AD________________

GRANT NUMBER DAMD17-97-1-7095

TITLE: The Elevated Breast Cancer Mortality in the Northeastern U.S. is Secondary to Poorer Survival Rather than Increased Incidence

PRINCIPAL INVESTIGATOR: James S. Goodwin, M.D.

CONTRACTING ORGANIZATION: University of Texas at Galveston
Galveston, Texas 77555

REPORT DATE: August 1998

TYPE OF REPORT: Annual

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

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 1998	Annual (1 Aug 97 - 31 Jul 98)

4. TITLE AND SUBTITLE

The Elevated Breast Cancer Mortality in the Northeastern U.S. is Secondary to Poorer Survival Rather than Increased Incidence

5. FUNDING NUMBERS

DAMD17-97-1-7095

6. AUTHOR(S)

Goodwin, James S., M.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Texas University at Galveston
Galveston, Texas 77555

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

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The study examines the relative contributions of incidence, treatment and survival to the elevated mortality from breast cancer in older women in the northeast region of the U.S. Algorithms to determine incident breast cancer, treatment, stage and date of diagnosis are being evaluated and refined using the SEER-Medicare linked data base. Applying these algorithms to the Medicare data in 1991, we will identify a cohort of all women aged 70 years and older newly diagnosed with breast cancer. Incidence rates for 1991 as well as one, three and five year survival for these women will then be calculated for each state and region over the 1991-1996 interval.

14. SUBJECT TERMS

Breast Cancer Geographic Variation, Aging, Survival, Incidence

15. NUMBER OF PAGES

10

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

✓ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

9/16/98

PI - Signature Date

(4) TABLE OF CONTENTS

(5) INTRODUCTION

Geographic differences in mortality from breast cancer

The existence of geographic variation in breast cancer mortality rates has been known for some time (1-4). Blot et al (2) reported that breast cancer mortality at the county level was approximately 25% higher in the northeast than in the south, central or west. This difference remained after controlling (by county) for mean income, ethnicity, urbanization, birth rate, and mortality from ovarian or colonic cancer.

Sturgeon and her associates at the NCI have recently published an analysis of the geographic differences in breast cancer mortality, taking into account differences in geographic distribution of risk factors for breast cancer (age, age at menarche, age at menopause, age at first live birth, first degree relative with breast cancer, prior benign breast disease, body mass index, alcohol intake, years of education, and menopausal estrogen use) (4). A 13% elevation in northeastern breast cancer mortality rate relative to the south remains after adjusting for risk factor differences.

One striking finding of these studies is that the regional variation in breast cancer mortality rates is entirely confined to older women (2,4). A number of hypotheses have been proposed to explain these geographic differences (2,4-7). These explanations all share the assumption that elevated mortality reflects elevated incidence. The concentration of the regional mortality differences in older women makes it less likely that they are secondary to regional differences in incidence, because one might expect that increase mortality secondary to increased incidence would show up in all age groups. More importantly, our preliminary data gives no evidence for increased breast cancer incidence in the northeast (see below).

Measuring geographic differences in breast cancer incidence

The NCI's Surveillance, Epidemiology and End Results (SEER) program collects cancer incidence, treatment and mortality data from nine population-based registries. However, the SEER data are not ideally suited for use in addressing questions involving geographic differences such as posed in this proposal. For example, only one SEER site Connecticut, is in the northeast, and only one, the Atlanta metropolitan area, is in the south.

There has been considerable recent interest in the use of Medicare billing data to obtain population-based cancer incidence for men and women $\geq$ 65 years of age for the entire country (8-10). Fisher et al (11) compared diagnoses obtained from Medicare Part A claims data for 1985 to those gathered in the National DRG Validation Study (12). Compared to the data from the reabstracted records, the sensitivity for the diagnosis of breast cancer based on Part A Medicare claims diagnoses was 0.97 with a specificity > 0.99. Breast cancer had the highest level of agreement of the 32 diagnoses examined. Whittle et al (9) compared incidence rates generated with algorithms applied to 1983-85 Medicare Part A claims data for various cancers to those obtained from SEER for 1981-85. There were no significant differences in overall or age-specific incidence between Medicare and SEER for breast cancer. McBean and his colleagues at HCFA compared incidence data generated from Medicare Part A claims files and SEER for 1986-87, with the comparison restricted to the five SEER states (8). The incidence rates for breast cancer using the two sources of data were within 1% of each other.

