factories.

Case Finding and Diagnostic Confirmation

All women who reported a suspicious breast lump through July 31, 2000, were initially evaluated by a factory medical worker. If the medical worker confirmed the presence of signs or symptoms compatible with breast cancer, the woman was referred to a surgeon for further evaluation. Medical workers were encouraged to refer women to one of three hospitals operated by the STIB where special breast clinics had been established for the trial, but women in some factories were referred to other hospitals with which the factories had contractual agreements for care of their workers. Further evaluation of the referred women by a surgeon, the decision to biopsy, the diagnosis by a pathologist, and all treatment of women found to have breast cancer were done without involvement of study personnel. The arm of the study to which a woman was assigned was not disclosed to the medical personnel at the referral hospitals, although in some instances they may have known this from their general knowledge of the trial or learned it from the referred woman.

Factory medical workers recorded the identity of each woman who had a breast biopsy or more extensive surgery, how the lump was detected, and the time intervals between initial lump discovery and first medical consultation (defined as patient delay) and between initial consultation and diagnosis (defined as system delay). This information was abstracted onto standardized forms by the BSE workers and constituted the primary means of active case finding. This method of case finding was supplemented by periodic visits by the BSE workers to the three STIB hospitals and, less frequently, to other hospitals to which women with suspected breast cancer were referred.

The STIB tumor and death registry received annual summaries from each factory of all deaths and of all cancers that developed in current and retired workers during the preceding year. These reports were reviewed manually by registry workers to supplement the active case finding procedures. The records of the Shanghai Cancer Registry *(37)* were similarly reviewed manually each year by these registry workers.

To assess completeness of case finding, data on women in the original census of those in the STIB who were not known to have breast cancer were matched by computer to the data from 1989 through 1998 of the Shanghai Cancer Registry *(37)*. This matching identified 28 previously unidentified women with breast cancer, 17 (nine in the instruction group and eight in the control group) of whom had completed a baseline questionnaire and are included in the total case count. Twenty-three of the 28 women with breast cancer were detected among the approximately 25 000 women who had severed ties with the STIB. By applying the rates of detection for the years 1989 through 1998 separately to women still associated and not associated with the STIB in 1999 and 2000, we estimate that about 15 additional women with breast cancer in the Shanghai area were missed by the case-finding process. The estimated number of missed breast cancers was similar in the instruction and control groups. The small but unknown number of additional breast cancers that were missed among the women who moved away from Shanghai is taken into account in the calculation of mortality rates by censoring.

The numbers of breast cancers diagnosed from 1989 through 1998 in the study in each relevant 5-year age group were compared with the expected numbers, calculated by applying the age-specific rates for those same years and age groups from the Shanghai Cancer Registry to the appropriate groups of women in the cohort. The resultant age-adjusted standardized incidence ratio (SIR) was 1.12 (95% confidence interval [CI] = 1.06 to 1.19), suggesting that our case finding was at least as complete as the Shanghai Cancer Registry.

The medical records of all women found to have a histologically confirmed benign or malignant breast lesion were reviewed by specially trained study personnel. The histologic diagnosis of all lesions was recorded and, for malignant cases, sufficient information was collected on tumor size, spread to regional lymph nodes, and distant metastases for classification according to the tumor–node–metastasis (TNM) scheme *(38)*. Carcinomas *in situ* (Tis) were included as breast cancers in the study analyses. Tumor size as recorded on pathology reports was used. Lymph node status was defined as follows: N0, at least one axillary lymph node was histologically examined and none had evidence of tumor involvement or, in the absence of histologic information, no clinically palpable lymph nodes were detected and the primary tumor was judged to be limited to the breast; N1, histologic evidence of tumor involvement in the axillary lymph nodes and the nodes clinically movable or, in the absence of histologic evidence of tumor involvement in the axillary lymph nodes, clinically palpable, but movable axillary lymph nodes; N2, palpable axillary lymph nodes fixed to each other or to surrounding structures on palpation; and N3, histologic evidence of spread to internal mammary lymph nodes. Information on first course of treatment was also ascertained for all women with breast cancer.

Histologic slides of sections from all lesions were collected from pathology laboratories, reviewed for quality by a local pathologist, and sent to Seattle, WA, for storage. A reference pathologist reviewed slides of sections from 1044 tumors diagnosed as breast cancer in Shanghai; only 10 were not confirmed

as breast cancer. Because absence of diagnostic confirmation may have been the result of insufficient sampling of the tissue for review, all tumors diagnosed as breast cancer by a pathologist in Shanghai were included in the study analyses.

Slides of sections from 1071 lesions diagnosed as benign breast disease in Shanghai were also reviewed. The reference pathologist who reviewed the slides diagnosed carcinoma in 17 lesions: six of 25 lesions originally diagnosed as atypical hyperplasia, eight of 601 lesions of proliferative disease without atypia, and three of 445 other benign breast conditions. Of these 17 additional lesions [three invasive ductal carcinomas, four invasive ductal carcinomas with a predominant intraductal component *(39),* and 10 ductal carcinomas *in situ*], three occurred in women with a previously diagnosed breast cancer. The other 14 women were included as additional women with breast cancer in the study analyses. By applying the above proportions of initially diagnosed benign lesions found on review to be breast cancer to the 3230 benign breast disease lesions with relevant histologic diagnoses in Shanghai that were not reviewed, we estimate that 11 invasive and 18 *in situ* carcinomas in the instruction group and nine invasive and 10 *in situ* carcinomas in the control group may have been missed.

Deaths From Breast Cancer

Because the primary end point for this trial is death from breast cancer, multiple methods were used to ensure that a high proportion of all such deaths were identified. The vital status of all women with a benign or malignant breast lesion was ascertained by matching the pathology data to the vital status data for the cohort. All women with breast cancer or benign breast disease who were not known to be dead were actively followed up, with home visits if necessary, through 2000. In addition, reports of all deaths from breast cancer were obtained from the STIB tumor and death registry and records of ongoing nested case–control studies that used the trial cohort were reviewed to identify deceased nonparticipants.

For each death identified by any of these means, a Chinese physician reviewed the relevant clinical and hospital records and interviewed family members of the deceased women, when necessary, to ascertain the cause of death. A death from breast cancer was defined as a death that would not have occurred when it did if the women had not had breast cancer. The physician was asked to make a judgment as to whether the woman's death was "very likely," "probably," "probably not," or "very likely not" because of breast cancer. The decision was recorded on a form along with a statement summarizing the evidence used in making the decision. This information was sent to Seattle, WA, translated into English by a second Chinese physician (W. Li), and reviewed jointly by her and the principal investigator (D. B. Thomas) to reach a consensus. Additional information was sought from the physician in Shanghai if needed, but this was rarely necessary.

Effects of Economic Reform

In 1994, China began a shift from a planned to a market economy that resulted in factory closures, layoffs of redundant workers, early retirements, and individuals leaving STIB to work in the private sector. Some factories in central Shanghai were moved to surrounding areas, and others were merged. To minimize the level of contamination when factories in different arms of the study merged, the factorywide activities of the larger of the two factories were continued after the merger. In all statistical analyses, however, individual women are retained in their original study arm.

