An Assessment of Occupation and Industry Data from Death Certificates and Hospital Medical Records for Population-Based Cancer Surveillance

G. Marie Swanson, PhD, Ann Grossbart Schwartz, MPH, and Robert W. Burrows, MPH

Abstract: This study analyzed 30,194 incident cases and 4,301 death certificates for completeness of occupational reporting. Analysis of data accuracy was based upon a comparison of more than 2,000 death certificates with incident abstracts and 352 death certificates with interview data. Death certificates had a higher proportion with occupation (94.3%) and industry (93.4%) reported than did incident abstracts of hospital medical records (39.0% and 63.5%, respectively). Compared with occupational history data obtained by interview, 76.1% of the death certificates were exact matches for usual occupation and industry. (*Am J Public Health* 1984; 74:464–467.)

Introduction

Specification of the relationship between a population's occupation and industry characteristics and the incidence of particular types of cancers in that population is an important element in cancer etiology. Delineation of occupational risk for cancer at a descriptive level can provide new etiologic leads, as well as complement existing data regarding occupations and industries that may have an elevated risk for specific cancers.[1,2] Furthermore, descriptive assessment of occupational risks can address the concerns of the National Institute for Occupational Safety and Health for continuous monitoring.[3] In addition to the public health issue of responding to concerns of workers and industry,[4,5] data are needed from which one can evaluate the frequent mass media reports of cancer cases or cancer risks among a group of workers or in a particular plant. Nevertheless, routine reporting of either cancer incidence or mortality by occupation and industry is in its formative stages.

The purpose of this study was to explore the potential for routine monitoring of cancer incidence by occupation and industry. Comparing hospital medical record data, death certificate data, and interview data, mortality surveillance is evaluated against incidence surveillance. While the methodologic issue is of particular concern in reference to cancer, the results of this study could also be applied to other areas of occupational health research. Few studies have assessed the quality of occupational data.[6] No previous assessment of hospital medical record occupation and industry data was found in the literature. Only a few studies of death certificate data evaluated completeness and accuracy of occupation and industry information,[7–11] while a larger group of studies report cancer mortality rates by occupation.[12–14]

Methods

The first phase of the study analyzed completeness of occupational reporting for the 30,194 incident cancer cases diagnosed in Metropolitan Detroit in 1980 and 1981 and recorded by the Cancer Surveillance Section of the Division of Epidemiology of the Michigan Cancer Foundation. This population-based cancer surveillance system encompasses Metropolitan Detroit (Wayne, Oakland, and Macomb Counties) and has been a participant in the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) program since its inception in 1973. A full description of the national SEER program[15] and the Metropolitan Detroit cancer reporting system[16] is available elsewhere.

An important factor in the consistency of data on the initial case abstracts is our data collection method. Abstractors trained and on the staff of the Michigan Cancer Foundation Cancer Surveillance Section collect data from the hospital medical records.* Occupation and industry data are then coded by other staff members.

The second phase of the study consisted of an analysis of 4,301 incident cases, diagnosed in 1979, who had died during either the year of diagnosis or the first year after diagnosis. Frequency of reporting occupation and industry by source was compared and an estimate was made of the total reporting obtained from both sources combined. The third segment of the study assessed the accuracy of usual occupation and industry reported on 352 death certificates for cases from whom occupational history data had been obtained by interviews. These cases included patients diagnosed with bladder cancer in 1978 or with multiple myeloma in 1978, 1979, or 1980. Death certificate data were evaluated as a potential element of an incidence reporting system, rather than as an independent mortality surveillance system.

The initial abstract of the hospital medical record for incident cancer cases has included each patient's occupation and industry since 1973. For cases newly diagnosed in 1980 and 1981, occupation and industry were coded and computerized. Cases diagnosed in 1979, who died before January 1, 1981, were also coded and the data compared to similar information on death certificates. The US Census Bureau codes for occupation and industry, as revised for the 1980 census of the US population, were utilized.[17] Death certificate data were obtained from microfilm copies of death certificates. Interview data had also been coded within Census Bureau guidelines.

