Identifying Factors Associated with Loss to Follow-up among Patients Reported to the New York State Cancer Registry

Baozhen Qiao[1], Maria J. Schymura[1,2], Xiuling Zhang[1], April A. Austin[1], Colleen G. Sherman[1]

[1] New York State Cancer Registry, New York State Department of Health, Albany, NY, United States; [2] Department of Epidemiology and Biostatistics, University at Albany School of Public Health, Rensselaer, NY, United States

INTRODUCTION

State cancer registries in the United States are resources for estimating population-based cancer survival. However, the completeness of patient follow-up might affect the accuracy of survival estimates. As one of the Surveillance, Epidemiology and End Results (SEER) Program registries, the New York State Cancer Registry (NYSCR) is required to meet certain patient follow-up standards. Like many registries, the NYSCR conducts patient follow-up largely through linkages with other data sources, including state vital records, National Death Index (NDI), Social Security Administration (SSA), Medicaid, and Statewide Planning and Research Cooperative System (SPARCS) administrative files. However, even after expending great effort on linkages, a small proportion of patients remain lost to follow-up (LTFU). In this study, we intended to identify factors that are associated with the likelihood of LTFU in the NYSCR.

MATERIALS AND METHODS

- First primary cancers (sequence number = 00 or 01, excluding DCO and autopsy cases) diagnosed during 2000-2018 among New York State residents were selected for study. All patients were followed through December 31, 2018, which is the follow-up date that was set and evaluated by the SEER program for the November 2021 data submission.
- Based on patients' vital status and date of last contact, follow-up status was categorized into two groups: patients not LTFU - including those deceased or those alive with a date of last contact of 12/31/2018 or after, and patients LTFU - including those not known to be deceased (referred to as "alive") with a date of last contact prior to 12/31/2018.
- The number and percentage of patients LTFU were examined by demographic and tumor characteristics among all patients and alive patients, respectively.
- For patients who were lost to follow-up, the timing of LTFU (within 1 year, 1-5, 5-10, or >10 years after cancer diagnosis) was further examined by demographic and tumor characteristics.
- Multivariate logistic regression analyses were performed to evaluate associations between demographic/tumor characteristics and likelihood of LTFU.
- Since 5-year survival rates have been commonly used for measuring disease prognosis among cancer patients, we also conducted secondary analyses to evaluate rates of LTFU within 5 years after cancer diagnosis. These analyses were restricted to patients diagnosed during 2000-2013 to ensure that all patients had at least 5 years of follow-up.

RESULTS

- Among 1,797,228 patients diagnosed during 2000-2018, 74,722 were lost to follow-up prior to December 31, 2018, representing 4.2% of all patients and 7.6% of alive patients. Among 1,304,137 patients diagnosed during 2000-2013, 37,714 were lost to follow-up within 5 years after cancer diagnosis, representing 2.9% of all patients and 4.7% of alive patients.

RESULTS

- The number and percentage of patients LTFU by demographic and tumor characteristics are shown in Table 1.
- About 60.3% of LTFU occurred within 1 year after cancer diagnosis (Figure 1). The percentage of LTFU within 1 year after cancer diagnosis was particularly higher for patients who were foreign born or uninsured, for cases that were reported by laboratories or physician offices, and for individuals with only one primary.

Table 1. Demographic/Tumor Characteristics of Patients LTFU prior to December 31, 2018 or within 5 Years after Cancer Diagnosis

