# Supporting Statement A

## *The Sister Study PHASE 2:*

## *Environmental and Genetic Risk Factors for Breast Cancer*

## *(NIH/NIEHS)*

***Submitted: 20 August 2009***


**Principal Investigator and Co-Project Officer:**

Dale P Sandler PhD
Chief, Epidemiology Branch
National Institute of Environmental Health Sciences
PO Box 12233
Research Triangle Park NC 27709
Phone: 919-541-4668
Fax: 919-541-2511
Email: sandler@niehs.nih.gov

**Co-Principal Investigator:**

Clarice Weinberg PhD
Chief, Biostatistics Branch
PO Box 12233
Research Triangle Park NC 27709
Phone: 919-541-4927
Fax: 919-541-4311
Email: weinberg@niehs.nih.gov

**Project Officer:**

Paula S Juras PhD
Epidemiology Branch
National Institute of Environmental Health Sciences
PO Box 12233
Research Triangle Park NC 27709
Phone: 919-541-4668
Fax: 919-541-2511
Email: juras@niehs.nih.gov

**Summary**

The Sister Study is a prospective cohort study of the environmental and genetic risk factors for breast cancer and other diseases among 50,000 sisters of women who have had breast cancer. The long-term design allows us to also study the impact of environmental and genetic factors on survival and outcomes following a breast cancer diagnosis.

In the United States, approximately 180,000 new cases of breast cancer are diagnosed each year. Breast cancer accounts for over 30% of all new cancer cases among women and 15% of cancer deaths. The etiology of breast cancer is complex, with both genetic and environmental factors playing a role. Currently established breast cancer risk factors account for less than half the variation in breast cancer risk across the United States, and known breast cancer genes are found in fewer than 10% of breast cancer patients. Although the concordance rate for monozygotic twins is less than 20% (underscoring the importance of environmental contributors), sisters of women with breast cancer have, on average, a 2-fold increase in risk for breast cancer themselves. By focusing on a genetically susceptible group, more precise estimates of the contribution of environmental and other non-genetic factors to disease risk may be possible and the power to study gene by environment interactions will be greatly enhanced.

Between August 2003 and August 2009, we enrolled just over 50,000 at-risk volunteers into a prospective cohort study known as the Sister Study. Sister Study participants were recruited from across the United States and Puerto Rico. They were between ages 35 and 74 at enrollment and had at least one full or half sister who was diagnosed with breast cancer. Participants were recruited through a variety of means including the media, breast cancer advocacy groups, medical practitioners, community partners, and Sister Study participants themselves. Recruitment strategies were designed to enroll a cohort of sisters that is ethnically, geographically, and socioeconomically diverse. At enrollment, participants provided complete histories of personal and family health, reproductive health, diet, and environmental and lifestyle exposures. They completed a home exam in which height, weight, waist circumference, and blood pressure were measured by an examiner who also collected a blood sample. Whole blood, serum,

plasma, blood clots, blood spots, and cryopreserved lymphocytes were stored for future use. Participants also provided a first morning urine sample, toenail clippings, and household dust wipe samples.

Now in Phase 2 of the Sister Study, the enrolled high-risk cohort will be followed actively for ten or more years for the development of breast cancer and other diseases. We anticipated, on average, 300 new cases of breast cancer to be diagnosed each year; thus after five years of follow-up, we will have sufficient power, with about 1,500 new breast cancer cases, to address hypotheses regarding gene-environment interactions that require large numbers of cases. Baseline data, banked blood, urine, and toenail samples, as well as banked environmental samples will provide a rich resource for testing current and future hypotheses regarding risks for breast cancer and a wide range of other medical conditions.

Sister Study participants will be contacted annually to track changes in their medical history. Every two to three years, participants will be asked to provide more comprehensive health and exposure updates and to provide additional information on other factors that may affect disease risk/survival such as stress phenotype or diet. Over time, participants may be asked to provide additional biological or environmental samples and will be invited to participate in more focused clinical studies of specific outcomes. Participants who develop breast cancer will be asked to provide information about their diagnosis and treatment and asked to sign medical release forms allowing us to request medical records, additional details about diagnosis and treatment, and tumor tissue and/or diagnostic H&E slides from their health care providers. Similarly, additional information to allow disease confirmation or validation of self-reported diagnoses will be sought from participants who report other diseases of interest such as asthma, uterine fibroids, diabetes, thyroid disease, osteoporosis, rheumatoid arthritis and other autoimmune diseases, neurodegenerative diseases, and other cancers. Nested case-control or case-cohort studies will be carried out among sisters who develop cancer or other clinical outcomes of interest and a sample of those who do not, to assess specific gene-environment interactions or other hypotheses. Patients who develop breast cancer during the follow-up period for the Sister Study will continue to be followed to study the role of environment and genes in survival following breast cancer diagnosis. Add-on studies

may collect serial biological samples from informative subgroups to evaluate preclinical biomarkers and assess changes in biomarkers over time.