Mergers of factories and transfers of individual women to different factories resulted in 5725 (4.3%) of the women in the instruction group being transferred to a control factory and 4783 (3.6%) of those in the control group being transferred to an instruction group factory within 5 years of entering the study. After 5 years, all instruction and active reinforcement activities ceased, so these percentages represent, respectively, the proportion of women in the instruction group who were not subjected to the full range of BSE activities and the proportion of women in the control group who may have received some information about BSE. After their fifth year, an additional 3156 (2.4%) women in the instruction group and 1544 (1.2%) women in the control group transferred to a factory in the opposite study arm; 2090 of the women in the instruction group worked in a single factory that merged with a control factory 8 years after the baseline activities were completed.

Blind Judgments

All judgments regarding the diagnosis and coding of breast lesions, the cause of death, and inclusions or exclusions of women from the study during data processing were made without knowledge of the study arm of the subjects under consideration.

Data Processing and Analysis

Forms with data on individual women in the cohort were optically scanned in Shanghai. Data on women who developed breast diseases were key entered into the database, primarily also in Shanghai. All data were sent to Seattle, WA, and edited, and identified errors were returned to Shanghai for correction. All data were ultimately merged into a database for analyses in Seattle.

Individuals within the same factory or group of factories may have similar risk factors or medical care that could influence their breast cancer incidence or mortality. Because the study participants were randomly assigned by factory rather than individually, this possible lack of independence must be taken into account in the statistical analyses *(40).* Variation in incidence rates among factories observed in preliminary analyses at baseline (data not shown) was completely accounted for by stratification on hospital of affiliation of the woman's factory at baseline. On the basis of this observation, we stratified the randomization of factories by hospital affiliation. Thus, the formal randomization-based test for intervention effects was similarly stratified on hospital affiliation of factory. Because of the large number of randomized factories and the complete homogeneity of rates among factories with each stratum, the stratified, randomization-based analyses of factory-specific rates would be virtually identical to similarly stratified analyses in which the individual is formally considered the unit of randomization. Therefore, statistical comparisons of the breast cancer mortality and survival in the instruction and control groups were performed as though the women had been individually randomly assigned with stratification on the hospital of affiliation of their factory. In each of these analyses, the assumption of homogeneity of factories within stratum was verified.

Mortality rates were estimated using standard life table methods, with time measured from date of entry into the study.

Women were censored at the time they left the STIB and were lost to follow-up or at time of death. Otherwise, follow-up was through the year 2000. Survival in women with breast cancer was calculated both from time of entry into the study and from time of diagnosis to death from breast cancer. Active follow-up was complete for all women with breast cancer through the end of the year 2000, including those who left the STIB, and only women who died of other causes were thus censored in these analyses. Survival probabilities were estimated using both life table and Kaplan–Meier *(41)* methods. Differences in survival between groups were evaluated using the log-rank test *(42),* stratifying on hospital affiliation of factory (four strata). Cumulative risk ratios were estimated using Cox proportional hazards models *(43),* similarly stratifying on hospital affiliation of factory. For each model, the proportional hazards assumption was evaluated graphically and was judged to hold in these analyses.

Distributions of categorical variables in different groups of women were compared using the chi-square test. Data management and analysis were conducted using SAS Software, version 6.12 (SAS Institute, Inc., Cary, NC). All statistical tests were two-sided.

RESULTS

Comparability of Study Groups

The randomization procedure yielded two groups that were similar with respect to risk factors for breast cancer and other variables, as ascertained from the baseline questionnaire (Table 2). The factories in the instruction and control groups were also similar with respect to hospital affiliation, number of employees, and the time of initiation of trial activities (data not shown).

Compliance

A high level of attendance was achieved for the baseline instruction and the first reinforcement session (Table 3). The lower attendance level at the second reinforcement session coincided with economic reforms initiated in 1994. Attendance by women in the control group at the concurrent low-back-pain prevention sessions was similar to that for the second reinforcement session at 84.2%. Similarly, attendance at two regularly scheduled supervised BSE sessions remained high through the first 4 years after the baseline instruction and then dropped in year 5. The total number of initial instruction, reinforcement, and supervised BSE sessions that current and retired workers could have attended was 15 and 13, respectively. The mean and median number of sessions actually attended were 12.3 and 13 for the current workers and 11.1 and 12 for the retired workers.

Only 10% of the women in both groups attended fewer than eight sessions. These women tended to be older than average (54.1% >45 years compared with 46.6% of the whole instruction group) but did not differ from the other women with respect to marital status, parity, family history of breast cancer, or use of tobacco or alcohol.

Proficiency

We determined the ability of randomly selected women to find breast lumps in silicone models (Table 4). Women in the instruction group consistently found a higher proportion of lumps, including both those that were easily palpable (10 mm in diameter, hard, and superficially placed) and those that were more difficult to feel (3 mm, soft, and deeply placed). Among women in the instruction group, lump-detecting ability was greatest immediately after the videos and declined to about the pre-video level in women assessed 1 year later. A similar pattern was not observed in the control group of women.

Women in the instruction group also had greater specificity in lump finding in the silicone models than women in the control group. The percentages of tested women in the instruction and control groups who erroneously reported finding one or more lumps that were not there in each of the six test models were 20.0% versus 33.7%; 16.8% versus 26.3%; 27.2% versus 43.1%; 28.1% versus 41.3%; 21.3% versus 36.0%; and 25.1% versus 39.8%. These percentages are all generated from approximately 1200 women per group, and the differences are all statistically significant ($P<.001$).

The proportions of sampled women in the instruction group who were observed to correctly perform various aspects of proper BSE technique varied from 66% (firm pressure) to 90% (use of pads of three fingers) before the first video and from 74% (use of a mirror) to 96% (hand over head) after the first video. These measures of competency declined to about pre-video levels in women assessed 1 year later. A similar pattern was ob-

Table 2. Percentages of women in the instruction and control groups in the randomized trial of breast self-examination (BSE) in Shanghai who had various risk factors for breast cancer and other characteristics at time of enrollment

	% women with risk factor	
	Instruction group	Control group
Risk factor	(N = 132 979)	(N = 133 085)
Age, y		
30–34	17.3	17.5
35–39	20.7	21.3
40–44	15.5	14.6
45–49	5.9	5.1
50–54	7.3	7.1
55–59	15.9	16.4
⩾60	17.4	18.0
No. of pregnancies		
0	4.2	4.3
1	17.5	18.2
2	26.3	26.1
⩾3	51.9	51.4
Ever induced abortion	51.9	50.0
Age at first live birth >30 y	17.2	16.5
Ever breast fed	80.7	79.6
Ever use of contraceptives		
IUD	48.6	48.3
Injectable	4.8	4.5
Oral	14.7	13.8
Tubal ligation	17.6	17.9
Age at menarche <13 y	10.3	10.5
Menopausal	39.8	40.6
Menopause after age 50	20.2	20.8
Ever breast lump	4.2	3.7
Clinical breast exam		
Ever	66.9	70.4
In past year	8.0	11.7
Practiced BSE in past year	11.9	13.3
Sister or mother with breast cancer	2.7	2.5
>1 alcoholic drink per month	3.9	4.3
Ever smoke cigarettes	2.8	3.0
Hospital affiliation of factory		
STIB hospital 1	22.3	21.1
STIB hospital 2	26.8	22.7
STIB hospital 3	7.4	9.5
Other hospital	43.5	46.7

*IUD = intrauterine device; STIB = Shanghai Textile Industry Bureau.

