

An assessment of selection bias in the *Cancer Incidence in Louisiana by Census Tract* report

Lauren Maniscalco, MPH¹; Carla Rosales, MPH¹; Lu Zhang, PhD, MPH¹; Julia Poynter²; Wendemi Sawadogo²; Xiao-Cheng Wu, MD, MPH¹

¹Louisiana Tumor Registry, ²LSUHSC School of Public Health

BACKGROUND

In 2017, lawmakers passed a bill requiring that the Louisiana Tumor Registry (LTR) release cancer incidence rates by census tract. Previously, cancer statistics were released at the parish (county) level. Thus, the LTR produced its first annual census tract cancer incidence report in March 2018. The second report was issued in March 2019 covering diagnosis years 2005 through 2015.

Given the small geographic area of a census tract, the importance of assigning a case to the correct census tract is highly important. To avoid assigning cases to the wrong census tract, cases were excluded in the calculation of census tract incidence rate if the census tract certainty of the residential address was low.



OBJECTIVE

This study aimed to assess selection bias by determining if the excluded cases differ by sociodemographic or clinical characteristics from the cases that were included in the report.

METHODS

For diagnosis years 2005-2015, cases were stratified by their census tract certainty: high (census tract certainty score = 1) and low (census tract certainty score > 1) to assess differences between the two groups with respect to demographics, cancer type, and stage at diagnosis.

Sociodemographic and clinical characteristics of the two groups were compared using the Chi-Square test or Students' t-test. All analyses were conducted using SAS 9.4.

LSU Health
NEW ORLEANS
Louisiana Tumor Registry

Lauren S. Maniscalco, MPH | lspezia@lsuhsc.edu | LTR, Registry Liaison
Carla B. Rosales, MPH | crosa@lsuhsc.edu | LTR, Research Epidemiologist
To access *Cancer Incidence in Louisiana by Census Tract*, please visit:
<https://sph.lsuhs.edu/louisiana-tumor-registry/>

RESULTS

Overall, 244,907 Louisiana cancer cases were included in the analysis for diagnosis years 2005-2015, with 2.5% of cases having low census tract certainty. The mean age of the low census tract certainty group was 62.4, which was significantly lower than the mean age (64.3) of the high census tract certainty group ($P < 0.0001$).

Compared to the patients with high census tract certainty, the patients with low census tract certainty were more likely to be male (58.3% vs. 53.5%, $P < 0.0001$), black (33.1% vs. 27.7%, $P < 0.0001$), younger than 69 years (68.1% vs. 62.6%, $P < 0.0001$), living in high poverty census tracts (39.5% vs. 37.7%, $P = 0.0056$) or living in rural areas (35.4% vs. 17.5%, $P < 0.0001$). With respect to stage at diagnosis, the patients with low census tract certainty were more likely to have localized disease (54.3% vs. 48.8%, $P < 0.0001$) (Table 1).

Table 1. Demographic and tumor characteristics between cancer cases with high vs. low census tract certainty

Characteristics	All		High Census Tract Certainty		Low Census Tract Certainty		P-Value ¹
	No.	%	No.	%	No.	%	
Total	244,907		238,835	97.5	6,072	2.5	
Sex							<0.0001
Male	131,350	53.6	127,809	53.5	3,541	58.3	
Female	113,546	46.4	111,015	46.5	2,531	41.7	
Race							<0.0001
White	174,208	71.2	170,200	71.3	4,008	66.1	
Black	68,184	27.9	66,177	27.7	2,007	33.1	
Other	2,314	1.0	2,267	1.0	47	0.8	
Age Group							<0.0001
≤69 years of age	153,700	62.8	149,565	62.6	4,135	68.1	
>69 years of age	91,207	37.2	89,270	37.4	1,937	31.9	
Census Tract Poverty ²							0.0056
High	92,472	37.8	90,101	37.7	2,371	39.5	
Low	152,339	62.2	148,705	62.3	3,634	60.5	
Urban vs. Rural ³							<0.0001
Urban	200,995	82.1	197,071	82.5	3,924	64.6	
Rural	43,912	17.9	41,764	17.5	2,148	35.4	
Stage ⁴							<0.0001
Localized	109,282	48.9	106,265	48.8	3,017	54.3	
Regional	50,837	22.8	49,580	22.8	1,257	22.6	
Distant	49,426	22.1	48,430	22.2	996	17.9	
Unknown/Unstaged	13,743	6.2	13,461	6.2	282	5.1	

¹ Significance of categorical variables assessed using the Chi-Square test.

