Racial Differences in the Intensity of Breast Cancer Treatment

Emma Bojinova*

ABSTRACT

Breast cancer has the highest mortality rate among females diagnosed with cancer in the United States. Racial discrimination along with the stigma of inferiority of the minority groups can adversely affect health. The goal of this paper is to investigate if there are racial differences in the intensity of treatment of breast cancer patients. Based on cross-sectional data from the Healthcare Cost and Utilization Project, different specifications are used to estimate if African-Americans, Hispanics, and other racial groups are treated less intensively in comparison with white Americans when they are admitted to hospitals in 35 states.

INTRODUCTION

Cancer is the second leading cause of death for Americans and breast cancer has the highest mortality rate among females diagnosed with cancer. Racial discrimination, both individual and institutional, along with the stigma of inferiority of the minority groups, can adversely affect health. Hence, it is important to understand whether there is racial discrimination in the intensity of treatment, which can explain the higher mortality rate for African-Americans, or if there are other factors that account for this disparity. Thus, the purpose of this paper is to investigate whether there are racial differences in the intensity of treatment for breast cancer, which is proxied by hospital expenditures recorded for an inpatient claim.

Based on cross-sectional data from the Healthcare Cost and Utilization Project (HCUP) for year 2002, different specifications are used to estimate if black and other racial groups such as Hispanics, Asians, Native Americans, and others are treated less intensively in comparison with white Americans when they are admitted to hospitals in 35 states. The results suggest that there is some evidence for the presence of racial discrimination in this particular year.

BACKGROUND AND PREVIOUS LITERATURE

According to the Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute the age-adjusted breast cancer incidence rate for the period 2002-2006 was 123.8 per 100,000 women per year. The breakdown by race is as follows: the incidence rate is highest for white females (127.8 per 100,000 women), followed by black females (117.7 per 100,000), Asian/Pacific Islanders (89.5 per 100,000), and Hispanics (88.3 per 100,000 women). The lowest incidence rate is recorded for American Indian/Alaska Native females (74.4 per 100,000). In contrast, black females have

* Department of Economics and Finance, Canisius College, 2001 Main St., Buffalo, NY 14208.
Phone: (716) 888-2721. E-mail: bojinove@canisius.edu

the highest age-adjusted mortality rate (33 per 100,000) as compared to other races for the same period. Breast cancer mortality rate is approximately 38.1% lower for white women (23.9 per 100,000) and between 87.5% and 264% lower for the other 3 racial groups mentioned above.

The early detection of breast cancer is very important for a patient's survival. Mammography is especially useful for identifying breast cancer at an early stage even before physical symptoms develop. Many studies have shown that early detection increases treatment options and decreases mortality. However, the decision to undergo breast cancer screening depends on whether the person has health insurance, as well as on the educational level and awareness of breast cancer symptoms. For instance, women who lack health insurance, who are poor, less educated, or live without a husband tend to have the lowest prevalence of mammography use due to their limited access to health care. According to the National Center for Health statistics, African-American, Hispanic, and American Indian women are more likely to be diagnosed with breast cancer at a later stage of disease development, which will affect their hospital expenditures.

Furthermore, these expenditures depend not only on the clinical status of patients, but also on the duration of stay, reason for admission, and whether this is first admission or one of the followed re-admissions to a hospital. For example, a woman in an advanced stage of breast cancer during her first admission to a hospital will undergo more diagnostic and therapeutic procedures and as a result her expenditures will be high as compared to say her third re-admission when she can have chemotherapy or radiation therapy only. A number of studies have shown that there is a direct relationship between length of stay and hospital charges but as the length of stay decreases expenditures decrease less than proportionately because the latter is associated with higher intensity of treatment during early stay in hospital. Also, for cancer patients it was found that the "cost of treatment may decrease with severity because of the futility of any further active intervention, while at the same time mortality rate goes up for each stage and substage" (Medstat Disease Staging Software Reference Guide, HCUP 2002).