With the exception of the analysis of Fisher et al (11), the other validation studies were ecological, comparing rates generated from Medicare to those generated by SEER. Obviously, the existence of the SEER-Medicare linked data base now allows for validation of algorithms to identify incident breast cancer from the Medicare claims data at the level of the individual. This permits innovative studies of patterns of individual cancer care and outcomes for the entire elderly population.

(6) BODY OF PROPOSAL

Hypothesis

We hypothesize that survival after the diagnosis of breast cancer in older women will vary by state and that women in the northeast region will have poorer survival compared to women in the south and west.

Proposed Technical Outcomes

1. To develop valid algorithms for utilizing Medicare Part A and B claims to determine incident breast cancer.
2. To determine breast cancer incidence by state for women age 70 and older identified in 1991 Medicare claims data.
3. To determine survival by state for breast cancer in women age 70 and older diagnosed in 1991 and followed for at least five years.
4. To determine the relative contributions of incidence and survival to the elevated breast cancer mortality in the northeast.

Methods

Study Cohort. Eligible women include all female Medicare beneficiaries age 70 and over in 1991 who were enrolled in Parts A and B over the study's period of medical use, 1986-1992. Excluded are members of HMOs (approximately 5% of beneficiaries in 1991) and enrollees for whom Medicare is a secondary payor, whose claims may not be included in the Medicare files (13). Also excluded are women residing outside the 50 states and District of Columbia. Our specific cohort of breast cancer subjects consists of women identified with incident breast cancer in 1991. For the analyses described in this proposal, each subject will be assigned to a state based on her residence at the time of diagnosis. States are further aggregated into four regions and nine divisions based on U.S. Bureau of the Census definitions.

Data

Medicare Files. The study uses data on Medicare services and expenditures that are collected as part of the Medicare Claims Data system. Enrollment information is obtained through annual denominator files, which contain one record per person enrolled over the year.

Diagnosis on the hospital inpatient, hospital outpatient and physician claims are coded in the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM). Procedures in the physician and hospital outpatient claims are coded in the HCFA Common Procedure Coding System (HCPCS), which includes Common Procedure Terminology (CPT) codes and other codes assigned by the Health Care financing

Administration and local carriers. Procedures on the hospital records are coded in ICD-9-CM.

Subjects with breast cancer will be identified in 1991 from the physician, hospital inpatient, and hospital outpatient claims. Data on inpatient hospitalizations from 1986-1991 will be used to exclude prevalent cases. Patient follow-up will be performed with the enrollment records for 1991 through 1997. An extra year is searched for late reported death dates falling within the period of follow-up (5 years).

In order to construct our cohort of breast cancer cases, we will obtain from HCFA all claims for eligible women with a breast cancer diagnosis code (ICD-9-CM codes 174.XX) or procedure code (85.XX) on any one claim during 1991 from the Part A inpatient or Part B outpatient files. A master analytic file for the study is developed that contains one record per claim (hospitalization, billed physician service/procedure, hospital outpatient facility service/procedure) for all beneficiaries in the cohort. The claims records is rewritten to a common format that includes type of claim, health insurance claim number, diagnoses, procedures/services, specialty and ate(s) of service.

SEER-Medicare Linked Data Base. Data from the Surveillance, Epidemiology and End Results (SEER) program were originally merged with the Medicare data from 1973 through 1989 (14). SEER registries collect information on incidence, extent of disease, initial treatment, and survival in addition to demographic characteristics such as age, race and sex. About 94 percent of the SEER Registry cases diagnosed at age 65 between 1973 and 1989 were linked with the claims. Our preliminary analyses with the linked data found that the linkage rate for the breast cancer bases is the same. The linkage process has now been expanded to include 1990-1993. The SEER-Medicare data base is being used to evaluate and refine our proposed algorithms. Cases from 1990 are used in the initial assessment and refinement. Revised algorithms are evaluated using the 1991 data.

Results

Geographic variation in breast cancer mortality rates is most pronounced in older women. As noted above, geographic variation in breast cancer mortality rates has been known for some time (1-4). These regional differences have been relatively sable over two or three decades. Table 1 presents breast cancer mortality rates for white females in 1990, by age category and geographic region that we derived from vital Statistics and U.S. Bureau of Census data (15-17).

