

AD _____

Award Number: DAMD17-00-1-0417

TITLE: Environmental Exposures at Birth and at Menarche and Risk
of Breast Cancer

PRINCIPAL INVESTIGATOR: Jo L. Freudenheim, Ph.D.

CONTRACTING ORGANIZATION: University of New York at Buffalo
Amherst, New York 14228-2567

REPORT DATE: June 2002

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

6/50
1118

REPORT DOCUMENTATION PAGEForm Approved
OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE June 2002	3. REPORT TYPE AND DATES COVERED Annual (1 Jun 01 - 31 May 02)	
4. TITLE AND SUBTITLE Environmental Exposures at Birth and at Menarche and Risk of Breast Cancer			5. FUNDING NUMBERS DAMD17-00-1-0417	
6. AUTHOR(S) Jo L. Freudenheim, Ph.D.				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of New York at Buffalo Amherst, New York 14228-2567 E-Mail: Jfreuden@buffalo.edu			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012			10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
20021118 049				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited			12b. DISTRIBUTION CODE	
13. Abstract (Maximum 200 Words) (abstract should contain no proprietary or confidential information) This is a population-based study that examines location of residence at birth and menarche in relation to proximity to industrial sites, gasoline stations, toxic waste sites and heavily trafficked roadways as risk factors for subsequent breast cancer. It also examines estimated exposure to benzene and to PAHs as risk factors and evaluates genetic susceptibility in relation to these exposures and breast cancer. There are 15,969 individual addresses, representing 3,091 participants in Erie and Niagara counties, in the study. A validation study was conducted to assess the positional accuracy of addresses geocoded in the GDT enhance TIGER file using a Global Positioning System (GPS). Blood samples from all participants with samples have been sent for DNA extraction and genotyping. Standard Industrial Classification (SIC) directories are being used to categorize exposure groups and to geocode point sources of pollution. To date, 13 uranium processing sites and 8 steel mill sites have been geocoded and proximity to birth residence established. Clustering analysis was conducted to identify geographic patterns of residence for breast cancer cases and controls at critical time periods.				
14. SUBJECT TERMS breast cancer, cancer etiology, carcinogenesis, environment, early childhood exposures, geography, molecular epidemiology, gene-environment interaction			15. NUMBER OF PAGES 19	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited	

Table of Contents

Cover	1
SF298	2
Introduction	4
Body	5
Key Research; Accomplishments	15
Reportable Outcomes	16
Conclusions	17
References	19

INTRODUCTION

In this population-based study we are examining environmental exposures experienced at birth and at menarche as risk factors for breast cancer. We will examine location of residence during these potentially sensitive time periods in relation to proximity to industrial sites, gasoline stations, toxic waste sites and heavily trafficked roadways as risk factors for subsequent disease. Residential histories were obtained from all participants during our case-control study, which ended May, 2001. This study included women, age 35-79 with incident, primary, histologically confirmed breast cancer living in Erie or Niagara counties. Controls are frequency matched to cases on age, race and county of residence. Residence at the time of birth and menarche, as well as the potential exposure sites will be geocoded into GIS. The primary objectives of this study are: 1) To investigate distance from steel mills, chemical factories, gasoline stations, toxic waste sites, other industrial sites and major roadways of the residence of cases and controls at the time of birth and at menarche as risk factors for pre- and postmenopausal breast cancer. 2) To examine estimated exposure to benzene and to PAHs as risk factors for pre- and postmenopausal breast cancer. 3) To evaluate genetic susceptibility in relation to these exposures and breast cancer risk by examining genetic variability in metabolism by NQ01, GST M1-1, GST P1-1 and CYP 1A1. Potential confounding factors will also be assessed. These include age, education, income, family history of breast cancer, Quetelet index, body fat distribution, having been breastfed, age at menarche, age at menopause, pregnancy history, lactation and contraceptive history, menstrual cycle length, birth weight, smoking and passive smoke exposure history, and diet and occupational history. There are no major results to report at this time. However, some preliminary findings are discussed in the text of this report.

BODY OF REPORT

Task 1: Investigate distance from steel mills, chemical factories, gasoline stations, toxic waste sites, and other industrial site of the residence of cases and controls at the time of birth and at menarche as risk factors for pre- and postmenopausal breast cancer.

A. Geocoding

The geocoding of residential addresses has been completed and preliminary analyses are currently underway for task 1. Data collection for a more extensive validation of the GDT TIGER reference theme is almost complete, where the actual latitude and longitude of 200 randomly selected addresses was measured with Global Positioning System (GPS) units. In addition, several preliminary analyses have been completed for 1) spatial clustering of cases and controls at birth and menarche, 2) the potential for selection bias, and 3) residential proximity to 13 industrial sites. The following report describes in more detail the current status of the geocoding and the preliminary analyses.