A lot of research has been done in explaining the differences in cancer survival rates among different socioeconomic groups. Some of the studies considered several cancer sites, whereas others concentrated on a single cancer site. A study by Kravdal (2000) found that the excess mortality was about 15 percent lower for patients who had a post-secondary education as compared to those with compulsory schooling after controlling for age, stage at the time of diagnosis, and registered differences in tumor characteristics. However, Kravdal did not find clear indication that host factors such as comorbidities and immune functions, as well as treatment and care differences matter for the differential in survival rates. Majority of breast cancer survival studies found some evidence suggesting that socially advantaged have better survival rates after controlling for possibly earlier detection of the disease among people corresponding to higher social classes (see for example LeMarchand et al., 1984; Bassett and Krieger, 1986; Karjalainen and Pukkala, 1990; Gordon et al., 1992; Ansell et al., 1993; Schrijvers et al., 1995). Figueroa and Breen (1995) analyzed cases of breast and cervical cancer diagnosed in the period 1989-1990 in San Francisco, Detroit, and Atlanta. They found that 87% of the breast cancer cases were

diagnosed late, when already tumor was malignant. They found that a significant part of the variation in diagnostic stage was explained by the residence in underclass area. The likelihood of late-stage diagnosis also increased with age and was higher for females living without a spouse. Katz and Hoffer (1994) found similar results for breast cancer patients living in Ontario, Canada.

Several studies found evidence that health insurance matters with regards to breast cancer screening, surgical procedures or other treatment procedures. Mitchell and Hadley (1997) considered nonelderly women diagnosed with breast cancer using hospital inpatient discharges for 1988 and 1991 in five states (CA, MD, MA, NJ, and NY). The authors found that the probability of breast-conserving surgery is 2.7% lower for females enrolled in HMOs, 4.8% lower for Medicaid and 6.6% lower for self-pay patients as compared to females having private insurance plans. A study by Thorpe and Howard (2003) found substantial differences in cancer spending by insurance status based on the Medical Expenditure Panel Survey for 1996-1999. They considered 5 big cancer types, among which is breast cancer. Their results showed that nonelderly cancer patients without health insurance have higher risk of being inadequately treated especially if they are of Hispanic origin.

In addition, there are a number of studies based on HCUP data, some of them were concentrated on the relationship between hospital volume and survival, and others looked at racial discrimination. Bach et al. (2001) studied patients 65 years old or older diagnosed with lung cancer between 1985 and 1996 and had surgery at a hospital that participated in the Nationwide Inpatient Sample (76 hospitals). Their results suggest that “patients who undergo resection for lung cancer at hospitals that perform large numbers of such procedures are likely to survive longer than patients who have such surgery at hospitals with a low volume of lung-resection procedures.” Dimick et al. (2003) found similar results for patients undergoing surgery for colorectal cancer based on 1997 HCUP data using logistic regressions. Andrews and Elixhauser (2000) examined whether there is difference in the rate of receiving therapeutic procedures between Hispanic and white non-Hispanic patients based on 1993 discharge data for California, Florida and New York (states that account for half of the Hispanic population in the USA). Their findings provide evidence that Hispanics are undertreated in a sense that they are less likely to receive major therapeutic procedures for 38% of the 63 conditions they examined and more likely for 6% of the conditions as compared to non-Hispanic whites. There are a number of other HCUP studies looking at racial or sex disparities for patients diagnosed with various diseases including other types of cancer (Ball and Elixhauser, 1996; Harris et al, 1997; Andrews and Elixhauser, 2000; Shenn, 2002). However, according to my knowledge there are no studies looking at racial differences in the intensity of treatment for breast cancer patients based on HCUP data. Therefore, my paper can be considered as a contribution in this less researched area.

DATA DESCRIPTION

The source of data is the Healthcare Cost and Utilization Project (HCUP) for year 2002. The nationwide inpatient sample (NIS) consists of approximately 7.85 million hospital stays from about 1,000

hospitals in the United States. It covers 35 states and is designed to approximate a twenty-percent sample of the U.S. community hospitals, which allows for making inferences for the country as a whole. The advantages of using HCUP data is the availability of large number of inpatient records, the good data on health insurance, hospital characteristics, and different disease diagnoses. It should be noted that the unit of observation in this data set is an inpatient claims record not the patient itself. Thus, there is a possibility that the same individual went to a hospital for treatment multiple times in a given year which can affect the estimates. The data is also censored because we do not observe the whole universe (population) - just the individuals that go to a hospital and are treated file a claim.

