Geographic Variations in Breast Cancer Survival Among Older Women: Implications for Quality of Breast Cancer Care

James S. Goodwin,1,2,3 Jean L. Freeman,1,2,3 Jonathan D. Mahnken,2,3 Daniel H. Freeman,2,3 and Ann B. Nattinger4

Departments of 1Internal Medicine and 2Preventive Medicine and Community Health, and 3Sealy Center on Aging, The University of Texas Medical Branch, Galveston.
4Department of Internal Medicine, Medical College of Wisconsin, Milwaukee.

Background. Breast cancer care, such as utilization of screening procedures and types of treatment received, varies substantially by geographic region of the United States. However, little is known about variations in survival with breast cancer.

Methods. We examined breast cancer incidence, survival, and mortality in the 66 health service areas covered by the Surveillance, Epidemiology, and End Results (SEER) program for women aged 65 and older at diagnosis. Incidence and survival data were derived from SEER, while breast cancer mortality data were from Vital Statistics data.

Results. There was considerable variation in breast cancer survival among the 66 health service areas ($\chi^2 = 202.7$, $p < .0001$). There was also significant variation in incidence and mortality from breast cancer. In a partial correlation weighted for the size of the health service area, both incidence ($r = .812$) and percent 5-year survival ($r = -.587$) correlate with mortality. In a Poisson regression analysis, the combination of variation in incidence and variation in survival explains 90.9% of the variation in mortality.

Conclusions. There is considerable geographic variation in survival from breast cancer among older women, and this contributes to variation in breast cancer mortality. Geographic variations in breast cancer mortality should diminish as the quality of breast cancer care becomes more standardized.

Variations in the rate of diseases across different populations and geographic areas provide powerful clues in studies of disease etiology. For example, Doll and Peto relied primarily on such information to conclude that the great majority of cancers were environmentally determined, for example, tobacco, diet, pollutants, and others (1). Much of the information on geographic variation in cancer rates comes from Vital Statistics mortality data, which are used as a proxy indicator of disease incidence. However, mortality rate from any disease is a function of both the incidence rate and the survival rate of persons with the disease. Diseases with high mortality and short survival exhibit a very close relationship between incidence rate and mortality rate (2–4). In contrast, in diseases with a substantial chance for cure, factors that affect survival, such as early diagnosis and treatment, should have an important effect on mortality rates.

There are well-described, stable geographic variations in mortality rates from breast cancer in the United States. These exist at a regional level (e.g., Northwest vs South) and also at the level of small areas (e.g., census tracts within a state) (5–10). The geographic differences have stimulated considerable research about their cause, such as regional differences in diet or differences in proximity to high-voltage electric lines. One factor common to these investigations is the assumption that variation in mortality from breast cancer reflects variation in incidence of breast cancer. We have previously argued that regional variations in breast cancer mortality may reflect differences in survival as well as differences in incidence (6). In this report, we test that hypothesis by examining the variations in incidence, survival, and mortality from breast cancer among different health service areas (HSAs). HSAs are aggregations of counties based on a cluster analysis of where Medicare patients obtained routine hospital care in 1988 (11). They are a good tool to examine variations in medical practice and outcomes across small areas (12). We assess the relation of any variation in incidence and survival to the variation in breast cancer mortality. We use data from the Surveillance, Epidemiology, and End Results (SEER) program to assess incidence and survival, and Vital Statistics data to determine breast cancer mortality rates.

Methods

Subjects
Eligible women for our study include those who were age 65 and older in 1985–1991 residing in one of the nine SEER areas. Our specific cohort of breast cancer subjects consists of women identified with incident breast cancer in 1985–1991.
Sources of Data

SEER tumor registry.—The National Cancer Institute’s SEER program supports population-based tumor registries in selected geographic areas. In 1985, these areas included the metropolitan areas of San Francisco/Oakland, Detroit, Atlanta, and Seattle, and the states of Connecticut, Iowa, New Mexico, Utah, and Hawaii (13). These areas cover approximately 14% of the U.S. population. SEER’s own quality assurance checks report a high degree of completeness of determining incident cancers. SEER incidence rates are used to project to the total U.S. population to estimate the national incidence rates for various cancers.

The SEER Public Use File is the source of data for estimating incidence and survival rates. For all incident breast cancer cases, the file includes month and year of diagnosis, county of residence, and period of follow-up. If the subject has died, the file also contains the date of death and cause of death as coded on the state death certificate with ICD-9 codes.

National Center for Health Statistics mortality data files.—Mortality data from the National Center for Health Statistics (1985–1991) were used to construct breast cancer mortality rates at the county level. Each record in these files represents one death and includes data elements on the decedent’s age, gender, county of residence, and underlying cause of death. The source of these data is information recorded on death certificates filed for deaths in each of the 50 states. In our analyses of the mortality data, the death is allocated to the county of the decedent’s residence, not where death occurred.

