The Breast Journal

ORIGINAL ARTICLE

Establishment of a Southern Breast Cancer Cohort

Kristina L. Bondurant, PhD,* Sarah Harvey, MPH,† Suzanne Klimberg, MD,‡ Susan Kadlubar, PhD,§ and Martha M. Phillips, PhD*,¶

*Department of Epidemiology, Fay W. Boozman College of Public Health, University of Arkansas for Medical Sciences, Little Rock, Arkansas; †Epidemiology Branch, Center for Public Health Practice, Arkansas Department of Health, Little Rock, Arkansas; ‡Division of Breast Surgical Oncology, Department of Surgery, College of Medicine, University of Arkansas for Medical Sciences, Arkansas; §Department of Medical Genetics, College of Medicine, University of Arkansas for Medical Sciences, Arkansas; and ¶Department of Psychiatry, College of Medicine, University of Arkansas for Medical Sciences, Arkansas

■ **Abstract:** Breast cancer continues to be among the most common cancers affecting women in the United States. Researchers investigating the area are turning their attention to novel prevention, detection, and treatment options. Recent molecular epidemiology research has highlighted the effects of both genetic and environmental exposures on an individual's risk of developing breast cancer and predicted response to treatment. Cohort designs are a potentially powerful tool that researchers can utilize to investigate the genetic and environmental factors affecting breast cancer risk and treatment options. This paper describes the recruitment of a community-based cohort of women in a southern state. The Spit for the Cure Cohort (SFCC), being developed by researchers at the University of Arkansas for Medical Sciences (Little Rock, AR), is designed to be representative of the female population of the state with oversampling of women with a history of breast cancer and women of color. To date, the SFCC includes more than 14,000 women recruited from all 75 counties of Arkansas and six neighboring states. Methods used to recruit and maintain the cohort and collect both questionnaire data and genetic material are described, as are the demographic characteristics of the cohort as it currently exists. The recruitment methods utilized for the SFCC are rapidly building a breast cancer cohort and providing a large biorepository for molecular epidemiology research. ■

Key Words: breast neoplasms, cohort studies, environment, genetic markers

INTRODUCTION

According to the American Cancer Society, an estimated 192,370 new cases of invasive breast cancer are predicted in US women and over 40,000 women are estimated to die from their disease in 2009 (1). Excluding skin cancer, breast cancer continues to be the most common cancer affecting women in the United States (1). The focus for breast cancer research investigators is novel prevention, detection, and treatment options for what is now thought to be a grouping of multiple diseases (2).

Recent molecular epidemiology research has highlighted the effects of both genetic and environmental exposures on an individual's risk of developing breast cancer and predicted response to treatment. To understand more fully the incidence, recurrence of, and survival after treatment for breast cancer, researchers have studied differences in genetic variants controlling metabolism of environmental exposures and chemotherapeutic agents, genetic variants offering protection against DNA damaging agents, and the linkage of biomarkers of exposure to carcinogens and dietary nutrients in relation to questionnaire data, single nucleotide polymorphisms (SNPs), and other biomarkers. For example, studies examining smoking as a risk factor for breast cancer have provided inconsistent results (3). However, when genetic variation in N-acetyltransferase 2, an enzyme responsible for the detoxification of carcinogens in cigarette smoke, was considered, genetic variants in this enzyme were strong modifiers of breast cancer risk (4). In addition, an increased risk of breast cancer has been associated with high levels of meat consumption (5). Genetic

Address correspondence and reprint requests to: Kristina L. Bondurant, PhD, Department of Epidemiology, Fay W. Boozman College of Public Health, University of Arkansas for Medical Sciences, 4301 West Markham, Slot #820, Little Rock, AR 72205-7199, USA, or e-mail: bondurantkristinal@uams.edu.

DOI: 10.1111/j.1524-4741.2011.01082.x

© 2011 Wiley Periodicals, Inc., 1075-122X/11

variants in enzymes responsible for the activation/detoxification of procarcinogens found in high temperature cooked meat have been described for other cancers, while the findings for breast cancer are inconsistent. Genetic variants in tamoxifen–metabolizing enzymes have been reported to contribute to overall survival and recurrence of disease in patients receiving adjuvant tamoxifen therapy (6,7). Others have shown an association between survival of breast cancer and glutathione-S-transferase P1 polymorphisms in patients treated with chemotherapy (8). Defects in enzymes such as BRCA1, which is involved in cell cycle checkpoints, as well as DNA repair are high penetrance mutations, although their population frequency is low (9). Low penetrance genetic variants in DNA repair genes that have high population frequencies have also been associated with susceptibility to breast cancer, and the impact of these variants vary by ethnicity (10). The interaction between genetic variants in enzymes modulating reactive oxygen species and fruit and vegetable consumption (foods that are rich in antioxidants) have also been described (11).

