EVALUATION OF OCCUPATION AND INDUSTRY IN THE NH STATE CANCER REGISTRY

APHA Annual Meeting

November 7, 2007

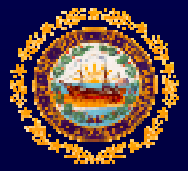
Karla Armenti, ScD*

S Cherala*, MO Celaya, J Rees**, BL Riddle****

Pamela Schumacher/NIOSH

**New Hampshire Department of Health and Human Services, Division of Public Health Services, Bureau of Disease Control and Health Statistics, Health Statistics and Data Management Section, and the New Hampshire State Cancer Registry, 2007.*

***Departments of Community and Family Medicine and the Norris Cotton Cancer Center, Dartmouth Medical School, and the New Hampshire State Cancer Registry, Lebanon, NH.*



Division of Public Health Services
New Hampshire Department of Health and Human Services

Background & Significance

There is evidence to suggest that a variety of cancers may result from occupation-related risk factors and exposures (Steenland K, et al, 2003).

Background & Significance

- While central cancer registries funded by the National Program of Cancer Registries are required to collect industry and occupation (I/O) information when available, it is well recognized that I/O data in cancer registries has many limitations.

- The quality of cancer registry data on industry and occupation (I/O) in New Hampshire has not been formally evaluated. However, a study by Massachusetts Cancer Registry suggested that considerable improvements in data quality for these variables may be possible by a careful review of medical records (Levy J et al. 2001).

Background & Significance

- For diagnosis year 2005, NHSCR received 6,713 analytic reports from 26 NH hospitals and one VA hospital. Approximately 85% of these cases lacked industry and occupation. I/O are required reportable text data items.
- We performed a reabstraction audit for a random selection of cancer case records, to evaluate current data quality and estimate the potential for improvement in data quality for industry and occupation.

Methods & Results

- Medical records for I/O history were reviewed for patients diagnosed with cancer using a sample of cases.
- A trained occupational coder from NIOSH classified I/O data according to the 2000 Bureau of Census which merges both NAICS (North American Industry Classification System) and the SOC (Standard Occupation Classification).
- Our analysis includes an evaluation of the frequency of records with no available I/O information.

Methods: Data (NHSCR)

- We randomly selected 20 cases diagnosed in 2005 for every cancer registrar in a registry hospital (n=14) and non-registry hospital (n=11).
  - 2 cases from each of the 5 major sites: female breast, prostate, lung, colon, and hematopoietic; and
  - 10 cases from other sites.
  - If there were not enough cases from the major sites, other sites were included to make up the 20.

Methods: Data (NHSCR)

- A total of 708 cases were eligible for medical record review.
- Final sample consisted of 474 reviewed cases (Table 1)

Some non-registry hospitals accession <20 cases/year, AND not all cases were available at the time of the review.

Methods: Study Sample

Reasonably representative of the NH population

- 98% Caucasians (NH incident cancer population 98%)
- 50% male (NH incident cancer population 48%)
- Mean age was 65
- Sampled cases included a small excess of bladder cancer and hematopoetic (leukemia) cancers, and a lower number of prostate, breast and lung cancers.

Methods: Medical Record Review

- I/O data were collected during NHSCR's routine data quality audit.
- Auditors were not provided with the original reported I/O data.
- I/O history was collected by reviewing all available documentation in the patient medical records and recording all I/O stated from the initial to the last hospital encounter provided in the medical record.

Methods: Medical Record Review

- The following information was collected:
  - First & last dates of hospital encounter*
  - # chart volumes reviewed, # chart volumes not reviewed*
  - Date of source where I/O was found
  - Source of I/O history
  - Exact text about I/O from medical record
  - Duration of occupation*
  - Record of hazardous exposure*

** Not used in any analysis*

I/O Coding & Analysis

- A NIOSH expert assigned I/O codes using the Census 2000 Industrial and Occupational Classification Systems to each I/O originally reported by the hospital and also to the I/O text collected during the audit.
- The Census 2000 Industry codes were grouped into 20 major sectors based on the 1997 North American Industry Classification System (NAICS), and;
- Occupation codes were grouped into 23 major occupation groups based on the 2000 Standard Occupational Classification (SOC) Manual.

I/O Coding & Analysis

Codes from both the original and the reabstracted data were then categorized according to the availability of I/O information:

- Both I/O available
- Industry or occupation only
- I/O coded to an unpaid workforce*
- Occupation available, but the industry was coded to unpaid workforce
- Industry available, but occupation was coded to unpaid workforce
- Industry coded to unpaid workforce, but the occupation was unknown
- I/O was unknown

**Unpaid workforce includes homemakers (farmers and housewives), volunteers, students, "retired", and "did not work" (never worked, child, infant, inmate, disabled, etc.).*

I/O Coding & Analysis

- **Third category: "Best I/O", which took the best I/O data from either the original or reabstracted categories.**

  Note: Best was defined as an I/O code other than "unknown", with I/O availability being the best, followed by the availability of industry or occupation only.