I concentrate my research on breast cancer inpatient stays and restrict the HCUP sample to discharges with principal diagnosis “breast cancer” (based on ICD-9-CM codes), which reduced the sample to 22,678 observations. Observations with missing values for the variables of interests, i.e. race and total charges, are deleted. There are 156 observations for breast cancer patients that are male, which are dropped from the sample to avoid potential unobserved gender differences with regards to treatment and disease development. Furthermore, Georgia does not report race due to confidentiality of reports. There is also missing race data for some of the other states. As a result, after deleting these observations the sample size decreases by 6,084 inpatient records.

The following variables from the NIS are used in the study: total charges (*totchg*), median household income category for patient's zip code (*zipinc*), length of stay (*los*), number of procedures on this record (*npr*), died during hospitalization (*died* – a dummy variable equal to one if died and zero otherwise), expected primary payer (*pay1*), age in years at admission (*age*), and whether the admission was elective (*elective*). The dependent variable is *lntotchg*, which is the natural logarithm of total charges for an inpatient stay. I use the logarithmic transformation to account for possible skewness of the expenditure distribution (to rule out big outliers in the sample). I create two dummy variables for race – black and other race (each variable is equal to one if the patient is black or other race respectively, and zero otherwise). The indicated category is white. I collapse Hispanics, Asians, Native Americans, and others into one dummy variable called *otherrace* because I am generally interested in the potential differential between African-Americans and non-Hispanic whites. I also create categorical variables for health insurance status and median household income for patient's zip code. The expected primary payer variables are Medicaid, Medicare, and private insurance (incl. HMOs and PPOs). The indicated category is other expected primary payer, which includes self-pay, charity, and the like. The median household income for patient’s zip code is not a continuous variable but instead income is reported in ranges or categories, i.e. from $1 to $24,999, from $25,000 to $34,999, from $35,000 to $44,999, and from $45,000 or more. Thus, the following dummy variables are created respectively: low income, below median income, and above median income. The indicated category is high income ($45,000 or more).

As mentioned in a previous section, hospital expenditures depend on disease severity. I control for this relationship by including variables such as disease staging and AHRQ comorbidity measures. I merge national inpatient sample with the HCUP severity data by a common code that is uniquely defined

in both datasets. The AHRQ comorbidity measures define thirty different coexisting medical conditions, which are likely to be present prior to the hospital stay and are not directly related to the reason for hospital admission or principal diagnosis. All these comorbidity measures are defined as categorical variables (equal to one if the patient has the disease and zero otherwise). The presence of comorbidities can increase the cost of treatment, so it is important to rule out their impact on total expenditures for breast cancer patients. Disease staging criteria, developed by Medstat, define the severity for different medical diseases. It is measured on a scale from 0 to 4 with stage one being a disease with no complications; stage 2 is a disease with local complication; stage 3 corresponds to an increased disease complexity – it involves multiple sites or has systemic complications; and stage 4 is death. The staging variable (*ds_stage*) is measured on a continuous scale (has substages) to better represent the severity of a particular disease.

The summary statistics by race are presented in Table 1. Whites seem to be a little older compared to breast cancer patients that are African-Americans (blacks) or belong to other race. The length of stay in a hospital is somewhat longer for African-Americans than for other racial groups on average (3.41 vs. 2.62 and 2.41). In addition, in this data set African-Americans have relatively higher hospital expenditures, lower rate for elective admission, higher mortality rate, more advanced stage of the disease at admission, and are poorer on average in comparison with whites and patients of other races.

