Published in final edited form as:
Cancer Causes Control. 2011 September ; 22(9): 1319–1331. doi:10.1007/s10552-011-9805-9.

The After Breast Cancer Pooling Project: Rationale, Methodology, and Breast Cancer Survivor Characteristics

Sarah J. Nechuta[1], **Bette J. Caan**[2], **Wendy Y. Chen**[3,4], **Shirley W. Flatt**[5], **Wei Lu**[6], **Ruth E. Patterson**[5], **Elizabeth M. Poole**[3], **Marilyn L. Kwan**[2], **Zhi Chen**[1], **Erin Weltzien**[2], **John P. Pierce**[5], and **Xiao Ou Shu**[1]

[1]Division of Epidemiology, Department of Medicine, Vanderbilt University School of Medicine, Nashville, TN, 37203-1738, USA

[2]Division of Research, Kaiser Permanente, 2000 Oakland, CA 94612, USA

[3]Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, 02115, USA

[4]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02115, USA

[5]Cancer Prevention and Control Program, Moores UCSD Cancer Center, University of California, San Diego, CA, 92093-0901, USA

[6]Shanghai Center for Disease Control and Prevention, Shanghai, China

Abstract

The After Breast Cancer Pooling Project was established to examine the role of physical activity, adiposity, dietary factors, supplement use, and quality of life (QOL) in breast cancer prognosis. This paper presents pooled and harmonized data on post-diagnosis lifestyle factors, clinical prognostic factors, and breast cancer outcomes from four prospective cohorts of breast cancer survivors (three US-based and one from Shanghai, China) for 18,314 invasive breast cancer cases diagnosed between 1976 and 2006. Most participants were diagnosed with stage I-II breast cancer (84.7%). About 60% of breast tumors were estrogen receptor (ER)+/progesterone receptor (PR)+; 21% were ER−/PR−. Among 8,118 participants with information on HER-2 tumor status, 74.8% were HER-2− and 18.5% were HER-2+. At 1–2 years post-diagnosis (on average) 17.9% of participants were obese (BMI ≥30 kg/m^2), 32.6% were overweight (BMI 25–29 kg/m^2) and 59.9% met the 2008 Physical Activity Guidelines for Americans (≥ 2.5 hours per week of moderate activity). During follow-up (mean=8.4 years), 3,736 deaths (2,614 from breast cancer), and 3,564 recurrences have been documented. After accounting for differences in year of diagnosis and timing of post-diagnosis enrollment, five-year overall survival estimates were similar across cohorts. This pooling project of 18,000 breast cancer survivors enables the evaluation of associations of post-diagnosis lifestyle factors, QOL, and breast cancer outcomes with an adequate sample size for investigation of heterogeneity by hormone-receptor status and other clinical predictors. The project sets the stage for international collaborations for the investigation of modifiable predictors for breast cancer outcomes.

Keywords

Breast neoplasm; Survival; Recurrence; Life style

Corresponding author: Xiao Ou Shu, M.D., Ph.D., Professor of Medicine, Vanderbilt Epidemiology Center, Vanderbilt University Medical Center, 2525 West End Avenue, Suite 600, IMPH, Nashville, TN 37203-1738, Xiao-ou.shu@vanderbilt.edu.

Conflicts of interest: The authors have no conflicts of interest to declare.

INTRODUCTION

Numerous studies have examined the role of lifestyle factors such as diet, adiposity, and physical activity in the etiology of breast cancer [1–4]. Less is known, however, about the role of lifestyle factors in relation to breast cancer prognosis, although research in this area has begun to accumulate in the past several years [5–8]. In addition, many studies on lifestyle and breast cancer prognosis have focused on pre-diagnosis lifestyle factors, and fewer studies have examined associations of post-diagnosis lifestyle factors with survival and cancer recurrence [5, 7, 9–11].

An estimated 4.4 million women are living with breast cancer worldwide [12]. The role of modifiable lifestyle factors during and after cancer treatment in cancer prognosis is of particular interest to cancer survivors [5, 6]. Non-clinical post-diagnosis risk factors, including physical activity [9, 11, 13–16], adiposity [9, 11, 13, 17–19], quality of life (QOL) [20, 21, 22, 23], and dietary factors [9, 13, 24–29] have been associated with breast cancer prognosis. However, while it is well recognized that breast cancer is a heterogeneous disease with outcomes varying greatly by tumor expression of estrogen receptors (ER), progesterone receptors (PR), and human epidermal growth factor receptor 2 (HER-2) [30], little research has been conducted to investigate whether effects of lifestyle factors on outcomes differ by the subtypes of breast cancer. This effort has been primarily hindered by the need for studies with a larger sample size.

The After Breast Cancer Pooling Project (ABCPP) was initiated in 2009 through funding from the NIH American Recovery and Reinvestment Act (ARRA) (3R01CA118229-03S1) to support collaborative research on post-diagnosis lifestyle factors, QOL, and breast cancer survival. The ABCPP, which includes data on over 18,000 breast cancer survivors from four population-based prospective cohorts recruited from multiple US sites and Shanghai, China, was designed to evaluate the role of modifiable post-diagnosis lifestyle factors (physical activity, overweight/obesity, weight change, dietary intake of soy and cruciferous vegetables, and dietary supplement use) and QOL in relation to breast cancer outcomes. Special emphasis will be placed on investigating potential interactions between lifestyle factors and QOL with clinical predictors, such as hormone receptor status and treatment status. This paper describes the methodology of the ABCPP and characteristics of participants (clinical characteristics, socio-demographics, and select lifestyle factors and co-morbidities) by cohort and combined.

MATERIALS AND METHODS

After Breast Cancer Pooling Project—The ABCPP includes pooled data from three prospective cohorts specifically designed to focus on breast cancer survivors, the Shanghai Breast Cancer Survival Study (**SBCSS**), the Women's Healthy Eating and Living (**WHEL**) Study and the Life After Cancer Epidemiology (**LACE**) Study, as well as breast cancer cases from the Nurses' Health Study I (**NHS**), an ongoing longitudinal cohort of initially healthy women. This collaborative effort began with three cohorts (LACE, SBCSS, NHS), and was expanded after the project was funded to include one additional cohort with a large number of cases and detailed post-diagnosis lifestyle data (WHEL). The initial aims are focused on the role of specific post-diagnosis diet and lifestyle factors in relation to prognosis; hence, included cohorts were required to have cancer treatment data, detailed post-diagnosis diet and lifestyle data (including weight gain, body mass index (BMI), physical activity and vitamin supplement use) and active follow-up for breast cancer outcomes. A total of 18,314 invasive first primary breast cancer cases were included in the ABCPP.

Cohort-specific methods—The methodology of each cohort has been described previously (see: SBCSS [26], WHEL [31], LACE [32], NHS [33]); hence, we provide below only a brief description of data collection methods, focusing on those factors relevant for the pooling project. In addition, an overview of study methodology for each cohort is shown in Table 1. Briefly, each cohort collected data on clinical factors (tumor characteristics, treatment status), reproductive factors, family history of breast cancer, QOL, medical history including co-morbidities, anthropometric data, smoking history, alcohol intake, recreational physical activity, supplement use, and used a food frequency questionnaire (FFQ) to assess dietary intake.