- **First, we compared the availability of I/O in the original data with data collected from reabstraction (Table 2).**

I/O Coding & Analysis

- We also calculated the difference in I/O availability between the original and the “best” I/O data obtained through both the original abstraction and our reabstraction.*
- The source within the medical record from which I/O was found was categorized into nine categories (Table 3).

I/O Coding & Analysis

- A rapid report is required 45 days from date of diagnosis and the completed case is required 180 days from the date of diagnosis
- Registrars should have all I/O information available at least 6 months from the diagnosis date.

I/O Coding & Analysis

- Under this assumption, the source date from which the I/O was collected during the review was classified according to the period of time surrounding diagnosis (Figure 7):
  - Source date was more than 6 months before the date of diagnosis
  - Source date was at the time of diagnosis
  - Source date was more than 6 months after the diagnosis date

I/O Coding & Analysis

- Hospitals were classified according to their annual caseload (total caseload, not just cancer patients)
  - Small = ≤105 cancer cases per year
  - Medium = 106-400 cancer cases per year
  - Large = 401+ cancer cases per year
- Hospitals were also classified according to whether or not they have a cancer program approved by the College of Surgeons Commission on Cancer (CoC).

Results:

Table 1. Characteristics of NHSCR vs. sample data

		NHSCR		Sample	
		No.	%	No.	%
Age	0-49	1142	17.0	65	13.7
	50-59	1454	21.7	106	22.4
	60-64	832	12.4	64	13.5
	65-69	861	12.8	43	9.1
	70-74	963	14.3	60	12.7
	75-79	569	8.5	58	12.2
	80+	892	13.3	78	16.5
	Total	6713	100.0	474	100.0
Rac					
	White	6575	97.9	465	98.1
	Non-white	68	1.0	4	0.8
	Unknown	70	1.0	5	1.1
	Total	6713	100.0	474	100.0
Sex					
	Male	3193	47.6	239	50.4
	Female	3520	52.4	235	49.6
	Total	6713	100.0	474	100.0

Table 1: Our final sample consisted of 474 reviewed cases (Table 1). The study sample was reasonably representative of the New Hampshire population with cancer from which it was selected.

	NHSCR		Sample	
	No.	%	No.	%
Primary Site				
Female Breast	1292	19.3	58	12.2
Lung & Bronchus	937	14.0	47	9.9
Prostate	770	11.5	36	7.6
Colo-Rectal	651	9.7	48	10.1
Hematopoetic	199	2.97	43	8.0
Melanoma of Skin	432	6.4	30	6.3
Female Genital System	335	5.0	24	5.1
Bladder	319	4.8	49	10.3
Other Digestive Organs	363	5.4	31	6.5
Oral Cavity & Pharynx	167	2.5	14	3.0
Larynx	60	0.9	3	0.6
Testis	31	0.5	1	0.2
Other Urinary Organs	143	2.1	11	2.3
Brain & Other Nerv. Syst	172	2.6	7	1.5
Thyroid	133	2.0	17	3.6
Lymphoma	288	4.3	27	5.7
Unknown & Other	421	6.3	28	6.9
Total	6713	100.0	474	100.0

Table 2. Availability of I/O information in cases reported originally, reabstracted, and both.

	Original		Reabst		Best		Difference Original vs.		Difference Original vs.	
	No.	%	No.	%	No.	%	No.	%	No.	%
I/O Available	72	15.2	257	54.2	278	58.6	206	43.5	185	39.0
Ind or Occ Only	15	3.2	62	13.1	57	12.0	42	8.9	47	9.9
I/O Unpd Workforce	19	4.0	91	19.2	86	18.1	67	14.1	72	15.2
Occ Avail, Ind Unpd Workforce	1	0.2	1	0.2	2	0.4	1	0.2	0	0.0
Ind Avail, Occ Unpd Wkfc	1	0.2	6	1.3	6	1.3	5	1.1	5	1.1
Ind Unpd Wkfc, Occ Unk	0	0.0	1	0.2	0	0.0	0	0.0	1	0.2
I/O Unknown	366	77.2	56	11.8	45	9.5				
Total	474	100.0	474	100.0	474	100.0	321	67.7	310	65.4

Table 3: Source of I/O on Record

	No.	%
FACE SHEET	156	36.4
H&P	86	20.0
CONSULT	69	16.1
ADMISSION NOTE	68	15.9
CLINIC or MD OFFICE NOTE	34	7.9
NOT AVAILABLE	10	2.3
DISCHARGE SUMMARY	3	0.7
ER FORMS/NOTES	2	0.5
OBITUARY, DEATH CERTIFICATE	1	0.2
Total	429	100.0

Figure 1. Availability of I/O information in cases reported by hospitals, reabstracted during NHSCR audit, and both combined.