Geocoding Status

Addresses for 3,091 participants were entered into database that permits geocoding of the addresses. There are total 20,862 individual addresses. Among them, 77% (15,969/20,862) addresses are within Erie and Niagara counties.

We have compiled 15,969 addresses (within Erie and Niagara Counties) that are ready for geocoding. Of these 15,969 addresses, 3891 had some missing information required for geocoding. Polk Directory searches were conducted for 43% (1669/3891) of these addresses; in some cases such searches were not possible because either there were no known dates of residence or there was no Polk Directory for the years/town when the subject resided at that particular address. The success rate for finding the missing address information in the Polk Directory was 48% (806/1669). To find more missing addresses, especially the address at birth, applications for copies of the birth certificates have been submitted to the health department.

Because our major interest is environmental exposures at the birth and at menarche, the completeness and accuracy of the data on year of residences are great concerns. Significant efforts were used to check the reported residential year moved in and out. A protocol was developed and all the years were checked. 1105 corrections were made on the recorded years, among which 78% (862/1105) were associated with residences in Erie and Niagara counties.

We are using New York State industrial directories to identify point sources of pollution and to estimate exposure to benzene or polycyclic aromatic hydrocarbons (PAHs). These directories are primarily based on information collected by the New York State Department of Commerce. According to New York State law, every firm is required to submit directory registration forms. Information in the directory includes the firm's name, address, products, number of employees, and standard industrial classification (SIC). This information is being used to geocode point sources of pollution, categorized

by exposure groups (for example, chemical factories, steel mills, solvent exposure), and to estimate benzene and PAHs exposure. We are currently doing data entry of industrial codes and geocoding these data. We will also collect historical information regarding toxic waste sites and heavily trafficked roadways and railways.

B. Validation of GDT enhanced TIGER file

The purpose of this validation was to assess the positional accuracy of the geocoded addresses in the Geographic Data Technology (GDT) and enhance Topologically Integrated Geographic Encoding and Referencing system (TIGER) file for use in epidemiologic studies. The GDT enhanced TIGER file is being validated by using a global positioning system (GPS) unit to measure the latitude and longitude of 200 randomly selected geocoded addresses. This validation is intended to assess the accuracy of the Arcview geocoding process with the TIGER file as the reference theme. Simply, how close to the actual latitude and longitude of participants addresses are the geocoded positions.

The Erie and Niagara county residential addresses were first geocoded using the standard ArcView geocoding commands with zipcode as the zone variable. Of the addresses that matched, a random sample of 200 addresses was selected for this validation. These addresses were randomized in such a way that there were 50 addresses of cases and 50 addresses of controls in urban areas (Buffalo, Kenmore, Lackawanna, and Niagara Falls) as well as 50 cases and 50 control addresses from outlying areas. This was done so that the positional accuracy between urban and outlying areas could be compared in addition to the case control comparisons.

The actual latitude and longitude were measured with 2 GPS units (Garmin's etrex and Magellan's Pioneer). Investigators drove to each address and recorded the latitude and longitude of each address from a street location closest to the address in question.

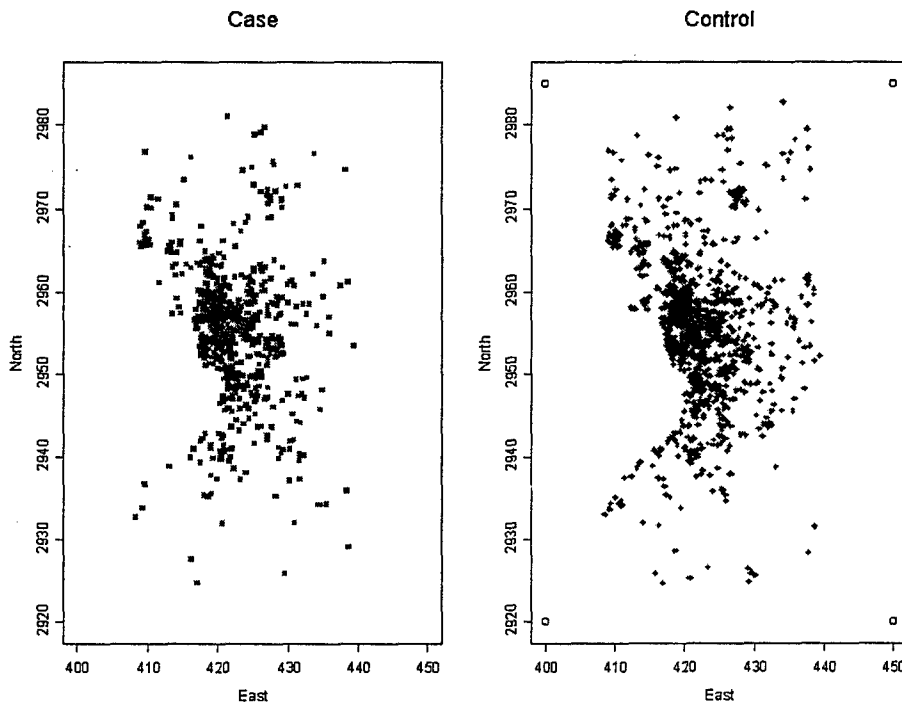
Observed latitudes and longitudes were then transcribed into a dataset and their coordinates will be located on the TIGER reference theme. The latitude and longitude for each geocoded point will be computed from the TIGER file. This will produce two points for each address, one observed latitude and longitude and one derived from the geocoding process. The distance between the 2 points for each address will be estimated with ArcView and the mean difference the standard deviation between the observed coordinates and the geocoded coordinated will be computed.