For SBCSS, diet was assessed at baseline, 18-, 36-, and 60-months post-diagnosis using a validated FFQ that was specifically designed to capture nutrient and major food intake among Chinese women living in Shanghai [34]. Diet was assessed in WHEL using the Arizona Food Frequency Questionnaire (AFFQ) at baseline, year one, and year four. The AFFQ is a modification of the food frequency component of the Block Health Habits and History Questionnaire developed by Block and colleagues at the National Cancer Institute (NCI), and earlier versions have been validated against food recalls [35, 36]. For LACE, dietary habits were assessed using the Fred Hutchinson Cancer Research Center Food Frequency Questionnaire (FHRCC-FQ) at baseline and the follow-up survey at about five to six years post-diagnosis. The FHRCC-FQ questionnaire is a validated, self-administered, semi-quantitative FFQ with approximately 120 items and is an adaptation of the 95-item Health Habits and Lifestyle Questionnaire developed by Block and colleagues at the NCI [35]. The NHS assessed dietary habits using validated semiquantitative FFQs about every four years [33].

QOL was assessed at baseline in the SBCSS using the 36-item Short Form Health Survey (SF-36) [37] for 55.8% participants and the General QOL Inventory-74 (GQOLI-74) [38] for 44.2% of participants (The GQOLI-74 was re-administered at the 36-month interview). For WHEL and the NHS, QOL was assessed at multiple time points after diagnosis using the SF-36 [39]. LACE used the Functional Assessment of Cancer Therapy – Breast (FACT-B) [40] to asses QOL at the baseline and follow-up surveys.

Data Harmonization for the ABCPP

Overview of data harmonization procedures—Data harmonization procedures were initially planned at an in-person meeting with study investigators. At this meeting, variable definitions were developed and analytic protocols were drafted. Subsequently, a data dictionary was created for the pooling study variables and was distributed to cohort investigators to enable creation of a cohort-specific dataset with the data items needed for the pooling project. Monthly conference calls and frequent email communications were used to refine standard definitions for each variable and discuss any modifications of the analytic protocols, including handling of missing data and recoding of variables. After creation of the cohort-specific datasets with agreed upon variables, each study investigator sent a dataset to Vanderbilt University (VU), the data coordinating center for the ABCPP, for data checking and additional standardization as needed. The VU team created a merged dataset with standardized variables for the analysis phase of the project. Below we discuss the main common variables to be used in all the ABCPP investigations.

Clinical Characteristics—Harmonized data on clinical characteristics include age at diagnosis, tumor-node-metastasis (TNM) stage based on the AJCC 6th edition [41], ER/PR status, HER-2 status, histological grade, surgery, chemotherapy, radiotherapy, and endocrine therapy. For the US cohorts, tumor characteristics were based on medical records from the local hospitals. For the SBCSS, tumor characteristics and clinical data were verified by

medical records for 98.1% of cases. The agreement rates between self-reported and medical chart information ranged from 94–98%.

Socio-demographics and Reproductive Factors—Harmonized data on socio-demographics, reproductive factors, and family history of cancer assessed at baseline/first post-diagnosis survey include education, race/ethnicity, menopausal status, parity, age at first birth, and family history of breast cancer.

Lifestyle Factors and Co-morbidities—Harmonized data on lifestyle factors and co-morbidities assessed at baseline/first post-diagnosis survey and presented here include: alcohol consumption, smoking, BMI, recreational physical activity, diabetes, and hypertension. BMI was calculated as weight in kilograms (kg) divided by height in meters squared (m^2) and categorized using the World Health Organization (WHO) international classifications [42]: underweight (<18.5 kg/m^2), normal weight (18.5 to 24.99 kg/m^2), overweight (25 to 29.99 kg/m^2), obese (≥30 kg/m^2). Instead of using population-specific BMI cut points for Asian breast cancer survivors, we used the same BMI cut points for all participants (the WHO international classifications) [42], as were used previously in the SBCSS [18], to allow comparisons by cohort. All cohorts provided recreational physical activity levels converted into metabolic equivalents (METs) [43] in MET-hours/week. For the present analysis, physical activity was categorized based on tertiles.

Outcomes—Harmonized outcome data include: total mortality (death from any cause), breast cancer-specific mortality (death from breast cancer), and a new breast cancer event, hereafter referred to as recurrence. Recurrence includes a local/regional recurrence, distant recurrence/metastasis, or development of new primary breast cancer in the ipsilateral or contralateral breast) [16, 44]. New primary breast cancers were recorded only in the SBCSS, WHEL, and LACE cohorts.

Data Analysis

Statistical definitions for clinical characteristics, socio-demographics, reproductive factors, lifestyle factors, and co-morbidities are shown in the Tables. Frequency distributions for categorical variables and means with standard deviations (SD) for continuous variables were calculated for harmonized data by study cohort and for the all cohorts combined. Women missing the date of the first survey/enrollment or with no follow-up time (n=22) were excluded from the analyses. Delayed entry Cox proportional hazards regression models were used to estimate five-year survival functions for each study separately, adjusting for age at diagnosis. Entry time began at baseline/first survey after diagnosis (mean of 6.5 months post-diagnosis for the SBCSS, 2 years post-diagnosis for WHEL and LACE, and 1 year post-diagnosis for the NHS). Follow-up ended at date of death or date of last contact (i.e., date of last follow-up survey or date of last registry linkage, whichever was most recent).

Investigators of each individual cohort received institutional review board approval from their home institution(s) to participate in the ABCPP. Data use agreements were also signed by the principal investigators for each institution.

Future Multivariable Analytic Plan for the ABCPP

A general analytic plan was developed for multivariable analyses of data from the ABCPP. First, we will analyze individual study data using delayed entry Cox proportional hazards models, with time since diagnosis as the time scale. The entry date will be the date of the baseline survey for the SBCSS, WHEL, and LACE or the date of the first survey after diagnosis with measurement of the exposure of interest for the NHS. Second, we will conduct meta-analyses with the study-specific hazards ratios using inverse-variance weights

in random-effects models [45]. The Q test statistic will be used to test for heterogeneity in risk estimates across studies [46]. If heterogeneity is present, study-specific estimates and the pooled hazard ratios from the random-effects models will be presented. In addition, the reasons for heterogeneity will be investigated, both between studies and for relevant subgroups overall and within a cohort, as appropriate. Sensitivity analyses will be conducted to exclude specific studies or subgroups, as appropriate, depending on the research question of interest. For example, to address potential heterogeneity due to changes in breast cancer treatment over time (diagnosis dates range from 1976–2006) and geographical differences (China as compared to US breast cancer survivors), sensitivity analyses that exclude cases diagnosed in earlier years (or Chinese breast cancer patients) will be conducted. If heterogeneity is not present, a pooled analysis using the combined individual data from each cohort will be conducted for the exposure-disease associations of interest using delayed entry Cox proportional hazards regression models stratified by study.

