

RESEARCH LETTER

Prevalence of Primary Epithelioid Sarcoma in the United States: A Cross-sectional Study Using the SEER Database

Shangyi Fu BS^{1,2}, Ramlah Khan BS³, Darien Hinson BS¹, Melissa Marchan-Martinez BS⁴, Diana Bonilla, MS⁴, Ibeth Caceres BS¹, Danny Huynh BS⁵

¹ School of Medicine, Baylor College of Medicine, Houston, TX

² Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX

³ College of Medicine, Texas A&M University Health Science Center, Temple, TX

⁴ John Sealy School of Medicine, The University of Texas Medical Branch, Galveston, TX

⁵ Cullen College of Engineering, University of Houston, Houston, TX

ABSTRACT

Primary epithelioid sarcoma (PES) is a rare form of sarcoma that usually affects the extremities and comprises less than 1% of all sarcomas. PES can occur in various parts of the body, including bone, adrenal glands, vulva, and pleura. Despite its rarity, there is a significant knowledge gap in the medical literature concerning diagnosing and treating PES. A PubMed search revealed only 27 articles on PES, highlighting the limited available options for patients with this condition. To address this gap, we investigated the prevalence of PES in the US population. Our study utilized data from the Surveillance, Epidemiology, and End Results (SEER) program to identify disease occurrence patterns from 2017 to 2018. We found that the highest prevalence was among individuals of White race between the ages of 25-59, with the second-highest prevalence in individuals of Black race between the ages of 30-64. Additionally, our findings suggest that the prevalence rate of PES is the highest in the 30-34 age group.

Primary epithelioid sarcoma (PES) is a rare, aggressive soft tissue malignancy characterized by epithelioid-appearing cells with both mesenchymal and epithelial features, classically forming nodules of epithelioid to spindle cells with central necrosis that resemble granulomas. PES mainly affects young adults, often in the hands and fingers¹. Due to its rarity and similarity to other tumors, it can be challenging to diagnose, with symptoms such as pain, swelling, and lumps². Our objective was to utilize the SEER database, an initiative recently launched by the Surveillance Research Program (SRP) within

the Division of Cancer and Population Sciences (DCCPS) at NCI³, to determine the prevalence of PES among various sociodemographic groups, identify those at the highest risk of developing the disease, and provide personalized patient care based on sociodemographic factors.

As our study had a retrospective design, it was determined to be exempt from review by the Institutional Review Board (IRB). We conducted a cross-sectional analysis of the SEER database by identifying patients diagnosed with primary epithelioid sarcoma using the ICD-10-CM code C49.49 and ICD-

9-CM code 171.4. Subsequently, we obtained and analyzed the electronic medical records of these patients to collect data on age, sex, and self-reported race. To calculate the prevalence of PES in the overall population, we utilized census data and estimations to account for populations that were lost to follow-up, which were provided by the SEER database.

As of 2020, the SEER database had enrolled 43,926,825 individuals (**Table 1**). For our study, we utilized the most recent available data, which includes census data from 2017-2018. Our analysis identified 217.2 cases of PES, corresponding to an overall prevalence rate of 0.49 per 1,000,000 individuals. The highest prevalence rate was observed among individuals in the 35-39 age group. Our findings indicate that the prevalence rate of primary epithelioid sarcoma is highest among the White population and lowest among the American Indian population for the years 2017-2018.

The SEER database predominantly comprises White (71%), Black (12%), American Indian (2%), and Asian-Pacific Islander (15%) individuals⁴, whereas the demographic composition of the United States population is different, with 76% White, 14% Black, 1% American Indian, and 7% Asian individuals⁵. Despite these differences, a chi-squared test of independence revealed no significant difference between the SEER and U.S. Census populations, $X^2(3, N = 304,167,848) = 3,404,209.89$, $p < 0.00001$ (**Table 2**), suggesting that the two populations are sufficiently comparable for estimating PES prevalence. Based on these demographics, our prevalence calculation for PES likely underestimates the incidence in White and Black populations while overestimating it in Hispanic and Asian populations. Moreover, it is likely that additional patients were not

captured due to factors such as residency status, limited access to healthcare, and gaps in census reporting.