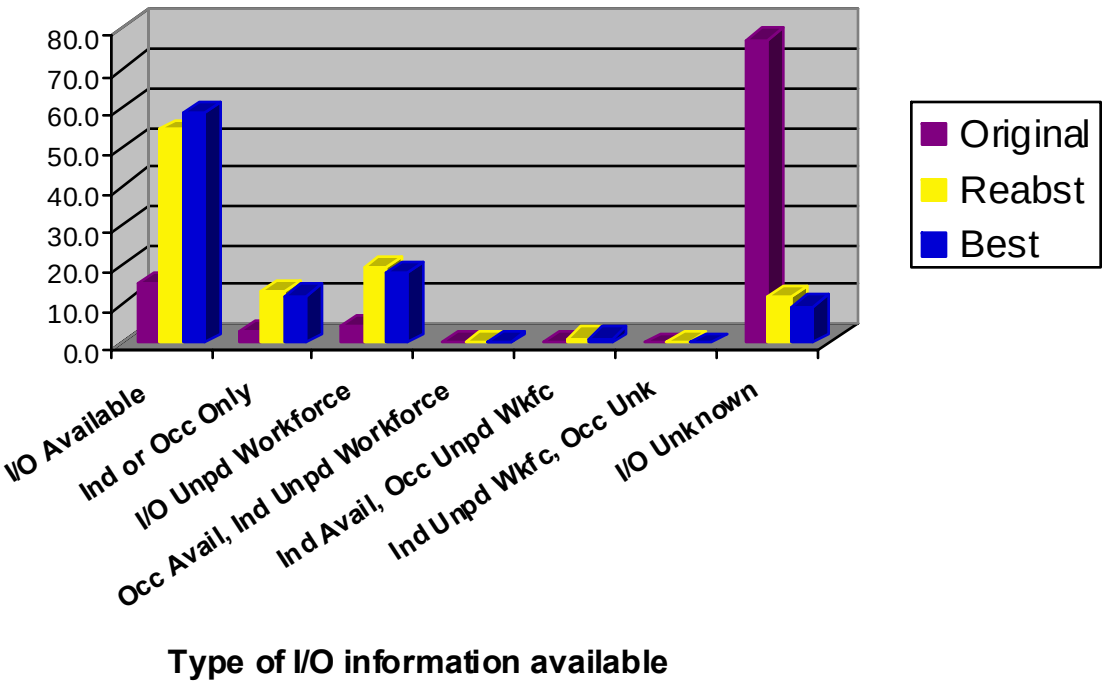
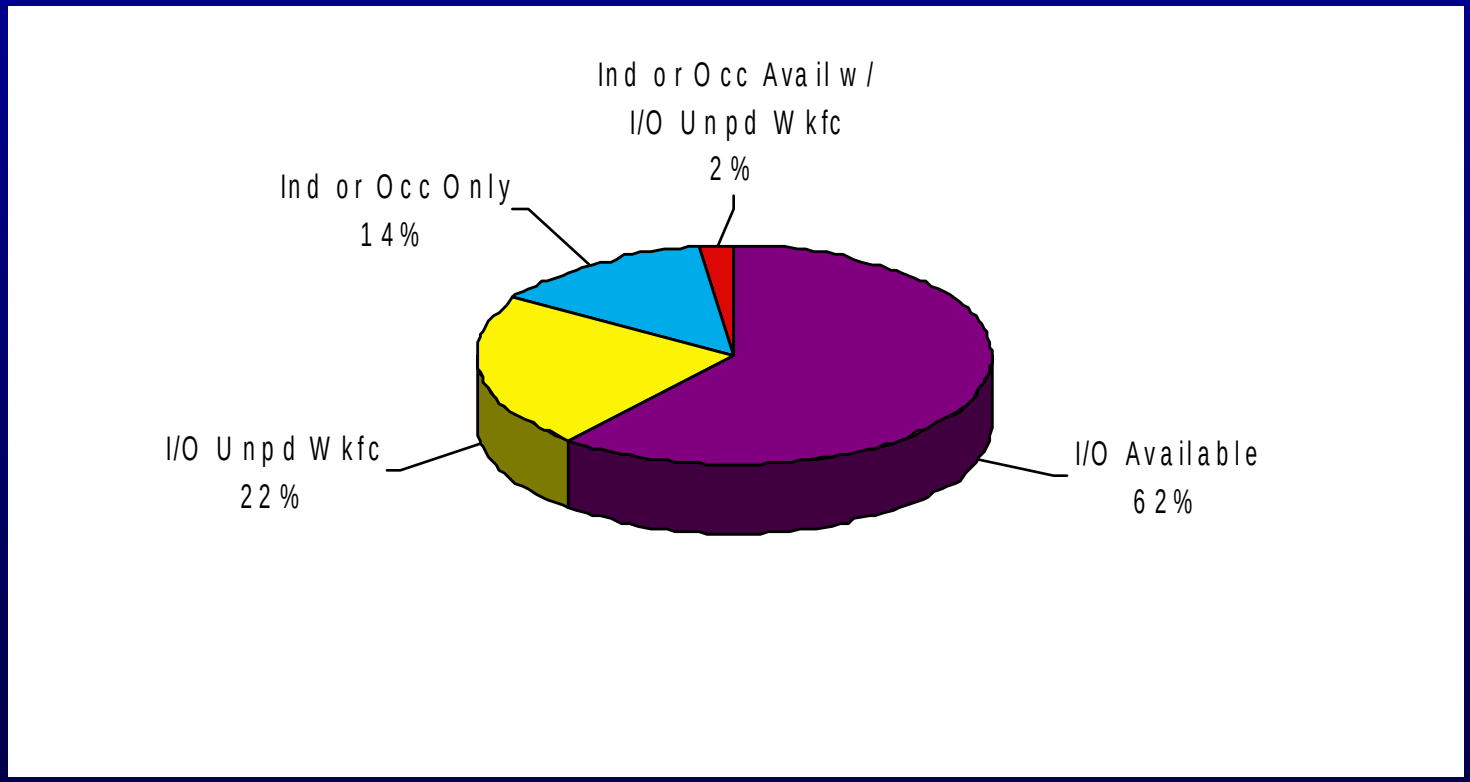


