RESEARCH ARTICLE

# Assessing the Impact of Socio-economic Variables on Breast Cancer Treatment Outcome Disparity

**Min Rex Cheung**

## Abstract

**Background: We studied Surveillance, Epidemiology and End Results (SEER) breast cancer data of Georgia USA to analyze the impact of socio-economic factors on the disparity of breast cancer treatment outcome. Materials and Methods: This study explored socio-economic, staging and treatment factors that were available in the SEER database for breast cancer from Georgia registry diagnosed in 2004-2009. An area under the receiver operating characteristic curve (ROC) was computed for each predictor to measure its discriminatory power. The best biological predictors were selected to be analyzed with socio-economic factors. Survival analysis, Kolmogorov-Smirnov 2-sample tests and Cox proportional hazard modeling were used for univariate and multivariate analyses of time to breast cancer specific survival data. Results: There were 34,671 patients included in this study, 99.3% being females with breast cancer. This study identified race and education attainment of county of residence as predictors of poor outcome. On multivariate analysis, these socio-economic factors remained independently prognostic. Overall, race and education status of the place of residence predicted up to 10% decrease in cause specific survival at 5 years. Conclusions: Socio-economic factors are important determinants of breast cancer outcome and ensuring access to breast cancer treatment may eliminate disparities.**

**Keywords:** SEER - breast cancer - race - neighborhood education attainment - disparity - cause specific survival

*Asian Pacific J Cancer Prev,* **14 (12)**, 7133-7136

## Introduction

Surveillance Epidemiology and End Results (SEER) (http://seer.cancer.gov/) is a public use cancer registry of United States of America (U.S.A.). Cancer is a major human burden. One out of three women and one out of two men in the U.S.A. will develop cancer in a lifetime (Siegel et al., 2012). Breast cancer is the most common cancer among women. It is estimated that 226,870 women will be diagnosed with and 39,510 women will die of cancer of the breast in 2012. The age-adjusted death rate was 23.0 per 100,000 women per year. The age-adjusted incidence rate was 124.3 per 100,000 women per year. These rates are based on cases diagnosed in 2005-2009 from 18 SEER geographic areas (http://seer.cancer.gov/statfacts/html/breast.html). SEER is funded by National Cancer Institute and Center for Disease Control to cover 28% of all oncology cases in U.S.A. SEER started collecting data in 1973 for 7 states and cosmopolitan registries. Its main purpose remains through collecting and distributing data on cancer, it strives to decrease the burden of cancer. SEER data are used widely as a bench-mark data source for monitoring cancer outcomes in U.S.A. and in other countries ( Shavers et al., 2003; Wampler et al., 2005; Gross et al., 2008; Lund et al., 2008; Schootman et al., 2009; Downing et al., 2010; Martinez et al., 2010; Rudat et al., 2012; Schlichting et al., 2012; Yao et al., 2012). Because of the uniformity and scope of the socio-economic data collected by SEER, their data are ideal for identifying the disparity in oncology outcome in different geographical and cultural areas ( Harlan et al., 1995; Shavers et al., 2003; Wampler et al., 2005; Gross et al., 2008; Lund et al., 2008; Downing et al., 2010; Martinez et al., 2010; Martinez et al., 2012; Schlichting et al., 2012; Yao et al., 2012). This study will focus on the disparity in breast cancer treatment outcome in the state of Georgia in U.S.A. Georgia is a relatively typical American state. By focusing on one state, this study aimed to find out if there were unintended socio-economic factors impacting on breast cancer outcome in a relatively homogeneous, medium sized state as a model. Similar epidemiology studies focusing on more homogeneous areas have also been done in Australia (Roder et al. 2012; Roder et al. 2013). We used SEER 18 data that covered the Atlanta and rural Georgia since 1974, and Greater Georgia since 2010. In particular, we sought to identify the socio-economic factors that contribute to worse outcome in breast cancer treatment with a goal to eliminate the disparity in the future.

