



Industry and Occupation Data from Registry and Death Certificates Celaya MO¹, Armenti KR², Riddle BL¹, Calvert GM³, Hosain GMW⁴, Rees JR¹







¹New Hampshire State Cancer Registry, Department of Epidemiology, Geisel School of Medicine at Dartmouth, Hanover, NH; ²NH Occupational Health Surveillance Program, Institute on Disability, University of New Hampshire, Durham, NH; ³Division of Surveillance, Hazard Evaluations and Field Studies, National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, Cincinnati, Ohio; ⁴Bureau of Public Health Statistics and Informatics, Department of Health and Human Services, New Hampshire Division of Public Health Services, Concord, NH.

ABSTRACT

BACKGROUND: Capturing data on industry and occupation (I&O) for cancer registration depends on the proper completion of several steps. First, the patient who is registered at the hospital or other facility must be asked about their usual I&O. Medical providers or clerical staff may instead ask about current occupation, or use another method that does not provide ideal information. Next, the registrar must find I&O data in the medical record, resolve any discrepancies between multiple entries, and interpret as needed. Completing death certificates involves a different process through questioning relatives or for deaths in hospital, through medical records. Because of the different processes in obtaining data, it is reasonable to expect differences between occupation and industry data in the cancer registry and in death certificates.

PURPOSE: To investigate whether the I&O data on death certificates and in the cancer registry are comparable, and whether death certificates could be used to supplement missing registry data.

METHODS: In this exercise, we will compare the data obtained from these two sources in subsets of New Hampshire residents with cancer. The subsets will include (i) common jobs such as housewife/homemaker, teacher, healthcare worker, etc., (ii) uncommon jobs, and (iii) patients in the registry with missing I&O data.

RESULTS: We will evaluate concordance and discrepancies, and examine the utility of death certificates for completing data on patients whose I&O data are missing in the

CONCLUSION: We will discuss potential biases that may result because death certificate I&O data can only be used for patients who have died, whereas missing I&O for those who are alive cannot be resolved.

BACKGROUND

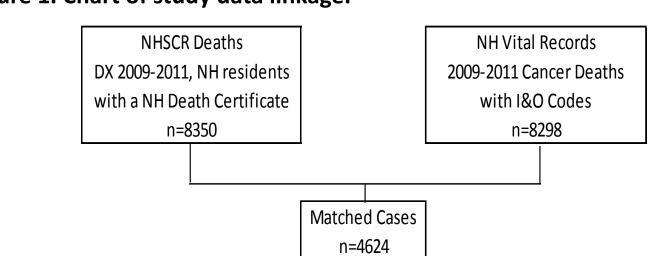
- There is evidence to suggest that a variety of cancers may result from occupationrelated risk factors and exposures.
- Two major public health data sources have been utilized for cancer surveillance and
- Death certificates have been better utilized for cancer surveillance of workers. Statewide cancer registries are designed to capture all incident cancer cases.
- The New Hampshire State Cancer Registry (NHSCR) is participating in a project with NIOSH (National Institute for Occupational Safety and Health) that aims to develop and maintain a multi-state population-based occupational cancer surveillance system utilizing cancer registry data that will be used to identify occupational cancer
- Industry and Occupation (I&O) are required reportable text data items. The quality of I&O data in the cancer registry was evaluated and found to be of high quality.1
- In this study, we investigate whether the I&O data on death certificates and in the cancer registry are comparable, and whether death certificates could be used to supplement missing registry data.

METHODS

Data (Figure 1)

- Data came from two sources: from the NHSCR and from NH death certificates. • I&O data in death certificates had been coded by NIOSH (National Institute for Occupational Safety and Health) experts for deaths occurring through year 2011. We selected cases with date of death in years 2009-2011 (n=8298).
- Cancer cases diagnosed in years 2009-2011 were selected if they were expired and had linked to a NH death certificate (n=8350). We decided to exclude cases diagnosed before 2009 because of known differences in data quality.
- Some cases that were diagnosed during 2009-2011 did not link with a death certificate if the death occurred after 2011.
- A total of 4624 cases were eligible for medical record review. They fulfilled the criteria: (i) cancers diagnosed 2009-2011 and (ii) patient expired between 2009-2011.

Figure 1. Chart of study data linkage.



