

# Missing Stage Information for Prostate Cancer Cases – Too Much Reliance on Collaborative Stage?

Maria J. Schymura, Baozhen Qiao, Amy R. Kahn, New York State Cancer Registry, Albany, NY

## Introduction

**Background:** The NAACCR Data Assessment Work Group is evaluating the quality of various data items submitted in response to NAACCR's *Call for Data*, including derived summary stage (dSS). Although we in the New York State Cancer Registry (NYSCR) have long been aware of the high proportion of our cases missing stage information, it is through our involvement in this NAACCR work group that we realized that ours ranks highest among U.S. registries in the percent of prostate cases with unknown dSS. We hypothesized that the high percentage of unknown dSS was primarily due to the high proportion of prostate cases reported to the NYSCR by non-hospital sources, sources which we do not require to use collaborative stage (CS) for reporting.

**Objectives:** Our objectives were to assess the reasons for this unfortunate distinction and to determine what, if anything, we could do to lower the percent of unknown stage for prostate cancer cases in the NYSCR.

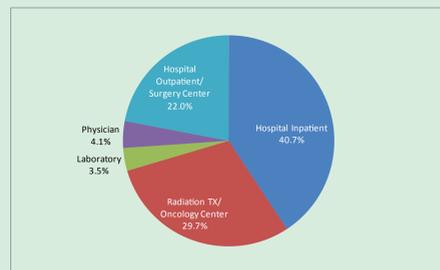
## Methods

Prostate cancer cases diagnosed from 2004 to 2008 among NY State (NYS) residents, excluding DCO and autopsy cases, were selected for analysis. We examined the "type of reporting source" (NAACCR item 500) distribution and calculated the percent missing dSS for each type of reporting source. Although the coding for this NAACCR item changed effective with 2006 diagnoses, we were able to apply the new codes to data for the entire period because of the granularity with which we collected the data. For prostate cancer cases with unknown dSS, we examined individual CS data elements required for deriving dSS to determine whether the percent unknown was attributable to a specific CS data item. For these cases, we also evaluated whether any other stage information was present in the records, i.e. directly coded summary stage (SS) or TNM stage [Note: the NYSCR requires the reporting of SS from hospital sources]. Experienced coders reviewed the pathology reports for all 2008 laboratory-only prostate cancer cases with unknown dSS to ascertain whether they included sufficient data for deriving summary stage. Finally, we examined NYSCR and SEER data for 2006 to 2008 by type of reporting source to evaluate source-type differences in prostate cancer reporting to the NYSCR and to SEER registries.

## Results

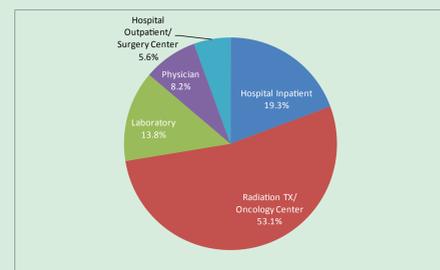
- Of the 76,443 NYS prostate cancer cases diagnosed from 2004 to 2008 (excluding DCO and autopsy cases), 17.6% had unknown dSS.
- Based on type of reporting source, which is assigned using a hierarchy based on likelihood of complete information, only 40.7% of prostate cancer cases were hospital inpatients; 29.7% were reported by radiation treatment centers, 22.0% by outpatient sources, 4.1% by physicians, and 3.5% by laboratories.

Figure 1. Distribution of reporting sources for prostate cancer, 2004-2008



- 53.1% of prostate cancer cases with unknown dSS were reported by radiation treatment centers; 19.3% were hospital inpatients, 13.8% were laboratory-only cases; 8.2% were reported by physicians and 5.6% by outpatient sources.

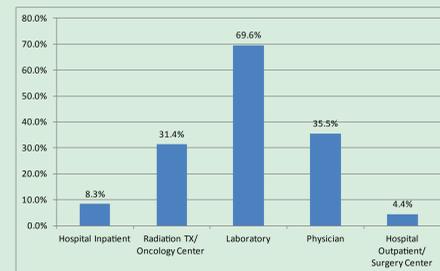
Figure 2. Distribution of reporting sources for prostate cancer with missing dSS, 2004-2008



## Results - continued

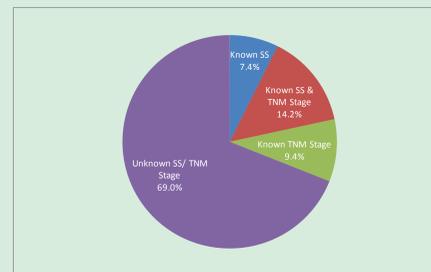
- The percent unknown dSS varied by type of reporting source. It was 8.3% for hospitals; 31.4% for radiation treatment centers; 4.4% for outpatient sources; 35.5% for physicians; and 69.6% for laboratories.