Table 3. Percentage of 132 979 women in the instruction group who attended baseline breast self-examination (BSE) instruction and two reinforcement sessions and two or more scheduled supervised BSE sessions per year

Sessions	Percentage of women
Instruction/reinforcement	
Baseline instruction	98.5
Reinforcement session 1	95.1
Reinforcement session 2	83.1
All three sessions	79.5
Any two sessions	17.7
Only one session	2.6
None	0.1
Supervised BSE (years)	
Year 1 (1989–1991)	91.6*
Year 2 (1990–1992)	81.4
Year 3 (1991–1993)	78.1
Year 4 (1992–1994)	73.6
Year 5 (1993–1995)	48.7

*Includes 82 776 current workers who were scheduled for four sessions and who attended two or more. The percentages of these women who attended all 4, 3, 2, 1, and no sessions were 64.1, 21.9, 7.9, 4.8, and 1.2, respectively.

served in relation to the second reinforcement session. About one third of the women assessed before the first video, and about two thirds of those assessed before the second video, did not examine the upper peripheral aspects of their breasts, and about one third of the women assessed before both videos did not examine their nipples. Only about one fourth of the women assessed right after the reinforcement sessions did not examine these areas, but the percentage of women tested 1 year later who missed these areas was near that for the women tested before the reinforcement sessions. Almost all women at all assessment sessions examined all other portions of their breasts.

Intermediate Variables

Slightly fewer women in the instruction group than in the control group were diagnosed with breast cancer, but the difference is not statistically significant ($P = .47$) (Table 5). With the possible exception of an increase in the detection of probable prevalent cancers in the instruction group during the first 6 months of trial participation, there is no evidence that the diagnoses of the breast cancers in the instruction group were brought forward in time; the numbers of cancers in the two groups are similar for each year of the trial after the first year (Table 6). Statistically significantly ($P<.001$) more women in the instruction group had one or more histologically confirmed benign breast lesions, and these lesions tended to be slightly smaller in the instruction group (47.0%, <2 cm) than in the control group (41.9%, <2 cm). The ratio of total biopsy specimens to histologically diagnosed carcinomas was 4.2 in the instruction group and 2.7 in the control group. This ratio was greater for women in the instruction group than for women in the control group during every year after entry into the trial (Table 6). It was highest during the first 6 months of trial participation in both groups. Thereafter, it gradually declined with time through year 6 in the instruction group, but not in the control group.

Table 4. Percentage of lumps with various characteristics detected in silicone breast models by sampled women in the instruction and control groups in relationship to the time of viewing the two video reinforcement sessions*

Lump characteristics	Relationship to time of video reinforcement sessions					
	Before		Immediately after		1 year after	
	Instruction group (%)	Control group (%)	Instruction group (%)	Control group (%)	Instruction group (%)	Control group (%)
			First video session			
Size, mm						
3	730/1302 (56.1)	523/1239 (42.2)	358/590 (60.7)	291/701 (41.5)	764/1285 (59.5)	556/1348 (41.3)
5	817/1309 (62.4)	576/1231 (46.8)	380/581 (65.4)	382/695 (55.0)	843/1292 (65.3)	731/1351 (54.1)
10	897/1306 (68.7)	726/1234 (58.8)	423/582 (72.7)	342/704 (48.6)	911/1285 (70.9)	777/1361 (57.1)
Hardness						
Soft	760/1305 (58.3)	551/1225 (45.0)	361/561 (64.4)	319/701 (45.5)	790/1299 (60.8)	650/1344 (48.4)
Intermediate	881/1302 (67.7)	661/1239 (53.3)	413/590 (70.0)	356/701 (50.8)	894/1285 (69.6)	733/1348 (54.4)
Hard	803/1310 (61.3)	613/1240 (49.4)	387/573 (67.5)	340/698 (48.7)	834/1278 (65.3)	681/1368 (49.8)
Depth						
Medium	1351/1964 (68.8)	1069/1849 (57.8)	627/878 (71.4)	592/1055 (56.1)	1334/1933 (69.0)	1169/2036 (57.4)
Deep	1094/1953 (56.0)	756/1855 (40.8)	534/875 (61.0)	423/1045 (40.5)	1184/1929 (61.4)	895/2024 (44.2)
			Second video session			
Size, mm						
3	668/1230 (54.3)	586/1294 (45.3)	745/1208 (61.7)	555/1281 (43.3)	696/1383 (50.3)	541/1298 (41.7)
5	729/1212 (60.2)	747/1311 (57.0)	878/1223 (71.8)	694/1285 (54.0)	870/1404 (62.0)	728/1298 (56.1)
10	835/1226 (68.1)	785/1299 (60.4)	874/1210 (72.3)	689/1277 (54.0)	928/1405 (66.1)	739/1302 (56.8)
Hardness						
Soft	635/1217 (52.2)	610/1291 (47.3)	781/1209 (64.6)	575/1283 (44.8)	695/1404 (49.5)	579/1301 (44.5)
Intermediate	825/1230 (67.1)	786/1294 (60.7)	864/1208 (71.5)	717/1281 (56.0)	929/1383 (67.2)	749/1298 (57.7)
Hard	772/1221 (63.3)	722/1319 (54.7)	852/1224 (69.6)	646/1279 (50.5)	870/1405 (61.9)	680/1299 (52.4)
Depth						
Medium	1160/1843 (62.9)	1116/1946 (57.4)	1283/1813 (70.8)	1030/1923 (53.6)	1337/2112 (63.3)	1107/1949 (56.8)
Deep	1072/1825 (58.7)	1002/1958 (51.2)	1214/1828 (66.4)	908/1920 (47.3)	1157/2080 (55.6)	901/1949 (46.2)

*The numerators represent the total number of lumps found by all tested women. The denominators are the total lumps of each type that could have been detected by the tested women. The denominators were calculated by summing the number of lumps of each type in the three models on which each woman was tested and then summing these numbers for all tested women. The percentage of lumps detected is given in parentheses.

