RESEARCH ARTICLE

Using SEER Data to Quantify Effects of Low Income Neighborhoods on Cause Specific Survival of Skin Melanoma

Min Rex Cheung

Abstract

Background: This study used receiver operating characteristic (ROC) curves to screen Surveillance, Epidemiology and End Results (SEER) skin melanoma data to identify and quantify the effects of socioeconomic factors on cause specific survival. Methods: 'SEER cause-specific death classification' was used as the outcome variable. The area under the ROC curve was to select best pretreatment predictors for further multivariate analysis with socioeconomic factors. Race and other socioeconomic factors including rural-urban residence, county level % college graduate and county level family income were used as predictors. Univariate and multivariate analyses were performed to identify and quantify the independent socioeconomic predictors. Results: This study included 49,666 patients. The mean follow up time (SD) was 59.4 (17.1) months. SEER staging (ROC area of 0.80) was the most predictive factor. Race, lower county family income, rural residence, and lower county education attainment were significant univariates, but rural residence was not significant under multivariate analysis. Living in poor neighborhoods was associated with a 2-4% disadvantage in actuarial cause specific survival. Conclusions: Racial and socioeconomic factors have a significant impact on the survival of melanoma patients. This generates the hypothesis that ensuring access to cancer care may eliminate these outcome disparities.

Keywords: Melanoma - SEER data - socio-economic disparity - cause specific survival

Asian Pacific J Cancer Prev, **14 (5)**, 3219-3221

Introduction

The Surveillance Epidemiology and End Results (SEER) cancer registry data have been extensively used to build prognostic models for skin melanoma (Weir et al., 2011; Brewer et al., 2012; Shaikh et al., 2012; Siegel et al., 2012). SEER is a particularly important source for identifying disparities in cancer treatment (Baine et al., 2011; Singal et al., 2012). National Cancer Institute and Center for Disease Control fund SEER to monitor the cancer epidemiology of U.S. SEER registers about 28% of all of the oncology cases in U.S. SEER started collecting data in 1973 for 7 state and cosmopolitan registries. Through collecting and distributing data on cancer, SEER strives to decrease the burden of cancer. SEER data have been used widely as a benchmark for studying cancer outcomes in U.S. and in other countries; (Ognjanovic et al., 2009; Sultan et al., 2009; Cheung et al., 2010; McDowell et al., 2010; Pappo et al., 2010; Bhatia, 2011; Perez et al., 2011). The vastness and uniformity of SEER data are ideal for identifying potential socioeconomic disparities in oncology outcome. In addition to the biological staging factors and the treatment factors, this database also contains a large number of county level socio-economic data. The nature of the socio-economic barriers to good outcome for skin melanoma has not been well characterized. This study used ROC analysis to select the best pretreatment prediction factor for multivariate analysis with the socioeconomic factors (Gimotty et al., 2007; Pradhan et al., 2011; Cheung, 2012; Lee et al., 2012). Furthermore, the effects of the socioeconomic factors were quantified. These data may be useful in designing future clinical trials for melanoma patients.

Materials and Methods

SEER registries contain public use data. These data can be used for analysis with no internal review board approval needed. SEER registry has massive amount of data available for analysis, however, manipulating the data could be challenging. SEER Clinical Outcome Prediction Expert (SCOPE) (Cheung, 2012). The data were obtained from SEER 18 database. SEER*Stat (http://seer.cancer.gov/seerstat/) was used for listing the cases. The filters used were: 'Year of diagnosis'='2004','2005','2006' AND 'Site rec with Kaposi and mesothelioma'='Melanoma of the Skin'. 'SEER cause-specific death classification' was used as the outcome variable. All of the statistics and programming of this study were performed in Matlab (www.mathworks.com). The absolute risk of cause specific death was used in ROC area calculations. For univariate and multivariate analysis the following coding was used:

*275 S Bryn Mawr Ave, K43, Bryn Mawr, PA, USA *For correspondence: cheung.r100@gmail.com*

Table 1. Univariate Kolmogorov-Smirnov 2-sample Tests and Multivariate Cox Proportional Hazard Regression of Cause Specific Survival of Skin Melanoma Cancers

	Kolmogorov-Smirnov tests			Cox Proportional Regression		
	h	p	k	beta	s.e.	p
SEER stage (0=local/regional, 1=distant/unstaged)	1.000	2.13E-28	0.967	2.569	0.034	0.000
Race (0=not African American, 1=African American)	1.000	6.14E-09	0.639	0.810	0.138	0.000
Rural-urban status of county of residence (0=urban, 1 =rural)	1.000	1.94E-05	0.402	0.072	0.050	0.145
County level %college graduate (0=more than 25%, 1=less/equal 25%)	1.000	1.14E-07	0.479	0.114	0.047	0.017
County level family income (0=higher/equal $50k/year, 1=lower than $50k/year)	1.000	6.60E-08	0.487	0.115	0.048	0.000

*The result h of Kolmogorov-Smirnov test statistics k was 1 if the test rejected the null hypothesis at the 5% significance level; 0 otherwise. S.e. are the standard errors of Cox proportional hazards coefficients beta. Test probability p<0.05 was statistically significant

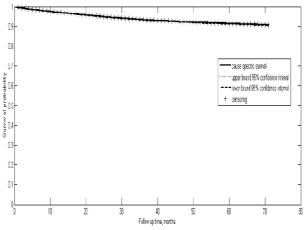

Figure 1. Kaplan Meier Cause Specific Survival Plot of SEER Skin Melanoma Cancer Patients

SEER stage (0=local/regional, 1=distant/unstaged); race (0=not African American, 1=African American); rural-urban status of county of residence (0=urban, 1 =rural); county level % college graduate (0=more than 25%, 1= less/equal 25%); County level family income (0=higher/ equal $50k/year, 1=lower than $50k/year). Kaplan-Meier method was used to plot the survival curves. This study used an earlier homemade version of Matlab Kolmogorov-Smirnov's test extended to compare two censored survival curves (posted on Matlab File Exchange). A more accurate version and without normalization to take into account of failures at zero followup times has been submitted to the same website.

Results

There were 49666 patients included in this study (Figure 1). The follow up duration (S.D.) was 59.4 (17.1) months. The mean (S.D.) age was 45.9 (17.5) years. The absolute risk of death from skin melanoma was 7.5% in this cohort. 392 SEER patents younger than 20 years old were diagnosed with melanoma from 2004 to 2006. They had a 1.8% risk of cause specific death compared with 7.5% for the older patients. There was a 5.8% risk of cause specific death for female and 8.7% for male patients. 98.9% of skin melanoma did not have histological grading completed. SEER stage (Figure 2a and Table 1) was the most predictive model for risk of cause specific death. There was about 40% difference in cause specific survival (Figure 2a) between the local/regional versus the distant and un-staged patients. SEER's relatively simple staging model had a higher ROC area (0.80) than that (0.76) of the AJCC (American Joint Committee on Cancer) stage. There were only 227 African American patients. However, race (Figure 2b) was highly significant predictor or cause specific survival. Rural residence (Figure 2c), and lower county education attainment (Figure 2d), and lower county family income (Figure 2e) were significant univariates (Table 1). But only lower county income and lower county education level were significant predictors of