Cohort designs are a powerful epidemiological tool that researchers can utilize to investigate the genetic and environmental factors affecting breast cancer risk and treatment options. Ongoing cohorts which have been particularly successful represent educated groups of professionals, e.g., the Nurses' Study (12) and the Health Professionals Follow-up Study (13). Several large cohort studies (including those outlined above) that obtained DNA from a portion of their study participants (14) have been assembled into a consortium, the Breast and Prostate Cancer and Hormone-Related Gene Variant Study (15). However, these cohorts are primarily clustered in the Northeastern United States or on the West Coast and Hawaii, e.g., Nurses' Health Study I and II (12), Health Professionals Follow-up Study (13), Multiethnic Cohort (16). Our literature search revealed only two cohorts with women from the Southern US: one community-based cohort directed toward health disparities (17); and the Sister's Study, which is recruiting women nationwide but is restricted to sisters of women who have had breast cancer, an exclusively high-risk group (18). A recent review of publications listed in the catalog of the National Human Genome Research Institute highlights the limited representation of minority groups in the majority of genome-wide association studies (GWAS) (19). Only 3% of GWAS participants were African American. The authors speculate that this lack of diversity in GWAS samples may be due to the fact that the majority of GWAS used existing cohorts.

Environmental and dietary factors that may contribute to breast cancer risk and/or response to treatment may differ substantially by geographic region. For example, dietary practices, unique environmental exposures, or the presence of novel breast cancer susceptibility genes in genetically heterogeneous subgroups, may be factors important for investigation. In addition, these factors may interact with other factors typically associated with poor prognosis after breast cancer diagnosis, e.g., belonging to a minority group, having lower socioeconomic status, and having less access to screening programs or health care (20). Thus, the omission of Southern women from cohorts may limit researchers' abilities to ask and answer important questions.

This paper describes the recruitment of a cohort of women that may help to fill these gaps in breast cancer research. The Spit for the Cure Cohort (SFCC) is a community-based cohort being developed by researchers at the University of Arkansas for Medical Sciences (Little Rock, AR) and designed to be representative of the female population of the state. To date, the SFCC includes more than 14,000 women recruited from all 75 counties of Arkansas, six neighboring states, and several distant states (e.g., New York, Virginia, Florida, California). This manuscript describes in some detail the methods used to recruit and maintain the cohort and collect both questionnaire data and genetic material, as well as the demographic characteristics of the cohort as it currently exists.

MATERIALS AND METHODS

The SFCC is being developed using an innovative, community-based recruitment strategy. Efforts to develop the cohort began in the fall of 2007, when UAMS researchers partnered with the Susan G. Komen for the Cure® Arkansas affiliate to recruit cohort participants from among women participating in the local Race for the Cure®. More than 45,000 women dedicated to breast cancer awareness, research and survivor support participated in the race that year. Since that time, we have established a wide range of community partnerships and expanded recruitment efforts throughout Arkansas, using the specific strategies described in some detail below. All

methods and materials have been reviewed and approved by the UAMS Institutional Review Board.

Staff and Training

Recruitment is completed by both paid and volunteer recruiters working in teams of two or more recruiters. Teams are deployed to ensure that sufficient staff are available to handle whatever participant volume may occur and to assure safety in recruitment venues. The first recruiters for the project were identified from among existing staff experienced in epidemiological research studies at UAMS. Paid staff were added as need increased. Important characteristics of recruiters include research experience, experience in working with patients, strong organizational skills, attention to detail, and an ability to travel within the state and to work nights and weekends. Volunteer recruiters are engaged to assist in recruitment in large events, when large numbers of potential participants are expected to be available within a limited amount of time. All SFCC recruiters, both paid staff and volunteers, undergo in-depth training, with topics including recruitment processes overall, procedures for obtaining informed consent, administration of the questionnaire, and collection of the saliva sample. Training sessions, conducted by the study director and the training coordinator, include a detailed presentation of the recruitment process, mock interviews, and a training packet highlighting important topics from the training presentation. In addition, all recruiters are required to complete the human subjects protection training course required by the UAMS Institutional Review Board. Volunteer recruiters are always partnered with experienced staff recruiters, who serve as recruitment team leaders. To assure fidelity with protocols, recruiters undergo routine observation during recruitment, and all completed materials are audited for completeness and accuracy after each recruitment event.