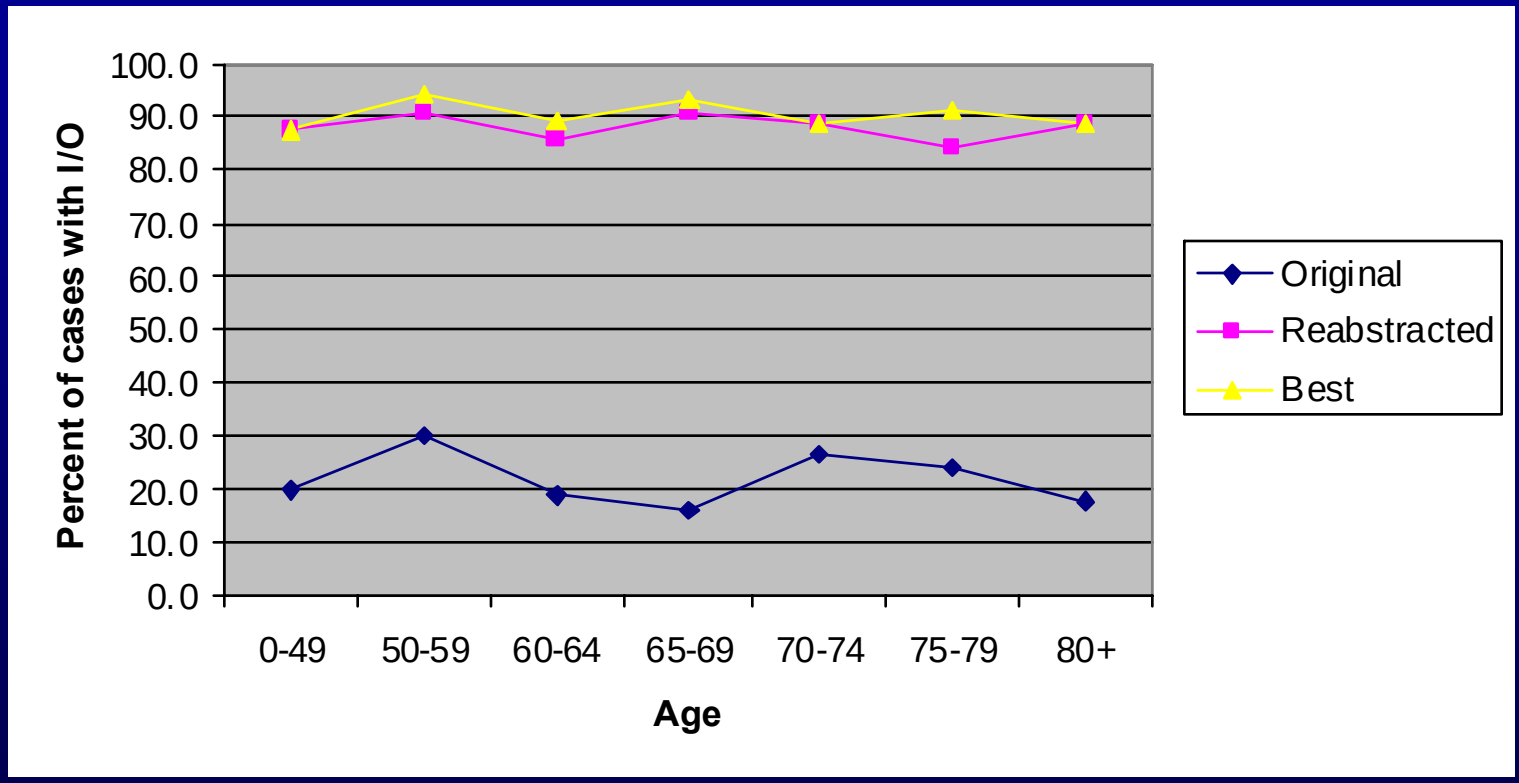
Figure 2. I/O information obtained on reabstraction for cases where I/O was originally reported as unknown or missing.