Table 5. Numbers of women with malignant and benign breast lesions in the instruction and control groups of the randomized trial of breast self-examination in Shanghai

Type of breast lesion	Instruction group	Control group	Total
Carcinoma*			
Invasive	823	862	1685
In situ	33	28	61
Unknown whether invasive†	1	0	1
Total histologically confirmed	857	890	1747
Clinical diagnosis only	7	6	13
Total carcinomas	864	896	1760
Other malignancies of breast	4	3	7
Uncertain whether malignant	5	0	5
Benign breast disease			
Affected No. of women	2387	1296	3683
Total No. of biopsy examinations	2761	1505	4266
Total No. of women with biopsy examinations‡	3253	2189	5442
Total No. of biopsy examinations	3627	2398	6025

*Women with multiple malignancies were counted only once.

†This breast lesion was histologically confirmed as carcinoma, but whether invasive or *in situ* was not recorded.

‡Excludes clinically diagnosed breast cancers.

A higher proportion of breast cancers were diagnosed when *in situ* (Tis, stage 0) and less than or equal to 2 cm in diameter (T1) by women in the instruction group (3.9% and 44.9%, respectively) than by those in the control group (3.2% and 41.6%, respectively) (Table 6). Although these differences were not statistically significant ($P = .60$, chi-square test with four degrees of freedom), such differences were observed during most of the years after entry into the trial. However, these differences are small and may be of limited clinical importance. There was evidence of digit preference, with many lesions recorded as exactly 2.0 cm in diameter. However, when the tumors were classified as less than 1.5 cm, 1.6–2.3 cm, 2.4–3.4 cm, and equal to or greater than 3.5 cm, the differences between malignancies in the two groups were even less than for T1 tumors.

Among women in the instruction group, a slightly higher percentage of those who had attended all BSE instruction, reinforcement, and supervised practice sessions were diagnosed at Tis or T1 than those who attended fewer sessions: 56 (52.3%) of 107 current workers who attended all 15 sessions compared with 78 (45.3%) of 172 current workers who attended 11 or fewer sessions ($P = .67$); and 56 (48.7%) of 115 retired workers who attended all 13 sessions compared with 36 (44.4%) of 81 retired workers who attended nine or fewer sessions ($P = .97$).

The extent of regional lymph node involvement was similar in the instruction and control groups. The percentages of women with N0, N1, N2, and N3 tumors were 47.0%, 44.4%, 7.6%, and 1.0%, respectively, in the instruction group and 48.3%, 44.1%, 6.7%, and 0.8% in the control group. Among the 797 (92.2%) women with breast cancer in the instruction group and 844 (93.8%) women with breast cancer in the control group with histologically ascertained nodal status, evidence of tumor was found in 314 (39.4%) and 342 (40.5%), respectively. Only 13 (1.5%) and 22 (2.5%) of the cancers in the instruction and control groups, respectively, presented with distant metastases (M1). The TNM Summary Stage *(39)* of the cancers in the women in the two groups also did not differ: 29.7% of the cancers in both groups presented at stage 0 or I.

Detection and Treatment of Breast Cancer

Information on how the breast cancer was first found was ascertained from 739 women in the instruction group and 748 women in the control group. Only 20 (2.7%) and 27 (3.6%) of the tumors in the two groups, respectively, were initially found by means of a clinical breast examination. In the instruction group, 605 (81.9%) were reported found when practicing BSE. Comparable information was not ascertained from women in the control group; lesions were reported by the remaining 721 (96.4%) women as being found "accidentally" or "by themselves."

The time from initial detection to first medical evaluation occurred within 1 week for 447 (65.0%) of 688 women in the

Table 6. Numbers of histologically diagnosed carcinomas and benign breast biopsy specimens, ratios of the total number of biopsy specimens to carcinomas, and proportions of carcinomas diagnosed as Tis and T1 tumors in women in the instruction and control groups by time since entry into the randomized trial of breast self-examination in Shanghai*

Time since entry into trial	No. of histologically diagnosed carcinomas†		No. of benign breast biopsy specimens		Total No. of biopsy specimens		Total No. of biopsy specimens to No. of carcinomas		% carcinomas‡ Tis plus T1	
	Instruction	Control	Instruction	Control	Instruction	Control	Instruction	Control	Instruction	Control
<6 mo	60	40	490	112	551	152	9.0	3.8	45.0	55.0
6–12 mo	35	39	193	70	228	109	6.5	2.8	55.9	44.7
Year 2	70	70	375	175	445	245	6.4	3.5	45.6	45.7
Year 3	77	62	293	160	370	222	4.8	3.6	42.9	53.2
Year 4	85	89	251	156	337	245	3.9	2.8	48.8	43.8
Year 5	90	86	229	138	319	224	3.5	2.6	55.2	44.7
Year 6	91	95	200	122	291	217	3.2	2.3	45.1	41.9
Year 7	95	98	191	134	286	232	3.0	2.4	47.8	40.4
Year 8	79	99	175	142	254	241	3.2	2.4	59.0	49.0
Year 9	83	102	179	159	262	261	3.2	2.6	49.4	41.4
Year 10	63	81	135	98	198	179	3.1	2.2	41.0	45.0
Year 11	29	29	50	39	79	68	2.7	2.3	53.6	31.0
Total	857	890	2761	1505	3620	2395	4.2	2.7	48.8	44.8

*Tumors were classified according to the tumor–node–metastasis (TNM) scheme *(38)*.

†Excluding seven cancers in the instruction group and six cancers in the control group that were clinically diagnosed only.

‡Based on cancers of known T-classification only and excluding 19 cancers in each group with unknown T-classification.

instruction group for whom this information was available and for 406 (58.7%) of 692 women in the control group. The time from referral to treatment was almost the same for women in the two groups, with a median of 14 and 13 days, respectively ($P = .48$, chi-square test with seven degrees of freedom).

Slightly more women with breast cancer in the instruction group than in the control group had breast-conserving surgery (4.4% compared with 2.7%) or a simple mastectomy without (2.8% compared with 2.5%) or with (40.9% compared with 36.6%) axillary lymph nodal dissection, and slightly fewer had radical mastectomies (50.7% compared with 56.7%) ($P = .06$, chi-square test with four degrees of freedom). Stratification by stage revealed that these differences were not a result of women in the instruction group having been diagnosed at a slightly less advanced stage than women in the control group. Comparing women in the instruction group with women in the control group, nearly equal proportions of women were treated with radiation (23.7% in each group), hormones or antihormones (predominantly tamoxifen) (84.1% versus 81.9%), chemotherapy with single agents (6.1% versus 7.1%) or combinations of agents (80.0% versus 80.9%), traditional Chinese herbal medicines (30.1% versus 31.3%), and various combinations of treatments. Nearly all women in both groups (93.8%) received systemic therapy with either hormones (predominantly tamoxifen) or chemotherapy or both, in addition to surgery.

Mortality

During the 10 to 11 years of follow-up, 5349 (4.0%) of women in the instruction group and 5939 (4.5%) of the women in the control group died, and 9873 (7.4%) and 9997 (7.5%), respectively, left the STIB. Survival in the two groups at the end of follow-up was 95.2% (95% CI = 95.1 to 95.3) and 94.9% (95% CI = 94.7 to 95.0), respectively ($P<.001$). The number of deaths was greater in the control group than in the instruction group during each year of the trial, whereas the numbers of women who left the STIB were similar each year.