## Materials and Methods

The Georgia cancer registry data were obtained from SEER 18 database. SEER registry has massive amount

*FROS Radiation Oncology CyberKnife Center, 40-20 Main St., 4 Fl, Flushing NY, USA *For correspondence: cheung.r100@gmail.com*

of data available for analysis, however, manipulating this data pipeline could be challenging. SEER Clinical Outcome Prediction Expert (SCOPE) is designed and implemented to mine SEER data and construct accurate and efficient prediction models (Cheung, 2012).

SEER is a public use database that can be used for analysis with no internal review board approval needed. Seer*Stat (http://seer.cancer.gov/seerstat/) was used for listing the cases. The filters used in Seer*Stat for the Georgia breast cancer case selection were: (Site and Morphology.Site rec with Kaposi and mesothelioma)=' Breast' AND Not (Age at Diagnosis. Age recode with <1 year olds)='00 years','01-04 years','05-09 years','10-14 years','15-19 years' AND (Race, Sex, Year Dx, Registry, County.SEER registry)='Atlanta (Metropolitan) - 1975+','Rural Georgia - 1992+','Greater Georgia - 2000+' AND (Race, Sex, Year Dx, Registry, County.Year of diagnosis)='2004','2005','2006','2007','2008','2009'. The last update of the SEER data was in November 2011 incorporating some of 2010 U.S. census data.

This study examined a long list of socio-economic factors (SEFs), staging and treatment factors that were available in SEER database with the goal of identifying the best SEFs to explain the disparity in breast cancer specific survival (COD='breast' in SEER). We used receiver operating characteristic curve to select the best pretreatment univariates for further analyses (Cheung et al., 2001a; 2001b). Similar strata were fused to make more efficient models if the ROC performance did not degrade (Cheung et al., 2001a; 2001b). Survival analysis was used to compute time to breast cancer specific survival data, Kolmogorov-Smirnov 2-sample test was used for comparing two survival curves and Cox multivariate analysis was performed to ascertain if SEFs were independently prognostic. To estimate the relative importance of SEFs versus traditional factors by Cox regression, the most significant pretreatment factors AJCC (American Joint Committee on Cancer) 2006 stage, estrogen receptor status (0=positive, 1=otherwise), progesterone receptor status (0=positive, 1=otherwise), race (0=non African American, 1=African American) and education attainment of county of residence (0>25% college graduate, 1=otherwise) of were used in a multivariate Cox model. These factors were scored as 1 for high-risk groups and 0 otherwise as indicated. All statistical analyses was performed in Matlab (http://www.mathworks.com/matlabcentral/fileexchange/authors/37883).

## Results

We analyzed 34,671 breast cancer cases diagnosed from 2004 to 2009 obtained from SEER database (Figure 1A). We used receiver operating characteristic curve (Hanley and McNeil, 1982) to study the performance of various univariate predictors. We have identified race/ethnicity, county percent graduating from college, American Joint Committee on Cancer staging (according to the AJCC 6th edition manual), estrogen receptor (ER) status and progesterone receptor (PR) status were discriminating models that could be used to build multivariate models. The American Joint Commission on Cancer (AJCC) staging had the highest ROC (S.D.) area of 0.83 (0.004) among the factors tested. The overall surviving fraction for the selected cases is 85% (Figure 1a). About one third of the deaths were not related to the diagnosis of the breast cancer. Therefore, this study explicitly built models to predict the cause specific survival as opposed to overall survival.

The AJCC model of the breast cancer was fed into SCOPE to be successively tested if adjacent strata could be merged (implemented as a subroutine of SCOPE) based on the ROC areas. The 7 risk strata of AJCC metastatic model were simplified to 3 strata without sacrificing the accuracy (ROC area (S.D.) is 0.82 (0.01). The second best model was AJCC non-metastatic model with (ROC area (S.D.) is 0.768 (0.01). The next tier of predictive models are biological ER/PR (ROC area (S.D.) is 0.656 (0.009); and treatment surgery/radiotherapy receipt (ROC area (S.D.) is 0.654 (0.005); socio-economic race/county% college bivariate (ROC area (S.D.) is 0.652 (0.004). Among the SEFs tested, Race/ethnicity and college education attainment in a county were the most predictive and the two SEFs combined has better ROC performance than the individual ones.