I/O information obtained on reabstraction for cases where I/O was originally reported as unknown or missing.

- Reabstraction found the same I/O information for 25% (n=116) of cases.
- Of the 366 cases that had been reported with unknown or missing I/O, 10% (n=45) still had missing or unknown I/O information even after reabstraction.
- Of the remaining 321, we found complete I/O for 196 (62%), and found either industry or occupation for 46 (14%).
- 2% of records included valid information on an industry or occupation that is not in the paid workforce.

Figure 3. Percent of cases with I/O information as originally reported by hospitals, reabstraction, and the best of both by age.



In all three groups, I/O information was most likely to be found for those aged 50-59.

Figure 4. Percent of cases with I/O information as originally reported by hospitals, reabstraction, and the best of both by major primary cancer sites.

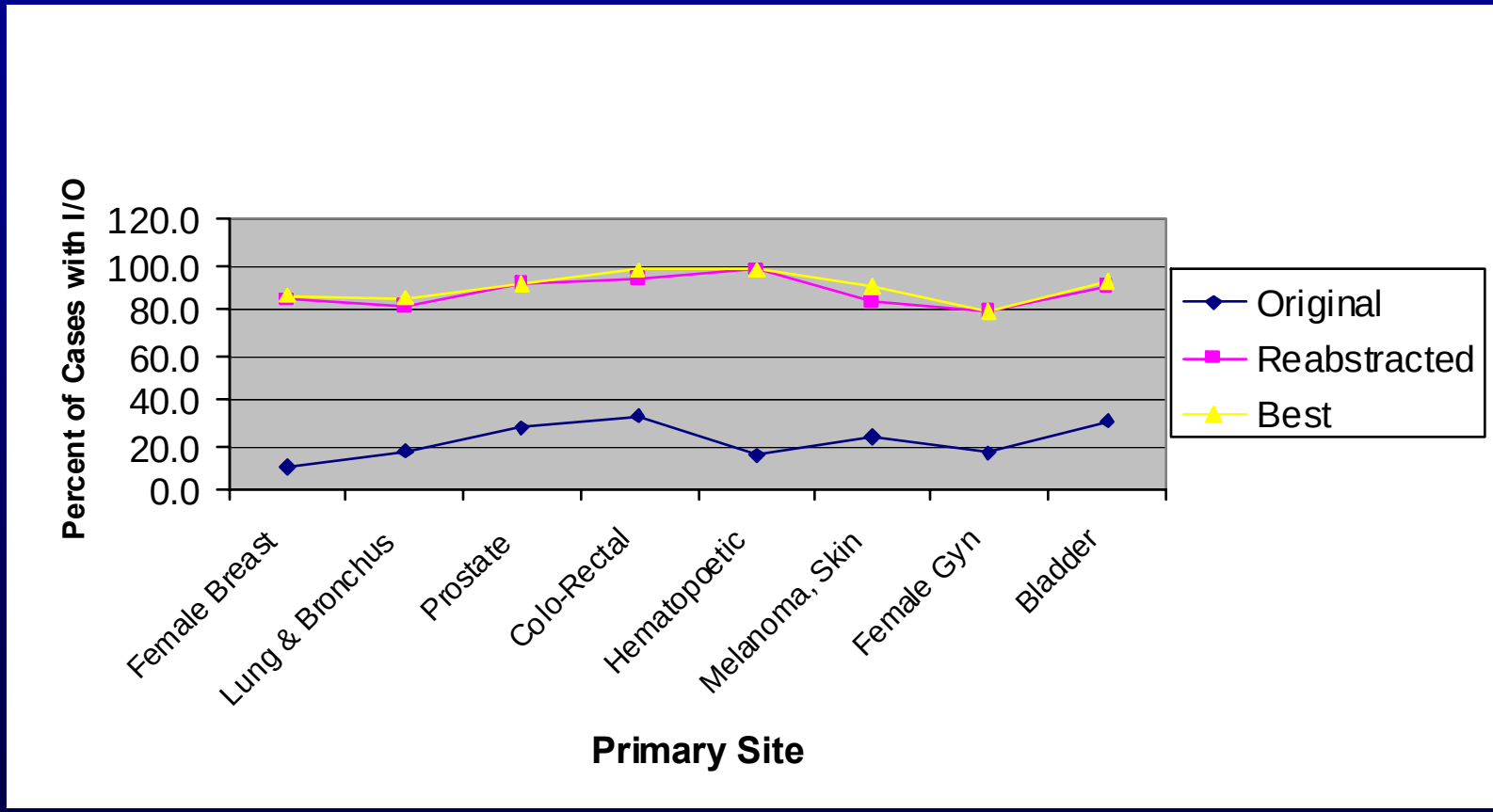
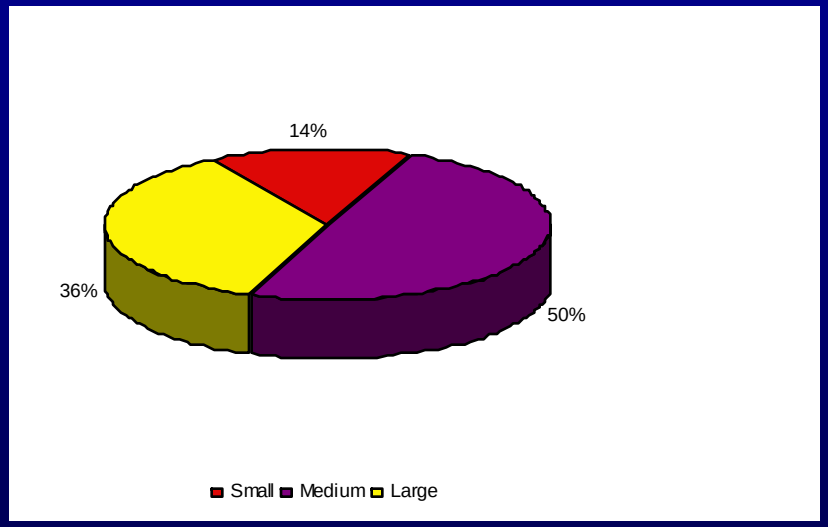