² High poverty was defined as at least 20% of persons with an income below the federal-defined poverty level.

³ Urban and rural areas were defined using data from the U.S. Department of Agriculture at the parish (county) level.

⁴ Derived Summary Stage 2000; leukemias and lymphomas were excluded.

Among the top five cancers in Louisiana, the proportion of cases with low census tract certainty was highest for colorectal cancer (4.7%) and lowest for breast cancer (2.0%) (Figure 1). Of all cancer cases within each region of the state, the proportion of cases with low census tract certainty was highest in the Central Louisiana region (4.9%) and lowest in the New Orleans region (0.9%) (Figure 2).

Figure 1. Percent of Total Cases with Low Census Tract Certainty for the Top 5 Cancers in Louisiana

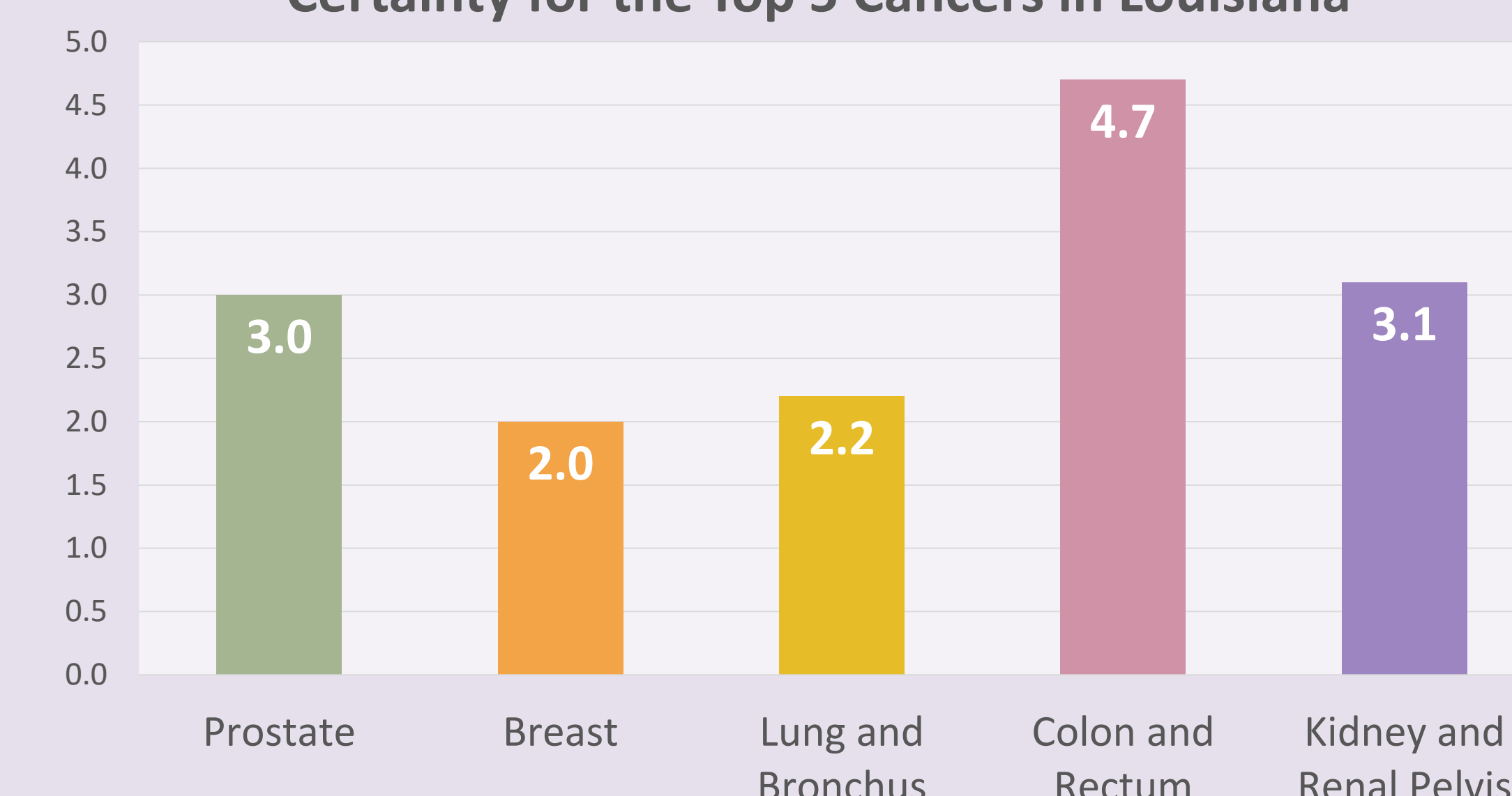
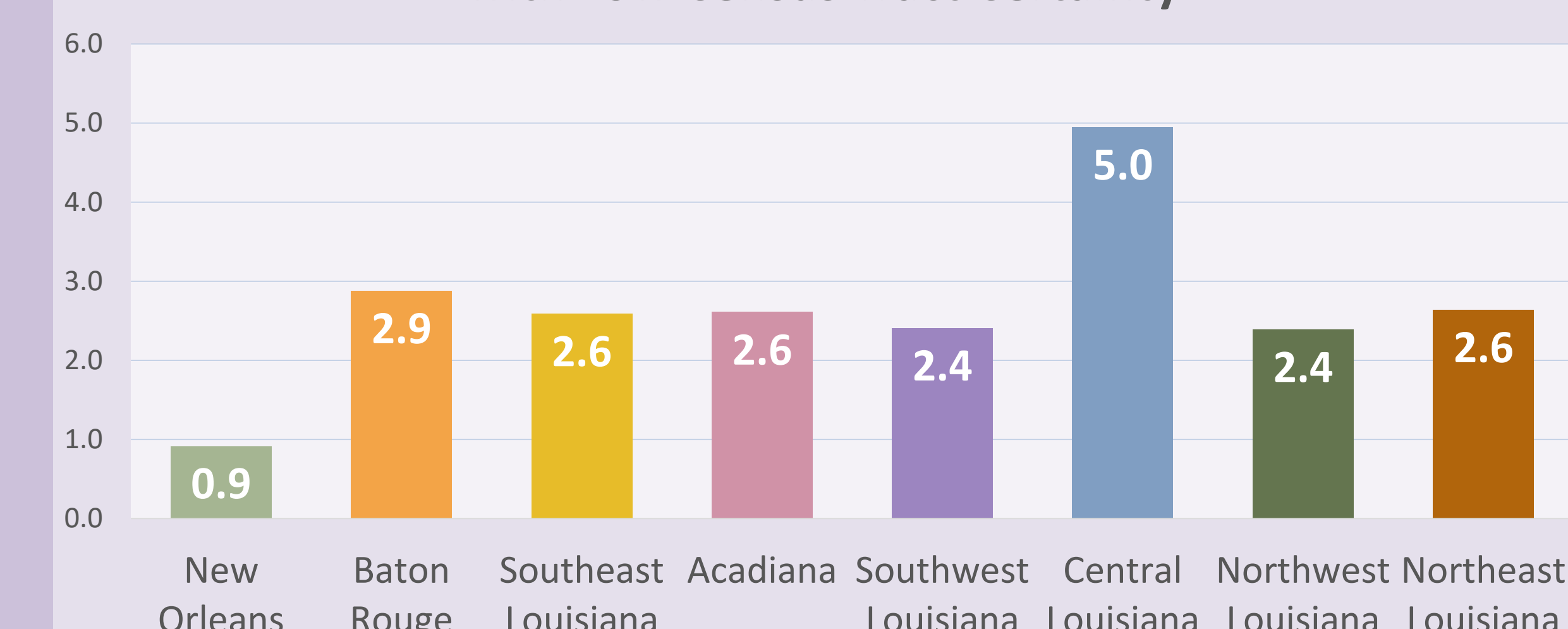


Figure 2. Percent of Total Cases in Louisiana Regions with Low Census Tract Certainty



CONCLUSIONS

Despite observing significant differences in sociodemographic characteristics and stage at diagnosis between cases with high versus low census tract certainty, the overall proportion of cases with low geographic certainty was less than 3%. The large sample size could result in highly statistically significant differences, although the change in the frequency of the characteristics between the two groups was still small.

The impact of selection bias may vary depending upon the percentage of excluded cases out of the total cases for individual census tracts. For census tracts with a large number of cases, excluding a few cases will not affect the results; for census tracts with a small number of cases, excluding a couple of cases will result in a large bias. However, given our publication criteria of at least 16 cases for each census tract included in the report, those census tracts with very few cases that would be highly sensitive to selection bias are likely already suppressed from publication.

ACKNOWLEDGEMENTS

The authors acknowledge the support of the LSUHSC School of Public Health, as well as the LTR's funding agencies: NCI-SEER (HHSN2612018000071/HHSN26100002), CDC-NPCR (NU58DP00632), and the State of Louisiana.