Health Service Area

Our study uses the HSA definitions that are used in the Atlas of United States Mortality to illustrate geographic differences in death rates for selected causes of death (12). These HSAs are based on those identified by the National Center for Health Statistics to define one or more counties that are “relatively self-contained with respect to the provision of routine hospital care” (11). HSAs were formed by applying a clustering algorithm on travel patterns of patients for their hospital care. The objective was to form areas where residents in these areas were more likely to use routine hospital services in the area than outside it. The SEER Public Use File contains population, incidence, and survival information at the county level. These counties were aggregated into HSAs using a file that maps U.S. counties by Federal Information Processing Standards codes (a two-digit state number followed by a three-digit county number) to its corresponding HSA. This file is included in the Atlas of United States Mortality CD-ROM.
The SEER areas analyzed in this study contained 72 of the 805 HSAs. Smaller HSAs were aggregated within SEER area to yield clusters with a minimum of five expected breast cancer deaths in 1985–1991, forming a total of 66 HSAs. The aggregated HSAs were 704 and 814 (New Mexico); 711, 755, and 799 (Utah); and 544, 552, 626, and 665 (Iowa).

Incidence, Mortality, and Survival Rates
Incidence rates by HSA were calculated for each age group using the number of incident cases and population estimates provided by the SEER program. Likewise, age-specific breast cancer mortality rates were generated from estimates of the number of breast cancer deaths using the mortality data files and the SEER population estimates. Information on the date of death, underlying cause of death, and period of follow-up in the SEER database allows us to generate 5-year Kaplan-Meier survival rates for each HSA. Events were comprised of breast cancer-specific deaths (ICD-9 174xx). Deaths from all other causes were censored. These 5-year survival rates were generated for each year from 1985–1991 and then averaged over these 7 years for analysis.

Analysis
Breast cancer survival among the HSAs was initially examined with Kaplan-Meier survival curves, and their homogeneity was assessed with the log rank test. Geographic variations in incidence and mortality were modeled using Poisson regression with HSA as an independent variable. The relationships between incidence and mortality and between survival and mortality were examined with bubble plots, Pearson correlation coefficients, partial correlation coefficients, and Poisson regression models. The correlation estimates were based on weighted least squares because the HSAs (the unit of observation) differ greatly in size. Treating each HSA with equal weight resulted in correlation coefficients being driven by HSAs with a small number of breast cancer cases. The statistical tests of significance and associated p values were derived from Poisson regression because the assumptions are more suitable for estimating vital rates. All computations and data management were done using the SAS version 6.12 statistical application (SAS Institute, Cary, NC) (14).

RESULTS
Figure 1 shows the distribution of incidence, mortality, and percent 5-year survival from breast cancer among the 66 HSAs in the nine SEER areas. There is considerable variation in all three measurements. There are a number of ways to demonstrate that this variation is not simply random. For example, when the incidence and mortality rates were modeled using Poisson regression, HSA was a significant predictor (p < .0001) for both models. Kaplan-Meier survival curves indicated significantly different survival times across the HSAs (χ² = 202.7, p < .0001).

Figure 2 shows the relationship between incidence of breast cancer and breast cancer mortality for each of the 66 HSAs. In this figure, the size of the bubble is proportional to the total number of women age 65 or older in each HSA. As expected, there is a clear correlation between incidence and mortality, with a weighted correlation coefficient $r = .733$ ($p < .0001$).

Figure 3 shows the relationship between breast cancer survival and breast cancer mortality for each of the 66 HSAs.
HSAs. They are significantly correlated with a weighted correlation coefficient \( r = -0.335 \) (\( p < 0.0001 \)).

Table 1 presents weighted partial correlations of both incidence and survival with breast cancer mortality. Both incidence and survival (adjusted for each other) are correlated with mortality. Because the over-65 age group is heterogeneous, we also repeated the weighted Pearson and partial correlation coefficients after stratifying the population into those 65 to 74 and those 75 and older. As shown in Table 1, incidence and survival are highly correlated with mortality in both age groups.

The weighted least square analysis employed in Table 1 did not allow for a test of statistical significance of the variation in breast cancer mortality attributable to the two variables, incidence and survival. To further investigate this relationship, we estimated Poisson regression models for the combined group and the two age strata (Table 2). The \( R^2 \) (generalized coefficient of determination) (15) for the model including both incidence and survival is .909, indicating that the combination of incidence and survival from breast cancer explains approximately 91% of the total variation in breast cancer mortality among the 66 HSAs. These associations are all statistically significant.

**DISCUSSION**

The results of this study can be summarized as follows. There is considerable variation among the 66 health service areas within SEER tumor registries in breast cancer mortality rates, as assessed by Vital Statistics data, and in incidence

### Table 1. Simple and Partial Correlations of Breast Cancer Incidence and Breast Cancer Survival to Breast Cancer Mortality in 66 HSAs*

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Variable</th>
<th>Simple Correlation</th>
<th>Partial Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>65+</td>
<td>Incidence</td>
<td>.733</td>
<td>.812</td>
</tr>
<tr>
<td></td>
<td>Survival</td>
<td>-.335</td>
<td>-.587</td>
</tr>
<tr>
<td>65–74</td>
<td>Incidence</td>
<td>.483</td>
<td>.562</td>
</tr>
<tr>
<td></td>
<td>Survival</td>
<td>-.251</td>
<td>-.405</td>
</tr>
<tr>
<td>75+</td>
<td>Incidence</td>
<td>.662</td>
<td>.672</td>
</tr>
<tr>
<td></td>
<td>Survival</td>
<td>-.413</td>
<td>-.435</td>
</tr>
</tbody>
</table>