I&O Coding

- NIOSH experts assigned I&O codes using the Census 2000 Industrial and Occupational Classification Systems to each I&O reported on the death certificates.
- NIOSH experts also assigned I&O codes using the Census 2010 Industrial and Occupational Classification Systems to each I&O reported to the NHSCR.
- We converted the 2010 Census codes to the 2000 Census codes so that both data sets would be comparable.
- The Census 2000 industry codes were grouped into major sectors based on the North American Industry Classification System (NAICS).²
- Occupation codes were grouped into major occupation groups based on the Standard Occupational Classification (SOC) major groups.³

METHODS, CONT.

by out-of-state central registries.

Data Items Used in Analysis

- The data items *Industry* and *Occupation* are defined identically as on death certificates.⁴ They are defined as the usual occupation and usual industry (i.e. the kind of work and workplace during most of a patient's working life before the cancer diagnosis).⁵
- I&O data from NHSCR and death certificates were grouped into *Known* and Unknown categories.

"different" (but known in both sources), and "unknown".

- Patients classified as Retired for I&O were grouped with the Unknown.
- grouped with Known. • I&O codes from both sources were then compared and categorized into "same",

Volunteers and Students were grouped with Unpaid Workforce, which was

• NHSCR case reports were assigned a reporting source. Because hospitals with an in-house cancer registry have better access to patient demographic data, priority was given to registry hospitals, so that cases reported by a registry hospital, even if another source reported that same case, were grouped into the registry hospital category. The same was done for those reported by a nonregistry hospital, then a non-hospital source. Non-NH cases were only reported

<u>Analysis</u>

- First, we compared the availability of I&O in the NHSCR and death certificates. • We also looked at what was the highest number of I&O in the death certificates
- compared to what was reported to the NHSCR.
- We also compared I&O from both sources and determined if there was a difference in I&O availability between the NHSCR and death certificates.
- Descriptive analyses and tabulations for all variables were performed using the Statistical Package for the Social Services (SPSS) v22 for Windows (SPSS Inc; Chicago, IL, USA; www.spss.com)

RESULTS

Table 1. Characteristics of expired cancer cases in the New Hampshire State Cancer Registry (NHSCR) that linked with NH vital records death certificates by occupation and industry for New Hampshire residents diagnosed 2009-2011.

	Total ———	Occupation				Industry			
		NHSCR		Vital Records		NHSCR		Vital Records	
		% Known	% Unknown	% Known	% Unknown	% Known	% Unknown	% Known	% Unknown
Age at DX									
0-49	245	82.4	17.6	95.5	4.5	85.3	14.7	95.9	4.1
50-59	654	81.5	18.5	97.9	2.1	85.3	14.7	98.8	1.2
60-64	471	76.4	23.6	98.3	1.7	79.4	20.6	98.7	1.3
65-69	615	71.1	28.9	97.7	2.3	72.7	27.3	98.4	1.6
70-74	606	69.3	30.7	97.5	2.5	68.6	31.4	99.2	0.8
75-79	680	64.9	35.1	98.2	1.8	66.9	33.1	99.1	0.9
80+	1353	60.2	39.8	98.9	1.1	58.4	41.6	98.9	1.1
Race									
White	4563	69.5	30.5	98.1	1.9	70.3	29.7	98.7	1.3
Non-white	53	64.2	35.8	98.1	1.9	71.7	28.3	98.1	1.9
Unknown	8	25.0	75.0	100.0	0.0	25.0	75.0	100.0	0.0
Gender									
Male	2492	74.6	25.4	97.3	2.7	76.6	23.4	98.4	1.6
Female	2132	63.3	36.7	99.0	1.0	62.9	37.1	99.1	0.9
DX YEAR									
2009	2078	69.8	30.2	98.0	2.0	70.2	29.8	98.8	1.2
2010	1679	67.6	32.4	98.3	1.7	68.6	31.4	98.7	1.3
2011	867	71.7	28.3	97.8	2.2	73.8	26.2	98.4	1.6
Number of Reporting Sources									
Single source	3333	65.9	34.1	98.1	1.9	65.8	34.2	98.6	1.4
Multiple sources	1291	78.2	21.8	98.1	1.9	81.8	18.2	99.1	0.9
Type of Reporting Source*									
Registry	3902	71.5	28.5	98.0	2.0	73.4	26.6	98.7	1.3
Non-registry	178	36.5	63.5	97.2	2.8	38.2	61.8	97.2	2.8
Non-hospital	174	86.8	13.2	98.9	1.1	62.1	37.9	98.3	1.7
Non-NH	370	55.1	44.9	98.6	1.4	55.9	44.1	99.5	0.5
Stage at Diagnosis									
In situ	81	51.9	48.1	97.5	2.5	55.6	44.4	100.0	0.0
Local	716	67.0	33.0	97.2	2.8	67.3	32.7	98.2	1.8
Regional	835	73.4	26.6	98.0	2.0	71.5	28.5	98.8	1.2
Distant	2393	69.4	30.6	98.3	1.7	73.1	26.9	98.9	1.1
NA	27	63.0	37.0	100.0	0.0	63.0	37.0	100.0	0.0
Unknown	572	69.2	30.8	98.3	1.7	62.8	37.2	98.3	1.7
Total	4624	69.4	30.6	98.1	1.9	70.3	29.7	98.7	1.3