**The Sister Study completed PHASE 1 enrollment in August 2009, concluding contact with the public at large, and as of 31 August 2009 will discontinue the data collection using Sister Study enrollment questionnaires and materials previously approved by OMB.**

**This application is to request renewed approval for PHASE 2 long-term clinical follow-up of the enrolled cohort, including currently approved long-term clinical follow-up activities. Currently, using the previously OMB-approved forms, 31,137 participants have completed or are in progress for completing their first Biennial Update. Also 911 participants have reported incident cases of breast cancer, and are in progress for completing incident breast cancer follow-up activities.**

## A. Justification

### A.1. Circumstances Making the Collection of Information Necessary

The National Institute of Environmental Health Sciences (NIEHS) is responsible for conducting research on chemical, physical, and biological factors in the environment that affect human health. The **Sister Study**, with its focus on potential environmental causes of breast cancer, is supported by the mandate of NIEHS as defined by US Code Title 42, Chapter 6A, Subchapter III, Part A, Section 281, as amended by the Health Research Extension Act of 1985, which is "the conduct and support of research, training, health information dissemination, and other programs with respect to factors in the environment that affect human health, directly or indirectly." *[Attachment 1 deleted from final package]***.**

Breast cancer is the most common cancer among women in developed countries. In the United States, 192,370 new cases of invasive breast cancer are anticipated in 2009. Many of the best-understood risk factors are endogenous or are related to lifestyle choices that are not easily modified. Family history of breast cancer is one of the most well established risk factors for the disease with an approximate 2-fold increase in risk in first-degree relatives. No known genetic or environmental breast cancer risk factor or

combination of risk factors has high enough relative risks to account for the observed association with family history.

The etiology of breast cancer is complex, with genetic and environmental factors likely playing a combined role. Identifying gene-environment or gene-gene interactions will be important in understanding breast cancer etiology and identifying prevention strategies. A number of genes already have been identified as candidates for study. Large-scale studies are needed to confirm reported gene-environment interactions and to test new biologically based hypotheses as both technology and our understanding of mechanisms improve over time. The Sister Study offers a unique opportunity to gather important epidemiological data that will make a difference in understanding the multi-factorial etiology of breast cancer.

There is suggestive evidence, including studies of disease patterns in migrants, to support a role of the environment in breast cancer risk, although specific environmental factors have not been clearly elucidated. With the exception of recent studies of organochlorine pesticides and earlier studies of irradiated populations, little work has been done in the area of non-lifestyle environmental factors. In part, this may be due to the difficulty of retrospectively measuring most environmental exposures. Few studies have examined occupational exposures among women, although there is evidence that some may play a role in breast cancer etiology. The National Toxicology Program has identified at least 35 mammary carcinogens, many of which have not been evaluated in human studies. Agents on this list include pesticides and fumigants, chemicals used in manufacturing rubber, vinyl, polyurethane foams, benzene-based dyes, and some pharmaceuticals, as well as solvents.

The Sister Study is designed to investigate the independent and joint effects of genetic susceptibility and environmental, biological and lifestyle factors on the risk of breast cancer and other diseases in a cohort of sisters of women with breast cancer. In addition, we will have the opportunity to efficiently assess risks for other diseases with similar risk factor profiles or which are otherwise important causes of morbidity in women, such as cardiovascular disease, autoimmune disease, osteoporosis, and diabetes.