Participant Eligibility, Recruitment, and Enrollment Processes

Women between the ages of 18 and 100 years of age who are able to give informed consent are eligible to enroll in the cohort. After informed consent processes are completed, participants must be able to complete a short questionnaire and provide a 2 mL saliva sample. Incomplete data collection (i.e., missing questionnaire or saliva sample) is considered a failed recruitment for that individual.

Participants are recruited in a number of venues, including Susan G. Komen Race for the Cure® events, American Cancer Society Relays for Life, community festivals and events, faith based events, and community and business sponsored health fairs. Events are identified through key partners, such as the Komen affiliates, the American Cancer Society, and the Winthrop P. Rockefeller Arkansas Cancer Community Network which specifically partners with underserved communities throughout Arkansas to empower communities, provide important cancer education, and bring screening services to rural areas. Recruiters also attend local community health fairs, breast cancer awareness programs and support groups for women, many of which are sponsored by churches or community-based organizations in rural communities. Prior to an event, we take advantage of opportunities to partner with event organizers to distribute brochures, flyers, and press releases describing the project.

On-site recruitment processes include: (a) project introduction; (b) informed consent; and (c) data collection. When recruiting in small event venues, recruiters complete these processes on an individualized basis. When recruiting in very large venues, the three processes may be completed sequentially but in separate stations, so that participants are introduced to the project and then directed to a table for completion of informed consent processes, from which they are directed to a data collection table. Experience has refined these large-venue processes to maximize quality control and participant satisfaction, while minimizing participant burden.

Regardless of venue size, participants provide written informed consent, fill out a short intake questionnaire and provide a saliva sample. During the informed consent process, each participant is asked to give specific permission for study personnel to contact them in the future to ascertain their breast cancer status and to solicit participation in research studies that may require additional information from participants. If the participant is willing to be contacted in the future, they are asked to provide their address, phone number and email address on their consent form. Failure to give consent to be recontacted does not exclude a woman from enrollment in the cohort. Once informed consent processes are complete, the recruiter gives the participant specific instructions concerning completion of the questionnaire and providing the saliva sample. Informed consent and data collection

processes take the average participant approximately 10–15 minutes.

Data Collection and Management

The intake questionnaire asks questions about personal history of breast cancer, including age at diagnosis and treatment received. Other questions address preventive treatments, demographics, lifestyle factors (e.g., exercise), and other risk factors (e.g., age at menarche and menopause, use of birth control or other hormones, family history of breast cancer). The intake questionnaire is designed to cover basic questions that may be useful in identifying subgroups of participants for research studies; in-depth information about specific risk factors or exposures (e.g., diet, environmental exposures) will be gathered in follow-up questionnaires. The questionnaire is scannable; participants indicate their preferred responses by filling in circles on a standard form.

Participants provide 2 mL of saliva for future DNA extraction using the Oragene®-DNA Self Collection Kit (21). The Oragene®-DNA kit is highly stable at a broad range of temperatures making it ideal for field-based sample collection. Completed kits can be stored at room temperature for up to 5 years before the isolation of DNA samples (22).

All materials collected from participants in the field (consent forms, questionnaires, saliva samples) are labeled with a barcode that serves as a unique 10-character participant identifier and placed in a sealed envelope for transport back to the project offices. Questionnaire data are scanned into an electronic database; data from the consent form (i.e., name, date of recruitment, contact information if provided) are entered into a separate electronic database. Saliva samples are stored in boxes; storage location (box number, column number, row number) for each sample is entered into a third electronic database by unique identifier so that any given sample can be readily retrieved. All databases and files are organized by unique identifier to protect participant confidentiality while facilitating retrieval of participant-specific materials if necessary.

RESULTS

Table 1 summarizes the number and types of events through which SFCC recruitment has taken place, as well as the recruitment outcomes by event type. The greatest numbers of events have been community related events, though Susan G. Komen for the Cure events has yielded the greatest number of participants. Overall, 50.7% of the cohort has derived from Komen affiliate sponsored events.