Table 1 presents information describing whether or not I&O was found in NHSCR case reports and in death certificates from NH Vital Records.

- NH Vital Records data were more complete than NHSCR data for I&O overall, only 2% of death certificates while 31% of NHSCR had Unknown Occupation, and 1%, 30% for Industry.
- Age groups 0-64 had the highest known I&O. This was expected given this is the majority of the working population.
- NHSCR had the best quality I&O data during diagnosis year 2011. As a Specialized Registry for Comparative Effectiveness Research, NHSCR collected more in-depth data for year 2011 than any other diagnosis year.
- I&O was more often available in NHSCR for cases reported by multiple sources than cases reported by a single source.
- Non-registry hospital data had a greater proportion of unknown I&O.

RESULTS, CONT.

Table 2. Occupations from NHSCR and NH death certificates

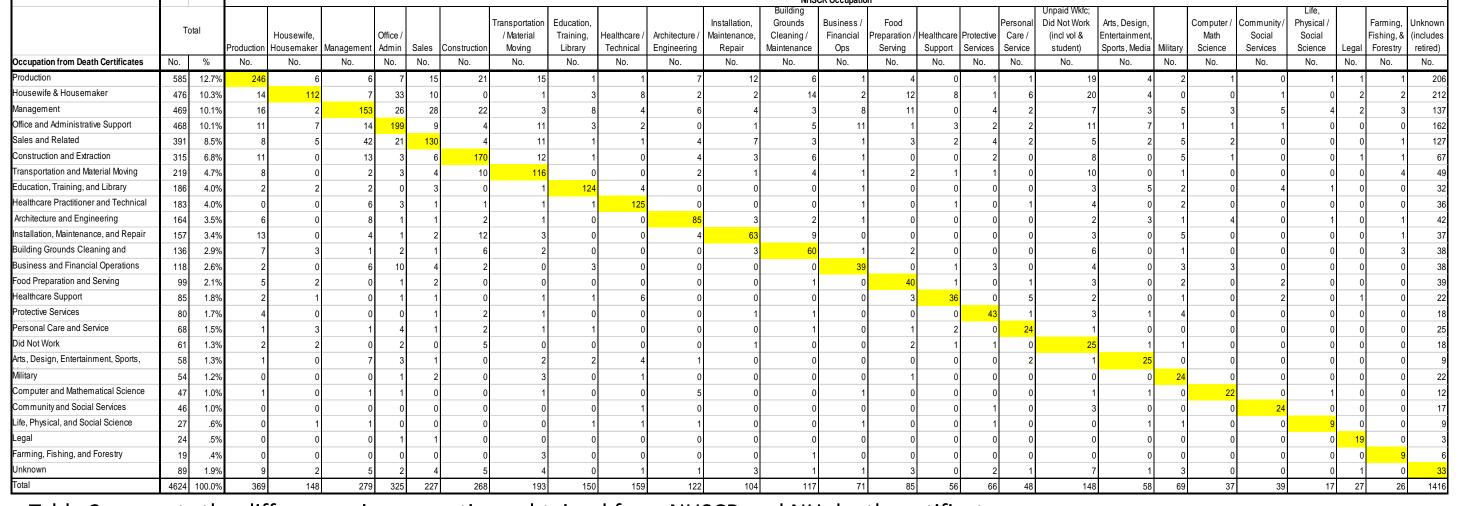


Table 2 presents the differences in occupations obtained from NHSCR and NH death certificates

- The most common occupation found in death certificates was Production. The majority of Unknown occupation in NHSCR records were found to be Production in the corresponding death certificate.
- The differences give an idea of differences in classification, e.g. where NHSCR has Production, 123 (33%) cases have something different in the death certificate.