A total of 154 (0.12%) of the women in the instruction group and 158 (0.12%) of those in the control group developed breast cancer and died. The numbers of women judged to have very likely, probably, probably not, and very likely not, died of their disease were 107, 28, 17, and 2 in the instruction group and 103, 28, 25, and 2 in the control group, respectively. These differences are not statistically significant ($P = .67$). A total of 135 (0.10%) and 131 (0.10%) in the two groups, respectively, were judged to have either very likely or probably died of their disease, and these are considered the deaths from breast cancer in all subsequent analyses presented. A total of 1 276 611 woman-years were accrued in the instruction group, and 1 286 678 woman-years were accrued in the control group. The cumulative breast cancer mortality rates by year since entry into the trial (Fig. 1) for the women in the two arms of the study are similar: 150.6 versus 120.6 per 100 000 woman-years, respectively ($P = .72$, log-rank test). The cumulative risk ratio for women in the instruction group relative to women in the control group is 1.04 (95% CI = 0.82 to 1.33). The cumulative risk ratio based on all deaths in women with breast cancer was similar ($P = .94$, log-rank test). The results are also similar for women of different ages. The cumulative risk ratios, based on deaths from breast cancer, were 1.23 (95% CI = 0.83 to 1.82) and 0.96 (95% CI = 0.70 to 1.31) for women aged less than 50 years and 50 years or older at baseline, respectively.

Survival

We assessed survival in women with breast cancer from the time of entry into the trial to eliminate any effect of lead-time bias (Fig. 2, A). No difference in survival for women in the two arms of the study was detected ($P = .50$, log-rank test). There was also no difference in survival from time of diagnosis ($P = .94$, log-rank test), clearly demonstrating no increase in lead-time resulting from teaching BSE (Fig. 2, B). When the analyses were restricted to women who had received systemic treatment with hormones or chemotherapy, the results were very similar. Mortality from breast cancer did not decrease statistically significantly with number of BSE instruction, reinforcement, and supervised BSE sessions attended (data not shown). The cumulative survival rate from entry into the trial was 94.6% (95% CI = 92.4% to 96.9%) for women with Tis or T1 tumors and 78.1% (95% CI = 76.9% to 82.1%) for women with more advanced disease ($P<.001$, log-rank test), which indicates that if teaching BSE had resulted in a sufficiently large shift toward earlier diagnosis, a difference in survival in the two arms of the study would have been observable.

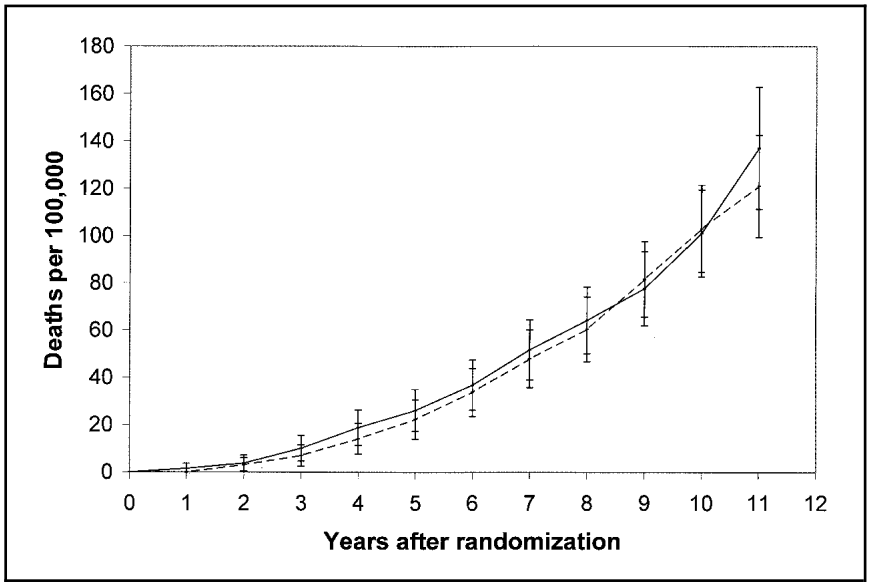

Fig. 1. Cumulative breast cancer mortality per 100 000 women in the instruction group (**solid line**) and control group (**broken line**) of the randomized trial of breast self-examination in Shanghai. **Error bars** represent 95% confidence intervals.