Figure 1B shows the surviving proportions when the patients were separated by race. African American (n=9240) had statistically worse survival outcome (Kolmogorov-Smirnov 2-sample test: h=1; p=1.6224e-09; k=0.5459). Table 1 shows the proportions of surviving patients separated by education attainment of county of residence. Patients lived in less educated (n=19764) areas had statistically worse survival outcome (Kolmogorov-Smirnov 2-sample test of college education attainment of county of residence: h=1; p=0.0300; k=0.2429). The difference in cause specific survival was about 2% for county education level and 10% for race/ethnicity at 60 months (Figure 1B and 1C). AJCC 2006 stage was the most predictive factor tested in this study for breast cancer specific survival (Figure 1D). Table 1 shows the Cox Proportional Hazard multivariate analysis. Multivariate analysis demonstrated that the SEF race and college degree attainment of the county of residence added independently significant information even when modeled with these

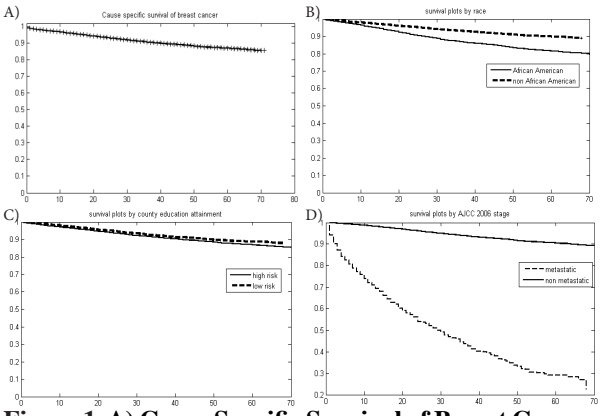


**Figure 1. A) Cause Specific Survival of Breast Cancer; B) Survival Plots by Race; C) Survival Plots by Country Education Attainment; and D) Survival Plots by AJCC 2006 Stage**

**Table 1. Multivariate Cox Proportional Hazard Analysis of AJCC 2006 Stage, Estrogen Receptor Status, Progesterone Receptor Status, Race and Education Attainment of the County of Residence on Breast Cancer Specific Survival***

| Predictor | Coefficient beta | Standard error | p value (x1.0e-04) |
|---|---|---|---|
| AJCC 2006 Stage | 2.43 | 0.043 | 0.00 |
| ER | 0.81 | 0.060 | 0.00 |
| PR | 0.53 | 0.064 | 0.00 |
| Race/ethnicity | 0.28 | 0.040 | 0.00 |
| % college graduates in county of residence | 0.15 | 0.039 | 0.92 |

*p values<0.05 represent statistically significant predictor

powerful biologic predictors.

For county with less than 25% of college graduate, there were 56% patients did not receive radiotherapy and more county with higher education attainment, there were 46% patients received radiotherapy. 35% and 34% were the comparable percentages of black patients receiving and not receiving radiotherapy respectively. Thus the predictive power of black/college bivariate could only be partially explained by whether the patients received radiotherapy or not. In terms of no receipt of radiotherapy, we hypothesize that living in a less educated county appears to fare less well because fewer patients received radiotherapy.

## Discussion

Using SEER data to study the effects of radiotherapy for breast cancer is an active area of investigation. The newest SEER data Nov 2011 have incorporated 2010 US census. However, the newly available socio-economic data have not been used in recent studies (Agarwal et al., 2012; Jagsi et al., 2012; Martinez et al., 2012; Sail et al., 2012; Yan et al., 2012). AJCC staging was found to be the best biological model to predict breast cancer specific survival. For comparison, the ROC area in predicting PSA failure based on Gleason Score, T-stage and PSA was about 0.75 in our previous studies (Cheung et al., 2001a; 2001a). Hormonal status (Davies et al., 2011) has been found to be predictive of breast cancer outcome in previous studies and was confirmed here. Race could predict in breast case about 10% cause specific survival decrement at 5 years (Figure 1B) and was significant in multivariate analysis with AJCC stage and hormonal receptor status (Table 1).