Table 3. Industries from NHSCR and NH death certificates

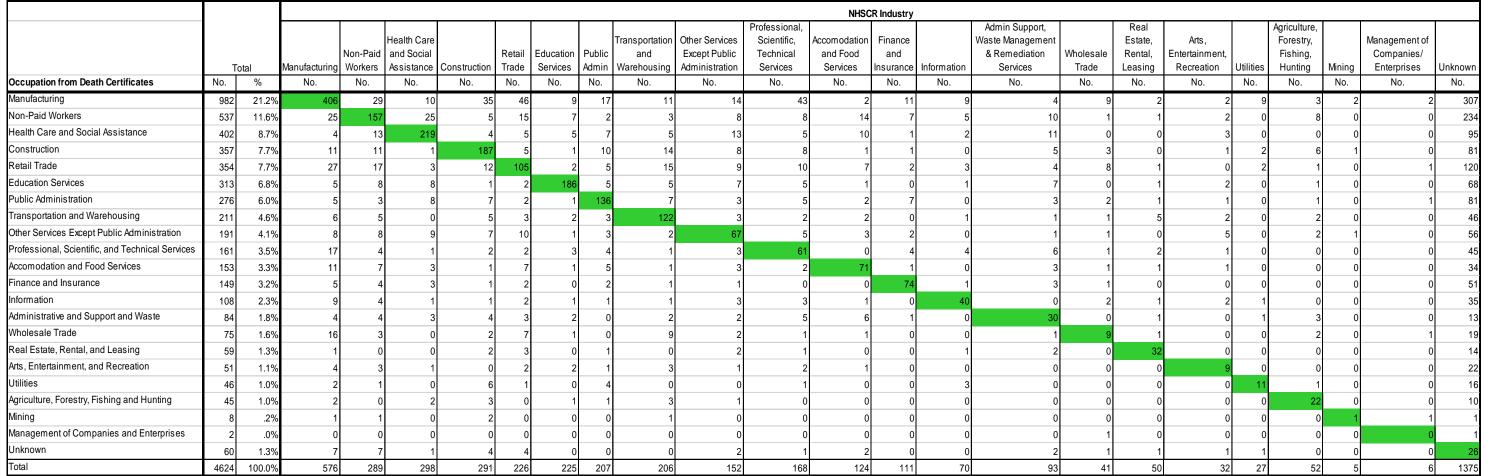


Table 3 presents the differences in industries obtained from NHSCR and NH death certificates.

- When industry was available in both the registry and the death certificate, 60% of death certificates coded as manufacturing were also coded that way in cancer registry data. In contrast, 71% of cancer registry cases coded as manufacturing were also coded that way in death certificates.
- The majority of Unknown industry in NHSCR records were also found to be these same classifications in the corresponding death certificate.

Figure 2. Occupation findings from NHSCR when compared to NH death certificates.