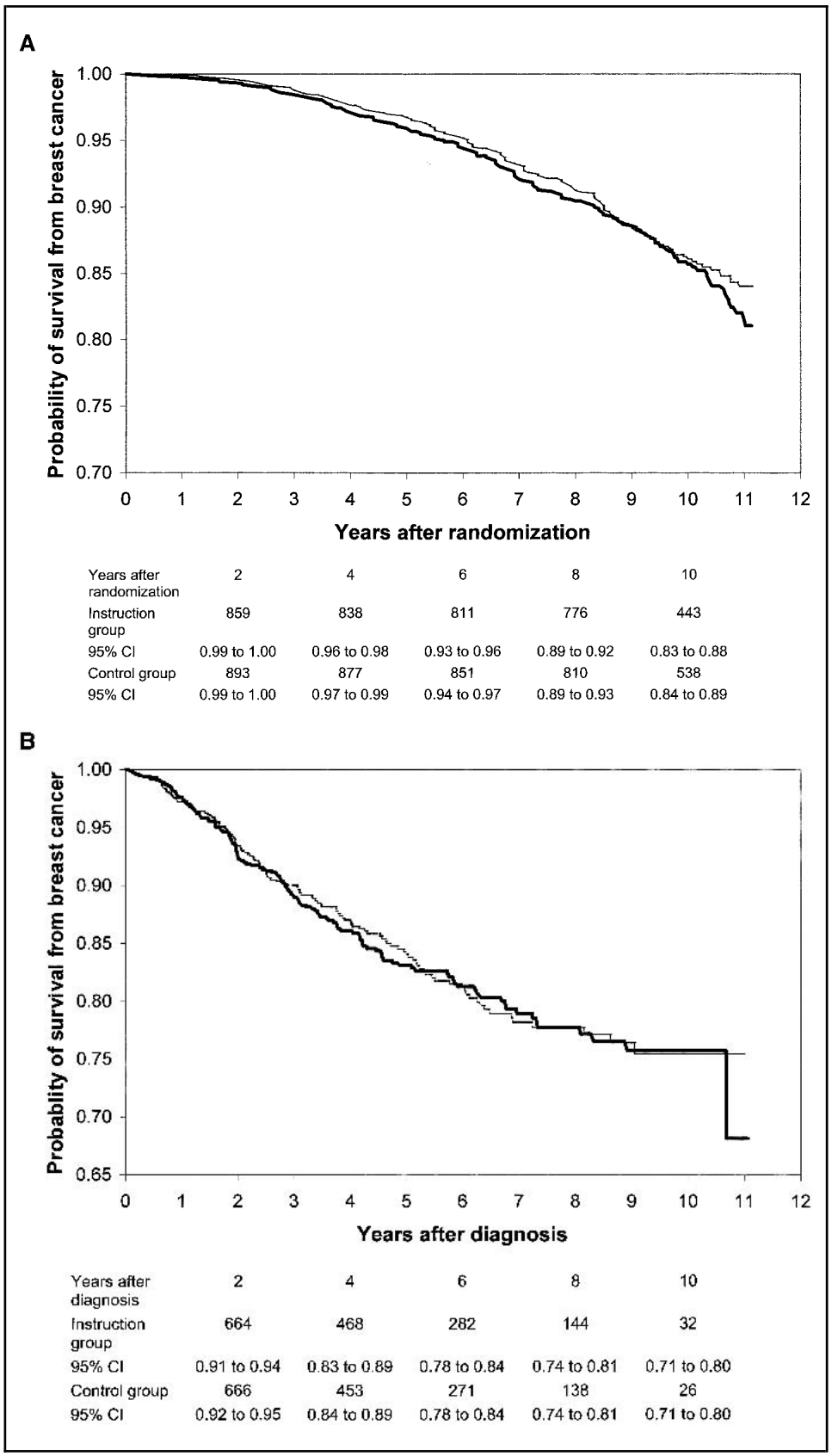

Fig. 2. A) Probability of survival from breast cancer in case patients by study group and time from entry into the randomized trial of breast self-examination in Shanghai. **B**) Probability of survival from breast cancer in case patients by study group and time from the time of diagnosis. **Thick line** = instruction group; **thin line** = control group. Tables show number of patients at risk and 95% confidence intervals (CIs) for the probabilities for years 2, 4, 6, 8, and 10 after randomization or diagnosis.

DISCUSSION

This randomized trial of a massive effort to teach and encourage approximately 133 000 Chinese women to practice BSE did not show a reduction in mortality from breast cancer over a 10- to 11-year period. These results are not likely the result of inadequate duration of the study. Randomized trials of mammography in women of about the same age as those in this investigation *(44)* showed a reduction in breast cancer mortality rates after about 5 years of follow-up. In addition, our study had

80% power to detect a true reduction in risk of about 30%, which is the level of reduction in risk observed for mammography. Although a smaller reduction could have readily been missed, a reduction of much less than this would have limited clinical value.

There were nearly three times as many exclusions after randomization from the instruction group (2.6%) as there were from the control group (1.0%), presumably because the instruction group women were asked to make a greater commitment to the study than simply completing the baseline questionnaire. However, it is virtually impossible that the mortality rates from breast cancer in the excluded women could have been so extreme that they could have appreciably altered the overall breast cancer mortality rates in the two groups if they had been included.

The two comparison groups were similar in age and prevalence of a variety of risk factors for breast cancer, and a similar proportion (0.5%) of the eligible women in both groups were excluded because they gave a history of breast cancer on the baseline questionnaire. Women in the two groups also worked in factories of similar size and with similar affiliations with various hospitals, thereby ensuring comparability of available diagnostic and treatment facilities.

More women died of causes other than breast cancer in the control group than in the instruction group for reasons that are unknown. Causes of death other than breast cancer were ascertained from records of the STIB Tumor and Death Registry but were not verified. There were no obvious differences in specific causes of death that contributed disproportionately to the overall differences in mortality rates. The differences could therefore be due to under ascertainment of deaths or to a generally healthier lifestyle in the instruction group. If applicable to breast cancer, these possibilities would result in either an under ascertainment of breast cancer deaths in the instruction group of about 10%, or a survival advantage of the same magnitude in the women with breast cancer in that group. Either of these possibilities would have the effect of making BSE seem slightly more efficacious than it actually was, not of obscuring a true beneficial effect of BSE on breast cancer mortality.

Only about 7.5% of the women severed their ties to their factory. Because this percentage is low, and about the same in the two comparison groups, it is unlikely that loss of these women to the full 10 years of follow-up could account for the trial outcome.

Inadequate identification of breast cancers is also not a plausible explanation for the results. Matching of the study cohort to the Shanghai Cancer Registry revealed few missed cancers, and independent reading of histologic slides from women with benign breast disease resulted in few additional cancers being found that had erroneously been classified as benign. Furthermore, the level of under ascertainment of breast cancers estimated from these two activities was similar for the instruction and control groups. In addition, the overall incidence rate of breast cancer was slightly higher in the study cohort than that reported in the Shanghai Cancer Registry. Data from this registry meet internationally accepted criteria for completeness *(37),* and our level of case finding appears to be at least as complete as that for the Shanghai Registry.

It is also unlikely that incomplete ascertainment of breast cancer deaths affected the outcome of this trial. All women with benign and malignant breast disease were actively followed up for vital status through the end of the year 2000, and deaths from breast cancer reported to the STIB tumor and death registry were followed back and included as cancers in the study if they were found to have been previously missed.

A spurious increase in the relative risk of dying from breast cancer that might have obscured a true beneficial effect of BSE on breast cancer mortality could have occurred if more deaths from breast cancer were erroneously ascribed to other causes in the missed cases in the control group than in the missed cases in the instruction group. This is also an unlikely explanation for our results. We estimated that only about 35 invasive cancers were missed (15 not found and 20 erroneously diagnosed as a benign condition). Furthermore, these missed cancers were equally distributed between the two arms of the study. Even under the unlikely assumptions that half of the missed cancers in the control group died, that all of these deaths were erroneously coded as dying of something other than breast cancer, and that the missed cancers in the instruction group all lived, this would add just nine breast cancer deaths to the 131 observed in the control group, reducing the observed relative risk of dying of breast cancer from 1.04 to about 1.

Breast cancer screening in the control women also cannot explain the results. No mammographic screening was available to women in the study population, and the level of clinical breast examination activity was low, with few women in either group having had their cancer detected by this means. By randomizing on factory, contact between women in the two arms of the study was minimized, and the women in the control group were not subjected to the reminder posters or other activities in the instruction group factories. The number of women in the control group who transferred into an instruction group factory was small, and their transfer occurred largely after most activities to encourage BSE had ceased. Over half the women in the control group presented with tumors more than 2 cm in diameter and with involvement of axillary lymph nodes. Because BSE instruction emphasizes detection of small, localized tumors, women in the control group were not bringing their cancer to medical attention when small and confined to the breast with such frequency that successful screening could not have effected an improvement in extent of disease at diagnosis.

The results of testing random samples of women on their ability to detect breast lumps in silicone models provided objective evidence that the instruction enhanced both the sensitivity and specificity of a woman's ability to identify lumps. Direct observation of a sample of the women in the instruction group while they were practicing BSE indicated that they had adequate technique and good coverage of most areas of the breast. However, these measures of proficiency were more favorable soon after reinforcement sessions than they were 1 year later, suggesting some decline in skill level with time (although their BSE skills were still greater than those of women in the control group). Whether the residual level of skill was sufficient to effect a decline in mortality with sufficiently frequent BSE practice is conjectural, but it is clear that the teaching program did enhance the overall skill level in the instruction group.