RESULTS

The ABCPP includes 18,314 women aged 20–83 years diagnosed with invasive primary breast cancer between 1976 and 2006. Table 1 displays mean years of follow-up and major endpoints by cohort. After a mean follow-up of 8.4 years, 3,736 total deaths, 2,614 breast cancer-specific deaths, and 3,564 breast cancer recurrences have been documented across the four cohorts. The majority of deaths (71%) were due to breast cancer; 11.4% were from other malignancies, 7.3% from cardiovascular disease, and 10.4% from other causes.

Clinical Characteristics

Table 2 displays age, tumor characteristics, and treatment data by study and combined. Mean age (SD) at diagnosis was younger for the SBCSS (53.5 (10.0)) and WHEL (51.2 (8.9)), than for LACE (58.3 (11.0)) and the NHS (60.4 (9.4)). Across studies, SBCSS participants had the lowest percentage of ER+/PR+ tumors (51.0%) and LACE participants had the highest (68.3%). About 41.0% of participants had HER-2 data available. SBCSS participants had the highest percentage of HER-2+ tumors (30.5%) and WHEL participants had the lowest (8.3%). The NHS had a large amount of missing data for HER-2 status, because HER-2 status was not routinely assessed until around 2005, when results from adjuvant trastuzumab trials were published [47].

Information on whether women received adjuvant chemotherapy, radiotherapy, and hormonal therapy was available for the majority of the ABCPP participants (Table 2). However, detailed treatment regimen data on dose and duration were not available for all cohorts and these data were not pooled. Most SBCSS participants had a mastectomy (94.0%) and received chemotherapy (92.2%), whereas only about one third received radiotherapy, reflecting current treatment trends in China [48]. Across the three US cohorts, mastectomy and radiotherapy percentages were similar; however, chemotherapy percentages were lowest for the NHS (41.5%), compared to WHEL (70.0%) and LACE (57.3%), reflecting the lower stage and earlier treatment period (i.e., 1980’s) among NHS participants. As shown in Table 2, the NHS had a large amount of missing data for many of the clinical characteristics, in particular for cases diagnosed in the 1970’s and 1980’s. Hence, we also display the distribution of this cohort for cases diagnosed from 1990 onward, which is more comparable to the diagnosis dates of the other cohorts. When we limited the NHS cohort to cases diagnosed from 1990 onward, the amount of missing data for clinical characteristics was reduced, in particular for cancer treatment.

Table 3 shows select clinical characteristics by race/ethnicity (Non-Hispanic White, Non-Hispanic Black, Hispanic, and Asian (Asian American and Chinese)). Non-Hispanic White and Asian American women had a higher proportion of stage I breast tumors and a lower

proportion of stage II breast tumors, compared with the other race/ethnicities. The percentage of ER−/PR− breast tumors was highest among Non-Hispanic Black women (32.0%) and lowest among Asian American women (13.5%). Chinese women had the highest percentage of HER-2+ tumors and Non-Hispanic White women had the lowest.

Socio-demographics, Reproductive Factors, and Family History

Table 4 shows selected non-clinical participant characteristics assessed at the baseline/first post-diagnosis survey (unless otherwise indicated) by study and combined. Education levels were lower for women in the SBCSS (only about 15.8% had an education level above high school). Education level was college graduate or higher for all participants of the NHS, who were all registered nurses [33]. For the US cohorts, the majority of participants were Non-Hispanic White. The percentage of breast cancer survivors with a first-degree family history of breast cancer was low for participants of the SBCSS (5.6%), compared with about 13–21% in the US cohorts. About half of participants were postmenopausal close to (within approximately six months) or at diagnosis for both SBCSS and WHEL. For LACE, 63.7% of women were postmenopausal at diagnosis and for NHS 81.6% were postmenopausal (assessment on average 1 year post-diagnosis).

Lifestyle Factors and Co-morbidities

Table 4 presents data separately by study and overall for select lifestyle factors and co-morbidities. Overall, 52.3% of US participants reported drinking alcoholic beverages at least monthly and 9.8% currently smoked, whereas 42.0% were former smokers. In contrast, few women in the SBCSS reported alcohol consumption (0.3%) or ever smoking (2.7%). For BMI and exercise, we included data from the 18-month post-diagnosis survey for SBCSS participants, rather than the baseline 6-month post-diagnosis survey, since the other cohorts assessed these factors on average one to two years post-diagnosis. Using the WHO international classifications for obesity and overweight, overall, 17.9% of breast cancer survivors were obese at approximately one to two years post-diagnosis and 32.6% were overweight. Percentages of overweight and obese women tended to be fairly similar for WHEL, LACE, and the NHS, although the percentage of obese women was lower for the NHS. Chinese women were leaner with an obesity prevalence of about 6.3%, compared with 19% or higher for US women. The NHS had the greatest proportion of women reporting any recreational physical activity at one to two years post-diagnosis. Among breast cancer survivors who were participating in recreational physical activity at least monthly (n=14,390), 59.9% met the 2008 Physical Activity Guidelines for Americans (at least 8.3 MET-hours per week) [49]. About 7% of SBCSS, LACE, and NHS participants reported diabetes at the baseline/first survey after diagnosis; only 2.2% of WHEL participants had diabetes at baseline.

Age-Adjusted Five-Year Survival Overall and by TNM Stage at Diagnosis

Table 5 gives age-adjusted five year survival estimates overall and by stage, before and after exclusions, to account for differences by cohort in the timing of enrollment in relation to cancer diagnosis and year of diagnosis. First, the start of follow-up was changed to begin two years after diagnosis for all four cohorts, to standardize approximately the timing of enrollment in relation to diagnosis, which resulted in the exclusion of 438 participants. Two years was selected because the mean time between diagnosis and enrollment was two years for WHEL and LACE cohorts. Second, because treatment regimens have changed over time [50, 51], we excluded 2,793 NHS participants diagnosed prior to 1991. After these exclusions, five-year survival estimates were similar across the cohorts for women with TNM stage I and II breast tumors. For stage III tumors, more variation in five-year survival rates was observed, which ranged from 77.3% in the SBCSS to 93.2% in WHEL. The higher

survival rates for stage III tumors in LACE and WHEL reflect the exclusion at recruitment of women with tumors with extension to the chest wall.

STRENGTHS AND LIMITATIONS

The ABCPP includes pooled and harmonized data on clinical characteristics, socio-demographics, reproductive factors, co-morbidities, and select post-diagnosis lifestyle factors for over 18,000 breast cancer survivors. We are currently using the ABCPP to comprehensively evaluate the associations of physical activity, adiposity, dietary intake of soy and cruciferous vegetables, dietary supplement use, and QOL in relation to breast cancer outcomes. We have established a methodology for pooling and harmonizing data that can be used to expand the ABCPP in the future with the addition of other breast cancer survivor cohorts. The long-term goal of this project is to promote sustained collaborations with current ABCPP investigators, and to initiate new collaborations with other breast cancer survivor cohorts to develop large-scale international investigations of modifiable predictors of breast cancer outcomes. As part of this effort, ABCPP investigators are members of the Breast Cancer Consortium for Outcomes and Survival (BC^2OS), which is supported by a NIH web portal. For more information or to become a member of BC^2OS see: http://epi.grants.cancer.gov/Consortia/single/bc2os.html.