Demographic/Tumor Characteristics	Patients LTFU prior to December 31, 2018[1]			Patients LTFU within 5 Years after Cancer Diagnosis[2]		
	Count	% of All Patients (n=1,797,228)	% of Alive Patients (n=989,924)	Count	% of All Patients (n=1,304,137)	% of Alive Patients (n=799,687)
Total	74,722	4.2	7.6	37,714	2.9	4.7
Gender						
Male	33,024	3.8	7.3	15,932	2.5	4.2
Female	41,698	4.5	7.8	21,782	3.3	5.2
Age						
<20	1,598	8.8	10.4	725	5.5	3.5
20-64	44,813	5.1	7.3	25,899	4.1	5.5
65+	28,311	3.1	7.9	11,090	1.7	6.5
Race						
White	46,395	3.2	5.9	22,447	2.1	3.5
Black	13,115	5.1	9.7	7,278	4.0	6.9
American Indian/Alaska Native	80	3.0	5.1	43	2.4	3.9
Asian/Pacific Islander	10,115	12.3	18.4	5,344	10.2	15.1
Unknown	5,017	28.0	30.9	2,602	27.1	30.0
Ethnicity						
Non-Hispanics	61,808	3.8	6.9	30,259	2.5	4.2
Hispanics	12,914	7.8	13.0	7,455	6.6	10.2
Birth Country						
U.S. Born	14,628	1.4	3.5	6,275	0.8	1.6
Foreign Born	22,813	7.9	15.8	14,231	6.6	11.5
Unknown	37,281	7.9	8.8	17,208	5.8	6.3
NYS Region						
NYC	43,350	6.5	11.7	23,658	4.9	8.0
NYS Excl NYC	31,372	2.8	5.1	14,056	1.7	2.8
Poverty Level						
0%-<5%	16,520	3.5	6.0	7,911	2.2	3.4
5%-<10%	17,562	3.7	6.6	8,310	2.4	3.9
10%-<20%	20,362	4.3	7.9	10,093	3.0	5.1
20%-100%	19,839	5.4	10.5	11,094	4.2	7.4
Unknown	439	8.3	16.5	306	6.7	10.7
Rural Urban						
Metropolitan Counties	72,099	4.4	7.9	36,694	3.1	5.0
Non-metropolitan Counties	2,623	1.8	3.5	1,020	0.9	1.6
Insurance Status						
Insured	32,960	2.8	4.8	12,303	1.6	2.5
Any Medicaid	15,448	7.0	12.9	7,406	5.4	10.1
Uninsured	4,005	18.6	31.7	2,791	16.5	26.5
Unknown	22,309	5.6	13.4	15,214	4.1	7.1
Year of Diagnosis						
2000-2004	15,446	3.5	9.9	13,352	3.0	5.2
2005-2009	16,721	3.5	7.3	11,142	2.4	3.8
2010-2014	20,895	4.3	7.1	13,220	3.4	5.3
2015-2018	21,660	5.5	7.0	-	-	-
Tumor Behavior						
In Situ	7,883	6.3	7.4	3,681	4.3	4.6
Invasive	66,839	4.0	7.6	34,033	2.8	4.7
Stage						
In Situ/Local	43,822	4.9	6.7	21,479	3.4	4.0
Regional	12,328	3.6	7.0	6,857	2.8	4.9
Distant	9,184	2.3	8.7	4,972	1.7	6.5
Unknown	9,388	5.8	16.6	4,406	3.4	8.4
Type of Reporting Source						
Hospital Inpatient	34,430	3.3	7.6	20,078	2.5	4.9
Radiation Tx or Medical Oncology Center	7,058	4.1	5.5	2,913	2.4	3.0
Laboratory Only	5,252	14.1	20.0	2,262	8.4	10.5
Physician	5,681	11.1	14.2	1,569	5.3	6.8
Other Hospital Outpatient/Surgery Center	22,301	4.6	6.6	10,892	3.3	4.4
Diagnostic Confirmation						
Microscopically Confirmed	72,677	4.3	7.5	36,516	3.0	4.7
Clinical Confirmed	1,674	2.2	11.8	954	1.7	8.8
Unknown	371	2.3	13.3	244	1.7	6.3
Sequence Number						
Only One Primary	71,292	4.6	8.3	36,557	3.3	5.7
First of Multiple Primaries	3,430	1.4	2.5	1,157	0.6	0.7
Cancer Site Group[3]						
Cancers with Best Survival	24,522	5.9	7.8	10,852	3.6	4.2
Cancers with Good Survival	26,282	4.5	6.6	14,150	3.4	4.4
Cancers with Bad Survival	16,997	3.7	8.3	8,719	2.6	5.0
Cancers with Worst Survival	6,921	2.0	9.8	3,993	1.6	8.3

Footnotes for Tables 1: 1. Cancer cases were diagnosed during 2000-2018; 2. Cancer cases were diagnosed during 2000-2013; 3. Cancers were grouped based on survival rates. Cancers with Best Survival include prostate, testis, thyroid and melanoma of the skin cancers; Cancers with Good Survival include breast, cervix, uterus, bladder, kidney cancers and lymphomas; Cancers with Bad Survival include oral cavity, colorectal, larynx, ovary, myeloma, leukemias and others not listed here; Cancers with Worst Survival include brain/CNS, esophagus, stomach, lung and bronchus, liver, and pancreas cancers.

Footnotes for Figure 2: 1. Patients with unknown values for race, birth country, insurance status, stage, or diagnostic confirmation were included in the multivariate logistic regression models, but the results for those unknown categories are not shown in the Figure 2; 2. Reference groups used in the multivariate logistic regression models are: male for gender; white for race; non-Hispanic for ethnicity; U.S Born for birth country; NYC for NYS region; 0%-<5% for poverty level; metropolitan counties for rural urban; insured for insurance status; 2000-2004 for year of diagnosis; in Situ for tumor behavior; in Situ/local for stage; hospital inpatient for type of reporting source; microscopically confirmed for diagnostic confirmation; only one primary for sequence number; cancers with best survival for cancer site group.

RESULTS (cont.)