As of November 30, 2009, 14,384 women, aged 18–98 years, have enrolled in the Spit for the Cure cohort. These women represent all 75 counties in Arkansas, six neighboring states and several distant states. As seen in Table 2, the demographic characteristics of the SFCC are very similar to the demographics of the Arkansas female population, with three exceptions: the current SFCC cohort is, first, older and, second, more highly educated than the female population of Arkansas overall and third, includes a smaller proportion of Latinas than might be expected.

Given that supporting research to understand disparities in breast cancer and to understand risk and treatment response in high-risk groups, we summarize in Table 3 key characteristics of the cohort by breast cancer status. Data indicate that the group of participants with breast cancer are somewhat more likely to be white and older, compared to the group of participants without a personal breast cancer history. As of September 30, 2009, 17.8% of participants have reported at least one other female relative (mother, daughter, sister, etc.) participating in the cohort.

Table 1. Recruitment by Type of Event

Event type	Events to date	Participants with breast cancer	Participants with no personal history of breast cancer
Clinic	1 (0.5%)	89 (5.9%)	175 (1.4%)
Community event	107 (50.0%)	256 (16.9%)	3310 (25.7%)
Community health fair	57 (26.6%)	91 (6.0%)	2195 (17.1%)
Other	7 (3.3%)	1 (0.1%)	85 (0.7%)
Race for the cure	21 (9.8%)	848 (56.0%)	6449 (50.1%)
Relay for life	7 (3.3%)	13 (0.9%)	132 (1.0%)
UAMS Winthrop P. Rockefeller Cancer Institute	14 (6.5%)	217 (14.3%)	523 (4.1%)

Table 2. Demographic and Breast Cancer Characteristics of the Spit for the Cure Cohort, Compared to the Arkansas Female Population Overall, as of November 30, 2009

Demographic characteristics	Cohort*	Arkansas†
Racial/ethnic group		
African American	15.8%	15.0%
Caucasian	82.5%	80.4%
Other racial/ethnic group	3.1%	4.6%
Women who identified themselves as Hispanic	1.8%	3.6%
Women living in rural area‡	25.4%	43.4%
Age		
18–39 years	37.6%	37.3%
40–64 years	49.2%	41.9%
65 years and older	10.7%	20.8%
Education		
Less than high school diploma	2.0%	17.1%
High school graduate/GED	15.9%	34.2%
Some college	32.5%	31.8%
Bachelor's degree or higher	49.3%	16.9%

*Percentages may not total 100% due to non-response on questionnaire.
†Data for all females in the Arkansas population according to the 2008 American Community Survey, 1-year estimates (23).
‡"Rural" defined as Non-Metropolitan area designation using the USDA Economic Research Service 2003 Rural/Urban Continuum Code.

Twenty percent of the cohort overall have indicated a family history of breast cancer (defined as a mother, sister, or daughter having a diagnosis of breast cancer). Participants with breast cancer were more likely to report both occurrences (a relative in the cohort and a family history of breast cancer) than other participants. More than 95% of participants have agreed to maintain contact with the cohort research team in the future. Proportions were similar for women with and without a personal history of breast cancer (97.4% and 95.1%, respectively).

DISCUSSION

As described above, we have been successful over the past 2 years in recruiting more than 14,000 women to enroll in a population-based cohort of women with and without breast cancer and securing both questionnaire data and genetic material from those women for future study. Initiating recruitment at the Arkansas Affiliate Susan G. Komen Race for the Cure® allowed the rapid accrual of a large number of women to serve as the base for this breast cancer cohort. We have found four main advantages to establishing the cohort in this manner. First, the race attracts a large number of women highly motivated to participate in breast cancer-related research activities. Second, the race attracts many of the same women on a yearly basis, which helps us maintain contact with these participants. Third, the race encourages family participation, which helps us enroll first and second degree relatives who have been diagnosed with breast cancer, forming a "high risk" cohort within the general population cohort. Finally, the high volume of advertising and media attention during the race leads to rapid dissemination of information about the cohort and facilitates the identification of future recruitment events.