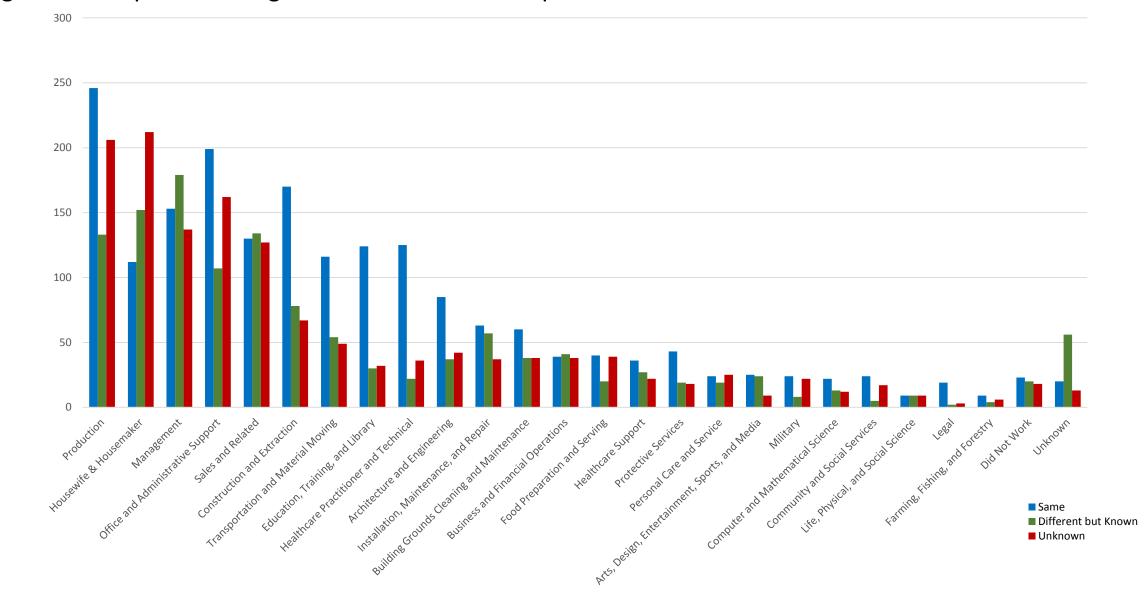
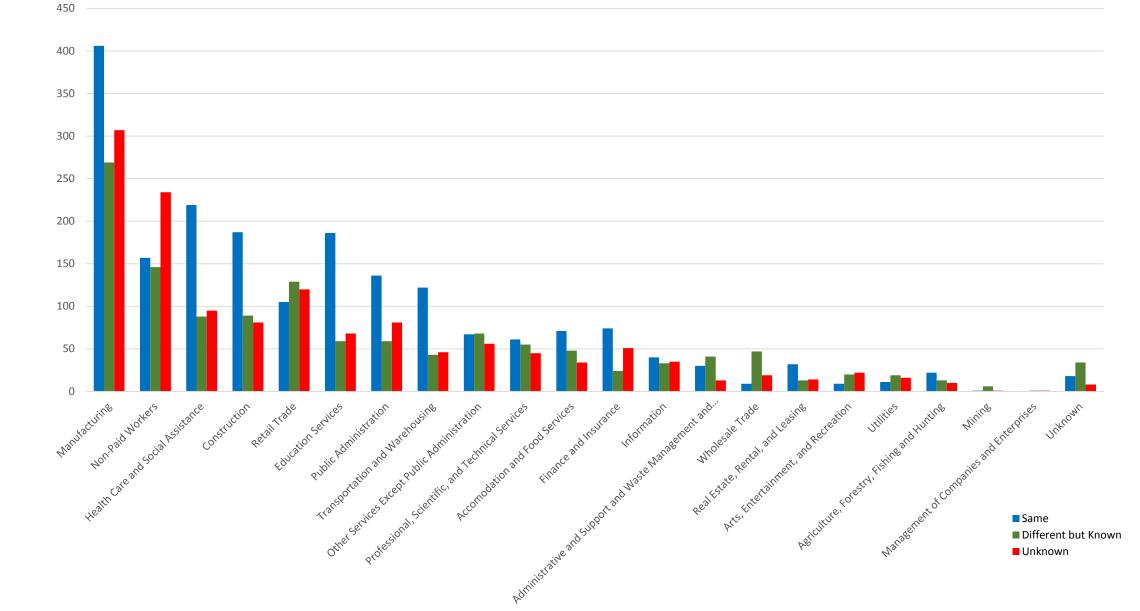


Figure 3. Industry findings from NHSCR when compared to NH death certificates.

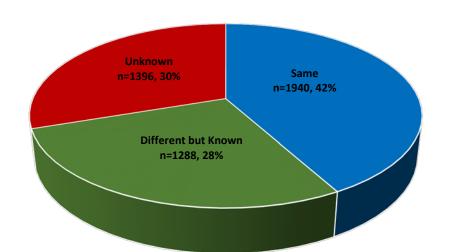


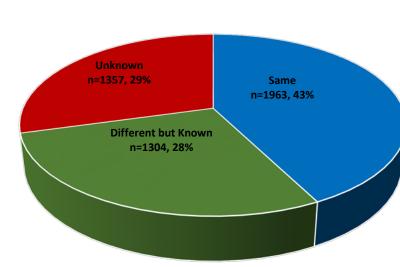
RESULTS, CONT.