By focusing on sisters, the study has several unique advantages. Not only are the sisters of women with breast cancer at greater risk for breast cancer themselves (nearly a two-fold risk based on other studies), they will have a higher prevalence of any genes that prove to be related to breast cancer risk and they should have a higher frequency of some risk factors due to behaviors and experiences they share with their sisters who have already developed breast cancer. The study is especially well suited to address hypotheses related to shared childhood and prenatal exposures, in addition to factors later in life. Our study will have greater statistical power than a similarly sized general population cohort (see **Sections B**). Furthermore, based on current experience, sisters of women with breast cancer are highly motivated to participate in breast cancer research. Participation and retention rates are higher than in other studies and the quality of data is high. These features will enhance both validity and statistical power to detect associations. The study is large enough to investigate many gene-environment interactions after five years of follow-up. Less common genes and rare exposures can be studied with continued follow-up of the cohort.

The prospective design offers several advantages over the retrospective case-control studies that have been most widely used to assess gene-environment interactions. The most often cited advantage of the case-control design is efficiency; in order to accrue the same number of cases in a population-based cohort, a substantially larger number of women must be studied. However, in retrospective studies, exposures must be ascertained after the occurrence of disease. Although certain exposures can be studied retrospectively using questionnaires or, occasionally, using occupational or medical records, numerous potential biases can limit the ability to make causal inferences. In the Sister Study, we ascertain exposures and collected biological and environmental samples prior to the onset of disease. By banking samples for later analysis, we minimize the costs of measuring exposure, since only new cases and a sub-sample of non-cases will be analyzed.

We enrolled a cohort of 50,000 high-risk women (ages 35 to 74) who are the cancer-free sisters of women who have had breast cancer. This cohort will be followed annually for at least ten years. Comprehensive baseline questionnaire data, fasting blood, first morning urine, household dust and toenail

samples were collected. Study questionnaires were designed to collect information on known, suspected, and speculative risk factors in order to maximize the chance of detecting environmental risk factors of concern to women. In order to address new hypotheses, it was necessary to also collect comprehensive data on what is currently known, to be able to account for these factors in other analyses. In contrast to previous cohort studies where environmental exposures are either ignored or addressed superficially, we aimed to collect sufficient data so that environmental risks can be conclusively identified, or in the event that risks are not confirmed, we can say that we have thoroughly investigated the possibility. Brief annual updates record changes in contact information, environment and health. Bi- or triennial questionnaires address new hypotheses that arise. Our study will assess hypotheses that derive from the scientific literature, but we will also collect data on poorly studied exposures of concern to women. The incident cases that develop over time will be followed for the duration of the study, allowing us to also study the role of lifestyle, environmental exposures, and genes in prognosis and survival among women with cancer.

The Sister Study is an NIEHS intramural study designed to allow for trans-NIH and extramural collaboration. The study was developed in response to the heightened concern expressed by numerous women's and environmental groups about the possibility of increased breast cancer risk due to environmental causes, and the perceived lack of serious attention that had been paid to such concerns. In addition to consulting with scientific colleagues both within the government and at universities, we solicited the input of groups such as the National Breast Cancer Coalition and other advocacy and interest groups and consumers as we planned the Sister Study and assessed its feasibility. Our Steering Committee, Scientific Advisory Board, and Recruitment and Retention Advisory Board are diverse groups of professionals in the various areas of this complex effort. They include experts in epidemiology, breast cancer, biological specimen management, laboratory science, as well as representatives from various organizations focused on minorities, breast cancer support, and women.

**A.2. Purpose and Use of the Information**

Information collected in this study will be used to further scientific understanding of the effects of environmental exposures in women who are genetically at risk for developing breast cancer and to address questions of concern to women at risk for breast cancer. Epidemiologists and biostatisticians at NIEHS and their collaborators at other institutions will be responsible for testing the hypotheses of interest and disseminating results through the scientific literature. Results will be published in medical and epidemiologic journals as well as basic science journals when appropriate. Results will be presented at scientific meetings and at meetings of breast cancer advocates and other interested groups. Data will be used to assess current hypotheses regarding risk factors for breast cancer, to generate new hypotheses for subsequent analyses in the Sister Study, and specifically to identify preventable risk factors or combinations of risk factors. In addition to scientists and clinicians who will use this information in developing prevention strategies and to advise their patients, results will be reported to the women who participate in the study and to other women through the media, our website, and other Sister Study publications and newsletters. Results of the study may figure in future risk assessments and evaluations of the carcinogenicity of specific environmental agents and could be used in the development of exposure guidelines or standards, should important environmental risks be uncovered. Thus the results will be of use to Public Health officials, other scientists, physicians, elected and appointed officials, and women and their families.