The differences between the geocoded location and the actual location will be statistically evaluated. The Student's T-test will be used to compare the mean differences in distance between:

- 1) cases and controls
- 2) urban and outlying areas
- 3) cases and controls stratified by urban/outlying areas.

C. Analysis of Geographic Clustering of Cases and Controls by Period of the Lifetime

To identify geographic patterns of breast cancer case and control in space and time, we have examined clustering patterns of residence at critical time periods. Figure 1 shows the location of cases and controls in the study area. A dot represents each residential location. We used a square boundary instead of a real county boundary to protect confidentiality. Using the spatial k-function comparing geographic distribution of breast cancer cases and controls over a range of spatial scale, we found evidence of clustered patterns of residence at critical time periods. There was slight evidence of clustering generally and a tendency of a high degree of clustering at small scale (geographic areas). The cases were more clustered on a small scale than controls. We found significant clustering of residence for breast cancer cases compared to controls in early life, especially for the place of residence at menarche. The evidence of clustering of residential locations at menarche and birth are stronger than those for those in later life. In particular, we have examined clustering at the time of the participants' first birth. The residential locations at menarche show strong evidence of clustering for cases compared to controls and suggest that there may be an influence of exogenous risk factors on breast cancer particularly at that time.



D. Examination of Geographic Selection Bias in the Data Set

To examine whether there is any selection bias among participating of breast cancer cases and controls, we compared the geographic distribution of cases and controls with the reference population (females age 35-79) by participation status (participants and non-participants). Mapping and centrographic measures were used for the comparison of geographic distribution and the chi-square test was used for aggregated data. We made case-control comparisons in relation to the reference population; controls were compared to the female population of Erie and Niagara counties based on census data, and participating cases were compared to incidence data from the New York state cancer registry. Second, participant-non-participant comparison by case-control status; participating versus nonparticipating controls and participating and nonparticipating cases.

Comparisons to the Reference Population

1) There was no remarkable difference in geographic distributions of controls compared to the census data for women of this age living in the two counties. Based on visual inspection and centrographic methods, we identified geographic patterns (central tendency and degree of dispersion) of controls and compared it with the reference population. The centrographic measures enable us to identify geographic concentration of point patterns such as geographic center point, as well as distributional patterns including the degree of dispersion or clustering, direction, orientation, and shape.

Table 1 and Figure 1 show the mean center and standard distance between control and the reference population. There was no remarkable difference in geographic concentration of participating controls compared to the reference population. The mean center of the two groups is slightly different; the distance between the mean center of control and population is 0.8 miles.

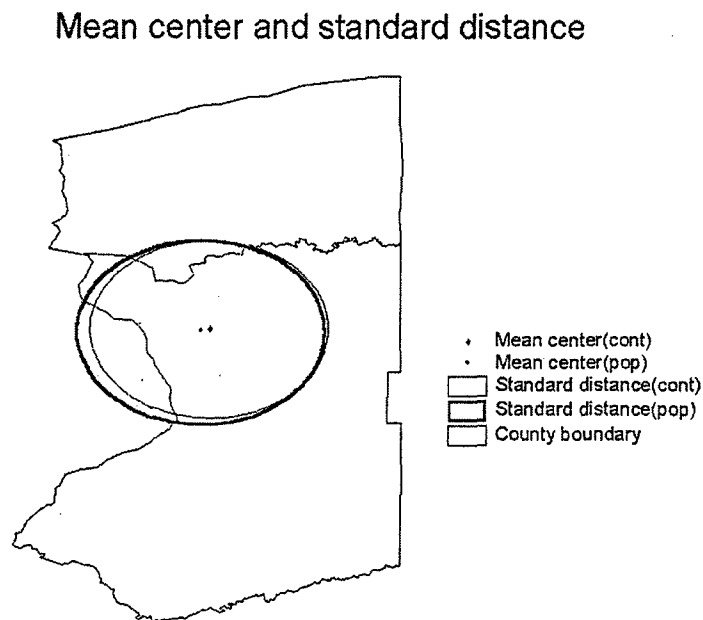
Table 1. Geographic concentration of control and reference population