Figure 2. Comparison of I&O data of 4,624 cancer cases from the NHSCR and NH death certificates.









In Figures 2-4, we took I&O data from cancer deaths in the NHSCR and compared it with I&O data obtained from NH death certificates.

- In Figure 2, we illustrate the total number/percent of cases that had the same I&O, how many had I&O but was different than what was reported in the death certificate, and how many were unknown.
- Of the 4624 cases included in the study, 42% (n=1940) of NHSCR had the same occupation reported in the death certificate. 43% (n=1963) cases had the same
- NHSCR had more unknown I&O (~30%) than what was reported on the death certificate (Figures 3-4).

DISCUSSION

- Vital Records data on I&O are more complete than NHSCR data
- For 30% of expired cases in the NHSCR, I&O were classified as retired, unknown or missing. It may be possible to improve I&O data for these cases by taking I&O from the death certificate (understanding that we cannot do the same for I&O data for living patients).
- In contrast, because of the larger proportion of cases with missing data in NHSCR, the proportion of NHSCR I&O that agreed with a given category in Vital Records was only 40%.
- For the cases where I&O is reported differently in NHSCR and Vital Records, more research is needed to determine which system is more likely to capture the longest held I&O

LIMITATIONS

- This comparison is based on a small sample of patients who died within a relatively short time of diagnosis, during 2009-2011. It is possible that these patients are systematically different from cancer patients who were diagnosed at other times, or who survived for longer after diagnosis.
- Further exploration of the data is needed to assess the impact of supplementing I&O data from death certificates for expired patients whose I&O is not known in NHSCR. When data are added to a database for one type of case but not another, there is a risk of introducing bias if those differences are not considered in future analyses. Until we fully understand the data and the potential for bias, one option is to retain both data sources in the registry database.
- Another option is to create separate fields for Vital Records I&O so that it is clear where the data originated.
- NHSCR text data contained more information than reflected here. NIOSH coders did not know the longest held I&O, and in this pilot effort, coded the first codeable response in the industry text field.
- The accuracy of death certificate and NHSCR I&O information is not known.

CONCLUSIONS

It may be feasible to improve I&O data in the NHSCR for cases who have died. However, data may not be comparable to those still alive. Therefore, it may be useful to introduce separate fields for I&O from Vital Records and NHSCR, so that the distinction is retained.

ACKNOWLEDGEMENTS

Funding for this project was provided in part by a contract with the New Hampshire Department of Health and Human Services and funding from the National Program of Cancer Registries (NPCR) of the Centers for Disease Control and Prevention through cooperative agreement UA/DP003930-02. The contents are solely the responsibility of the authors and do not necessarily reflect the official policies of the Centers for Disease Control and Prevention or the New Hampshire Department of Health and Human Services, nor does the mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government

The findings and conclusions in this presentation have not been formally disseminated by the National Institute for Occupational Safety and Health and should not be construed to represent any agency determination or policy.

We would like to thank our NH registrars for the hard work and dedication in collecting and reporting cancer data to the NHSCR, and Ms. Pamela Schumacher of NIOSH for coding our I&O data and for her guidance throughout this project.

http://www.cdc.gov/nchs/data/dvs/pt19manB1.pdf

1. Swanson GM, Schwartz AG, Burrows RW. An assessment of occupation and industry data from death certificates and hospital medical records for population-based cancer surveillance. Am J Public

2. Industry_Census_2000. The National Institute for Occupational Safety and Health. I&O Classifications. Available at: http://www.cdc.gov/niosh/topics/coding/pdfs/2000_Census_Industry.pdf. 3. Occupation_Census_2000. The National Institute for Occupational Safety and Health. I&O Classifications. Available at: http://www.cdc.gov/niosh/topics/coding/pdfs/2000 Census Occupation.pdf 4. Instruction Manual Part 19: Industry and Occupation Coding for Death Certificates, Effective 2003. Hyattsville, Md.: National Center for Health Statistics; 2002. Available at:

5. Thornton M. (ed). Standards for Cancer Registries Volume II: Data Standards and Data Dictionary, Record Layout Version 12.1, 15th ed. Springfield, III.: North American Association of Central Cancer Registries, June 2010.