*The simple and partial correlation coefficients are derived from weighted least squares regression. A positive simple correlation coefficient between incidence and mortality indicates that as the incidence increases, so does the mortality, with more populous HSAs having greater influence on the coefficient of correlation than smaller ones. Conversely, the negative coefficient values between survival and mortality indicate an increase in survival corresponded with a decrease in mortality. The weighted partial correlation coefficients have a similar interpretation, except that the relationship with mortality adjusts for the effects of the other variable (either incidence or survival). Here again, the larger HSAs have a greater impact on the simple and partial correlation coefficients.

### Table 2. Effect of Breast Cancer Incidence and Survival on Mortality From Breast Cancer in 66 HSAs, From Poisson Regression Model

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Coefficient of Determination* Variable</th>
<th>Estimated Coefficients</th>
<th>Standard Error</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>65+</td>
<td>Incidence</td>
<td>234.8375</td>
<td>20.8363</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>Survival</td>
<td>-1.9747</td>
<td>0.3250</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>65–74</td>
<td>Incidence</td>
<td>178.9161</td>
<td>28.9193</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>Survival</td>
<td>-1.7058</td>
<td>0.4064</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>75+</td>
<td>Incidence</td>
<td>187.8247</td>
<td>24.9422</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td></td>
<td>Survival</td>
<td>-1.2399</td>
<td>0.3107</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

of and survival from breast cancer, as assessed by SEER data. In a partial correlation, both variation in 5-year breast cancer-specific survival as well as variation in breast cancer incidence are highly correlated with breast cancer mortality among the 66 health service areas. Together, variation in incidence and survival explained >90% of the variation in breast cancer mortality.

Why might there be variation in survival after a diagnosis of breast cancer among different HSAs? Survival is affected by tumor biology, stage at diagnosis, and choice of treatment. Tumor characteristics, such as histologic grade, vary somewhat among different races. Thus, differences in survival among HSAs could be explained in part by the relative racial composition of the HSAs. Use of screening mammography, a major determinant of early diagnosis, has been shown to vary by region (8,16). In addition, several investigators have shown large geographic variations in the quality of treatment received for breast cancer, particularly among older women (17–21). Substantial proportions of older women with breast cancer are receiving breast-conserving surgery without axillary dissection or adjuvant irradiation (17–27). Such treatment is associated with a two-fold higher mortality rate in population-based studies, and it is not recommended by any authority or expert panel (23,25).

Perhaps the best-known example of differences in survival rates driving differences in mortality rates in breast cancer is in black-white differences; age-adjusted mortality rates from breast cancer are approximately 20% higher among African American women compared with non-Hispanic whites, while the incidence rates for blacks are approximately 12% lower (2,28). Had only the mortality rates been examined, it might have been erroneously concluded that breast cancer incidence is greater in black than in white women, but in reality the higher mortality rates are entirely secondary to the poorer survival experience of African American women with breast cancer (29).

There are a number of limitations to this study. First, the analyses were all ecological, at the level of the HSA and not at the level of the individual. Second, the associations among breast cancer incidence or survival on the one hand and breast cancer mortality on the other do not prove a causal relationship. Third, the findings were obtained in data from women aged 65 and older and may not reflect the pattern found in younger women.

Most studies of geographic differences in breast cancer mortality have assumed that such differences reflect differences in the incidence of breast cancer in different areas (5,7,9,10). These findings have in turn stimulated considerable research into the possible causes of such differences in breast cancer incidence (9,30–32). Our results confirm that mortality differences do indeed reflect differences in incidence, but they also show that an additional contributor to differences in mortality from breast cancer are geographic differences in survival. If the large geographic differences in breast cancer mortality are considered entirely secondary to differences in incidence of breast cancer, then there are no obvious steps to reduce those differences; that is, there are no generally accepted measures to reduce breast cancer incidence. On the other hand, the recognition that survival plays a role in geographic differences in breast cancer mortality would suggest that cancer control efforts directed to standardize the quality of breast cancer care would also diminish the geographic variations in breast cancer mortality.

Acknowledgments

This work was supported by Grants CA-871773 and CA-08554 from the National Cancer Institute. The data on breast cancer mortality utilized in this publication were made available in part by the Inter-university Consortium for Political and Social Research. The data for MORTALITY DETAIL PILES 1985 (VOLUME V) were originally collected by the National Center for Health Statistics of the U.S. Department of Health and Human Services. Neither the collector of the original data nor the Consortium bears any responsibility for the analyses or interpretations presented here.

Address correspondence to James S. Goodwin, MD, Sealy Center on Aging, The University of Texas Medical Branch, 301 University Blvd., Galveston, TX 77555-0460. E-mail: jsgoodwi@utmb.edu

References


Received September 28, 2001
Accepted January 24, 2002