Figure 1. Frequency Distribution of Timing of Patients LTFU by Demographic/Tumor Characteristics

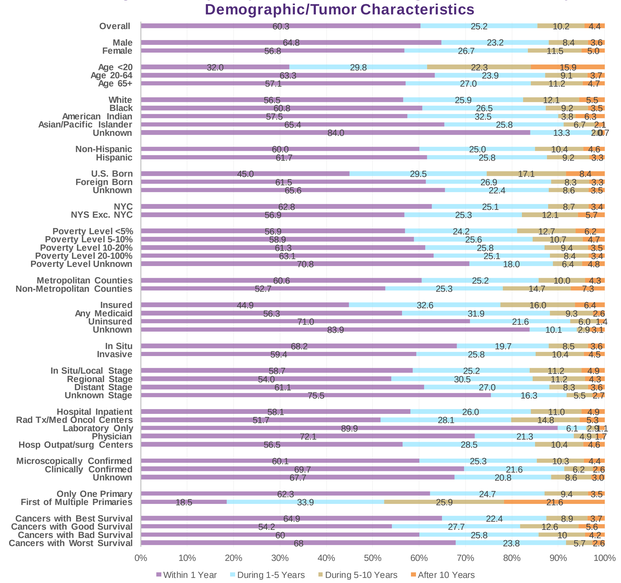
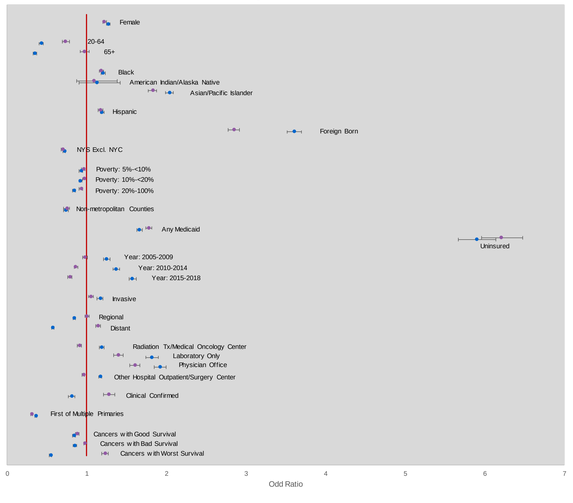


Figure 2. Multivariate Logistic Regression Results – Odd Ratios and 95% Confidence Intervals of LTFU among All Patients and among Alive Patients, Diagnosis Year of 2000-2018, Followed Through December 31, 2018



RESULTS (cont.)

- Multivariate logistic regression analyses indicated that all factors evaluated were statistically significantly associated with the likelihood of LTFU. Notably, patients who were female, black or Asian/Pacific Islander, Hispanic, foreign born, insured by Medicaid or uninsured, aged <20 years at cancer diagnosis, or living in NYC or metropolitan counties were more likely to be LTFU compared to their counterparts. In addition, cases reported by laboratories and physician offices also had a higher likelihood of LTFU compared to those reported by hospitals. Odd ratios and 95% confidence intervals are displayed in Figure 2 for patients followed through December 31, 2018. Similar findings were observed when evaluating 5-year LTFU except that among alive patients, 5-year LTFU rate was lower among patients < 20 years of age at diagnosis compared to older patients; in addition, the effect of type of reporting source on 5-year LTFU seemed to be relatively small (results not shown).

DISCUSSION AND CONCLUSION

- The current study found that LTFU was not random, rather certain patient groups have higher LTFU rates than others. The findings can be partially explained by varying data completeness and quality among different patient groups. For instance, higher percentages of missing or inaccurate social security numbers among children, changes in surnames after marriage among women, and different naming conventions among Asians and Hispanics can make record linkages more challenging; out-migration to the country of birth after a cancer diagnosis among the foreign born and lack of medical care among the uninsured can result in an absence of patient records in data sources that are commonly used for patient follow-up. In addition, cancer cases reported by laboratories and physician offices are more likely to have missing demographic information including social security numbers.
- Although some of the factors associated with LTFU are not amenable to change by cancer registries, more intense follow-up of cases reported by laboratories and physician offices may be a possible, albeit resource-intensive solution. Linkage to LexisNexis to obtain social security numbers and additional address information that may facilitate linkage to follow-up sources could also be helpful.
- How non-random LTFU affects survival estimates will be investigated in future studies.

AKNOWLEGMENTS

This project was funded in part by the Centers for Disease Control and Prevention's (CDC) National Program of Cancer Registries through cooperative agreement 6NU58DP006309 awarded to the New York State Department of Health and by the National Cancer Institute, National Institutes of Health, Department of Health and Human Services, under Contract 75N91018D00005. The contents are solely the responsibility of the New York State Department of Health and do not necessarily represent the official views of the CDC or NIH.