We have objective evidence that, during the first 4–5 years of the trial, the women practiced BSE under the supervision of medical workers an average of 12 times, or roughly every 4–5 months. Although women were strongly urged to practice BSE monthly on their own, the frequency of practice outside the clinic setting is unknown. An attempt was made to ask a sample of women how frequently they practiced BSE, but the responses

were uniformly "monthly" and hence not considered reliable. It can, however, be concluded with certainty that practicing BSE a minimum of every 5 months for 4–5 years did not have an effect on breast cancer mortality.

The women in the instruction group detected more benign breast lesions than women in the control group did during every year of the trial, and the lesions detected were of slightly smaller size, suggesting that the teaching of BSE did enhance the women's level of awareness and ability to find lumps in their own breasts. In addition, a slightly higher percentage of breast cancers in the instruction group than in the control group were diagnosed when they were *in situ* (Tis) or no more than 2 cm in diameter (T1), and diagnosis of tumors at these early stages was weakly associated with level of attendance at study activity sessions. Furthermore, a high proportion of women in the instruction group reported that they had found their breast cancer as a result of practicing BSE, and they sought medical attention after detecting their tumor in somewhat less time than did women in the control group. However, in spite of these apparent successes of the BSE teaching program, as many cancers in the instruction group as in the control group were diagnosed with axillary lymph nodal metastases. There is also no evidence that the diagnosis was brought forward in time by teaching BSE: the number of cases that occurred in the two groups was similar during all years of the trial (except for the first 6 months, when a few more, presumably prevalent cases, were diagnosed in the instruction group), and survival time from diagnoses was similar in the two groups, indicating the absence of a lead-time resulting from BSE screening.

Most women with breast cancer received surgery that was as extensive as or more extensive than women in many countries more economically advanced than China, and nearly all of the women in both groups additionally received systemic treatment with hormones or chemotherapy. Furthermore, there was no difference in survival from breast cancer between women in the instruction and control groups who had received these systemic treatments. Inadequate treatment is thus not a likely explanation for the absence of a beneficial effect of BSE on breast cancer mortality or survival.

It thus seems reasonable to conclude that a reduction in deaths from breast cancer was most likely not observed in this trial because the teaching of BSE did not result in breast cancer being diagnosed at a sufficiently less advanced stage of progression for appropriate therapy to have altered the course of the disease. Consequently, in developing countries, where mammographic screening is not available, it would not seem to be a good use of the limited funds available for preventive services to promote the practice of BSE. Although we do not know the level of BSE practiced by women in this trial outside the supervised sessions, the high level of attendance at most program activities, the ability to post reminders in showers and other conspicuous areas in the factories, and the availability of facilities for BSE practice in the factories, suggest that the level of BSE activity achieved by the women in this trial was as high as one could reasonably expect in a mass program directed at a large general population of women. It is unlikely that routine public health programs in developing countries could achieve this level of intensity, and a lower level of instruction and encouragement would almost certainly have no meaningful impact on breast cancer mortality rates. Furthermore, the results of this trial show that such programs would result in an increased rate of benign breast biopsies, further burdening the health care system.

For women in more developed countries with access to mammographic screening, the results of this trial should serve to emphasize that BSE is not a substitute for regular screening by mammography. The implications of our results for women who do have periodic mammograms are unclear. This was a trial of the teaching of BSE, not of the practice of BSE. It should not be inferred from the results of this study that there would be no reduction in risk of dying from breast cancer if women practiced BSE competently and frequently. It is possible that highly motivated women could be taught to detect cancers that develop between regular screenings, and that the diligent practice of BSE would enhance the benefit of a screening program. It would be useful to conduct a randomized trial in which all women receive regular mammographic examinations and half are randomly allocated to also receive intensive instruction in BSE. Observational studies have shown tumor size at diagnosis to be inversely associated with the frequency of BSE practice *(10,11,13–18),* suggesting that if proficient, sustained BSE practice could be achieved, such a trial might show a beneficial effect of BSE on breast cancer mortality. It is, however, unlikely that the level of BSE activity necessary to effect a change in breast cancer mortality could be achieved in a general population of women. The trial would have to be conducted in highly motivated individuals. Until such a trial is conducted, there is no reason to discourage women who choose to practice BSE from doing so. However, it should be emphasized to such women that they must practice BSE regularly and with a high degree of proficiency. They should also be informed that if they do this, they have an increased chance of having a breast biopsy that does not reveal a cancer and that it is not known whether practicing BSE to detect interval cancers that develop between mammographic screenings will reduce a woman's chance of dying from breast cancer.

References

(1) Greenwald P, Nasca PC, Lawrence CE, Horton J, McGarrah RP, Gabriele T, et al. Estimated effect of breast self-examination and routine physician examinations on breast-cancer mortality. N Engl J Med 1978;299:271–3.

(2) Smith EM, Francis AM, Polissar L. The effect of breast self-exam practices and physician examinations on extent of disease at diagnosis. Prev Med 1980;9:409–17.

(3) Gould-Martin K, Paganini-Hill A, Casagrande C, Mack T, Ross RK. Behavioral and biological determinants of surgical stage of breast cancer. Prev Med 1982;11:429–40.

(4) Philip J, Harris WG, Flaherty C, Joslin CA, Rustage JH, Wijesinghe DP. Breast self-examination: clinical results from a population-based prospective study. Br J Cancer 1984;50:7–12.

(5) Owen WL, Hoge AG, Asal NR, Anderson PL Jr, Cucchiara AJ. Self-examination of the breast: use and effectiveness. South Med J 1985;78:1170–3.

(6) Smith EM, Burns TL. The effects of breast self-examination in a population-based cancer registry. Cancer 1985;55:432–7.

(7) Kuroishi T, Tominaga S, Ota J, Horino T, Taguchi T, Ishida T, et al. The effect of breast self-examination on early detection and survival. Jpn J Cancer Res 1992;83:344–50.

(8) Auvinen A, Elovainio L, Hakama M. Breast self-examination and survival from breast cancer: a prospective follow-up study. Breast Cancer Res Treat 1996;38:161–8.

(9) McPherson CP, Swenson KK, Jolitz G, Murray CL Survival of women ages 40–49 years with breast carcinoma according to method of detection. Cancer 1997;79:1923–32.