	Control	Population
Mean center	422.3444, 2955.5567	421.5869, 2955.2234
Distance*	0.823*	
Standard deviation	5.484, 8.891	5.563, 9.255
Standard distance	10.45	10.85

* Mean center distances between control and population

** All units in the table are UTM miles

Figure 1. Mean Center and standard distance



2) We compared geographic distribution of cases to the incidence data from the New York state cancer registry. Because detailed locational information on the incidence number is not available, a chi-square test was used for the comparison of cases to breast cancer incidence data from the New York state cancer registry at the zip-code level. Chi-square test results were statistically significant suggesting different geographic distribution at the aggregate level. More detailed analysis on geographic distributions of cases in each zip area suggested that there may be some over-representation or under-representation in the sampling process. Since it is more likely to participate for women living closer to the university, we focused on the nearby area by drawing 3-5 mile circular windows around the university. We identified several zip-code areas overrepresented, (slightly higher number of cases than average, based on incidence numbers from New York State Cancer Registry), underrepresented, and both over and under-represented.

Comparisons of Participants with Those Refusing to Participate

3) There was no significant difference between respondents and non-respondents. Table 2 summarizes information on the 1,170 participating and 788 non-participating cases and 2,116 participating and 1,575 non-participating controls. The chi-square goodness-of fit statistic shows no significant difference between two groups (chi square of 1.56 significant at 0.01).

Table 2. Total number of respondent and non-respondent by case-control status

	Participant	Non-participant	Total
Case	1170	788	1958
Control	2116	1575	3691
Total	3286	2363	

4) Finally, aggregate-level comparison was performed to compare both cases and controls by participation status. Chi-square test for participating and nonparticipating case does not show any difference between two groups, while there are different patterns of control by participation status.

E. Examination of Breast Cancer Risk in Relation to Residential Proximity to Industrial Sites Contracted by the U.S. Atomic Energy Commission

Thirteen industrial sites in Erie and Niagara counties contracted with the U.S. Atomic Energy Commission in the 1940's and 1950's to process uranium for the U.S. nuclear program. A preliminary analysis of proximity to these sites revealed a more than three fold increase in premenopausal breast cancer for women born within 2.45 miles of one of these sites. An abstract based on these analyses has been accepted to the American College of Epidemiology Conference in September of 2002.

We are interested in the possible impact of residential proximity to these sites, particularly exposure in early life. We are examining residence at birth and at menarche. Descriptive characteristics of the subjects included in these analyses who lived in the two counties at the time of birth and at menarche are shown in Tables 3 and 4. Odds ratios (OR) and 95% Confidence Intervals (95% CI) for proximity to industrial sites at birth are shown in Table 5. Proximity of less than 2.45 miles to the closest industrial site at birth was associated with a nearly four-fold increase in risk of subsequent breast cancer (OR= 3.7; 95%CI 1.9-7.7) in premenopausal women compared to women whose birth address was more than 7.99 miles from the closest of these sites. Adjustment for age, parity, age at menarche, age at first birth, and education did not substantially alter the crude ORs. For the post menopausal women, there was no association between residential proximity to these sites and risk.

When the residence at menarche (Table 6) was examined, weaker associations between proximity and subsequent breast cancer were seen. For Premenopausal women residing within 2.45 miles of the closest site there was a 2-fold increase in risk (OR= 1.99; 95% CI 1.2-3.3) compared to women residing greater than 7.99 miles. Again, no association was seen for the postmenopausal women.

These findings are provocative but preliminary. In particular, we are attempting to determine whether there is biological plausibility and whether there would have been contamination at these residences. We will investigate the radiological surveys of five of the sites. Eight of these sites have been listed for current or past remediation and the US

Army Corps of Engineers maintains administrative files with radiological surveys for these sites.