Figure 5. Characteristics of the hospitals where complete I/O information was found.

Hospital Size



Small = <105 cases per year (non-registry hospitals)
Medium = 106-400 cases per year
Large = 400+ cancer cases per year

Hospital with a CoC-approved Cancer Program

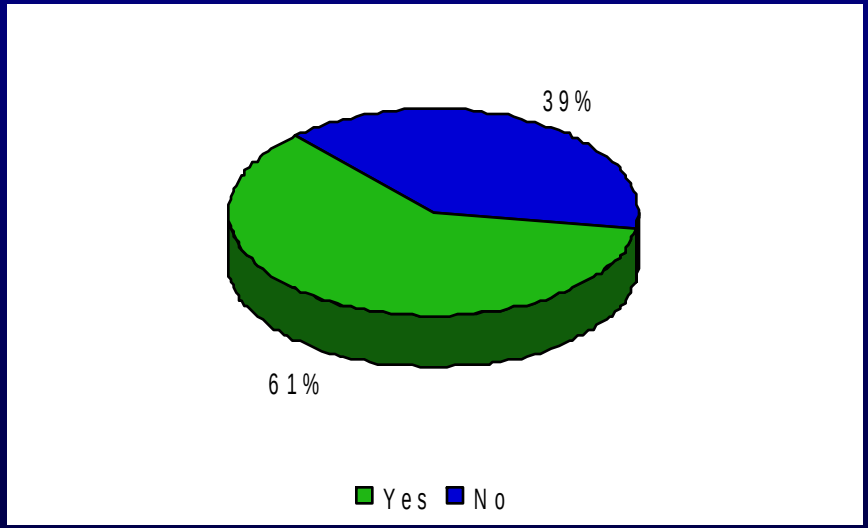
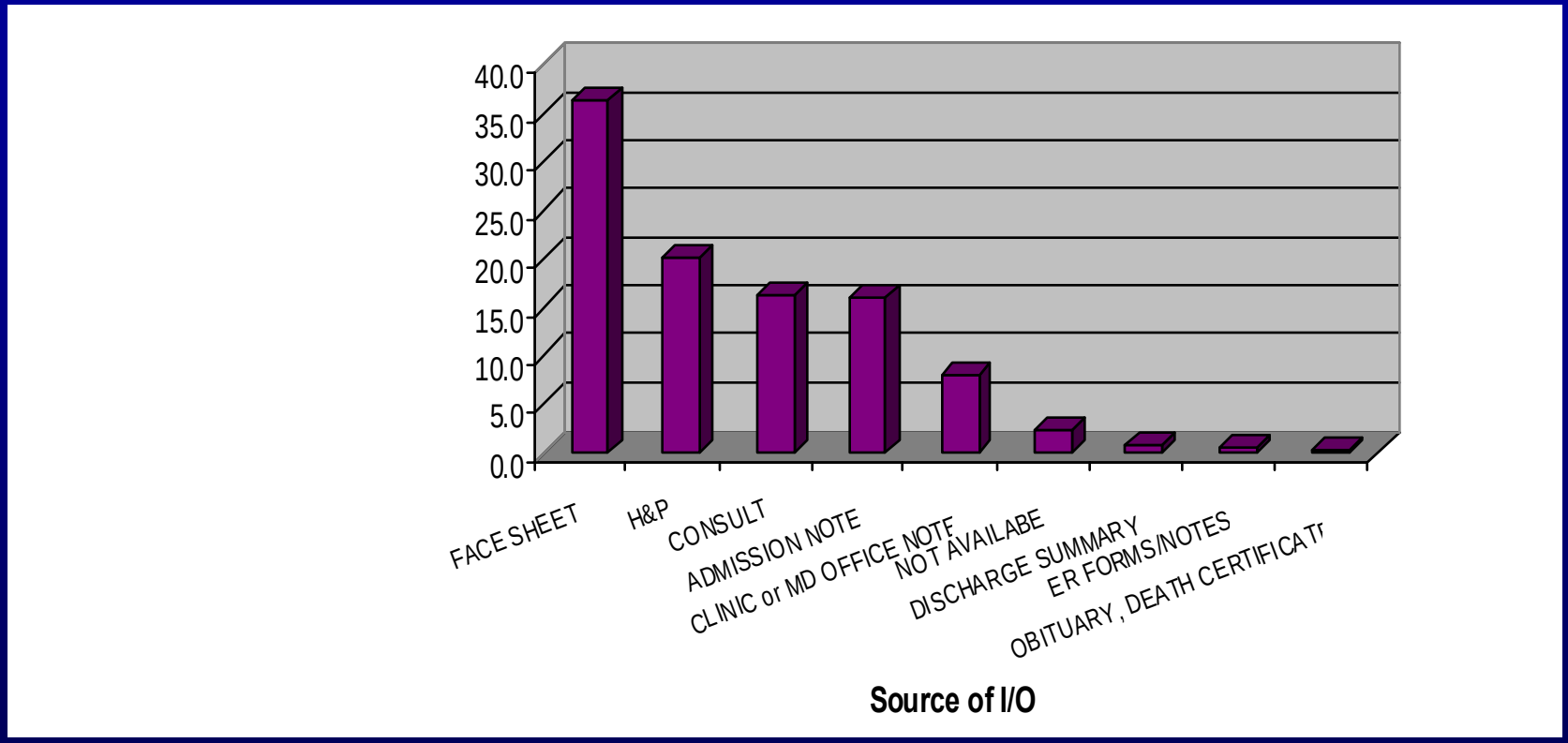
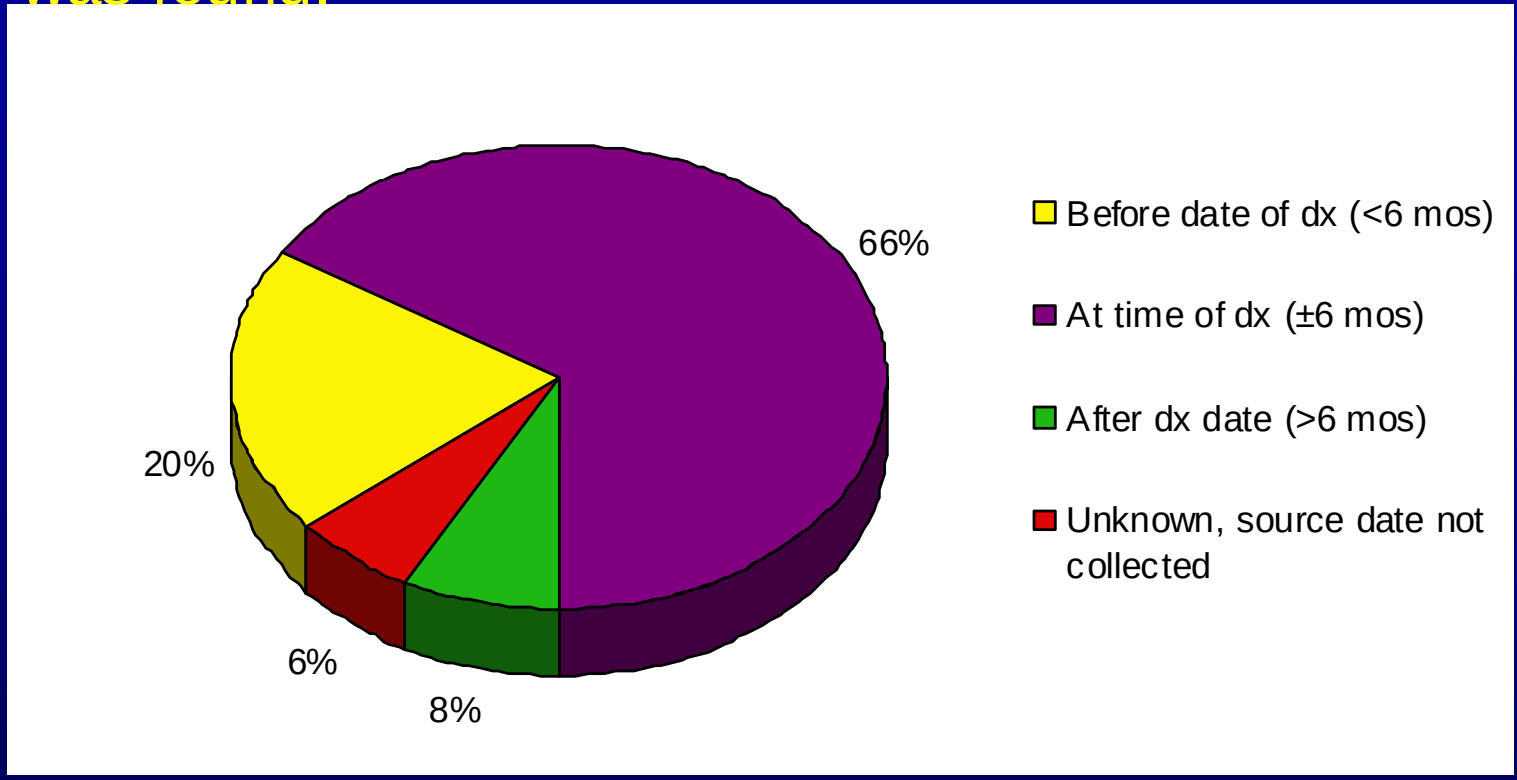