(10) Koibuchi Y, Iino Y, Takei H, Maemura M, Horiguchi J, Yokoe T, et al. The

effect of mass screening by physical examination combined with regular breast self-examination on clinical stage and course of Japanese women with breast cancer. Oncol Rep 1998;5:151–5.
(11) Interdisciplinary Group for Cancer Care Evaluation (GIVIO). Practice of breast self-examination: disease extent at diagnosis and patterns of surgical care. A report from an Italian study. J Epidemiol Community Health 1991; 45:112–6.
(12) Gastrin G, Miller AB, To T, Aronson KJ, Wall C, Hakama M, et al. Incidence and mortality from breast cancer in the Mama Program for Breast Screening in Finland, 1973–1986. Cancer 1994;73:2168–74.
(13) Feldman JG, Carter AC, Nicastri AD, Hosat ST. Breast self-examination, relationship to stage of breast cancer at diagnosis. Cancer 1981;47:2740–5.
(14) Tamburini M, Massara G, Bertario L, Re A, Di Pietro S. Usefulness of breast self-examination for an early detection of breast cancer. Results of a study on 500 breast cancer patients and 652 controls. Tumori 1981;67: 219–24.
(15) Hislop TG, Coldman AJ, Skippen DH. Breast self-examination: importance of technique in early diagnosis. Can Med Assoc J 1984;131:1349–52.
(16) Foster RS Jr, Costanza MC. Breast self-examination practices and breast cancer survival. Cancer 1984; 53:999–1005.
(17) Mant D, Vessey MP, Neil A, McPherson K, Jones L. Breast self examination and breast cancer stage at diagnosis. Br J Cancer 1987;55:207–11.
(18) Ogawa H, Tominaga S, Yoshida M, Kubo K, Takeuchi S. Breast self-examination practice and clinical stage of breast cancer. Jpn J Cancer Res 1987;78:447–52.
(19) Huguley CM Jr, Brown RL, Greenberg RS, Clark WS. Breast self-examination and survival from breast cancer. Cancer 1988;62:1389–96.
(20) Locker AP, Caseldine J, Mitchell AK, Blamey RW, Roebuck EJ, Elston CW. Results from a seven-year programme of breast self-examination in 89,010 women. Br J Cancer 1989;60:401–5.
(21) Le Geyte M, Mant D, Vessey MP, Jones L, Yudkin P. Breast self-examination and survival from breast cancer. Br J Cancer 1992;66:917–8.
(22) Senie RT, Lesser M, Kinne DW, Rosen PR. Method of tumor detection influences disease-free survival of women with breast carcinoma. Cancer 1994;73:1666–72.
(23) Holmberg L, Ekbom A, Calle E, Mokdad A, Byers T. Breast cancer mortality in relation to self-reported use of breast self-examination. A cohort study of 450,000 women. Breast Cancer Res Treat 1997;43:137–40.
(24) 16-year mortality from breast cancer in the UK Trial of Early Detection of Breast Cancer. Lancet 1999;353:1909–14.
(25) Muscat JE, Huncharek MS. Breast self-examination and extent of disease: a population-based study. Cancer Detect Prev 1991;15:155–9.
(26) Newcomb PA, Weiss NS, Storer BE, Scholes D, Young BE, Voigt LF. Breast self-examination in relation to the occurrence of advanced breast cancer. J Natl Cancer Inst 1991;83:260–5.
(27) Harvey BJ, Miller AB, Baines CJ, Corey PN. Effect of breast self-examination techniques on the risk of death from breast cancer. CMAJ 1997;157:1205–12.
(28) Baxter N. Preventative health care, 2001 update: Should women be routinely taught breast self-examination to screen for breast cancer? CMAJ 2001;164:1837–46.
(29) U.S. Preventive Services Task Force. Screening for breast cancer: recommendations and rationale. Ann Intern Med 2002;137:344–6.
(30) Miller AB, Chamberlain J, Tsechkovski M. Self-examination in the early detection of breast cancer. A review of the evidence, with recommendations for further research. J Chronic Dis 1985;38:527–40.
(31) Miller AB, Chamberlain J, Day NE, Hakama M, Prorok PC. Report on a Workshop of the UICC Project on Evaluation of Screening for Cancer. Int J Cancer 1990;46:761–9.
(32) Semiglazov VF, Moiseyenko VM, Bavli JL, Migmanova NS, Seleznyov NK, Popova RT, et al. The role of breast self-examination in early breast cancer detection (results of the 5-years USSR/WHO randomized study in Leningrad). Eur J Epidemiol 1992;8:498–502.
(33) Semiglazov VF, Sagaidak VN, Moiseyenko VM, Mikhailov EA. Study of the role of breast self-examination in the reduction of mortality from breast cancer. The Russian Federation/World Health Organization Study. Eur J Cancer 1993;29A:2039–46.
(34) Semiglazov VF, Moiseyenko VM, Manikhas AG, Protsenko SA, Kharikova RS, Ivanov VG, et al. Role of breast self-examination in early detection of breast cancer: Russia/WHO prospective randomized trial in St. Petersburg. Cancer Strategy 1999;1:145–51.
(35) Thomas DB, Gao DL, Self SG, Allison CJ, Tao Y, Mahloch J, et al. Randomized trial of breast self-examination in Shanghai: methodology and preliminary results. J Natl Cancer Inst 1997;89:355–65.
(36) Fletcher SW, O'Malley MS, Pilgrim CA, Gonzalez JJ. How do women compare with internal medicine residents in breast lump detection? A study with silicone models. J Gen Intern Med 1989;4:277–83.
(37) Gao YT, Fan J. In: Parkin DM, Whelan SL, Ferlay J, Raymond L, Young J, editors. Cancer in five continents. Vol. VII. IARC Sci Publ 1997;(143): 326–29.
(38) American Joint Committee on Cancer. AJCC cancer staging manual. 5th ed. Phildelphia (PA): Lippincott-Raven; 1997.
(39) World Health Organization. Histological typing of breast tumors. Tumori 1982;68:181–98.
(40) Cornfield J. Randomization by group: a formal analysis. Am J Epidemiol 1978;108:100–2.
(41) Kaplan EL, Meier P. Nonparametric estimates from incomplete observations. J Am Stat Assoc 1958;53:457–81.
(42) Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemother Rep 1966;50:163–70.
(43) Breslow NE, Day NE. Statistical methods in cancer research. Vol II—The design and analysis of cohort studies. IARC Sci Publ 1987;(82):1–406.
(44) Nystrom L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Ryden S, et al. Breast cancer screening with mammography: overview of Swedish randomised trials. [published erratum appears in Lancet 1993;342:1372]. Lancet 1993;341:973–8.

NOTES

Supported by Public Health Services grant ROI-CA46823 from the National Cancer Institute, National Institutes of Health, Department of Health and Human Services.

We thank Judith Calman, Dawn Fitzgibbons, Richard Gandolfo, Georgia Green, Ted Grichuhin, Jan Kikuchi, Jan Maloch, April Richardson, Laura Robins-Morris, Karen Wernli, and Shirley Zhang for their technical and administrative support; Drs. Yu-Tang Gao, Fan Jin, and Yong-Bing Xiang for access to the records and data of the Shanghai Cancer Registry; Dr. Cong-Zhi Liu for assessing cause of death; Dr. Benjamin O. Anderson for his helpful suggestions; and the approximately 5000 medical workers and more than 100 BSE workers of the Shanghai Textile Industry Bureau, without whose cooperation and efforts this trial could not have been conducted.

Manuscript received March 4, 2002; revised July 16, 2002; accepted July 26, 2002.