The Georgia cancer registry (one the SEER 18 registries) was used in our current studies. This state was used as a model for this study for several reasons. The entire state of Georgia is now covered by the SEER registry. While the state is a medium size state that provides relative social and economic uniformity, there are also significant social and economic variations that could be used as a model to study the impact of socio economic factors on oncology outcome.

It has been suggested that the disparity of using post-operative radiotherapy may have an impact on the survival of advanced breast cancer patients (Martinez et al., 2012). More studies are needed as related to the impact of socio-economic factors. For example, half of the breast cancer patients in this cohort (Table 1) did not receive radiotherapy. We investigated the disparity in outcome due to Race/% college graduate of County in relation to receipt of post-operative radiation treatment. Lack of radiotherapy has been shown to be associated with inferior outcomes as shown in other studies (Dragun et al., 2012; Feltner et al., 2012; Yao et al., 2012). We found that the breast cancer patients lived in less educated counties were at a disadvantage in terms of cause specific survival. Based on our data, we suggest that educating the public and patient about the utility of radiotherapy in the treatment of breast cancer may improve the frequency of receipt of radiotherapy and potentially the cause specific survival. Of note, a 10% improvement in 5 year actuarial cause specific survival would be more than the benefits of most chemotherapy regiments and the same as the benefit of post-operative radiotherapy after breast conservation therapy for patients with 3 or fewer positive lymph nodes at 15 years (Clarke et al., 2008; Voordeckers et al., 2009; Beal et al., 2010).

## References

Agarwal J, Agarwal S, Pappas L, Neumayer L (2012). A population-based study of breast cancer-specific survival following mastectomy and immediate or early-delayed breast reconstruction. *Breast J*, **18**, 226-32.

Beal SH, Martinez S R, Canter RJ, et al (2010). Survival in 12,653 breast cancer patients with extensive axillary lymph node metastasis in the anthracycline era. *Med Oncol*, **27**, 1420-1424.

Cheung R (2012). Poor treatment outcome of neuroblastoma and other peripheral nerve cell tumors may be related to under usage of radiotherapy and socio-economic disparity: A US SEER data analysis. *Asian Pac J Cancer Prev*, **13**, 4587-91.

Cheung R, Altschuler MD, D'Amico AV, et al (2001a). ROC-optimization may improve risk stratification of prostate cancer patients. *Urology*, **57**, 286-90.

Cheung R, Altschuler MD, D'Amico AV, et al (2001b). Using the receiver operator characteristic curve to select pretreatment and pathologic predictors for early and late post-prostatectomy PSA failure. *Urology*, **58**, 400-5.

Clarke M, Coates AS, Darby SC, et al (2008). Adjuvant chemotherapy in oestrogen-receptor-poor breast cancer: patient-level meta-analysis of randomised trials. *Lancet*, **371**, 29-40.

Davies C, Godwin J, Gray R, et al (2011). Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. *Lancet*, **378**, 771-84.

Downing S, Ahuja N, Oyetunji TA, et al (2010). Disparity in limb-salvage surgery among sarcoma patients. *Am J Surg*, **199**, 549-53.

Dragun A E, Huang B, Gupta S, Crew J B, and Tucker T C (2012). One decade later: trends and disparities in the application of post-mastectomy radiotherapy since the release of the american society of clinical oncology clinical practice guidelines. *Int J Radiat Oncol Biol Phys*, **83**, e591-6.

Feltner FJ, Ely GE, Whitler ET, Gross D, Dignan M (2012). Effectiveness of community health workers in providing outreach and education for colorectal cancer screening in Appalachian Kentucky. *Soc Work Health Care*, **51**, 430-40.