Figure 3. Percent of unknown dSS for prostate cancer cases, 2004-2008, by reporting source



- Of the hospital inpatient cases reported with unknown dSS, 69.2% were reported solely as non-analytic cases.
- Laboratory-only cases with known dSS tended to be submitted by hospitals. A review of all pathology reports for 2008 diagnosed prostate cancer cases missing dSS and reported as laboratory-only cases (n=533) found that only 14 cases (2.6%) contained enough information to compute dSS.
- Unknown dSS was not attributable to any one CS data item required to compute dSS. When dSS was unknown, all component CS data items tended to be unknown as well.
- For the prostate cancer cases missing dSS, directly coded summary stage (SS) was available for 21.6% of cases and TNM stage for 23.6%. In total, 31.0% of prostate cancer cases missing dSS had some usable stage information, thus reducing the percent of cases missing any stage information to 12.1%, which is still fairly high but better than 17.6%.

Figure 4. Percent of prostate cancer cases missing dSS for which other stage information was available.



- Prostate cancer cases missing all stage information were considerably older than cases with any stage information (mean age=73.3 [interquartile range 66 to 80] vs. mean age= 66.3 [interquartile range 59 to 73]).
- The type of reporting source distribution for prostate cancer cases in the NYSCR differs markedly from the distribution in all SEER registries. Compared to prostate cancer cases reported to the NYSCR, prostate cancer cases reported to the SEER registries are much more likely to have been hospital inpatients and much less likely to have been reported exclusively by radiation treatment centers or by outpatient sources. The percent of cases reported by physicians in NY and the percent of cases reported by laboratories in NY are within the range observed for the SEER registries.

Table 1. Type of reporting source by registry, prostate cancer, 2006-2008 (excluding DCO & autopsy cases)

Registry	Hospital Inpatient		Radiation TX / Oncology Center		Laboratory		Physician Office		Hospital Outpatient/ Surgery Center		Total Count
	Count	%	Count	%	Count	%	Count	%	Count	%	
San Francisco-Oakland	8,928	94.2%	10	0.1%	95	1.0%	444	4.7%	0	0.0%	9,477
Connecticut	7,706	88.0%	255	2.9%	461	5.3%	230	2.6%	93	1.1%	8,760
Detroit (Metropolitan)	9,409	88.2%	3	0.0%	852	8.0%	403	3.8%	0	0.0%	10,667
Hawaii	1,822	70.3%	351	13.6%	263	10.2%	111	4.3%	44	1.7%	2,591
Iowa	5,277	80.7%	363	5.6%	676	10.3%	177	2.7%	44	0.7%	6,538
New Mexico	2,452	63.7%	64	1.7%	32	0.8%	1,303	33.8%	0	0.0%	3,851
Seattle (Puget Sound)	8,130	78.7%	231	2.2%	926	9.0%	101	1.0%	939	9.1%	10,327
Utah	3,811	79.0%	7	0.2%	984	20.4%	23	0.5%	0	0.0%	4,825
Atlanta (Metropolitan)	5,103	76.2%	675	10.1%	602	9.0%	215	3.2%	98	1.5%	6,695
San Jose-Monterey	4,371	90.9%	5	0.1%	7	0.2%	424	8.8%	0	0.0%	4,807
Los Angeles	12,505	81.2%	19	0.1%	80	0.5%	2,763	17.9%	0	0.0%	15,400
California excluding SF/SJM/LA	31,823	87.4%	550	1.5%	506	1.4%	3,520	9.7%	7	0.0%	36,421
Kentucky	7,574	85.6%	0	0.0%	717	8.1%	196	2.2%	321	3.6%	8,848
Louisiana	8,413	81.2%	154	1.5%	298	2.9%	1,466	14.2%	30	0.3%	10,364
New Jersey	16,744	75.4%	879	4.0%	1,683	7.6%	2,652	11.9%	261	1.2%	22,221
New York	18,987	39.4%	15,280	31.7%	1,639	3.4%	2,268	4.7%	10,033	20.8%	48,207

1: Alaska Natives and Rural Georgia registries not shown due to small numbers.  
2: Nursing/Convalescent Home/Hospice category not shown but included in total count.

## Results - continued

- The percent of unknown dSS by reporting source in the NYSCR also differs markedly from what is seen in the SEER registries. It is much higher in the NYSCR than in the SEER registries for hospital inpatient cases and for cases reported by radiation treatment centers or by laboratories. The percent unknown dSS for cases reported by physicians or by outpatient sources in NY falls into the range observed for SEER registries.