Table 3. Descriptive Characteristics for residence at birth

Characteristic	Cases (n=266)	Controls (n=411)	P-value
Pre menopausal			
	Mean (SD)	Mean (SD)	
Age (yrs.)	43.75 (3.99)	44.00 (3.89)	0.529
Education (yrs.)	13.82 (1.93)	14.14 (2.41)	0.132
Age at Menarche (yrs.)	12.37 (1.76)	12.45 (1.73)	0.650
Parity	1.86 (1.28)	1.81 (1.23)	0.656
Age at First Birth (yrs.)	20.53 (10.89)	20.41 (11.35)	0.914
Birth Year	1955.95 (4.04)	1953.04 (3.94)	0.000
Menarche Year	1954.27 (160.70)	1965.49 (4.36)	0.394
First Birth Year	1962.43 (161.71)	1966.85 (115.30)	0.740
Post menopausal			
Age (yrs.)	54.83 (2.71)	53.40 (1.90)	0.000
Education (yrs.)	13.87 (2.91)	13.54 (2.17)	0.328
Age at Menarche (yrs.)	12.09 (1.52)	12.46 (1.99)	0.118
Parity	2.01 (1.21)	2.29 (1.37)	0.103
Age at First Birth (yrs.)	19.66 (9.51)	20.18 (9.13)	0.673
Birth Year	1944.12 (2.78)	1943.38 (2.09)	0.024
Menarche Year	1956.22 3(06)	1938 (182.41)	0.310
First Birth Year	1963.78 (10.14)	1946.69 (183.36)	0.317

Table 4. Descriptive Characteristics for residence at Menarche

Characteristic	Cases (n=593)	Controls (n=975)	P-value
Pre menopausal			
	Mean (SD)	Mean (SD)	
Age (yrs.)	44.17 (3.85)	44.11 (4.03)	0.858
Education (yrs.)	14.00 (2.04)	14.12 (2.35)	0.476
Age at Menarche (yrs.)	12.57 (1.64)	12.52 (1.66)	0.751
Parity	1.75 (1.18)	1.90 (1.25)	0.144
Age at First Birth (yrs.)	20.61 (11.06)	20.87 (10.85)	0.775
Birth Year	1954.45 (3.90)	1952.85 (4.14)	0.000
Menarche Year	1967.01 (4.28)	1965.37 (4.52)	0.000
First Birth Year	1975.05 (11.97)	1969.22 (95.43)	0.380
Post menopausal			
Age (yrs.)	60.52 (6.04)	61.34 (6.64)	0.053
Education (yrs.)	13.44 (2.72)	13.00 (2.17)	0.010
Age at Menarche (yrs.)	12.45 (1.53)	12.84 (1.67)	0.000
Parity	2.42 (1.70)	3.05 (1.86)	0.000
Age at First Birth (yrs.)	19.36 (9.83)	21.03 (7.99)	0.005
Birth Year	1938.19 (6.06)	1935.80 (6.24)	0.000
Menarche Year	1950.64 (5.94)	1948.65 (6.01)	0.000
First Birth Year	1957.55 (11.35)	1956.83 (9.63)	0.300

Table 5. Odds ratios and 95% confidence intervals for pre and post menopausal women by quartiles of distance from nearest industrial site contracted by the Atomic Energy Commission at birth

Distance (miles)	Cases (n=266)	Controls (n=411)	OR (95%CI)*	OR (95%CI)**
1 (>7.99)	12	54	1.00	1.00
2 (4.46-7.99)	16	69	1.04 (0.46-2.39)	1.07 (0.46-2.45)
3 (2.45-4.45)	53	90	2.65 (1.30-5.40)	2.67 (1.30-5.46)
4 (<2.45)	69	83	3.74 (1.85-7.55)	3.77 (1.86-7.65)
Postmenopausal women				
1 (>7.99)	44	48	1.00	1.00
2 (4.46-7.99)	44	35	1.37 (0.75-2.51)	3.50 (1.63-7.51)
3 (2.45-4.45)	14	12	1.27 (0.53-3.05)	2.64 (0.98-7.16)
4 (<2.45)	14	20	0.76 (0.34-1.69)	1.71 (0.69-4.22)

*Crude model

** Adjusted for age, education, age at first birth, parity, and age at menarche.

Table 6. Odds ratios and 95% confidence intervals for pre and post menopausal women by quartiles of distance from nearest industrial site contracted by the Atomic Energy Commission at menarche.