Table 1. Breast Cancer Mortality Rates per 100,000 Females by Region and Age (with Ratio to South Atlantic Region): White Females, 1990.

Region	25-54	55-65	65-74	75+
East North Central	19.03 (1.02)	81.68 (1.09)	117.74 (1.12)	172.93 (1.16)
East South Central	19.13 (1.03)	72.95 (0.97)	91.5 (0.87)	129.68 (0.87)
Mid Atlantic	19.75 (1.06)	93.82 (1.25)	134.69 (1.218)	179.39 (1.21)
Mountain	16.34 (0.88)	76.27 (1.02)	109.36 (1.04)	158.45 (1.06)
New England	17.38 (0.93)	84.57 (1.13)	121.84 (1.16)	183.63 (1.23)
Pacific	19.08 (1.02)	78.56 (1.05)	115.32 (1.10)	160.97 (1.08)
South Atlantic	18.65 (1.00)	74.86 (1.00)	104.83 (1.00)	148.83 (1.00)
West North Central	17.71 (0.95)	72.26 (0.97)	114.37 (1.09)	158.46 (1.06)
West South Central	18.78 (1.01)	74.28 (0.99)	100.83 (0.96)	137.75 (0.93)

The numbers in the parentheses in Table 1 express the mortality rate as a ratio compared to the South Atlantic rate. The rates for all the southern regions cluster around the South Atlantic rate, for each age category. The rates for the Mid-Atlantic and New England regions are not higher than the southern regions for women aged 25-54 but, starting with the 55-64 age groups and continuing with older women, there is an increasing disparity between the northeast and south. The rates in the other regions fall between those for the south and those for the northeast. A similar pattern is seen for black women, (not shown).

Variation in incidence of survival with and mortality from breast cancer among health service areas in the SEER sites. Health Service Areas (HSAs) are aggregations of counties and independent cities based on a cluster analysis of where Medicare patients obtained routine hospital care in 1988 (18). They are a good tool to examine variations in medical practice and outcomes across small areas (18).

We calculated breast cancer mortality, incidence and survival for each of the 73 HSAs in the 9 SEER areas. We present below the mean, median, 25th and 75th percentiles for breast cancer mortality, incidence and survival. The survival was expressed as five year cancer-specific death rate, calculated from the SEER data. Incidence was also from SEER. Mortality was calculated from 1990 Vital Statistics data. All data were for women aged 65-74 or 75+ for the years 1985-90 9so that there would be five years of follow-up).

Table 2. Incidence, Mortality and Cancer-specific Death Rate for Breast Cancer in 73 Health Service Areas Contained within the SEER Sites.

	65-74 Years			**75+ Years**		
	Incidence	Mortality	Five Year Cancer-specific Death Rate (%)	Incidence	Mortality	Five Year Cancer-specific Death Rate (%)
Mean	490	109	13.1%	531	139	14.7%
Median	510	109	12.6%	579	141	14.6%
75% Q3	563	127	9.7%	626	172	12.4%
25% Q1	432	86	15.5 %	463	112	18.2%

Incidence and mortality were expressed as per 100,000 adult women. Five year cancer-specific death rate was percent breast cancer deaths at five years.

There is variation among the HSAs in all three measurements: incidence, mortality and five year cancer-specific death rate. For breast cancer incidence, the interquartile range represents approximately 25-30% of the median in the two age categories; for mortality the interquartile range is approximately 40% of the median, while for 5 year cancer-specific death rate the interquartile range is approximately 50% of the median in the 65-74 year olds and 40% in the 75+ group. Thus, survival with breast cancer varies by geographic area as much as does incidence and mortality.

The variations are stable over time; for all three measurements. HSAs with high or low values for incidence, survival or cancer-specific death rate in 1985-87 also tended to have similar rates in 1988-90. The most compelling evidence that variation in survival contributes to variation in mortality rate is found in correlation between five year survival rate (calculated as 1 minus cancer-specific death rate) and mortality rate for breast cancer. In a partial correlation controlling for disease incidence, the coefficient of correlation between mortality rate and survival was $r = 0.37$ ($p = 0.001$) for women aged 65-74 and $r = 0.52$ ($p = 0.0001$) for women aged 75+.

It is important to remember that mortality rate and survival rate come from two entirely different sources of data. Mortality rate by HSA is derived from U.S. Vital Statistics data, while the survival rate was calculated for incident cases identified by SEER in 1985-90 and followed for at least five years. This provides strong support for the underlying hypothesis of this proposal, that geographic variations in survival from breast cancer contribute to the geographic variation in breast cancer mortality.