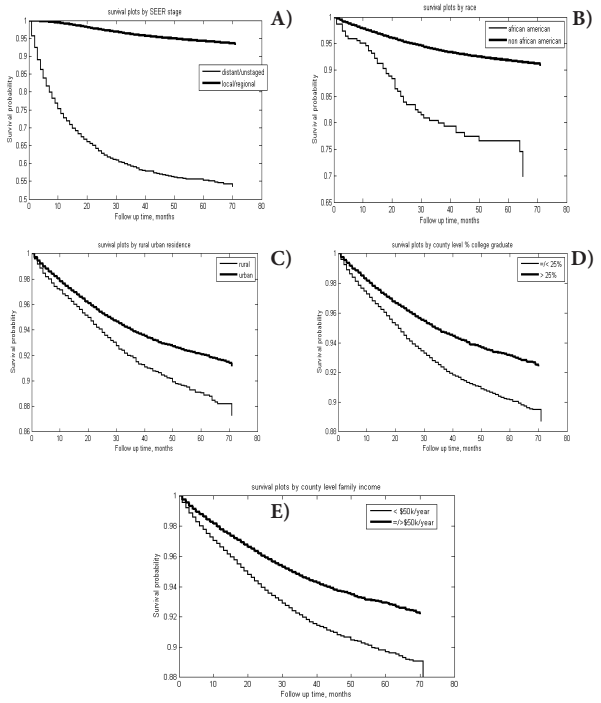

Figure 2. The Kolgomorov-Smirnov's Tests of Predictors for Skin Melanoma Specific Death. A) SEER stage, B) Race, C) Rural urban residence, D) County level % college graduate and E) County level family income

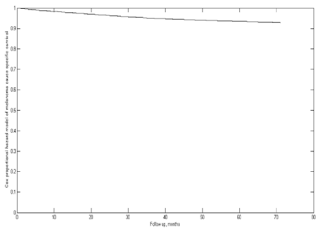

Figure 3. Cox Proportional Hazard Model for Skin Melanoma Cause Specific Survival

cause specific survival under multivariate analysis (Table 1). The corresponding Cox proportional hazard model is shown in Figure 3. Living in neighborhoods with lower social status overall decreased the cause specific survival by about 2-4%.

Discussion

This study was interested in identifying and measuring the effects of socioeconomic factors on the cause specific survival of skin melanoma patients (Figure 1). To that end, this study used ROC area to screen potential explanatory

factors. When measuring the accuracy of prediction models, ROC analysis takes into account both sensitivity and specificity of the prediction model. Ideal model would have a ROC area of 1 and a random model is expected to have an area of 0.5 (Hanley and McNeil, 1982). This study found SEER stage (Figure 2a) was the most predictive of all pretreatment models and was used for further multivariate analysis. SEER staging defines extent of disease as localized, regional, distant or incompletely staged/others. There was no information on why some patients were un-staged. Similarly no information on why most skin melanoma cancers were not graded. This kind of deficiency is expected as SEER collect data from very large areas of contributing cancer registries and there were likely to be various reasons for not staging the patients making it difficult to ascertain. For a ROC area, SEER stage's 0.80 could be considered high. For example, the prostate risk model had a ROC area of 0.75 in its accuracy of predicting biochemical failure (Cheung et al., 2001a; 2001b).

This study found race to be a significant univariate and multivariate predictor of cause specific survival favoring non-African American race/ethnicity by about 20% (Figure 2b). Therefore, African American patients had a very low incidence of skin melanoma, but once diagnosed they have poorer outcomes. The etiology for this could be biological or correlated with living in the counties with lower socioeconomic status. County level low family income and county level low % of college graduates decreased the cause specific survival by about 2-4% actuarially (Figure 2d-2e). These factors were independent predictors under multivariate Cox proportional hazard model analysis (Table 1, Figure 3). The data here did not offer the reasons for these observations. Although, long distance of travel and limited medical access to high-expertise centers could be the reasons. This hypothesis could be further studied in the future and could form a testable hypothesis to see if removing these barriers could improve the outcome of melanoma patients. In an interesting study, simply moving patients out from low income neighborhoods improved their diabetes and obesity (Ludwig et al., 2012). A study as this has never been done in oncology, however, it will test a very significant hypothesis for cancer treatment in general. The data presented here may be useful for designing such trials in the future.