Distance (miles)	Cases (n=638)	Controls (n=1034)	OR (95%CI)*	OR (95%CI)**
Pre menopausal				
1 (>7.99)	28	90	1.00	1.00
2 (4.46-7.99)	30	78	1.24 (0.68-2.25)	1.24 (0.68-2.26)
3 (2.45-4.45)	87	173	1.62 (0.98-2.65)	1.60 (0.97-2.63)
4 (<2.45)	71	114	2.00 (1.19-3.36)	1.99 (1.18-3.34)
Post menopausal				
1 (>7.99)	87	150	1.00	1.00
2 (4.46-7.99)	99	153	1.12 (0.77-1.61)	1.13 (0.77-1.65)
3 (2.45-4.45)	121	119	1.75 (1.22-2.53)	1.63 (1.08-2.45)
4 (<2.45)	115	160	1.24 (0.87-1.77)	1.13 (0.76-1.70)

*Crude model

** Adjusted for age, education, age at first birth, parity, and age at menarche.

F. Proximity to Chemical Factories

We have collected historical information about chemical factories operating in this region by industrial directories. We have done some preliminary analysis on risk of breast cancer and residential proximity to chemical facility 10 years before the diagnosis date (for cases) or interview date (for controls). We will also do these analyses for the childhood time periods of interest. The 1988 New York State Industrial Directory was used to extract the information on chemical facilities, including address of plant, number of employees and Standard Industrial Classification (SIC). There were all together 99 chemical facilities (SIC groups 28 for Chemicals and allied products, 29 for Petroleum refining and related industries, and 30 for Rubber and miscellaneous plastics products) in Erie and Niagara counties in 1988, which has 20 or more employees. At this time, we have been able to geocode 79 of these factories. The others had incomplete address information. Results are shown in the Table 7. This preliminary analysis does not suggest that living close to a chemical facility 10 years prior to diagnosis is associated with breast

cancer risk. However, among postmenopausal women, there is a nonsignificant 8% increased risk of breast cancer for those who lived within 0.68 miles of a chemical facility.

Table 7. Residential proximity to chemical facility 10 yrs ago and Risk of Breast Cancer

Categories of Distant	Cases	Controls	Crude OR	Adjusted OR* (CI)	P for trend
<u>Pre-menopausal</u>					
1: >2.64 miles	64	122	1.00	1.00	0.33
2: 1.24~2.64	45	117	0.73 (0.46-1.16)	0.75 (0.47-1.18)	
3: 0.68~1.24	65	108	1.15 (0.75-1.77)	1.16 (0.75-1.79)	
4: <=0.68miles	58	125	0.89 (0.57-1.37)	0.89 (0.57-1.38)	
<u>Post-menopausal</u>					
1: >2.64 miles	173	293	1.00	1.00	0.84
2: 1.24~2.64	154	296	0.88 (0.67-1.16)	0.89 (0.68-1.17)	
3: 0.68~1.24	158	307	0.87 (0.67-1.14)	0.89 (0.68-1.17)	
4: <=0.68miles	184	296	1.05 (0.81-1.37)	1.08 (0.83-1.41)	

* Odds ratios and 95% confidence intervals adjusted for age and education.

Task 2: To examine estimated exposure to benzene and to PAHs as a risk factor for pre- and postmenopausal breast cancer, with control for appropriate confounders.

Environmental benzene is the simplest form of aromatic hydrocarbon. According to the Total Exposure Assessment Methodology (TEAM) studies conducted by the U.S. Environmental Protection Agency, 99% of the total personal exposure to benzene was from inhalation, and benzene exposure through ingestion of water and food was minimal. The major source of benzene exposure was smoking. For smokers, 89% of the benzene exposure was from mainstream cigarette smoke. For nonsmokers, 10% of the benzene exposure was from passive smoke. Other sources of benzene exposure include auto exhaust or gasoline vapor emission, as well as industrial emissions. In the TEAM studies in the Los Angeles area, although autos and industry accounted for 96% of the outdoor airborne level of benzene, the personal benzene level was determined primarily by smoking, indoor air and personal factors. In the TEAM studies in Beaumont TX, Bayonne and Elizabeth NJ, Los Angeles, Antioch and Pittsburg, CA, they found no effect between personal exposure and living close to major fixed sources of benzene, such as oil refineries, storage tanks and chemical plants. In our study, we have lifetime smoking histories for all study participants, which will be utilized as one of the resources to estimate benzene exposure. For earlier environmental benzene exposure, especially at birth, since study participants were unlikely to be smokers (although some may be exposed to passive smoking), we will estimate the benzene exposure from proximity to major traffic road ways and relevant industrial sites.