The type and amount of information we collected at baseline *before* women develop breast cancer, and at subsequent yearly or bi/triennial intervals, fulfill many scientific and clinical needs. For breast cancer, many of the exposures of interest, including endogenous hormone levels, micronutrients, and even some environmental exposures, are measured most accurately in biological samples collected before the onset of disease or treatment and their associated symptoms and biological and lifestyle changes. The cohort design allows us to collect data on exposures, including biological exposure measures, diet and lifestyle, *before* the onset of disease. Forms for the currently approved long-term clinical follow-up data collection activities are in **Attachment 2.1 and 2.2**.

Brief self-administered forms are used annually to update changes in contact information and health status. Bi/triennial updates record changes in health, lifestyle, occupational and environmental exposures and address new hypotheses. The first Biennial Update focused on stress. As technology for self-collected environmental samples improves, women may be asked to provide such samples (for example, using in-home water test kits) as part of the bi/triennial follow-up.

Women who develop breast cancer or other conditions of interest during the course of follow-up are asked to allow us to obtain medical records and tissue samples from their health care providers.

Since the last review, we have completed full enrollment activities (data and sample collection) for the cohort, having enrolled 50,000 women whose sisters had breast cancer. Participants have completed Annual and Biennial Updates according to schedule. Average response rates for Update activities remain >90%. In addition, several reports have been published or submitted, describing factors under investigation:

Kim S, Parks CG, DeRoo LA, Chen H, Taylor JA, **Sandler DP**. Obesity and weight gain in adulthood and telomere length. Cancer Epid Biom Prev 2009;18:816-20.

Parks CG, Miller DB, Andrew ME, Cawthon RM, DeRoo LA, **Sandler DP**. Telomere length, current perceived stress, and urinary stress biomarkers in women. Cancer Epidem Biom Prev 2009;18:551-60.

Spector D, Mishel M, Skinner CS, DeRoo LA, VanRipper M, **Sandler DP**. Breast cancer risk perception and lifestyle behaviors among white and black women with a family history. Cancer Nursing 2009, in press.

Weinberg CR, Shore DL, Umbach DM, **Sandler DP**. Using risk-based sampling to enrich cohorts for endpoints, genes and exposures. Am J Epidemiol, 2007;166:447-455.

Xu Q, Parks CG, DeRoo LA, RM Cawthon, **Sandler DP,** Chen H. Multivitamin use and telomere length among women. Am J Clin Nutrition 2009;89:1-7.

Crowder K, **Sandler DP**, Parks C. Performance of a case classification algorithm for self-reported rheumatoid arthritis in relation to demographic factors, healthcare utilization, and smoking. Submitted

D'Aloisio AA, Baird DD, DeRoo LA, **Sandler DP**. Association of intrauterine and early life exposures with uterine leiomyomata in young women in the Sister Study. Submitted

Kim S, DeRoo LA, **Sandler DP**. Sociodemographic and lifestyle factors associated with sleep duration. Submitted

Kim S, DeRoo LA, **Sandler DP**. Eating patterns and nutritional characteristics associated with sleep duration. Submitted

Parks CG, DeRoo LA, Miller DA, McCanlies EM, Cawthon RM, **Sandler DP**. Work schedule and telomere length in The Sister Study. Submitted

**A.3. Use of Information Technology and Burden Reduction**

Web-based completion of annual updates is available. Bi/triennial follow-up forms may also be made available on the Web.

Computer Assisted Telephone Interview, or CATI, is a special data collection approach designed to reduce the burden to respondents and improve quality control. The Sister Study uses this technology to complete bi/triennial updates. This technology allows several advantages over the traditional pencil and paper method. First, it requires less paper. Second, there is no "mail wait" to get the information from participants. Also, the telephone interview requires little reading for the participant (only of help guides included in the enrollment Kit), an important factor when a segment of the population has low educational level or poor eyesight. Last, data extraction is more efficient with the CATI system as compared to the keyed entry method because skip patterns are automated and response inconsistencies can be queried at the time of the interview.

At enrollment during Phase 1 (now completed) name, address, SSN, date of birth, and medical information were collected. Personal identifiers are stored encrypted and separately from all other data. Now in Phase 2, PII are used to address Update materials, and to request medical records from physicians (with participant authorization). A Privacy Impact Assessment was completed for the Sister Study information management system.