Table 1: Summary Statistics by Race

	Black			***Other Race***			***White***		
Variable	**Obs**	**Mean**	**Std. Dev.**	**Obs**	**Mean**	**Std. Dev.**	**Obs**	**Mean**	**Std. Dev.**
age	1672	57.31	14.42	1634	56.75	14.15	12053	62.79	14.45
los	1672	3.41	4.73	1634	2.62	3.8	12053	2.41	2.77
npr	1672	1.85	1.22	1634	1.91	1.26	12053	2.03	1.24
totchg	1672	18162.04	22945.78	1634	17930.95	18109.24	12053	15409.9	14147.13
lntotchg	1672	9.47	0.78	1634	9.51	0.73	12053	9.39	0.69
died	1670	0.03	0.17	1633	0.02	0.14	12039	0.01	0.12
elective	1664	0.72	0.45	1628	0.79	0.41	12025	0.84	0.37
lowincome	1672	0.12	0.33	1634	0.07	0.26	12053	0.02	0.13
belowmedian	1672	0.28	0.45	1634	0.17	0.37	12053	0.18	0.38
abovemedian	1672	0.26	0.44	1634	0.21	0.41	12053	0.25	0.43
ds_stage	1580	1.32	0.71	1614	1.2	0.58	10936	1.16	0.51

Hospital expenditures may be correlated with hospital characteristics. Therefore, I control for this possible relationship using dummy variables for hospital ownership, location, region, and size. Such data is provided in the supplemental HCUP hospital data set, which I merge with the NIS and severity data set. It is important to mention that the perception of a patient about the disease and the social support she receives from her family and friends can influence the timing of hospitalization and length of stay, which indirectly affects hospital expenditures. However, they cannot be easily measured and are not available in the HCUP data.

ECONOMETRIC MODEL AND ESTIMATION TECHNIQUES

The incurred hospital expenditures serve as a proxy of the intensity of breast cancer treatment. Using expenditures, however, should be done with caution because there may be issues with co-insurance and health insurance reimbursement to doctors for government provided insurance such as Medicaid and Medicare. In addition, there is a possibility that a part of these expenditures may be due to defensive medicine. Therefore, I employ various specifications to estimate the possibility of racial discrimination in terms of intensity of treatment for breast cancer patients controlling for health insurance status. To avoid the possibility for highly skewed expenditures or having big outliers that change dramatically the mean, I use the logarithm of expenditures as the dependent variable and estimate several model specifications via ordinary least squares (OLS). Since there can be omitted variable bias that causes heteroskedasticity I employ the Huber-White correction to the OLS regressions. Thus, standard errors will be consistent and inferences can be made. The basic model can be specified as:

$$\text{lntotchg} = \alpha + \beta X + \gamma Race + \theta HI + \varepsilon$$

where *X* is a vector of inpatient claims' characteristics such as age in years at admission, length of stay in the hospital, number of procedures on record, dummy variables for patient's median income, whether the person died in the hospital and whether the admission was elective. The coefficients on the race dummy variables (*black* and *otherrace*) show the difference between the respective base category and the indicated category, white female patients with breast cancer, in terms of log of total expenditures. I control for differences based on the health insurance status using three dummy variables (*Medicaid, Medicare, and Private Insurance*).

In the next specifications, I include controls for hospital characteristics such as location (urban or rural), ownership/control (public, private for profit, and private not-for-profit), size approximated by the number of beds (small, medium, and large), and region (Northeast, Midwest, South, and West). Furthermore, to take into account that expenditures depend on the disease severity, I include dummy variables for disease staging and comorbidities. In addition, I expand the model by adding an interaction term between age and race to account for possible differences in hospital expenditures for women of different ethnicities at different ages. Thus, the expanded model is as follows:

$$\text{lntotchg} = \alpha + \beta X + \gamma Race + \theta HI + \varphi Hospital + \delta Severity + \lambda(XRace) + \varepsilon$$

I also estimate the expanded model via OLS for Medicaid and Medicare claims separately, which solves the problem with having various out-of-pocket expenditures and prices for given procedures provided to patients with different types of health insurance. Finally, I estimate quantile regression models introduced by Koenker and Bassett (1978) that look at the median or different percentiles of expenditures (25% or 75%). As a result, I can make inferences for a possible racial differential at more similar expenditure levels.