Polycyclic aromatic hydrocarbons (PAHs) include more than 200 congeners, produced from incomplete combustion of organic material. Common sources of PAHs include: cigarette smoke, certain industrial processes and smoked food products, such as charbroiled fish or meats. People are exposed to PAHs through air, food, water and skin

contact. PAHs in the air may be found attached to dust particles and as solids in soil or sediment. Cigarette smoke may increase exposure to PAHs 30-fold. Occupational exposure to PAHs commonly occur in coke, coal tar, aluminum/iron/steel production, municipal trash incineration, asphalt production and smokehouse operations. Moreover, according to the report by the Agency for Toxic Substances and Disease Registry (ATSDR), PAHs have been found in at least 600 hazardous waste sites nationwide. PAHs in the air generally degrade within days or weeks, by reacting with sunlight and other atmospheric compounds. PAHs in the soil or water generally degrade within weeks or months by microorganisms. When PAHs enter the human body they may be stored in the kidneys, liver, and tissues that contain fat. A small amount of PAHs may also be stored in the spleen, adrenal glands and ovaries. Usually, most PAHs are stored in the human body for a few days and leave through feces and urine. Since PAHs widely exist in the environment and may enter the human body through different pathways, it is difficult to accurately measure total PAHs exposure. In our study, to estimate PAHs exposure we will use the information collected on lifetime smoking, food consumption (24-12 months prior to the interview), traffic roadways and industrial PAHs emission sources (coke ovens, refineries, carbon graphite facilities, coal combustion boilers and medical waste/industrial incinerators).

We are currently collecting and analyzing the benzene and PAHs exposure information from lifetime smoking histories, traffic roadways and industrial and toxic waste sites. We will use this information to see if there is an association between exposure to benzene and PAHs and breast cancer risk.

Task 3: To evaluate genetic susceptibility in relation to these exposures and breast cancer risk by examining genetic variability in metabolism by NQ01, GST M1-1, GST P1-1 and CYP1A1.

A total of 2,985 subject blood samples have been shipped to Dr. Peter Shields' laboratory (Lombardi Cancer Center) for DNA extraction and genotyping analysis. The DNA extraction and genotyping is currently in progress.

KEY RESEARCH ACCOMPLISHMENTS

- A validation study was conducted to assess the positional accuracy of addresses geocoded in the GDT enhance TIGER file using a Global Positioning System.
- All remaining biological samples have been sent for DNA extraction and genotyping.
- We were able to use Polk directories for 1,669 of the 3,891 individual addresses with missing information. Our success rate for finding missing information using the Polk Directory was 48%, approximately 800 addresses.
- A protocol was developed to check the reported dates of residences (year moved in and year moved out). We were able to make 862 corrections associated with residences in Erie and Niagara counties.
- 13 sites of processed uranium and 8 steel mill sites have been geocoded and proximity to birth residence and residence at menarche established.
- Standard Industrial Classification (SIC) codes have been used to identify point sources of benzene or polycyclic aromatic hydrocarbons (PAH's).
- Clustering analysis was conducted to identify geographic patterns of residence for breast cancer cases and controls at birth, menarche, and age at first birth.
- A presentation, "Geographical Epidemiology of Breast Cancer in Western New York" was given at the Annual Meeting of the Association of American Geographers in Los Angeles, CA, March, 2002.
- A presentation, "Environmental Exposures associated with Lifetime Residential History: A GIS-based clustering Analysis of breast cancer," will be presented at the Annual Meeting of the Society for Epidemiologic Research, Palm Desert, CA in June, 2002.
- An abstract, "Residential Proximity at Birth to Industrial Sites and Subsequent Risk of Breast Cancer," was accepted to the American college of Epidemiology Conference, Albuquerque, NM in September, 2002.

REPORTABLE OUTCOMES

PRESENTATIONS:

Geographical Epidemiology of Breast Cancer in Western New York: Migration and Disease Clustering. Daikwon Han, Department of Social and Preventive Medicine, University at Buffalo. Presented at the Annual Meeting of the Association of American Geographers, Los Angeles, CA, March, 2002.

Migration has a significant effect on geographic variations of disease and health outcomes. The complex process of human movement is one of the complicating factors in explaining the causal relationships between disease and environment, but also an important determinant of human health due to the exposure to disease through movement. This study explores the migration effects on disease clustering to assess; 1) the importance of residential locations to the risk of breast cancer, 2) the statistical significance of clustering with migration effects. To identify the reasons for geographic variations of disease, the study presents hypotheses associated with migration and disease risks. Exploratory analyses in a GIS environment are used to detect the spatial-temporal patterns of residential locations and clustering of case-controls in Western New York. The overall effects of migration on disease clustering are identified by comparing the lifetime residential history of case-controls, after controlling for the known risk factors such as age and history of breast cancer. The investigation on the role of migration on disease clustering processes provide explanations on the consequences of in- and out-movement of people diagnosed with disease on the risk of disease as well as on the spatial variations of disease. Once significant clusters are identified, further work is required to investigate the relationships between residential changes and environmental exposures in explaining unknown etiology of breast cancer.