Table 2. Percent of unknown stage by reporting source, prostate cancer, 2006-2008 (excluding DCO & autopsy cases)

Registry	% Unk dSS	Hospital Inpatient		Radiation TX / Oncology Center		Laboratory		Physician Office		Hospital Outpatient/ Surgery Center	
		Total Count	% Unk dSS	Total Count	% Unk dSS	Total Count	% Unk dSS	Total Count	% Unk dSS	Total Count	% Unk dSS
San Francisco-Oakland	1.31%	8,928	1.11%	10	0.00%	95	4.21%	444	4.73%	0	0.00%
Connecticut	2.28%	7,706	1.05%	255	4.31%	461	12.58%	230	15.65%	93	5.38%
Detroit (Metropolitan)	0.67%	9,409	0.63%	3	0.00%	852	1.29%	403	0.50%	0	0.00%
Hawaii	6.87%	1,822	2.03%	351	0.57%	263	45.63%	111	17.12%	44	0.00%
Iowa	1.94%	5,277	0.99%	363	0.55%	676	8.73%	177	7.34%	44	2.27%
New Mexico	3.43%	2,452	3.22%	64	0.00%	32	28.13%	1,303	3.38%	0	0.00%
Seattle (Puget Sound)	1.14%	8,130	0.87%	231	0.43%	926	0.97%	101	9.90%	939	2.88%
Utah	1.04%	3,811	0.89%	7	0.00%	984	0.30%	23	56.52%	0	0.00%
Atlanta (Metropolitan)	1.09%	5,103	0.92%	675	0.00%	602	3.82%	215	0.47%	98	1.02%
San Jose-Monterey	1.29%	4,371	1.24%	5	0.00%	7	57.14%	424	0.94%	0	0.00%
Los Angeles	9.31%	12,505	1.90%	19	0.00%	80	48.75%	2,763	40.86%	0	0.00%
California excluding SF/SJM/LA	4.33%	31,823	1.84%	550	0.36%	506	34.58%	3,520	22.76%	7	0.00%
Kentucky	2.42%	7,574	1.86%	0	0.00%	717	0.56%	196	12.76%	321	1.25%
Louisiana	1.01%	8,413	0.87%	154	1.30%	298	0.34%	1,466	1.64%	30	6.67%
New Jersey	4.50%	16,744	3.43%	879	1.37%	1,683	12.77%	2,652	7.35%	261	0.77%
New York	16.95%	18,987	6.21%	15,280	31.18%	1,639	71.20%	2,268	28.92%	10,033	4.01%

1: Alaska Natives and Rural Georgia registries not shown due to small numbers.  
2: Nursing/Convalescent Home/Hospice category not shown but included in total count.

## Discussion

Although stage information for prostate cancer cases in the NYSCR can be improved by not relying solely on CS, the percent unknown is likely to remain high. We have already taken steps to address the high proportion of unknown stage attributable to radiation treatment centers by developing customized training. However, a survey of these centers revealed that staging information is frequently not available in their records and reporting is not likely to be conducted by a CTR (Houser et al, NAACCR 2012). In terms of hospital reporting, only approximately a third of NY hospitals are CoC accredited; these account for approximately 60% of analytic reports. Additionally, a substantial proportion of inpatient prostate cancer reports are non-analytic. In our effort to capture all cases we may be sacrificing data-item completeness. We are confident that our prostate cancer completeness is fairly high; NY prostate cancer rates rank in the 78<sup>th</sup>, 91<sup>st</sup>, and 93<sup>rd</sup> percentile for all races combined, blacks, and Hispanics respectively (NAACCR Call for Data 2011, Data Submission Summary Quality Report).

Our evaluation of source-type differences in prostate cancer reporting to the NYSCR and to SEER registries suggests some non-standard coding of the "type of reporting source" variable. It is highly unlikely that 6 of 15 SEER registries have no cases for which the best reporting source should have been "hospital outpatient/surgery center cases" and that all four California registries have virtually no "radiation treatment center" cases. That limitation aside, one finding was striking; namely that the percent unknown stage is extremely low in laboratory-only reported cases among SEER registries. Based on the information we found in our laboratory reports, we find this puzzling.

The cost per case is relatively low in NY, which comes at a price (i.e., lower data quality). In order for us to substantially improve the percent of cases missing stage, we would have to abstract the cases ourselves.

## Conclusions

In order to obtain more complete stage information for prostate cancer, all stage information, not only CS, should be considered and consolidated into a composite stage variable. To further improve the completeness of stage data for prostate cancer would require more active work on the part of the NYSCR and consequently more resources.

## Acknowledgement

This work was funded in part by CDC Cooperative Agreement 5U58DP000783 awarded to the New York State Department of Health.

New York State Cancer Registry (NYSCR) have long been aware of the high proportion of our cases missing dSS information, it is through our involvement in this NAACCR work group that we realized that ours ranks high among U.S. registries in the percent of prostate cases with unknown dSS. We hypothesized that the high percentage of unknown dSS was primarily due to the high proportion of prostate cases reported to the non-hospital sources, sources which we do not require to use collaborative stage (CS) for reporting.

**Objectives:** Our objectives were to assess the reasons for this unfortunate distinction and to determine anything, we could do to lower the percent of unknown stage for prostate cancer cases in the NYSCR.

## Methods

---

Prostate cancer cases diagnosed from 2004 to 2008 among NY State (NYS) residents, excluding DC and autopsy cases, were selected for analysis. We examined the “type of reporting source” (NAACCR item 10) distribution and calculated the percent missing dSS for each type of reporting source. Although the NAACCR item changed effective with 2006 diagnoses, we were able to apply the new codes to data from the entire period because of the granularity with which we collected the data. For prostate cancer cases with unknown dSS, we examined individual CS data elements required for deriving dSS to determine whether the percent