**A.4. Efforts to Identify Duplication and Use of Similar Information**

The information we collect is not available from other sources. There is little consensus in the scientific community on how the environment impacts breast cancer. While some studies have addressed

environmental factors like diet, pesticides, and electromagnetic fields, no conclusive evidence exists because of limits in sample size and/or study design. Although there are cohort studies, such as Harvard's Nurses Health Study, that do address risk factors for breast cancer, none of these cohorts has collected substantial information on environmental and occupational exposures, and none includes biological and environmental samples for all of the participants. All of the existing cohorts focus on diet and other lifestyle factors. Large-scale prospective studies such as the Sister Study are needed to validate some of the already reported gene-environment interactions and to test new biologically-based hypotheses as both the technology and our understanding of synergistic mechanisms improve over time.

As noted above, there are important advantages to the Sister Study design. First, these sisters are at increased risk of breast cancer, likely due to shared genetic and environmental factors. Their family history increases the expected number of new cancers, the frequency of relevant gene polymorphisms, and the frequency of relevant exposures, making the study efficient compared to unenriched cohort designs. Second, because the study is prospective, blood samples and risk factor information were collected prior to diagnosis. A third advantage is that sisters are highly motivated, which should improve data quality and completeness and reduce loss to follow-up.

We are unaware of any duplication of this project with any other project now underway at other organizations. Several prospective cohort studies--for example, The Nurses' Health Study, Canadian National Breast Screening Study, New York University Women's Health Study, Iowa Women's Health Study—have investigated breast cancer in women, but none have focused primarily on the gene-environment link, especially in terms of the broader external environment.

**A.5. Impact on Small Businesses or Other Small Entities**

We are asking physicians of women who develop breast cancer or certain other conditions to provide confirmation of the diagnosis (with written permission from the women). We are minimizing the time and effort this will require by asking only for answers to key questions about procedures and findings and/or simply sending copies of the medical record. We are also requesting samples of tissue from stored diagnostic pathology specimens. In order to standardize the process and reduce the burden, we ask

physicians to mail the tissue samples to us, and we offer to store the samples, to be returned to the physician on request.

Although the number of women in the study is large, only a fraction (approximately 300 per year) will develop breast cancer, and the diagnoses will be spread over many years. Since the study is nationwide, it is not likely that any one medical practice will be contacted more than once or twice. The chances of being the physician of a woman in our study is greater for doctors whose practice is located in large cities and/or is affiliated with major medical centers. We estimate that the information we are requesting will take only 15 minutes to provide and in many cases can easily be provided by assistants in the physician's office.

**A.6. Consequences of Collecting the Information Less Frequently**

Annual updates take ~10 minutes, and bi/triennial self-administered questionnaires are 75 minutes or less. Annual contact cannot be done less frequently because the analysis relies on exposure and health-status changes over time, and ascertaining cases close to the time of diagnosis is important. A participant's recall diminishes greatly with time, and death may occur. Annual contact is necessary to preserve reliability and completeness and will facilitate maintenance of the cohort and tracing of those few who are lost to follow-up.

**A.7. Special Circumstances Relating to the Guidelines of 5 CFR 1320.5**

There are no special circumstances relating to the guidelines of 5 CFR 1320.5 and the project fully complies.

**A.8. Comments in Response to the Federal Register Notice and Efforts to Consult Outside Agency**

A 60-Day Federal Register Notice was published 10 July 2009 on page 33259. There were no public comments.

Early discussions were held with special interest groups to gain information on to what extent women who are at higher risk of developing breast cancer were willing to participate in a new longitudinal study requiring a long-term commitment and intense initial data collection. These meetings involved advocacy groups including cancer-free sisters, breast cancer survivors, Latina women, African-American women,

lesbian women, and other minorites. Support for this study was overwhelmingly enthusiastic and the majority of comments were favorable. These discussions gave us specific and valuable feedback that was incorporated into our study protocol.