Figure 6. Characteristics of source from which I/O was found.



While the I/O could have been provided in multiple documents, we only accounted for the first source found during the time of diagnosis.

Figure 7. Characteristics of source date from which I/O was found.*



Of the 429 cases that had I/O information, only 8% were recorded after they would have been reported (>6mos from dx).

**The date when I/O data was found corresponds to the source from which the I/O was provided as shown in Figure 5.*

Discussion – Outcomes

- Demonstrated the potential to improve data collection for I/O in NH
- Existing registry data included both industry and occupation for only 15% of the original data records, compared to 54% of reabstracted records.

Discussion – Outcomes

- Records that contained partial I/O data comprised 8% of cancer cases in the original dataset and 34% in reabstracted records.
- A careful reabstraction process improved the proportion with no industry or occupation data from 77% to 12%.

Discussion

- Industry and occupation data were most commonly identified from the admission face sheet, or from the clinician's notes (history and physical, consult or admission notes).
- Combining data from the original cancer case report and the study reabstraction audit gave marginal additional improvements in data quality.

Conclusion

- It is feasible to achieve substantial improvements in industry and occupation data quality in New Hampshire by medical record review, with specific attention to face sheets and clinician's notes.
- Results of this study will be shared with reporting sources. It is our goal to increase accurate reporting of I/O data to the central registry through education and training.

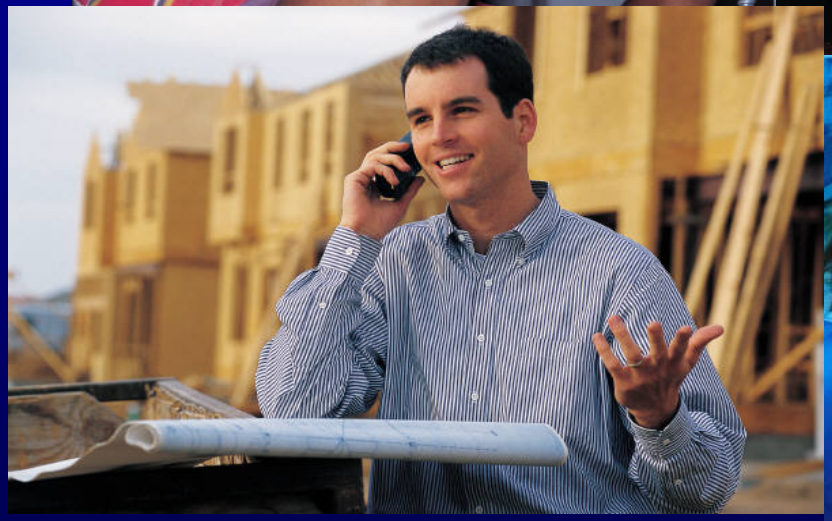
Conclusion

- Further analysis of data collected from the NH State Cancer Registry may be helpful in identifying occupational risk factors for cancer in New Hampshire.
- The findings of this study are useful in deciding if central cancer registries should expand the collection and coding of I/O data.

References

Steenland K, Burnett C, Lalich N, Ward E, Hurrell J, Dying for Work: the magnitude of US mortality from selected causes of death associated with occupation, Am J Ind Med, 2003.

Levy J, Brooks D, Davis L. Availability and quality of industry and occupation information in the Massachusetts Cancer Registry. Am J Ind Med. 2001 Jul;40(1):98-106.



Development of this program was supported by Centers for Disease Control and Prevention, PA02060 under cooperative agreement number U55/CCU-121912.