Gross C P, Smith B D, Wolf E, Andersen M (2008). Racial disparities in cancer therapy: did the gap narrow between

1992 and 2002? *Cancer*, **112**, 900-8.
Hanley JA, McNeil BJ (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, **143**, 29-36.
Harlan L, Brawley O, Pommerenke F, Wali P, Kramer B (1995). Geographic, age, and racial variation in the treatment of local/regional carcinoma of the prostate. *J Clin Oncol*, **13**, 93-100.
Jagsi R, Abrahamse P, Hawley ST, et al (2012). Underascertainment of radiotherapy receipt in Surveillance, Epidemiology, and End Results registry data. *Cancer*, **118**, 333-41.
Lund MJ, Brawley OP, Ward KC, et al (2008). Parity and disparity in first course treatment of invasive breast cancer. *Breast Cancer Res Treat*, **109**, 545-57.
Martinez SR, Beal SH, Chen SL, et al (2010). Disparities in the use of radiation therapy in patients with local-regionally advanced breast cancer. *Int J Radiat Oncol Biol Phys*, **78**, 787-92.
Martinez S R, Tseng W H, Canter R J, et al (2012). Do radiation use disparities influence survival in patients with advanced breast cancer? *Cancer*, **118**, 196-204.
Roder D, Webster F, Zorbas H, Sinclair S (2012). Breast screening and breast cancer survival in Aboriginal and Torres Strait Islander women of Australia. *Asian Pac J Cancer Prev*, **13**, 147-55.
Roder D, Zorbas H, Kollias J, et al (2013). Risk factors for poorer breast cancer outcomes in residents of remote areas of Australia. *Asian Pac J Cancer Prev*, **14**, 547-52.
Rudat V, Brune-Erbe I, Noureldin A, et al (2012). Epidemiology of breast cancer patients at a tertiary care center in the Eastern Province of Saudi Arabia. *Gulf J Oncology*, **1**, 45-9.
Sail K, Franzini L, Lairson D, Du X (2012). Differences in treatment and survival among African-American and Caucasian women with early stage operable breast cancer. *Ethn Health*, **17**, 309-23.
Schlichting JA, Soliman AS, Schairer C, et al (2012). Association of inflammatory and noninflammatory breast cancer with socioeconomic characteristics in the Surveillance, Epidemiology, and End Results database, 2000-2007. *Cancer Epidemiol Biomarkers Prev*, **21**, 155-65.
Schootman M, Jeffe DB, Lian M, Gillanders WE, Aft R (2009). The role of poverty rate and racial distribution in the geographic clustering of breast cancer survival among older women: a geographic and multilevel analysis. *Am J Epidemiol*, **169**, 554-61.
Shavers VL, Harlan LC, and Stevens JL (2003). Racial/ethnic variation in clinical presentation, treatment, and survival among breast cancer patients under age 35. *Cancer*, **97**, 134-47.
Siegel R, Desantis C, Virgo K, et al (2012). Cancer treatment and survivorship statistics, 2012. *CA Cancer J Clin*, **62**, 220-41.
Voordeckers M, Vinh-Hung V, Lamote J, Bretz A, Storme G (2009). Survival benefit with radiation therapy in node-positive breast carcinoma patients. *Strahlenther Onkol*, **185**, 656-62.
Wampler NS, Lash TL, Silliman RA, Heeren TC (2005). Breast cancer survival of American Indian/Alaska Native women, 1973-1996. *Soz Praventivmed*, **50**, 230-7.
Yan W, Christos P, Nori D, Chao KS, Ravi A (2012). Is there a cause-specific survival benefit of postmastectomy radiation therapy in women younger than age 50 with T3N0 invasive breast cancer? A SEER database analysis: outcomes by Receptor Status/Race/Age: Analysis Using the NCI Surveillance, Epidemiology, and End Results (SEER) Database. *Am J Clin Oncol*,
Yao N, Lengerich E J, and Hillemeier M M (2012). Breast cancer mortality in Appalachia: reversing patterns of disparity over