Table 3. Number and Percentage of SFCC Participants With Selected Characteristics by Breast Cancer Status, as of November 30, 2009*

	Participants with personal history of breast cancer (n = 1,240)	Participants with no history of breast cancer (n = 10,000)
Racial/ethnic group†		
African American or Black	165 (10.9%)	2107 (16.4%)
Caucasian or White	1335 (88.1%)	10527 (81.8%)
Other	35 (2.3%)	408 (3.2%)
Women who identified themselves as Hispanic	15 (1.0%)	250 (1.9%)
Age		
18–39 years	74 (4.9%)	5328 (41.4%)
40–64 years	1001 (66.1%)	6075 (47.2%)
65 years and older	404 (26.7%)	1135 (8.8%)
Relative also enrolled in cohort‡	287 (23.1%)	1714 (17.1%)
First-degree relative with breast cancer	424 (28.0%)	2370 (18.4%)
Hormonal treatment		
Preventive	N/A	123 (1.0%)
Breast cancer treatment	839 (55.4%)	
Age at breast cancer diagnosis		
<50 Years	675 (44.6%)	N/A
≥50 Years	773 (51.0%)	
Years since Breast Cancer Diagnosis		
≤5 Years	739 (48.8%)	N/A
6–10 Years	350 (23.1%)	
>10 Years	339 (22.4%)	
Breast cancer recurrence	121 (8.0%)	N/A

*Percentages may not total 100% due to non-response on questionnaire.
†Percentages may sum to greater than 100% due to multiple responses given by single participants, indicating identification with multiple race/ethnic categories.
‡Values calculated as of September 30, 2009.

The Arkansas population includes several important sub-populations traditionally underrepresented in well-established cohorts. Many of the characteristics thought to be associated with poor prognosis are more prevalent in the state, i.e., minority status, lower socioeconomic status, lack of access to screening programs or health care, rural residence, diabetes, and obesity. Despite these disadvantages, population mobility is low. The US Census Bureau estimates that 97% of the population remained in the state between 2007 and 2008; 93% were living in the same county, compared to the previous year; and 82% remained in the same house (23). This population stability bodes

well for efforts to follow the SFTC cohort over time. In fact, recent efforts to re-contact a subset of the cohort have resulted in successful contact with 87% of the cohort.

Our community-based recruitment efforts have generated a cohort whose demographics mirror those of the state of Arkansas. We have been particularly successful in recruiting women of color, who have traditionally been less likely to participate in research studies. We recognize, however, that we need to use more targeted methods for identifying venues that will enable us to reach our goal of recruiting additional numbers of women with breast cancer and women of color. Thus, we are now systematically attempting to identify events in communities and venues that will more likely engage larger numbers of these target groups. For example, we are working to establish collaborations with breast oncology clinics throughout the state. We are also working with community partners in the southeastern part of the state, which has larger proportions of African American residents, to identify appropriate recruitment venues in those communities.

The unintentional oversampling of certain age groups and educational levels and undersampling of Latinas is thought to be related to the venues in which recruitment takes place. Women ages 45–64 are commonly affected by breast cancer diagnoses and may be more interested in research participation than women as a whole. Further, the SFCC's higher education level compared to the state overall is likely due to more frequent recruitment in venues which attract professional women. Recent active recruitment in areas of Arkansas underrepresented in the cohort has increased our ability to reach women in different age groups and of varying education levels to mirror more effectively state population demographics. Regarding the undersample of Latinas, it is important to know that, while the number of Latinas residing in Arkansas has grown in the past 10 years, they still represent only a small proportion of the overall female population and are substantially younger (median age: 21.4 years) than their white or African American counterparts (median ages: 41.0 and 32.6 years, respectively) (23). With limited resources, we have made a strategic decision to focus on recruiting the oversamples described above.

There are several advantages to the recruitment process as well. Providing a saliva sample is less invasive than providing a blood sample, poses less risk to participants, and is more acceptable for many participants, maximizing the likelihood that a potential participant will consent (24). Venipuncture also requires trained phlebotomists, additional materials for on-site collection, and may prove a timely process depending on the donor, all of which increase the cost of data collection. Recent developments in the use of buccal cells or lymphocytes from saliva as a source for DNA have provided a quick, non-invasive alternative method for DNA collection. Further, the saliva collection methods allow for the rapid collection of large numbers of samples simultaneously (24). While there is an increased possibility for the introduction of bacterial contamination to the DNA sample when it is obtained from saliva, product developers maintain that, compared to buccal swabs or mouthwash methods, the saliva collection kit used by the SFCC has significantly less bacterial contamination (25). An additional benefit of this collection system is the long-term stability of the saliva samples under a range of temperatures, which is particularly useful when recruiting and collecting samples in the field.