Efforts to consult both within and outside NIEHS are summarized in **Attachment 3**. Also listed in **Attachment 3** are the current members of the Sister Study Scientific Advisory Board. Numerous researchers both within the NIH and in the extramural community were consulted during the planning of this study. Among the scores of scientists who were consulted during the early planning for the Study were, from the Division of Cancer Etiology and Genetics at the NCI: Drs. Patricia Hartge, Sholom Wacholder, Shelia Zahm, Aaron Blair, Jeffrey Streuwing, Louise Brinton, Rashmi Sinha, and Patricia Stewart; from the CDC: Drs. Thomas Sinks, Michelle Marcus (currently at Emory University), Heidi Blank, and Elizabeth Whelan (NIOSH); and from the National Institute of Aging, Dr. Tamara Harris. Other outside consultants include Drs. Pamela Horn Ross (Northern California Cancer Center -dietary phytoestrogens), Gladys Block (NCI - food frequency questionnaire), Julia Brody (Silent Spring Institute – exposure assessment) and Patricia Moorman (Duke University - medications and hormone replacement).

Our study was formally reviewed within the NIEHS intramural community, NIEHS advisory boards, and externally by federal and academic experts. Many continue to serve on the Sister Study Scientific Advisory and/or Recruitment and Retention Advisory Boards. As the study progresses, the archived data and samples will facilitate collaborations both within the NIH and in the extramural community.

Advisory to the study investigators in regular monthly meetings are the Sister Study Steering Committee, composed of NIEHS research staff Honglei Chen PhD (919-541-3782), Lisa DeRoo PhD (919-541-0799), Jane Hoppin ScD (919-541-7622), Stephanie London MD DrPH (919-541-5772), Christine Parks PhD (919-541-2577), and Jack Taylor MD PhD (919-541-4631), together with SSS senior staff, Polly Armsby, Deborah Bittner, Cynthia Kleeberger.

## A.9. Explanation of Any Payment or Gift to Respondents

During PHASE 1, participants received a prepaid 120-minute phone card valued at less than $5 as an incentive to complete the interviews and specimen collection at enrollment. While the <$5 value was not equivalent to the value of the effort required of participants, it helped to convey our appreciation of that effort.

Other tokens such as newsletters, magnets and bookmarks with the study logo are provided periodically to enhance participation and retention. Increasing response rates and retention will improve the quality of the scientific data we are collecting by minimizing response bias. Monetary and other incentives have been shown to significantly increase response rates.

## A.10. Assurance of Confidentiality Provided to Respondents

Procedures to protect the confidentiality of the study population and the data collected include the following:

- The data constitute a system of records under the Privacy Act System (#09-25-0134). A copy of the Federal Register Notice of System of Records of December 29, 1993 is included in **Attachment 4.**
- Each participant was assigned a study ID number. The ID alone is used to identify biological samples and all data forms. Only the ID number is entered into the database and used in the analysis of data. Subjects' names and addresses are stored separately.
- Any information linking subject's ID number to subject's name are kept in locked physical files or password-protected, restricted access electronic files at the North Carolina office of Social and Scientific Systems, Inc. (SSS), the NIEHS Epidemiology Branch Support Services Contractor.
- Employees of SSS undergo background checks, ethics training and sign a Pledge of Confidentiality (**see Attachment 5**).
- Only Sister Study research personnel have access to study data.
- Study results will be published in summary form only — no individual results will be published or shared.
- Shared samples and data will be provided without identifiers and study ID numbers will be scrambled to prevent accidental identification of participants.
- A **Certificate of Confidentiality** has been secured for this study (**see Attachment 6**).

- The proposal was initially reviewed by the Institutional Review Board of NIEHS on March 14, 2002 (see **Attachment 7**)**.** Also attached is an informal communication of approval for the most recent IRB Continuing Review (13 August 2009; **Attachment 8;** formal documentation not yet received) and Informed Consent form (completed in PHASE 1; **Attachment 9**).
- Informed consent forms spelled out the steps taken to protect privacy. Similar information was provided verbally at the time of enrollment, again prior to the CATI interview, and on the website.

The biological and environmental samples collected will be stored indefinitely in a secure building for future testing and may be disposed of at any time at the Investigator's discretion. Specimens are labeled with ID number only. These and related issues were explained in the Informed Consent documents. Specimens shared with outside researchers will be assigned a new identification number; the code linking the new and the old identification number will be known only to the NIEHS contractor responsible for the Sister Study field work. This new identification number will not be linked to any identifying information. Identifying information, such as name, social security number, or address will not be shared with other researchers. Samples will only be shared for scientifically valid studies that meet approved scientific and ethical standards. Samples and data that are shared can be used only for the specific research described in an approved research proposal and may not be used for other purposes without approval from the Sister Study investigators.