Several strengths of the ABCPP should be considered. First, the pooling of cohorts into one aggregate dataset will provide adequate statistical power to investigate heterogeneity in associations of lifestyle factors and breast cancer outcomes by specific tumor subtypes and other potential effect modifiers (e.g., stage, treatment status, age, and menopausal status), which has not been possible in previous single cohort investigations. Second, as described in the Results, we have successfully pooled and harmonized data from the four cohorts, providing the foundation for testing study hypotheses using a standardized protocol and setting the stage for future collaborative research. Third, by pooling the individual data, rather than conducting a meta-analysis of published effect estimates, we were able to standardize the definitions of exposures, potential confounders, and effect modifiers, as well as the analytic approach, which will decrease the heterogeneity by study and provide more precise estimations of associations [45].

Limitations of the ABCPP should also be considered. First, although the sample size is large, 10% or more of participants have missing data for several key clinical characteristics (e.g., cancer treatment and ER/PR status). However, a large portion of this missing data is due to the inclusion of NHS participants diagnosed in the 1970's and 1980's, when information on tumor characteristics and treatment history was less complete. Hence, depending on the specific aims under study, investigators can decide to exclude these earlier cases to reduce the amount of missing data. Regardless of the analytic exclusions, sensitivity analyses to investigate the influence of missing data on results will be conducted. Second, while we were able to compare results for Chinese women to US women of all race/ethnicities, the sample size for other racial/ethnic minority groups among the US participants was small, which limits our ability to investigate the associations of lifestyle factors, QOL, and breast cancer prognosis among women of these races/ethnicities. Third, we found differences in clinical characteristics, reproductive factors, BMI, physical activity, smoking, alcohol intake, and select co-morbidities across cohorts. However, we will be able to adjust for these factors in multivariable analyses using a standardized protocol and we also plan to estimate study-specific effect estimates, as well as pooled estimates, from either an aggregate analysis (pooled analysis of individual cohort data) or meta-analysis (pooled study-specific effect estimates), depending on the results of tests for heterogeneity by study. In addition, as mentioned in the Materials and Methods, sensitivity analyses will be conducted to exclude specific studies or subgroups, as appropriate. For example, to address

potential heterogeneity due to changes in breast cancer treatment over time or cultural differences, sensitivity analyses that exclude cases diagnosed in earlier years (i.e., before 1990) or are limited to the US cohorts, will be conducted. Fourth, some potential confounders and modifying factors will not be included in analyses, because the factors are not available in all four cohorts (e.g., income, total energy intake).

As is inherent in pooling projects of individual studies [45, 52], the data collection methods were different for each cohort included in the ABCPP. These differences resulted in challenges in data harmonization. Specifically, one challenge we faced was the standardization of the definition of breast cancer recurrences. Both LACE and WHEL had information on local, regional, and distant recurrences and new breast primaries. However, in the SBCSS, data were not available on local or regional recurrences as the proportion of women who had a mastectomy was high (94%), unlike in the US studies where only about 50% had a mastectomy. Furthermore, for the NHS, although some data was collected on second malignancies after cancer diagnosis, this information was not complete; hence, we do not have information on new breast primaries from the NHS in the ABCPP. To address these data harmonization issues, we will include two sensitivity analyses in future analytic protocols. These will include: (1) the exclusion of local and regional recurrences; and (2) the exclusion of women with new breast primaries.

In summary, the ABCPP will allow a comprehensive evaluation of the associations of select post-diagnosis lifestyle factors and QOL in relation to breast cancer prognosis, with a large sample size to enable the consideration of heterogeneity by tumor subtypes and other clinical factors such as treatment status and age. Further, we have established an infrastructure that will allow additional breast cancer survivor cohorts to be added to the ABCPP, and will promote large-scale international collaborations on modifiable factors and breast cancer outcomes research.

Acknowledgments

Financial Support: This study was supported by the National Cancer Institute (Grant number 3R01CA118229-03S1). The parent grants for each individual cohort included in the After Breast Cancer Pooling Project are as follows: SBCSS (Department of Defense, DAMD 17-02-1-0607 and National Cancer Institute, R01 CA118229), WHEL (Susan G. Komen Foundation, #KG100988), LACE (National Cancer Institute, R01 CA129059), and NHS (P01 CA87969).

The authors thank investigators and staff members of the research teams and study participants for their contributions and support for the study.

References

1. Michels KB, Mohllajee AP, Roset-Bahmanyar E, Beehler GP, Moysich KB. Diet and breast cancer: a review of the prospective observational studies. Cancer. 2007; 109:2712–2749. [PubMed: 17503428]
2. Hamajima N, Hirose K, Tajima K, Rohan T, Calle EE, Heath CW Jr, et al. Alcohol, tobacco and breast cancer--collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease. Br J Cancer. 2002; 87:1234–1245. [PubMed: 12439712]
3. Hankinson SE, Colditz GA, Willett WC. Towards an integrated model for breast cancer etiology: the lifelong interplay of genes, lifestyle, and hormones. Breast Cancer Res. 2004; 6:213–218. [PubMed: 15318928]
4. Friedenreich CM. The role of physical activity in breast cancer etiology. Semin Oncol. 2010; 37:297–302. [PubMed: 20709211]
5. Kushi LH, Kwan ML, Lee MM, Ambrosone CB. Lifestyle factors and survival in women with breast cancer. J Nutr. 2007; 137:236S–242S. [PubMed: 17182833]