Recruitment processes have been designed to minimize each participant's time commitment (5–10 minutes for consent, 10–15 minutes for data collection) while maximizing the amount of information collected. This minimized participant burden, in terms of invasiveness and time, has, we believe, contributed to our recruitment successes. In addition, because recruitment takes place primarily in the context of community events, and because our participation occurs at the invitation of community residents, participants are likely more comfortable with the context and thus more likely to agree to participate.

As with any recruitment effort, we strive to minimize potential bias by recruiting participants in the broadest range of venues. While we continue to accrue the largest number of participants in cancer-related events, we build on those interactions to identify other events, such as those sponsored by community-based organizations, churches, academic institutions, sports teams, and others. Word-of-mouth is an important method of generating community interest and identifying additional recruitment events in communities. Community residents attending a specific recruiting event are often motivated to host another recruitment event in their community and become valuable partners in reaching out to groups otherwise unaware of this breast cancer research opportunity in Arkansas. These activities help to broaden our population base and assure interaction with diverse individuals.

Our ability to recruit a cohort that represents the female population of the state of Arkansas is dependent on our ability to recruit in a diverse range of venues.

Our recruitment process has been successful thus far in reaching diverse areas of Arkansas. However, we continue to build upon strategies by increasing community-based outreach. We are actively working to recruit younger and less-highly educated individuals, to bring our cohort demographics into greater alignment with the population demographics. In addition, in order to provide a cohort that will allow researchers to address disparities and questions of survivorship or responses to treatment, we are also working to oversample minority women and breast cancer survivors.

We recognize that, while 95% of our cohort has indicated willingness to be re-contacted in the future, loss to follow-up will inevitably occur and losses will be greater without a comprehensive retention plan (14,26). We have developed such a plan and are in the process of the first comprehensive re-contact of the cohort. As noted above, first efforts have indicated a successful rapid contact with 87% of the cohort; we continue to attempt to contact the remaining 13%. We anticipate that future contacts will include follow-up questionnaires that elicit more detailed information about risk factors and prevention or treatment experiences.

Employing the Susan G. Komen Arkansas Race for the Cure® venue to recruit women into a breast cancer cohort study provided a rapid biorepository with sufficient numbers to address specific research questions. Acquisition of this sample for future studies will accelerate the process of selecting and recruiting study subjects for more specific research projects. For example, if an investigator is exploring the relationship between obesity and breast cancer risk in an underserved population, the data acquired in the original recruitment will provide a ready pool of study participants who can be selected based on body mass index calculated from data obtained in the intake questionnaire. These participants have agreed to be contacted for future studies and are thus much more likely to be willing to participate in an in-depth study that addresses environmental exposures, lifestyle choices and the underlying contribution of genetic variation to breast cancer. An area that will benefit greatly from this cohort is the field of pharmacogenomics (the study of genetic variability in relation to response to treatment). To date, pharmacogenomic studies have been performed within the context of controlled clinical trials with relatively small numbers of participants. By using SFCC participants, investigators have the opportunity to explore ''real world'', community-based pharmacogenomics. This capability will be invaluable in translating findings from clinical trials, validating them in a population-based study, thus furthering the field of personalized medicine.

ACKNOWLEDGMENTS

The authors would like to acknowledge the following individuals and organizations for their significant contributions to this study: Dr. Carol Sweeney; the Winthrop P. Rockefeller Cancer Institute at UAMS; the Arkansas and Ozark Affiliates of Susan G. Komen for the Cure®; the Arkansas Cancer Community Network; and our dedicated staff and volunteer recruiters. This work is funded by the Arkansas Breast Cancer Research Program (ABCRP) of the Winthrop P. Rockefeller Cancer Institute at UAMS.

REFERENCES

1. American Cancer Society. Cancer Facts and Figures 2009. 2009; Available at: http:/www.cancer.org/acs/groups/content/@nho/documents/document/500809webpdf.pdf (accessed August 10, 2009).

2. Newman B, Moorman PG, Millikan R, *et al.* The Carolina Breast Cancer Study: integrating population-based epidemiology and molecular biology. *Breast Cancer Res Treat* 1995;35:51–60.

3. Morabia A. Smoking (active and passive) and breast cancer: epidemiologic evidence up to June 2001. *Environ Mol Mutagen* 2002;39:89–95.

4. Ambrosone CB, Freudenheim JL, Graham S, *et al.* Cigarette smoking, N-acetyltransferase 2 genetic polymorphisms, and breast cancer risk. *JAMA* 1996;276:1494–501.