Participants may elect to leave the study at any time. As explained in the Informed Consent documents, no new data will be collected from patients who elect to drop out, but the data already provided will continue to be used in some analyses unless a written request to destroy data and specimens is received. Screening data on women who are found to be ineligible were not retained by the Sister Study.

**A.11. Justification for Sensitive Questions**

Breast cancer is a complex disease likely caused by many factors. The currently approved first Biennial Update includes a form that focuses on psychological stress and related measures. Many of the

questions address sensitive and personal issues such as personal experience with violence, physical and psychological abuse, and discrimination, and stress related to the sister's breast cancer diagnosis. In order to carry out a comprehensive analysis of stress, other measures focus on social support, personality traits and depression. The justification for each of these scales and their source and derivation is fully described in **Attachment 10**. In contrast to other updates, these forms are self-administered; no telephone option is available except in the rare case when a participant has difficulty reading or completing forms. As described in **Attachment 10**, stress may play an important direct or indirect role in risk for breast cancer and other health outcomes. It is likely however, that the relationship is complex and a comprehensive approach, including assessing acute and chronic stress at different time periods and accounting for factors that modify response to stress, is required. In addition to compelling scientific evidence, we learned through our early focus groups that many women believe that stress plays a role in risk for breast cancer and other diseases, and we have been urged, by the various constituencies that have endorsed the Sister Study, to study this topic of high interest to women.

Information is collected directly from participants. Participation is voluntary, and respondents can withdraw from the study at any time. Participants may refuse to answer specific individual questions, including those they find to be too sensitive or personal. All information is kept confidential to the extent provided by law. At no time will any individualized genetic results be given out. We have a Certificate of Confidentiality in place for this study. Participant informed consent forms are attached in **Attachment 9.**

**A.12. Estimates of Hour Burden Including Annualized Hourly Costs**

For the remainder of the study, women will be contacted once each year to update contact information and health status (5-10 minutes per response); and asked to complete short (60-75 minutes, total) updates every two-to-three years. Women diagnosed with breast cancer or other health outcomes of interest are asked to provide additional information about their diagnosis (20 minutes per response) and their doctors will be contacted to provide medical records related to diagnosis and treatments (15 minutes per response). The annual reporting burden is as follows: *Estimated Number of Respondents:* 50,000 study

participants and 2100 medical office staff. *Estimated Number of Responses per Respondent:* See table below:

| Activity (3-yrs) Estimated | Estimated Number of Respondents | Average Burden Responses per Respondent | Hours per Response | Estimated Total Burden Hours Requested |
|---|---|---|---|---|
| Annual Updates | 50,000 | 2 | 0.085 | 8,500 |
| Bi/Trienniel Follow-Up | 50,000 | 1 | 1.25 | 62,500 |
| Incident BC Case Follow-Up | 1800 | 1 | 0.33 | 594 |
| Incident Other Case Follow-Up | 300 | 1 | 0.33 | 99 |
| Incident Case Medical Office Contact | 2100 | 1 | 0.25 | 525 |
| **TOTAL** | | | | **72,218** |

*Average Burden Hours Per Response:* 0.7 hour; and *Estimated Total Burden Hours Requested:* 72,218 (over 3 years). The average annual burden hours requested is 24,073. The annualized cost to respondents is estimated at $14 (assuming $20 hourly wage X 0.7 hour). There are no Capital Costs to report. There are no Operating or Maintenance Costs to report.

**A.13. Estimate of Other Total Annual Cost Burden to Respondents or Recordkeepers**

There is no other total annual cost burden to respondents or recordkeepers.

**A.14. Annualized Cost to the Federal Government**

The estimated cost of contracting out information collection is $95,159,703. Federal Salary Support is $300,000 per year. Cost of archiving samples is about $350,000 per year, including purchase of freezers. Therefore, total cost is $104,259,703 over 14 years, with an annualized cost to the Federal Government of $7,447,122.