6. Doyle C, Kushi LH, Byers T, Courneya KS, Demark-Wahnefried W, Grant B, et al. Nutrition and physical activity during and after cancer treatment: an American Cancer Society guide for informed choices. CA Cancer J Clin. 2006; 56:323–353. [PubMed: 17135691]
7. Norman SA, Potashnik SL, Galantino ML, De Michele AM, House L, Localio AR. Modifiable risk factors for breast cancer recurrence: what can we tell survivors? J Womens Health (Larchmt). 2007; 16:177–190. [PubMed: 17388734]
8. Rock CL, Demark-Wahnefried W. Nutrition and survival after the diagnosis of breast cancer: a review of the evidence. J Clin Oncol. 2002; 20:3302–3316. [PubMed: 12149305]
9. McTiernan A, Irwin M, VonGruenigen V. Weight, Physical Activity, Diet, and Prognosis in Breast and Gynecologic Cancers. J Clin Oncol. 2010; 28:4074–4080. [PubMed: 20644095]
10. Kellen E, Vansant G, Christiaens MR, Neven P, Van Limbergen E. Lifestyle changes and breast cancer prognosis: a review. Breast Cancer Res Treat. 2009; 114:13–22. [PubMed: 18389367]
11. Wei EK, Wolin KY, Colditz GA. Time course of risk factors in cancer etiology and progression. J Clin Oncol. 2010; 28:4052–4057. [PubMed: 20644083]
12. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin. 2005; 55:74–108. [PubMed: 15761078]
13. Patterson RE, Cadmus LA, Emond JA, Pierce JP. Physical activity, diet, adiposity and female breast cancer prognosis: a review of the epidemiologic literature. Maturitas. 2010; 66:5–15. [PubMed: 20097494]
14. Carmichael AR, Daley AJ, Rea DW, Bowden SJ. Physical activity and breast cancer outcome: a brief review of evidence, current practice and future direction. Eur J Surg Oncol. 2010; 36:1139–1148. [PubMed: 20947287]
15. Ibrahim EM, Al-Homaidh A. Physical activity and survival after breast cancer diagnosis: meta-analysis of published studies. Med Oncol. 201010.1007/s12032-010-9536-x
16. Bertram LA, Stefanick ML, Saquib N, Natarajan L, Patterson RE, Bardwell W, et al. Physical activity, additional breast cancer events, and mortality among early-stage breast cancer survivors: findings from the WHEL Study. Cancer Causes Control. 2010; 22:427–35. [PubMed: 21184262]
17. Protani M, Coory M, Martin JH. Effect of obesity on survival of women with breast cancer: systematic review and meta-analysis. Breast Cancer Res Tr. 2010; 123:627–35.
18. Chen XL, Lu W, Zheng W, Gu K, Chen Z, Zheng Y, et al. Obesity and weight change in relation to breast cancer survival. Breast Cancer Res Tr. 2010; 122:823–833.
19. Ewertz M, Jensen MB, Gunnarsdottir KA, Hojris I, Jakobsen EH, Nielsen D, et al. Effect of Obesity on Prognosis After Early-Stage Breast Cancer. J Clin Oncol. 2011; 29:25–31. [PubMed: 21115856]
20. Falagas ME, Zarkadoulia EA, Ioannidou EN, Peppas G, Christodoulou C, Rafailidis PI. The effect of psychosocial factors on breast cancer outcome: a systematic review. Breast Cancer Res. 2007; 9:R44. [PubMed: 17640330]
21. Gotay CC, Kawamoto CT, Bottomley A, Efficace F. The prognostic significance of patient-reported outcomes in cancer clinical trials. J Clin Oncol. 2008; 26:1355–1363. [PubMed: 18227528]
22. Saquib N, Pierce JP, Saquib J, Flatt SW, Natarajan L, Bardwell WA, et al. Poor physical health predicts time to additional breast cancer events and mortality in breast cancer survivors. Psychooncology. 2011; 20:252–9. [PubMed: 20878837]
23. Patterson RE, Saquib N, Natarajan L, Rock CL, Parker BA, Thomson CA, et al. Improvement in self-reported physical health predicts longer survival among women with a history of breast cancer. Breast Cancer Res Treat. 2011; 127:541–7. [PubMed: 21042931]
24. Thomson CA, Rock CL, Thompson PA, Caan BJ, Cussler E, Flatt SW, et al. Vegetable intake is associated with reduced breast cancer recurrence in tamoxifen users: a secondary analysis from the Women's Healthy Eating and Living Study. Breast Cancer Res Treat. 2011; 125:519–27. [PubMed: 20607600]
25. Guha N, Kwan ML, Quesenberry CP Jr, Weltzien EK, Castillo AL, Caan BJ. Soy isoflavones and risk of cancer recurrence in a cohort of breast cancer survivors: the Life After Cancer Epidemiology study. Breast Cancer Res Treat. 2009; 118:395–405. [PubMed: 19221874]

26. Shu XO, Zheng Y, Cai H, Gu K, Chen Z, Zheng W, et al. Soy food intake and breast cancer survival. JAMA. 2009; 302:2437–2443. [PubMed: 19996398]
27. Kang X, Zhang Q, Wang S, Huang X, Jin S. Effect of soy isoflavones on breast cancer recurrence and death for patients receiving adjuvant endocrine therapy. CMAJ. 2010; 182:1857–62. [PubMed: 20956506]
28. Fleischauer AT, Simonsen N, Arab L. Antioxidant supplements and risk of breast cancer recurrence and breast cancer-related mortality among postmenopausal women. Nutr Cancer. 2003; 46:15–22. [PubMed: 12925299]
29. Nechuta S, Lu W, Chen Z, Zheng Y, Gu K, Cai H, et al. Vitamin Supplement Use During Breast Cancer Treatment and Survival: A Prospective Cohort Study. Cancer Epidem Biomar. 2011; 20:262–271.
30. Blows FM, Driver KE, Schmidt MK, Broeks A, van Leeuwen FE, Wesseling J, et al. Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. PLoS Med. 2010; 7:e1000279. [PubMed: 20520800]
31. Pierce JP, Faerber S, Wright FA, Rock CL, Newman V, Flatt SW, et al. A randomized trial of the effect of a plant-based dietary pattern on additional breast cancer events and survival: the Women's Healthy Eating and Living (WHEL) Study. Control Clin Trials. 2002; 23:728–756. [PubMed: 12505249]
32. Caan B, Sternfeld B, Gunderson E, Coates A, Quesenberry C, Slattery ML. Life After Cancer Epidemiology (LACE) study: A cohort of early stage breast cancer survivors (United states). Cancer Cause Control. 2005; 16:545–556.
33. Colditz GA, Hankinson SE. The Nurses' Health Study: Lifestyle and health among women. Nat Rev Cancer. 2005; 5:388–396. [PubMed: 15864280]
34. Shu XO, Yang G, Jin F, Liu D, Kushi L, Wen W, et al. Validity and reproducibility of the food frequency questionnaire used in the Shanghai Women's Health Study. Eur J Clin Nutr. 2004; 58:17–23. [PubMed: 14679362]
35. Block G, Hartman AM, Dresser CM, Carroll MD, Gannon J, Gardner L. A databased approach to diet questionnaire design and testing. Am J Epidemiol. 1986; 124:453–469. [PubMed: 3740045]
36. Ritenbaugh C, Aickin M, Taren D, Teufel N, Graver E, Woolf K, et al. Use of a food frequency questionnaire to screen for dietary eligibility in a randomized cancer prevention phase III trial. Cancer Epidem Biomar. 1997; 6:347–354.
37. Li L, Wang HM, Shen Y. Chinese SF-36 Health Survey: translation, cultural adaptation, validation, and normalisation. J Epidemiol Commun H. 2003; 57:259–263.
38. Li L, Wei H, Young D. The development of the General Quality of life Inventory. China Ment Health J. 1995; 9:227–231.
39. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992; 30:473–483. [PubMed: 1593914]
40. Brady MJ, Cella DF, Mo F, Bonomi AE, Tulsky DS, Lloyd SR, et al. Reliability and validity of the Functional Assessment of Cancer Therapy-Breast quality-of-life instrument. J Clin Oncol. 1997; 15:974–986. [PubMed: 9060536]
41. AJCC. American Joint Committee on Cancer staging manual. 6. Springer-Verlag; New York: 2002.
42. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet. 2004; 363:157–163. [PubMed: 14726171]
43. Ainsworth BE, Haskell WL, Leon AS, Jacobs DR Jr, Montoye HJ, Sallis JF, et al. Compendium of physical activities: classification of energy costs of human physical activities. Med Sci Sports Exerc. 1993; 25:71–80. [PubMed: 8292105]
44. Caan BJ, Kwan ML, Hartzell G, Castillo A, Slattery ML, Sternfeld B, et al. Pre-diagnosis body mass index, post-diagnosis weight change, and prognosis among women with early stage breast cancer. Cancer Cause Control. 2008; 19:1319–1328.
45. Smith-Warner SA, Spiegelman D, Ritz J, Albanes D, Beeson WL, Bernstein L, et al. Methods for pooling results of epidemiologic studies: the Pooling Project of Prospective Studies of Diet and Cancer. Am J Epidemiol. 2006; 163:1053–1064. [PubMed: 16624970]

46. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986; 7:177–188. [PubMed: 3802833]
47. Guarneri V, Barbieri E, Dieci MV, Piacentini F, Conte P. Anti-HER2 neoadjuvant and adjuvant therapies in HER2 positive breast cancer. Cancer Treat Rev. 2010; 36S3:S62–S66. [PubMed: 21129613]
48. Fan L, Zheng Y, Yu KD, Liu GY, Wu J, Lu JS, et al. Breast cancer in a transitional society over 18 years: trends and present status in Shanghai, China. Breast Cancer Res Treat. 2009; 117:409–416. [PubMed: 19153831]
49. USDHHS. 2008 physical activity guidelines for Americans. Washington, DC: USDHHS; 2008.
50. Gradishar WJ, O'Regan RM. Progress in systemic adjuvant therapy of early-stage breast cancer. Int J Clin Oncol. 2003; 8:239–247. [PubMed: 12955580]
51. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet. 2005; 365:1687–1717. [PubMed: 15894097]
52. Gallicchio L, Helzlsouer KJ, Chow WH, Freedman DM, Hankinson SE, Hartge P, et al. Circulating 25-hydroxyvitamin D and the risk of rarer cancers: Design and methods of the Cohort Consortium Vitamin D Pooling Project of Rarer Cancers. Am J Epidemiol. 2010; 172:10–20. [PubMed: 20562188]

Table 1

Cohort study characteristics, follow-up, and events in the After Breast Cancer Pooling Project (N=18,314)

	SBCSS	WHEL	LACE	NHS
	n=4,886	**n=3,088**	**n=2,265**	**n=8,075**
Study Design	Prospective cohort	Prospective follow-up of participants of a randomized controlled clinical trial	Prospective cohort	Prospective cohort
Location	Shanghai, China	Seven sites in the western United States	California, Utah, and WHEL sites	United States
Age at diagnosis, range	20–75	26–70	25–79	30–83
Year of diagnosis, range	2002–2006	1991–2000	1997–2000	1976–2004
Year of baseline recruitment, range	2002–2006	1995–2000	2000–2002	1976
Months between diagnosis and baseline survey/first survey after diagnosis, mean (range)	6.5 (3.6 to 11.4)	23.5 (1.9 to 48.3)	22.7 (11.1 to 38.9)	12.3 (0.91–45.9)
Recruitment methods	Cases were identified from the population-based Shanghai Cancer Registry	Tumor registry, community outreach, physician referral	Cases were identified from the Kaiser Permanente Northern California Cancer Registry (83%), Utah Cancer Registry (12%), and WHEL non-participants (5%)	Cases were identified through self-report or through death records
Baseline data collection methods	In-person interview at ~ 6 months after cancer diagnosis, medical records	Telephone interviews, clinic visits, medical records	Mailed questionnaire on average 2 years after cancer diagnosis	Mailed questionnaire to nurse or next of kin after report of diagnosis
Follow-up data collection methods	In-person interviews at ~18 months, 36 months, and 60 months post-diagnosis, linkage to Shanghai Vital Statistics Registry	4 follow-up clinic visits over 6 years; semi-annual telephone calls throughout surveillance period, medical records	Mailed questionnaire 5–6 years post-baseline, and mailed semi-annual or annual (after April 2005) health status update questionnaire	Mailed questionnaire every 2 years
Years of follow-up, mean (range)	4.2 years (0.5–6.2)	9.3 years (0.35–15.2)	7.8 years (0.10–10.3)	10.7 years (0.08–30.0)
Major Endpoints, n				
Deaths	485	452	376	2,423
Breast Cancer Deaths	423	356	207	1,628
Recurrence[a]	624	583	366	1,991

SBCSS Shanghai Breast Cancer Survival Study, *WHEL* Women's Healthy Eating and Living Study, *LACE* Life After Cancer Epidemiology Study, *NHS* Nurses' Health Study

[a]Defined as a local/regional recurrence, distant recurrence/metastasis or development of a new breast primary.

Table 2

Clinical characteristics by study and combined, After Breast Cancer Pooling Project[a]

	SBCSS (n=4,886)		WHEL (n=3,088)		LACE (n=2,265)		NHS				Combined (n=18,314)	
							All cases (n=8,075)		Cases diagnosed ≥ 1990 (n=5,115)			
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
Age at diagnosis (years)												
Mean (SD)	53.5 (10.0)		51.2 (8.9)		58.3 (11.0)		60.4 (9.4)		64.5 (7.8)		56.6 (10.4)	
TNM Stage (AJCC 6th edition)[b]												
I	1,679	(36.1)	1,191	(38.6)	1,057	(46.7)	4,044	(53.8)	2,915	(59.7)	7,971	(45.5)
II	2,209	(47.4)	1,410	(45.7)	949	(41.9)	2,295	(30.5)	1,389	(28.5)	6,863	(39.2)
III	740	(15.9)	487	(15.8)	258	(11.4)	935	(12.4)	459	(9.4)	2,420	(13.8)
IV	28	(0.6)	0		0		248	(3.3)	119	(2.4)	276	(1.6)
Missing	230		0		1		553		233		784	
ER/PR status												
ER+/PR+	2,439	(51.0)	1,908	(63.1)	1,530	(68.3)	3,848	(61.6)	2,913	(64.8)	9,725	(59.7)
ER+/PR−	635	(13.3)	368	(12.2)	319	(14.2)	1,068	(17.1)	745	(16.6)	2,390	(14.7)
ER−/PR+	362	(7.6)	129	(4.3)	41	(1.8)	219	(3.5)	132	(2.9)	751	(4.6)
ER−/PR−	1,350	(28.2)	618	(20.4)	350	(15.6)	1,112	(17.8)	705	(15.7)	3,430	(21.1)
Missing	100		65		25		1,828		620		2,018	
HER2												
Positive	837	(30.5)	180	(8.3)	320	(16.3)	161	(12.9)	161	(13.0)	1,498	(18.5)
Negative	1,689	(61.4)	1,775	(82.2)	1,643	(83.6)	967	(77.7)	965	(77.7)	6,074	(74.8)
Indeterminate	223	(8.1)	205	(9.5)	2	(0.1)	116	(9.3)	116	(9.3)	546	(6.7)
Missing	2,137		928		300		6,831		3,873		10,196	
Histological grade												
Well differentiated	435	(17.1)	484	(17.1)	430	(21.0)	943	(20.6)	837	(22.3)	2,292	(19.1)
Moderately					943	(46.0)	1,954	(42.7)	1,675	(44.5)	5,441	(45.3)
differentiated	1,304	(51.2)	1,240	(43.8)								
Poorly differentiated	806	(31.7)	1,108	(39.1)	679	(33.1)	1,682	(36.7)	1,250	(33.2)	4,275	(35.6)
Missing	2,341		256		213		3,496		1,353		6,306	
Surgery												