ESTIMATION RESULTS

The signs of the OLS estimated coefficients make sense except for the ones on race, private insurance, and elective admission in the specifications where these variables are statistically insignificant. As expected the coefficients on *npr* and *los* are positive and statistically significant at 1% level of significance showing that the log of hospital expenditures goes up as the number of procedures and length of stay increase. The coefficient on *age* is negative, which is consistent with the theory that older patients are treated less intensively on average (makes sense for elderly patients only). The results show that patients with low income, below and above median income have lower hospital expenditures as compared to those coming from high-income zip codes (all coefficients are significant at the 1% or 5% level in all specifications but the last two). The coefficients on *died* and *disease staging* are negative implying that at more advanced stage of the disease, incl. dying in the hospital, patients will have not too many opportunities for treatment and as a result their expenditures will be lower. *Died* is statistically significant in all specifications and *ds_stage* is significant in specifications 1 through 6. Regression results also show that elective admissions lead to higher hospital expenditures on average. This variable is only insignificant in specification 7 with controls for hospital ownership. According to the estimates from the regressions, government (Medicaid or Medicare) or firm provided insurance (private insurance) are associated with higher expenditures per inpatient record (excl. specification 7 for all three types of insurance and specification 6 for Medicaid) as compared to self-pay, charity, or other types of insurance. This result seems plausible considering the possibility that uninsured people will tend to spend less on treatment procedures. The coefficients on health insurance variables are significant at 1% level of significance for specifications 1 to 5 (with controls for hospital size and location). When I add controls for hospital region and ownership some of the signs change and some of the coefficients become insignificant. Specifications 3 and 7 have different comorbidity measures as controls. It turns out that only five comorbidities are significantly affecting expenditures in specification 3 (deficiency anemias, uncomplicated diabetes, metastatic cancer, obesity, and peptic ulcer disease) and only three in specification 7 (metastatic cancer, obesity, and peptic ulcer disease).

The main variables of interest in the regressions are the race variables and the interaction terms of the race variables with age: *black, otherrace, ageblack, and ageotherrace*. The interaction terms are positive and statistically significant in specifications 1 through 6. The coefficients on the two race variables are negative and statistically significant in all but specification 7 which provides some evidence for the presence of racial discrimination in the intensity of breast cancer treatment of African-Americans and other races compared to white Americans.

When I estimate the model for Medicare and Medicaid claims separately with all necessary controls I ignore issues of having difference in prices for various procedures and variation in out-of-pocket expenditures that complicate the analysis. The results from the regression based on Medicare claims illustrate that age, number of procedures, length of stay, disease staging, hospital size, location and ownership are significant predictors of hospital expenditures. Some of the income variables and

comorbidities are also statistically significant. The coefficients on black and other race, as well as on the interaction terms between age and race, are statistically insignificant from zero. The same can be said for the race variables estimated via OLS using Medicaid claims from HCUP. A possible explanation for these results is that patients of different races with such insurance plans are more alike/have similar characteristics and are less likely to be treated differently.

With regards to the quantile regressions, the results seem to be mixed. The coefficients on the health insurance variables and comorbidities, which are statistically significant in the previous specifications estimated by OLS, in all quantile regressions are insignificant. Race variables are also insignificant for the median and 25th percentile. For the 75th percentile, the regression coefficient estimates for *black, otherrace, ageotherrace* have the expected signs and are significant at 1% level of significance. The results suggest that racial discrimination is present only for the inpatient claims in the 75th percentile (claims with higher expenditures) and not for the 50th and 25th percentiles.

CONCLUSION

This paper tries to investigate whether there are disparities in the intensity of breast cancer treatment among different racial groups such as whites, blacks, and others. Based on cross-sectional data from the Healthcare Cost and Utilization Project for year 2002, I estimate different specifications using ordinary least squares and quantile regressions. The coefficients on race variables are negative and statistically significant in most of the specifications providing some evidence for racial discrimination. This result is also confirmed by the quantile regression for inpatient claims in the 75th percentile of hospital expenditures. However, further evidence is necessary to prove that African-Americans and other races are treated less intensively when admitted to hospitals for breast cancer procedures. A possibility for future research is to investigate this question using time-series data, as well as analysis by regions or states.