| Budget Category | Total Cost | Average Annual Cost |
|---|---|---|
| Labor | 25,759,744 | 1,839,982 |
| Federal Salary Support | 4,200,000 | 300,000 |
| Equipment | 743,831 | 53,131 |
| Subcontracts | 23,411,052 | 1,672,218 |
| Telephone | 1,322,739 | 94,481 |
| Printing | 2,773,604 | 198,115 |
| Supplies | 4,950,037 | 353,574 |
| Postage | 5,325,269 | 380,376 |
| Local Travel | 37,846 | 2,703 |
| Clinics/Hospital | 728,750 | 52,054 |
| Respondent Incentives | 1,762,628 | 125,902 |
| Other (recruiting, travel, consultants, rent) | 8,460,263 | 604,305 |
| Honoraria | 31,500 | 2,250 |
| Sample Storage | 4,900,000 | 350,000 |
| Indirect Costs and Fees | 19,852,440 | 1,418,031 |
| *Total Cost to Federal Government* | $104,259,703 | 7,447,122 |

### A.15. Explanation for Program Changes or Adjustments

Changes in burden and annualized cost merely reflect the anticipated progress and the next scheduled phases of the longitudinal study. With enrollment now complete, total burden request decreases from 194,131 hours over 3 years (in 2006 revision) to 72,218 hours over 3 years (for 2009 revision). This **does not** however represent a change in Protocol — merely completion of the Enrollment Phase of the Protocol, and normal **as-planned** progression of follow-up.

### A.16. Plans for Tabulation and Publication and Project Time Schedule

The primary goal of the study is to identify environmental and familial risk factors for breast cancer and other diseases by studying a cohort of sisters of women who have had breast cancer. The Sister Study is not designed around one particular *a priori* hypothesis, but is designed to allow us to address a number of hypotheses regarding gene-environment interactions and risk for breast cancer. Current hypotheses regarding environment-gene interactions will be addressed in the early years of the study. The reports

generated from this study will include the following risk factors that may be influenced by the action of genes with known polymorphisms:

**A.16-1**

| Factor | Genetic Marker of Interest |
|---|---|
| Cigarette smoke | CYP1A2, NAT2, GSTM1, CYP1A1, DNA repair polymorphisms, CYP2A6, CYP2C9 |
| Exogenous hormones | CYP17, CYP1A2, CYP1A1, estrogen receptor polymorphisms |
| Hormonal risk factors | CYP17, aromatase, hormone receptor polymorphisms |
| Oxidative stress | Genes involved in oxidative stress |
| Melatonin | Polymorphism screening |
| Nutritional and dietary factors | Vitamin D metabolism and receptor polymorphisms, CYP1A2, NAT2 |
| Grilled meat consumption | GSTM1 |
| Alcohol | ADH, ALDH, CYP2E1 |
| Sunlight exposure | Vitamin D metabolism and polymorphisms |
| Calcium, calcium channel blockers | Vitamin D metabolism receptor, estrogen metabolism and receptor polymorphisms |

Clearly, by the time enough cases have accrued, scientific understanding of biologic mechanisms and genes will have advanced considerably. A prospective study such as the Sister Study is designed to respond to new hypotheses as they emerge.

**A.16 - 2**

**Project Time Schedule**

| Activity | Time Schedule |
|---|---|
| Began vanguard phase of enrollment | August 2003 |
| Began nationwide enrollment | October 2004 |
| Completed PHASE 1 Enrollment questionnaires and collect specimens on 50,000 women | August 2009 |
| Began first Annual Update (self-administered questionnaire) | June 2005 |
| Began first Bi/Triennial Update | March 2008 |
| Analyses | Began mid-2007 |
| Publication | Began 2007 |

**A.17. Reason(s) Display of OMB Expiration Date is Inappropriate**

We request continued approval to display the OMB control number without the expiration date on printed forms. This request is based on the precedent (see OMB No. 0935-0104: Medical Expenditure Panel Survey: Survey About Your Diabetes Care) that this is a longitudinal study scheduled to last for 14 or more years, and for which individual enrollment and follow-up activities span across a number of years, thus across OMB expiration dates. Annual update forms as well as two entire parts of the 3-part bi/triennial update will be used throughout the length of the study. Therefore, the cost of multiple printing cycles, merely to change OMB date, rather than taking advantage of cost savings realized with larger batch printing of approved materials (hundreds of thousands of pages) that undergo little or no contextual change is inordinately costly to the government. Nonetheless, these items would continue to be included in each 3-year revision package sent to OMB for continuation of approval.

**A.18. Exceptions to Certification for Paperwork Reduction Act Submissions**

There are no exceptions to certification for this submission.