	SBCSS (n=4,886)		WHEL (n=3,088)		LACE (n=2,265)		NHS				Combined (n=18,314)	
							All cases (n=8,075)		Cases diagnosed ≥ 1990 (n=5,115)			
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
None	9	(0.2)	1	(0.03)	0	(0.0)	92	(1.4)	40	(0.9)	102	(0.6)
Lumpectomy	45	(0.9)	1,474	(47.7)	1,146	(50.6)	2,835	(43.8)	2,476	(52.8)	5,500	(32.9)
Mastectomy	4,592	(94.0)	1,613	(52.2)	1,119	(49.4)	3,548	(54.8)	2,178	(46.4)	10,872	(65.1)
Unknown type	240	(4.9)	0		0		0		0		240	(1.4)
Missing	0		0		0		1,600		421		1,600	
Chemotherapy[c]												
No	382	(7.8)	927	(30.0)	967	(42.7)	3,308	(58.5)	2,722	(60.9)	5,584	(35.1)
Yes	4,504	(92.2)	2,159	(70.0)	1,297	(57.3)	2,350	(41.5)	1,748	(39.1)	10,310	(64.9)
Missing	0		2		1		2,417		645		2,420	
Radiotherapy												
No	3,286	(67.3)	1,185	(38.4)	837	(37.0)	2,520	(43.1)	1,849	(39.7)	7,828	(48.7)
Yes	1,600	(32.8)	1,899	(61.6)	1,428	(63.1)	3,330	(56.9)	2,810	(60.3)	8,257	(51.3)
Missing	0		4		0		2,225		456		2,229	
Endocrine therapy[c]												
No	2,315	(47.5)	978	(31.7)	451	(20.1)	2,374	(36.7)	1,270	(27.1)	6,118	(36.7)
Yes	2,555	(52.5)	2,106	(68.3)	1,793	(79.9)	4,102	(63.3)	3,422	(72.9)	10,556	(63.3)
Missing	16		4		21		1,599		423		1,640	

SBCSS Shanghai Breast Cancer Survival Study, *WHEL* Women's Healthy Eating and Living Study, *LACE* Life After Cancer Epidemiology Study, *NHS* Nurses' Health Study

[a] All percentages exclude missing data.

[b] WHEL and LACE did not include women diagnosed with stage IIIB or IV breast cancer.

[c] Based on one measure at baseline/first post-diagnosis survey.

Table 3

Select clinical prognostic factors by race/ethnicity, After Breast Cancer Pooling Project[a]

	Non-Hispanic White (n=12,015)		Non-Hispanic Black (n=357)		Hispanic (n=387)		Asian American (n=292)		Chinese (n=4,886)	
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
TNM Stage										
I	5,698	(49.5)	126	(36.5)	145	(38.3)	134	(47.0)	1,679	(36.1)
II	4,059	(35.3)	161	(46.7)	182	(48.0)	121	(42.5)	2,209	(47.4)
III	1,515	(13.2)	53	(15.4)	51	(13.5)	29	(10.2)	740	(15.9)
IV	239	(2.1)	5	(1.5)	1	(0.3)	1	(0.4)	28	(0.6)
Missing	504		12		8		7		230	
ER/PR status										
ER+/PR+	6,521	(63.6)	142	(46.9)	229	(63.1)	190	(69.1)	2,439	(51.0)
ER+/PR−	1,582	(15.4)	48	(15.8)	40	(11.0)	38	(13.8)	635	(13.3)
ER−/PR+	338	(3.3)	16	(5.3)	14	(3.9)	10	(3.6)	362	(7.6)
ER−/PR−	1,810	(17.7)	97	(32.0)	80	(22.0)	37	(13.5)	1,350	(28.2)
Missing	1,764		54		24		17		100	
HER2										
Positive	532	(11.7)	33	(15.6)	37	(13.5)	41	(20.4)	837	(30.5)
Negative	3,736	(82.1)	164	(77.4)	225	(82.1)	155	(77.1)	1,689	(61.4)
Indeterminate	284	(6.2)	15	(7.1)	12	(4.4)	5	(2.5)	223	(8.1)
Missing	7,463		145		113		91		2,137	
Chemotherapy[b]										
No	4,716	(48.4)	104	(33.9)	125	(34.8)	93	(33.3)	382	(7.8)
Yes	5,027	(51.6)	203	(66.1)	234	(65.2)	186	(66.7)	4,504	(92.2)
Missing	2,272		50		28		13		0	
Radiotherapy										
No	4,030	(40.6)	111	(35.9)	140	(38.8)	124	(44.9)	3,286	(67.3)
Yes	5,897	(59.4)	198	(64.1)	221	(61.2)	152	(55.1)	1,600	(32.8)
Missing	2,088		48		26		16		0	
Endocrine therapy[b]										
No	3,392	(32.3)	123	(39.2)	103	(28.3)	63	(22.5)	2,315	(47.5)

	Non-Hispanic White (n=12,015)		Non-Hispanic Black (n=357)		Hispanic (n=387)		Asian American (n=292)		Chinese (n=4,886)	
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
Yes	7,103	(67.7)	191	(60.8)	261	(71.7)	217	(77.5)	2,555	(52.5)
Missing	1,520		43		23		12		16	

[a] Table excludes 371 women of other races/ethnicities.

[b] Based on one measure at baseline/first post-diagnosis survey.

Table 4

Select non-clinical participant characteristics assessed at baseline/first post-diagnosis survey by study and combined, After Breast Cancer Pooling Project[a]

	SBCSS (n=4,886)		WHEL (n=3,088)		LACE (n=2,265)		NHS (n=8,075)		Combined (n=18,314)	
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
Education										
< High school	2,278	(46.6)	25	(0.81)	125	(5.5)	0		2,428	(13.3)
High school	1,839	(37.6)	441	(14.3)	494	(21.9)	0		2,774	(15.2)
Some college/technical school	433	(8.9)	948	(30.7)	841	(37.3)	0		2,222	(12.1)
≥ College graduate	336	(6.9)	1,674	(54.2)	796	(35.3)	8,075	(100.0)	10,881	(59.4)
Missing	0		0		9		0		9	
Race/ethnicity										
Non-Hispanic White	0		2,634	(85.3)	1,809	(80.1)	7,572	(93.8)	12,015	(65.6)
Non-Hispanic Black	0		118	(3.8)	120	(5.3)	119	(1.5)	357	(2.0)
Asian	4,886	(100)	96	(3.1)	140	(6.2)	56	(0.69)	5,178	(28.3)
Hispanic	0		165	(5.3)	150	(6.6)	72	(0.89)	387	(2.1)
Other	0		75	(2.4)	40	(1.8)	256	(3.2)	371	(2.0)
Missing	0		0		6		0		6	
Family history of breast cancer										
No	4,614	(94.4)	2,445	(80.6)	1797	(79.4)	6,923	(86.7)	15,779	(86.9)
Yes	272	(5.6)	588	(19.4)	465	(20.6)	1,059	(13.3)	2,384	(13.1)
Missing	0		55		3		93		151	
Menopausal status[b]										
Premenopausal	2,384	(48.8)	1,570	(52.6)	514	(22.7)	1,123	(14.6)	5,591	(31.3)
Postmenopausal	2,502	(51.2)	1,416	(47.4)	1,441	(63.7)	6,298	(81.6)	11,657	(65.3)
Unclear	0		0		307	(13.6)	296	(3.8)	603	(3.4)
Missing	0		102		3		358		463	
Parity										
Nulliparous	198	(4.1)	683	(22.3)	362	(16.0)	559	(7.1)	1,802	(9.9)
1	961	(19.7)	465	(15.2)	266	(11.8)	655	(8.3)	2,347	(13.0)
2	1,616	(33.1)	1,071	(34.9)	713	(31.5)	2,160	(27.3)	5,560	(30.7)

