Cancer Epidemiology 49 (2017) 112–117

Contents lists available at ScienceDirect

Cancer Epidemiology

The International Journal of Cancer Epidemiology, Detection, and Prevention

journal homepage: www.cancerepidemiology.net

Socioeconomic measures influence survival in osteosarcoma: an analysis of the National Cancer Data Base

Benjamin J. Miller, Yubo Gao, Kyle R. Duchman*

Department of Orthopaedics and Rehabilitation, University of Iowa Hospitals and Clinics, 200 Hawkins Drive, 01008 JPP, Iowa City, IA 52242, USA

ARTICLE INFO

Article history:
Received 22 December 2016
Received in revised form 29 April 2017
Accepted 31 May 2017
Available online 9 June 2017

Keywords:
Osteosarcoma
survival
overall survival
socioeconomic status

ABSTRACT

Background: While previous studies have identified low socioeconomic status as a risk factor for metastatic disease in patients with high-grade osteosarcoma, the influence of socioeconomic status on overall survival remains unclear. The present study aims to investigate the relationship between survival and socioeconomic status in patients with high-grade conventional osteosarcoma.

Methods: The National Cancer Data Base (NCDB) was queried from 1998-2012 to identify all patients <40 years of age with a diagnosis of high-grade conventional osteosarcoma. A total of 3,503 patients were identified that met inclusion and exclusion criteria. Univariate relationships were investigated using Kaplan-Meier survival analysis and associated log-rank tests in order to determine patient, socioeconomic, tumor, and treatment variables associated with overall survival. Multivariate analysis was performed to determine independent predictors of survival.

Results: In order of decreasing magnitude, metastatic disease (Hazard Ratio [HR] 3.28, 95% Confidence Interval [CI] 2.82-3.82), primary site in the pelvis or spine (HR 2.15, 95% CI 1.79-2.59), positive surgical margins (HR 1.82, 95% CI 1.46-2.27), tumor size >8 cm (HR 1.47, 95% CI 1.24-1.74), age ≥18 years (HR 1.30, 95% CI 1.14-1.48), lowest quartile of composite socioeconomic status (HR 1.23, 95% CI 1.02-1.51), and Medicaid insurance (HR 1.18, 95% CI 1.02-1.38) were predictors of decreased survival at 5 years.

Conclusion: Treating providers should be aware that some of their patients may have challenges unrelated to their diagnosis that make timely presentation, adherence to treatment, and continued close surveillance difficult. This investigation suggests that socioeconomic variables influence overall survival for osteosarcoma in the United States, although not as dramatically as established tumor- and treatment-related risk factors.

1. INTRODUCTION

Osteosarcoma is the most common primary sarcoma of bone, typically occurring in adolescents and young adults [1,2]. With modern chemotherapy and surgical techniques, 5-year overall survival in high-grade conventional osteosarcoma approaches 70% for non-metastatic disease [2–4] and 30% for metastatic disease at diagnosis [5–8]. Several clinical risk factors have been established that predict a poor prognosis, including metastatic disease, poor response to chemotherapy, large tumor size, axial tumor location, positive surgical margins, and older patient age [3,9–14].

Socioeconomic factors, including household income, insurance status, education, and poverty, have been established as poor prognostic indicators in several other cancers [15–17]. While broad and difficult to quantify, socioeconomic measures reflect influences distinct from tumor biology and treatment. They are representative of individual, local, and regional disparity that may affect time to diagnosis, access to specialty care, adherence to treatment protocols, and ability to comply with long-term surveillance [18,19]. Prior work has established that osteosarcoma patients with lower socioeconomic status have an increased risk of presentation with metastatic disease [20], but no definitive difference in overall survival has been shown to date.

The National Cancer Data Base (NCDB) is a joint effort of the American College of Surgeons Commission on Cancer and the American Cancer Society. The NCDB collects treatment and outcomes data from more than 1,500 hospitals, representing 70% of all new cancer diagnoses in the United States [21,22]. This data source has been used to investigate risk factors for cancer survival in many types of malignancy, but never in osteosarcoma. We sought to utilize the NCDB to investigate patient, tumor, and treatment factors associated with diminished 2-, 5-, and 10-year

* Corresponding author at: Department of Orthopaedic Surgery, University of Iowa Hospitals and Clinics, 200 Hawkins Drive, 01015 JPP, Iowa City, IA 52242, USA.
E-mail addresses: benjamin-j-miller@uiowa.edu (B.J. Miller), yubo-gao@uiowa.edu (Y. Gao), kyle-duchman@uiowa.edu (K.R. Duchman).

http://dx.doi.org/10.1016/j.canep.2017.05.017
1877-7821/

overall survival in patients with high-grade conventional osteosarcoma.

2. MATERIALS AND METHODS

2.1. Human Subjects Determination

This study was exempt from Institutional Review Board review as it contains only deidentified data. A methodological review of the project proposal was performed by the NCDB prior to providing the requested information. The views expressed herein are those of the authors and are not necessarily reflective of the NCDB.