In our calculations of survival we used breast cancer specific survival (or 100% minus breast cancer-specific five year death rate). We used that figure rather than total survival because breast cancer mortality death rates in the Vital Statistics data measure only cases where breast cancer is listed as the underlying cause of death. In our future analyses, we will examine both breast cancer-specific and total death rates in the SEER-Medicare data. Also, in these preliminary analyses, we used a relatively crude measure of five year survival (or death rate), while in the future analyses we will estimate survival in HSAs by hazard ratios, allowing for multivariate analyses with the inclusion of other variables.

(7) CONCLUSIONS

This report is preliminary, relating to the work performed on this project over the initial 12 months of funding. What we have demonstrated thus far is that breast cancer mortality and also breast cancer survival clearly do vary by geographic area. Our goal for the remainder of this project is to test the hypothesis that decreased survival from breast cancer in the northeast U.S. contributes to the increased breast cancer mortality.

(8) REFERENCES

1. Deaths from breast cancer - United States, 1991. MMWR 43: 273-281, 1994.
2. Blot WJ, Fraumeni JF, Stone BJ. Geographic patterns of breast cancer in the United States. J Natl Cancer Instit 59: 1407-1411, 1977.
3. Ries LAG, Hankey BF, Miller BA, et al. Cancer statistics review, 1973-88. Bethesda, MD: National Cancer Institute, 1991. (NIH publication no 91-2789).
4. Sturgeon SR, Sahairer C, Mitchell G, Mcadams M, Brinton L, Hoover R. Geographic variation in mortality from breast cancer among white women in the United States. J Natl Cancer Inst 87: 1846-53, 1995.
5. The Northeast/Mid-Atlantic Study: Environmental factors and breast cancer in high-risk areas. Cancer Facts, Cancer Net, Natl Cancer Institut, NIH, May, 1994.
6. Bulbrook RD. Geographical variation in endorring function and its relation to breast cancer incidence: some special considerations. Breast Cancer Res Treat 18: 537-540, 1993.
7. Garland FC, Garland CF, Gorham ED, Young JF. Geographic variation in breast cancer mortality in the United States: a hypothesis involving solar radiation. Preventive Med 19: 614-622, 1990.

8. McBean AM, Warren JL, Babish JD. Measuring the incidence of cancer in elderly Americans using Medicare claims data. Cancer 73: 2417-25, 1994.
9. Whittle J, Steinburg EP, Anderson EF, Herbert R. Accuracy of Medicare claims data for estimation of cancer incidence and resection rates among elderly Americans. Medical Care 29: 1226-36, 1991.
10. McBean AM, Babish JD, Warren JL. Determination of lung cancer incidence in the elderly using Medicare claims data. Am J Epidemiol 137: 226-34, 1993.
11. Fisher ES, Whaley FS, Krushat WM, Malenka DJ, Fleming C, Baron JA, et al. The accuracy of Medicare's hospital claims data: progress has been made, but problems remain. Am J Public Health 82: 243-8, 1992.
12. Hsia DC, Krushat M, Fagen AB, Tebbett JA, Kusserow RP. Accuracy of diagnostic coding for Medicare patients under the prospective payment system. N Engl J Med 318: 352-55, 1988.
13. General Accounting Office Medicare PRO review does not assure quality fo care provided by risk HMOs (HRD-91-48). Washington.
14. Potosky AL, Riley GF, Lubitz JD, Mentnech RM, Kessler LG. Potential for cancer related health services research using a linked Medicare-tumor registry data base. Medical Care 31: 732-748, 1993.
15. National Center for Health Statistics. Vital Statistics of the United States, 1990, Vol. 11, Mortality, Part B. Washington Public Health Services, 1994.
16. U.S. Bureau of the Census, 1990 Census of Population, General Population Characteristic, United States (CP-1-1).
17. Goodwin JS, Freeman JL, Freeman D, Nattinger AB. Geographic variations in breast cancer mortality: do higher rates imply elevated incidence or poorer survival? Am J Public Health 88: 458-460, 1998.
18. Makuc DM, Haglund B, Ingram DD, et al. Health Service Areas for the United States. National Center for Health Statistics. Vital Health Stat 2:112, 1991.