	SBCSS (n=4,886)		WHEL (n=3,088)		LACE (n=2,265)		NHS (n=8,075)		Combined (n=18,314)	
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
3	1,135	(23.2)	550	(17.9)	467	(20.7)	2,234	(28.2)	4,386	(24.2)
≥ 4	976	(20.0)	296	(9.7)	453	(20.0)	2,305	(29.1)	4,030	(22.2)
Missing	0		23		4		162		189	
Age at first birth (years)[c]										
< 20	150	(3.2)	323	(13.6)	378	(20.0)	44	(0.60)	895	(5.5)
20–24	866	(18.6)	903	(38.0)	774	(41.0)	3,328	(45.3)	5,871	(36.1)
25–29	2,572	(55.4)	699	(29.4)	475	(25.2)	2,996	(40.8)	6,742	(41.5)
30–34	912	(19.6)	317	(13.3)	191	(10.1)	749	(10.2)	2,169	(13.3)
≥ 35	147	(3.2)	136	(5.7)	70	(3.7)	232	(3.2)	585	(3.6)
Missing	41		4		11		5		61	
Alcohol consumption										
Never/minimal	4,871	(99.7)	1,130	(36.6)	912	(48.2)	2,794	(40.5)	9,707	(57.9)
1–2 drinks/month	0	0	335	(10.9)	259	(13.7)	721	(10.4)	1,315	(7.8)
1–2 drinks/week	0	0	491	(15.9)	206	(10.9)	1,130	(16.4)	1,827	(10.9)
≥ 3 drinks/week	15	(0.3)	1,130	(36.6)	517	(27.3)	2,258	(32.7)	3,920	(23.4)
Missing	0		2		371		1,172		1,545	
Smoking										
Never	4,758	(97.4)	1,643	(53.8)	1,196	(52.9)	3,331	(42.4)	10,928	(60.5)
Past	96	(2.0)	1,276	(41.7)	891	(39.4)	3,514	(44.8)	5,777	(32.0)
Current	32	(0.7)	138	(4.5)	173	(7.7)	1,008	(12.8)	1,351	(7.5)
Missing	0		31		5		222		258	
Body mass index (kg/m^2)[d]										
< 18.5	80	(1.8)	31	(1.0)	19	(1.0)	150	(2.0)	280	(1.6)
18.5–24.99	2,579	(58.3)	1,301	(42.1)	723	(37.9)	3,584	(46.7)	8,187	(47.9)
25.0–29.99	1,483	(33.5)	954	(30.9)	649	(34.0)	2,484	(32.4)	5,570	(32.6)
≥ 30	280	(6.3)	802	(26.0)	516	(27.1)	1,453	(18.9)	3,051	(17.9)
Missing	464		0		358		404		1,226	
Recreational Physical Activity (MET h/week)[d]										
None	1,172	(26.5)	419	(14.0)	285	(14.7)	40	(0.6)	1,916	(11.8)
< 4.2	354	(8.0)	516	(17.3)	385	(19.8)	2,270	(32.7)	3,525	(21.6)

	SBCSS (n=4,886)		WHEL (n=3,088)		LACE (n=2,265)		NHS (n=8,075)		Combined (n=18,314)	
	n	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**	**n**	**(%)**
4.2–16.3	1,384	(31.3)	1,033	(34.6)	611	(31.5)	2,409	(34.7)	5,437	(33.3)
≥ 16.3	1,519	(34.3)	1,017	(34.1)	661	(34.0)	2,231	(32.1)	5,428	(33.3)
Missing	455		103		317		1,100		1,975	
Diabetes	357	(7.3)	56	(2.2)	166	(7.5)	634	(7.9)	1,213	(6.9)
Missing	1		546		58		0		605	
Hypertension	1,138	(23.3)	395	(15.5)	714	(32.1)	3,334	(41.3)	5,581	(31.5)
Missing	1		539		40		0		580	

SBCSS Shanghai Breast Cancer Survival Study, *WHEL* Women's Healthy Eating and Living Study, *LACE* Life After Cancer Epidemiology Study, *NHS* Nurses' Health Study

[a] All percentages exclude missing data. Exercise distributions exclude women with implausible values (met-hours/week ≥ 120).

[b] For WHEL and LACE menopausal status is at diagnosis.

[c] Among parous women.

[d] For SBCSS, BMI and exercise are for the 18-month follow-up survey.

Table 5

Five-year age-adjusted survival estimates overall and by stage at diagnosis for each cohort study, After Breast Cancer Pooling Project[a]

	SBCSS (n=4,628)		WHEL (n=3,088)		LACE (n=2,264)		NHS (n=7,274)		
	Follow-up begins at first survey after diagnosis	Follow-up begins two years after diagnosis[b]	Follow-up begins at first survey after diagnosis	Follow-up begins two years after diagnosis[b]	Follow-up begins at first survey after diagnosis	Follow-up begins two years after diagnosis[b]	Follow-up begins at first survey after diagnosis	Follow-up begins two years after diagnosis[b]	Follow-up begins two years after diagnosis[c]
TNM Stage (AJCC 6th edition)									
I	95.5	96.5	98.2	98.7	98.4	98.4	93.8	95.6	95.9
II	91.4	93.3	95.1	96.5	96.7	96.7	86.7	89.6	91.4
III[d]	67.0	77.3	89.7	93.2	88.2	88.5	68.2	75.4	83.4
Overall	89.0	92.2	95.4	96.9	96.6	96.7	88.3	91.1	93.3

SBCSS Shanghai Breast Cancer Survival Study, *WHEL* Women's Healthy Eating and Living Study, *LACE* Life After Cancer Epidemiology Study, *NHS* Nurses' Health Study

[a] Table excludes 1,060 participants with stage IV disease or missing stage. Survival estimates are adjusted for age at diagnosis.

[b] Follow-up begins two years after diagnosis for all four cohorts (excluded 167 women for SBCSS, 9 from LACE, 17 from WHEL, 245 women for NHS). Two years was selected, because this was the mean time between diagnosis and enrollment for WHEL and LACE participants.

[c] Further excluded cases diagnosed prior to 1991 (n =2,793).

[d] LACE and WHEL cohorts excluded women with breast tumors with extension to the chest wall.