2.2. Data Elements

We queried the NCDB from 1998-2012 and identified all cases of high-grade conventional osteosarcoma in patients younger than 40 years of age. Patients with low-grade osteosarcoma, non-conventional subtypes of osteosarcoma, not treated with multi-agent chemotherapy, not treated with surgical intervention, or unknown vital status were excluded. We limited the investigation to patients with only one known malignancy (osteosarcoma) in order to eliminate confusion with the survival analysis. We recorded patient age, sex, race, distance from ZIP code centroid of the patient's residence to the hospital that reported the case, population density (metro, urban, or rural), tumor location, tumor size, metastatic disease at initial presentation, and final surgical margins directly from the database. The NCDB collects socioeconomic measures at an individual level (insurance status) and ZIP code level (median annual household income, percent of population without a high school degree). Insurance status is reported by the NCDB as "not insured," "private insurance/managed care," "Medicaid," "Medicare," and "other government" (e.g. TRICARE, Veteran's Affairs). In order to participate in the NCDB, institutions must complete a data use agreement and are required to submit survival data to the NCDB annually, which is reported as all-cause survival. Specific causes of death are not reported, so calculations of cause-specific survival are not possible.

2.3. Composite Socioeconomic Status Measure

In order to account for several socioeconomic factors, we combined two socioeconomic variables, income and education, to create a composite measure of socioeconomic status (SES composite), similar to methods used in prior analyses [13,15,20,23]. The median household income was listed within the NCBD by matching the ZIP code of the patient at the time of diagnosis to data sourced from the 2012 American Community Survey, reporting years 2008-2012, and adjusted for 2012 inflation [24]. The NCDB reports this variable in quartiles (1 – <$38,000, 2 – $38,000–47,999, 3 – $48,000-62,999, 4 – ≥$63,000) based on equally distributed income ranges for all United States ZIP codes. Education, similarly, was derived from the 2012 American Community Survey to report the number of people in the ZIP code of residence at the time of diagnosis who did not graduate from high school. In the NCDB, this was reported relative to equally proportioned quartiles in the United States population (1 – ≥21%, 2–13-20.9%, 3–7-12.9%, 4 – <7%) (Table 1). The quartile assignments of the two measures were added, and new categories were created for a combined score of 2-3, 4-5, 6-7, and 8.

2.4. Statistical Analysis

Kaplan-Meier survival analysis was performed at 2, 5, and 10 years, and a log-rank test used at each time point in order to determine variables of interest associated with decreased survival.

Table 1
Kaplan-Meier Survival Estimates at 2, 5, and 10 Years for Each Individual Education and Income Quartile.

Education					
Year	Quartile				p value
	1	2	3	4	
2-year	77.8	82.1	82.2	85.1	0.001
5-year	59.1	63.0	66.1	67.4	0.001
10-year	50.9	57.0	58.0	61.5	<0.001
Income					
Year	Quartile				p value
	1	2	3	4	
2-year	77.5	82.0	82.6	84.1	0.016
5-year	57.8	62.7	65.6	67.1	0.001
10-year	50.8	55.2	58.4	60.9	0.001

Measures that demonstrated a level of association of at least $p < 0.1$ at 10 years were used to create a multivariate Cox proportional hazards model and ultimately included age, sex, race, insurance status, SES composite score, metastatic disease, site, tumor size, and tumor margins. For the multivariate models, there were a substantial number of patients with missing size (1131/3503) and margin status (962/3503). Rather than exclude these patients, we elected to create an additional unknown variable to represent missing data for these two characteristics, as has been done in similar analyses [13,25]. Additionally, for the multivariate analysis, we excluded any patient who had missing data for race, insurance status, income, percent with high school degree, metastatic disease, or site, leaving 3107 patients (88.7%) available for multivariate analysis.

2.5. Missing Data

The multivariate analysis was repeated while excluding all missing data for size and margins. We found similar hazard ratio estimates for all variables, although an increase in the 95% confidence intervals for our socioeconomic and insurance variables, likely due to a substantial reduction in the size of our cohort. The hazard ratios and confidence intervals for size and margins, specifically, were similar to our primary analysis.

3. RESULTS

3.1. Univariate Analysis

Univariate analysis revealed improvements in overall survival estimates at 5 years for patients with localized disease (69% vs. 26%, $p < 0.001$), extremity tumors (67% vs. 36%, $p < 0.001$), tumors ≤8 cm (72% vs. 59%, $p < 0.001$), negative margins (70% vs. 43%, $p < 0.001$), age <18 years (68% vs. 58%, $p < 0.001$), white race (65% vs. 60%, $p = 0.018$), private insurance (67% for private compared to 53% for uninsured and 58% for Medicaid, $p < 0.001$), and higher SES composite (68% in highest quartile compared to 58% in lowest quartile, $p < 0.001$) (Table 2). These differences were apparent at 2 years and maintained out to 10 years after diagnosis. When further evaluating the SES composite score and insurance status in patients with localized and metastatic disease, both SES composite score and insurance status had a significant influence on survival in patients with localized disease ($p = 0.001$ for both variables), but not in patients with metastatic disease ($p = 0.061$ and $p = 0.148$, respectively) (Fig. 1). There were no substantial differences in survival with variation in population density or distance to the treating center.