Environmental exposures associated with lifetime residential history: A GIS-based clustering analysis of breast cancer. Daikwon Han, Jing Nie, Matthew R. Bonner, and Jo L. Freudenheim. Department of Social and Preventive Medicine, University at Buffalo. Abstract presented at the Annual Meeting of the Society for Epidemiologic Research, Palm Desert, CA, June, 2002.

There is increasing evidence that early exposures may be related to risk of breast cancer. We were interested in whether there was clustering of breast cancer based on their residence in early life and identified spatio-temporal clustering of cases and controls at critical time periods, residential locations at birth, at menarche, and at the women's first birth. Data used here were part of the Center for Preventive Medicine case control study of incident, pathologically confirmed breast cancer (1996-2001) in Erie and Niagara counties. Controls were frequency matched on age and county of residence; controls less than 65 were randomly selected from the New York State Department of Motor Vehicles list and those greater than 65 from the Health Care Finance Administration list. All cases and controls provided lifetime residential histories. The spatial k-function method was used to calculate the distance between each residence within a certain search radius and

to compare observed with expected patterns over pre-specified distances. We found a general tendency of spatial clustering for cases for these time periods, especially at small geographic scales, compared with the simulated theoretical distribution of expected patterns. The evidence for clustered residence at birth and at menarche was stronger than that for first birth. This study provides additional evidence that early environmental exposures may be related to breast cancer risk.

ABSTRACT:

Residential Proximity at Birth to Industrial Sites and Subsequent Risk of Breast Cancer, MR Bonner, D Han, J Nie, JL Freudenheim, JE Vena, Department of Social and Preventive Medicine, University at Buffalo. Accepted to the American College of Epidemiology Conference, Albuquerque, NM, September, 2002.

Purpose: To investigate the relationship between residential proximity at birth to industrial sites contracted by the Atomic Energy Commission (AEC) to process radioactive material and the subsequent development of breast cancer (BC) in pre and post menopausal women.

Methods: We used a completed case-control study (n=3,335) and restricted subjects to lifetime residents of Western New York (n=1,181). Subjects were further restricted to those born in 1940 and later because the first industrial sites began operating under the AEC contract in 1940. A total of 266 primary incident breast cancer cases and 411 controls frequency matched by age were included in this analysis. Exposure was assessed as distance (in miles) of residence at birth to the 13 industrial sites. The closest site was then selected for each subject as a surrogate for environmental exposure. The distance to the closest site was categorized into quartiles based on the distribution in the controls. Odds ratios (OR) and 95% confidence intervals (95% CI) were used to estimate the association between residential proximity and subsequent BC. The ORs were adjusted for age, education, age at menarche, parity, and age at first birth.

Results: We observed an adjusted OR of 3.8 (95% CI 1.9-7.7) for premenopausal women residing less than 2.45 miles from the closest site when compared to women residing greater than 8 miles from the closest industrial site. No such associations were observed in post menopausal women.

Conclusion: These preliminary findings suggest that relatively close residential proximity to industrial sites involved in uranium processing may increase the risk of premenopausal BC. However, it is unclear whether this association can be attributed to the environmental contamination with radioactive material, or some other environmental contaminate also produced at these industrial sites.

RELATED GRANT APPLICATION:

Source: Postdoctoral Traineeship Awards of the Department of Defense Breast Cancer Research Program.

Title of Proposal:

“Integrating Geographic Information System (GIS) into Breast Cancer Epidemiologic Research” submitted by Daikwon Han, graduate student, Department of Geography, University at Buffalo.

CONCLUSIONS

There are no major conclusions to report at this time. Preliminary findings are discussed in the text of this report.

REFERENCES

Bailey, TC and Gatrell, AC, Interactive Spatial Data Analysis, Harlow: Longman, 1995.

Diggle, PJ and Chetwynd, AD, Second-order Analysis of Spatial Clustering in Homogeneous Populations, *Biometrics* 47:1155-1163, 1991.

Ebdon, D, Statistics in Geography, Oxford: Blackwell, 1988.

Ripley, BD, Modeling Spatial Patterns, *Journal of the Royal Statistical Society Series B* 39: 172-212, 1977.

Rowlingson, B and Diggle, PJ, SPLANCS: Spatial Point Pattern Analysis Code in S-Plus, *Computers and Geosciences* 19: 627-655, 1993.