Collectively, our data suggests that the prevalence of PES is least prevalent in American Indian populations and most prevalent among White populations. Further epidemiological studies that do not rely solely on billing code are needed to validate these findings and provide a clearer understanding of PES prevalence across diverse populations.

Conflict of Interest Disclosures: None

Funding: None

Corresponding Author:

Melissa Marchan-Martinez B.S.
301 University Blvd
4.112 McCullough Building
Galveston, TX, USA 77550
Email: memarcha@utmb.edu

References:

1. Sobanko JF, Meijer L, Nigra TP. Epithelioid sarcoma: a review and update. *J Clin Aesthet Dermatol.* 2009 May;2(5):49-54. PMID: 20729965; PMCID: PMC2924131.
2. Czarnecka AM, Sobczuk P, Kostrzanowski M, Spalek M, Chojnacka M, Szumera-Cieckiewicz A, Rutkowski P. Epithelioid Sarcoma-From Genetics to Clinical Practice. *Cancers (Basel).* 2020 Jul 29;12(8):2112. doi: 10.3390/cancers12082112. PMID: 32751241; PMCID: PMC7463637.
3. SEER Cancer Statistics Review, 1975-2018 Available online: https://seer.cancer.gov/csr/1975_2018/index.html (accessed on 2 March 2023).
4. Bureau, U.C. 2018 Population Estimates by Age, Sex, Race and Hispanic Origin Available

Table 1. Prevalence of primary epithelioid sarcoma in the United States, stratified by age and race.

	Estimated Prevalence Percent	Estimated Prevalence Count	Populations at Prevalence Date	Known Alive	Lost
White	0.00%	179.2	31,240,899.00	164	19
Black	0.00%	18.3	5224726.00	17	2
American Indian	0.00%	0	897,021.50	0	0
Asian or Pacific Islander	0.00%	17.9	6,564,178.50	17	1.9
Unknown	~	1.8	0.00	0	~
00 years	0.00%	0	522,150.00	0	0
01-04 years	0.00%	1	2,159,900.50	1	0
05-09 years	0.00%	4	2,732,193.50	3	1
10-14 years	0.00%	1	2,786,556.50	1	0
15-19 years	0.00%	4	2,781,405.00	4	0
20-24 years	0.00%	16	2,954,646.50	16	0
25-29 years	0.00%	22.6	3,391,241.00	21	2
30-34 years	0.00%	11	3,191,280.50	11	0
35-39 years	0.00%	26.6	3,049,214.50	23	4
40-44 years	0.00%	21	2,779,268.50	18	4
45-49 years	0.00%	17.3	2,889,978.50	16	2
50-54 years	0.00%	17.9	2,837,532.50	16	2
55-59 years	0.00%	21.4	2,873,440.00	20	2
60-64 years	0.00%	13.3	2,591,284.50	10	4
65-69 years	0.00%	13.8	2,124,581.00	12	2
70-74 years	0.00%	10	1,598,623.50	10	0
75-79 years	0.00%	7	1,077,416.00	7	0
80-84 years	0.00%	4	736,037.00	4	0
85+ years	0.00%	5.3	850,075.50	5	1

Table 2. Chi-square test of independence comparing SEER and USA estimated race proportions in 2018.

	Estimated SEER Population in 2018	Estimated USA Population in 2018	Row Totals
White	31240899 (33049304.38) [98953.07]	197606407 (195798001.62) [16702.57]	228847306

Black	5224726 (6661487.98) [309883.47]	40902223 (39465461.02) [52306.12]	46126949
American Indian	897022 (478652.71) [365678.20]	2417371 (2835740.29) [61723.87]	3314393
Asian-Pacific Islander	6564179 (3737380.93) [2138071.41]	19315021 (22141819.07) [360891.18]	25879200
Column Totals	43926826	2.6E+8	304167848 (Grand Total)

The observed population is depicted in each cell, expected in parentheses, and chi-square statistics in brackets. The Chi-square statistic is 3404209.8855. The p-value = < 0.00001.