In conclusion, racial and socioeconomic factors have a significant impact on the survival of melanoma patient. It generates the hypothesis that ensuring access to cancer care may eliminate these outcome disparities.

References

Baine M, Sahak F, Lin C, et al (2011). Marital status and survival in pancreatic cancer patients: a SEER based analysis. *PLoS One*, **6**, 21052.

Bhatia S (2011). Disparities in cancer outcomes: lessons learned from children with cancer. *Pediatr Blood Cancer*, **56**, 994-1002.

Brewer JD, Shanafelt TD, Otley CC, et al (2012). Chronic lymphocytic leukemia is associated with decreased survival of patients with malignant melanoma and Merkel cell carcinoma in a SEER population-based study. *J Clin Oncol,* **30**, 843-9.

Cheung MC, Zhuge Y, Yang R, et al (2010). Incidence and outcomes of extremity soft-tissue sarcomas in children. *J Surg Res*, **163**, 282-9.

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.

Gimotty PA, Elder DE, Fraker DL, et al (2007). Identification of high-risk patients among those diagnosed with thin cutaneous melanomas. *J Clin Oncol*, **25**, 1129-34.

Hanley JA, McNeil BJ (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology,* **143**, 29-36.

Lee M, Cronin KA, Gail MH, Feuer EJ (2012). Predicting the absolute risk of dying from colorectal cancer and from other causes using population-based cancer registry data. *Stat Med,* **31**, 489-500.

Ludwig J, Duncan G J, Gennetian LA, et al (2012). Neighborhood effects on the long-term well-being of low-income adults. *Science*, **337**, 1505-10.

McDowell HP, Foot AB, Ellershaw C, et al (2010). Outcomes in paediatric metastatic rhabdomyosarcoma: results of The International Society of Paediatric Oncology (SIOP) study MMT-98. *Eur J Cancer,* **46**, 1588-95.

Ognjanovic S, Linabery AM, Charbonneau B, Ross JA (2009). Trends in childhood rhabdomyosarcoma incidence and survival in the United States, 1975-2005. *Cancer,* **115**, 4218-26.

Pappo A S, Krailo M, Chen Z, Rodriguez-Galindo C, Reaman G (2010). Infrequent tumor initiative of the children's oncology group: initial lessons learned and their impact on future plans. *J Clin Oncol*, **28**, 5011-6.

Perez EA, Kassira N, Cheung MC, et al (2011). Rhabdomyosarcoma in children: a SEER population based study. *J Surg Res*, **170**, 243-251.

Pradhan TS, Stevens EE, Ablavsky M, et al (2011). Figure O staging for carcinosarcoma: can the revised staging system predict overall survival? *Gynecol Oncol*, **123**, 221-4.

Shaikh WR, Weinstock MA, Halpern AC, et al (2012). The characterization and potential impact of melanoma cases with unknown thickness in the United States' Surveillance, Epidemiology, and End Results Program, 1989-2008. *Cancer Epidemiol,* **37**, 64-70.

Siegel R, Desantis C, Virgo K, et al (2012). Cancer treatment and survivorship statistics, 2012. *CA Cancer J Clin,* **62**, 220-41.

Singal V, Singal AK, Kuo YF (2012). Racial disparities in treatment for pancreatic cancer and impact on survival: a population-based analysis. *J Cancer Res Clin Oncol*, **138**, 715-22.

Sultan I, Qaddoumi I, Yaser S, Rodriguez-Galindo C, Ferrari A (2009). Comparing adult and pediatric rhabdomyosarcoma in the surveillance, epidemiology and end results program, 1973 to 2005: an analysis of 2,600 patients. *J Clin Oncol,* **27**, 3391-7.

Weir HK, Marrett LD, Cokkinides V, et al (2011). Melanoma in adolescents and young adults (ages 15-39 years): United States, 1999-2006. *J Am Acad Dermatol,* **65**, 38-49.