ENDNOTES

1. The regression results are available from the author upon request.

REFERENCES

Altekruse S.F., C.L.Kosary, M. Krapcho, N. Neyman, R. Aminou, W. Waldron, J. Ruhl, N. Howlader, Z. Tatalovich, H. Cho, A. Mariotto, M.P. Eisner, D.R. Lewis, K. Cronin, H.S. Chen, E.J. Feuer, D.J. Stinchcomb, B.K. Edwards (eds). SEER Cancer Statistics Review, 1975-2007, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2007/, based on November 2009 SEER data submission, posted to the SEER web site, 2010.

Andrews, R.M. and A. Elixhauser. 2000 “Use of major therapeutic procedures: are Hispanics treated differently than non-Hispanic Whites” *Ethnicity & Disease*, 10(3): 384-394.

Ansell, D, S. Whitrnan, R. Lipton and R. Cooper. 1993. “Race, Income, and Survival from Breast Cancer at Two Public Hospitals” *Cancer*, 72(10): 2974-2978.

Bach, P.B. , L.D. Cramer, D. Schrag, R.J. Downey, S.E. Gelfand, and C.B. Begg. 2001. “The influence of hospital volume on survival after resection for lung cancer” *New England Journal of Medicine*, 345(3): 181-188.

Ball, J.K. and A. Elixhauser. 1996. “Treatment differences between blacks and whites with colorectal cancer” *Medical Care*, 34(9): 970-984.

Bassett, M.T. and N. Krieger. 1986. “Social Class and Black-White Differences in Breast Cancer Survival” *American Journal of Public Health*, 76(12): 1400-1403.

Dimick, J.B., J.A. Cowan Jr., G.R. Upchurch Jr., and L.M. Colletti. 2003. “Hospital volume and surgical outcomes for elderly patients with colorectal cancer in the United States” *Journal of Surgical Research*, 114(1): 50-56.

Dowell, M.A., B. Rozell, D. Roth, H. Delugach, P. Chaloux, and J. Dowell. 2004. “Economic and clinical disparities in hospitalized patients with type 2 diabetes” *Journal of Nursing Scholarship*, 36(1): 66-72.

Figueroa, Janis B. and Nancy Breen (1995) “Significance of Underclass Residence on the Stage of Breast or Cervical Cancer Diagnosis” *The American Economic Review*, 85(2): 112-116.

Gordon, N.H., J.P. Crowe, J. Brumberg, and N.A. Berger. 1992. “Socioeconomic Factors and Race in Breast Cancer Recurrence and Survival” *American Journal of Epidemiolog,* 135(6): 609-618.

Harris, D.R., R. Andrews, and A. Elixhauser. 1997. “Racial and gender differences in use of procedures for black and white hospitalized adults” *Ethnicity & Disease*, 7(1): 91-105.

Karjalainen *S.* and E. Pukkala. 1990. “Social class as a prognostic factor in breast cancer survival” *Cancer*, 66(4): 819–826.

Katz, S.J. and T.P. Hofer. 1994. “Socioeconomic Disparities in Preventive Care Persist Despite Universal Coverage” *The Journal of the American Medical Association*, 272(7): 530-534.

Kravdal, Oystein. 2000. “Social Inequalities in Cancer Survival” *Population Studies,* 54(1), 1-18.

LeMarchand L., L.N. Kolonel, A. Nomura. 1984. “Relationship of ethnicity and other prognostic factors to breast cancer survival patterns in Hawaii” *Journal of the National Cancer Institute*, 73:1259-65.

Mitchell, Jean M. and Jack Hadley. 1997. “The Effect of Insurance coverage on Breast Cancer Patients’ Treatment and Hospital Choices” *The American Economic Review*, 87(2): 448-453.

Shen J.J. 2002. “Severity of illness, treatment environments and outcomes of treating acute myocardial infarction for Hispanic Americans” *Ethnicity & Disease*, 12(4):488-98

Thorpe, Kenneth E. and David Howard. 2003. “Health Insurance and Spending among Cancer Patients” *Health Affairs*, 189-198.

Schrijvers C.T., J.W. Coebergh, L.H. Van der Heijden, and J.P. Mackenbach. 1995. “Socioeconomic status and breast cancer survival in the southeastern Netherlands, 1980-1989” *European Journal of Cancer*, 31A(10):1660–1664.