5. Boyd NF, Martin LJ, Noffel M, Lockwood GA, Trichler DL. A meta-analysis of studies of dietary fat and breast cancer risk. *Br J Cancer* 1993;68:627–36.

6. Nowell S, Sweeney C, Winters M, *et al.* Association between sulfotransferase 1A1 genotype and survival of breast cancer patients receiving tamoxifen therapy. *J Natl Cancer Inst* 2002;94:1635–40.

7. Nowell SA, Ahn J, Rae JM, *et al.* Association of genetic variation in tamoxifen-metabolizing enzymes with overall survival and recurrence of disease in breast cancer patients. *Breast Cancer Res Treat* 2005;91:249–58.

8. Sweeney C, McClure GY, Fares MY, *et al.* Association between survival after treatment for breast cancer and glutathione S-transferase P1 Ile105Val polymorphism. *Cancer Res* 2000;60:5621–4.

9. Huen MS, Sy SM, Chen J. BRCA1 and its toolbox for the maintenance of genome integrity. *Nat Rev Mol Cell Biol* 2010; 11:138–48.

10. Li H, Ha TC, Tai BC. XRCC1 gene polymorphisms and breast cancer risk in different populations: a meta-analysis. *Breast* 2009;18:183–91.

11. Ahn J, Nowell S, McCann SE, *et al.* Associations between catalase phenotype and genotype: modification by epidemiologic factors. *Cancer Epidemiol Biomarkers Prev* 2006;15:1217–22.

12. Belanger CF, Hennekens CH, Rosner B, Speizer FE. The nurses' health study. *Am J Nurs* 1978;78:1039–40.

13. Rimm EB, Giovannucci EL, Willett WC, *et al.* Prospective study of alcohol consumption and risk of coronary disease in men. *Lancet* 1991;338:464–8.

14. Willett WC, Colditz GA. Approaches for conducting large cohort studies. *Epidemiol Rev* 1998;20:91–9.

15. National Cancer Institute. Breast and Prostate Cancer and Hormone-Related Gene Variant Study. 2008; Available at: http://epi.grants.cancer.gov/BPC3/cohorts.html (accessed March 25, 2009).

16. Kolonel LN, Henderson BE, Hankin JH, *et al.* A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. *Am J Epidemiol* 2000;151:346–57.

17. Signorello LB, Hargreaves MK, Steinwandel MD, *et al.* Southern community cohort study: establishing a cohort to investigate health disparities. *J Natl Med Assoc* 2005;97:972–9.

18. National Institute of Environmental Health Services, National Institute of Health, Department of Health and Human Services. The Sister Study. 2009; Available at: http://www.sisterstudy.org/English/index1.htm (accessed June 1, 2009).

19. Haga SB. Impact of limited population diversity of genome-wide association studies. *Genet Med* 2010;12:81–4.

20. Bradley CJ, Given CW, Roberts C. Race, socioeconomic status, and breast cancer treatment and survival. *J Natl Cancer Inst* 2002;94:490–6.

21. Rogers NL, Cole SA, Lan HC, Crossa A, Demerath EW. New saliva DNA collection method compared to buccal cell collection techniques for epidemiological studies. *Am J Hum Biol* 2007;19:319–26.

22. DNA Genotek. Long-term stability of DNA from saliva samples stored in Oragene®•DNA. 2004; Available at: http://www.dnagenotek.com/pdf_files/PDWP005_Stability.pdf (accessed June 1, 2009).

23. US Census Bureau. American Community Survey, 2008; Available at: http://www.census.gov (accessed April 1, 2010).

24. Hansen TV, Simonsen MK, Nielsen FC, Hundrup YA. Collection of blood, saliva, and buccal cell samples in a pilot study on the Danish nurse cohort: comparison of the response rate and quality of genomic DNA. *Cancer Epidemiol Biomarkers Prev* 2007;16:2072–6.

25. Rylander-Rudqvist T, Hakansson N, Tybring G, Wolk A. Quality and quantity of saliva DNA obtained from the self-administrated oragene method–a pilot study on the cohort of Swedish men. *Cancer Epidemiol Biomarkers Prev* 2006;15:1742–5.

26. Colditz GA, Hankinson SE. The Nurses' Health Study: lifestyle and health among women. *Nat Rev Cancer* 2005;5:388–96.