In addition to occupations defined by the US Census Bureau, the current study includes persons reported as housewife, student, disabled, or unemployed in the category of cases with occupational information. Persons listed as retired, with no prior occupation reported, have been treated as having no occupational information. An estimate was made of the total proportion of incident cases for which

From the Division of Epidemiology, Michigan Cancer Foundation. Address reprint requests to Dr. G. Marie Swanson, Director, Division of Epidemiology, Michigan Cancer Foundation, 110 E. Warren, Detroit, MI 48201. This paper, submitted to the Journal May 2, 1983, was revised and accepted for publication August 19, 1983.

*While this brief report cannot fully describe abstracting procedures, a copy of the abstractors manual may be obtained from the authors upon request.

TABLE 1—Occupational Information Obtained from Incident Case Abstracts for Cancers Diagnosed in 1980 and 1981 in Metropolitan Detroit: Age and Sex Distribution

Age (years)	N	% Cases with Occupation Reported*	% Cases with No Information‡
		Males	
≤24	336	48.2	51.8
25–34	470	52.6	47.4
35–44	642	54.8	45.2
45–54	1752	60.2	39.8
55–64	3889	49.1	50.9
65–74	4189	24.9	75.1
75–84	2751	20.4	79.6
85+	785	18.9	81.1
All Ages	14,814	37.0	63.0
		Females	
≤24	437	44.6	55.4
25–34	921	51.6	48.4
35–44	1105	55.4	44.6
45–54	2132	53.7	46.3
55–64	3799	45.0	55.0
65–74	3657	28.7	71.3
75–84	2512	22.9	77.1
85+	812	21.2	78.8
All Ages	15,375	38.6	61.4

*Includes cases with valid census codes for occupation plus housewives, students, disabled, and unemployed.

‡Includes retired persons for whom no previous occupation was reported, as well as those with no information.

occupation could be obtained for a specific year of incident cases by combining information from initial case abstracts with death certificates of those cases who had died within the subsequent year.

Results are presented by sex and age groups. Analyses of completeness of occupational reporting were also made by hospital size, race, and cancer site.

Results

Table 1 presents information describing whether or not occupation was reported on the initial cancer case record, according to sex and age of patients diagnosed in 1980 and 1981. Overall, there is a slight difference between males and females, with occupation reported for 38.6 per cent of women and 37.0 per cent of men. The maximum rate of occupational reporting attained for men was 60.2 per cent at ages 45–54, while for women, it was 55.4 per cent at ages 35–44. A person is most likely to be employed between the ages of 25 and 64. For this age group, the rate of occupational reporting on the initial case abstract was 52.8 per cent for males and 49.5 per cent for females.

Analyses of the rate of occupation and industry reporting by hospital size, by race and sex of patients, and by primary site revealed minimal differences. Therefore, only age differences are presented in the following tables.

In Table 2, occupational reporting rates for death certificates are compared to rates for initial case abstracts among patients diagnosed with cancer in 1979 who died prior to January 1, 1981. Large differences are seen in every age group except 25–34. Overall, 94.3 per cent of the death certificates had occupation reported, compared with 39.0 per cent of the initial case abstracts. Among those aged 25–64, the abstract had occupation reported for 57.8 per cent, in contrast to 95.4 per cent for death certificates.

Reporting of industry information on death certificates was complete for 93.4 per cent of the cases overall, and for 63.5 per cent of the initial case abstracts (Table 2). For persons ages 25–64, industry was reported by 71.7 per cent of the abstracts and 91.3 per cent of the death certificates.

The total rate of occupational reporting that could be obtained by utilizing both incident case abstract data and death certificate data for incident cases is estimated in Table 3. Overall, occupation could be obtained for 53.7 per cent of annual incident cases. Again, there is considerable variation across age groups, with occupation available for 58.5 per cent of persons age 25 through 64.

Death certificate occupation was compared with initial case abstract occupation for the 1,676 1979 cases which had occupation reported from both sources. Agreement between these two sources was very good, ranging from 77.3 per cent to 100 per cent (Table 4). Overall, occupation matched between the two sources for 83.2 per cent of all cases and for 83.4 per cent of those aged 25–64. The same type of comparison was made for the 2,643 1979 cases with industry reported from both sources. Accurate matches for industry were seen for 96.7 per cent of all cases and for 96.8 per cent of persons aged 25–64 (Table 4).

For a third set of cases, a comparison was made between death certificates and interview data. Of those interviewed, 352 died by the end of 1982 (243 multiple myeloma patients and 109 bladder cancer patients). Both occupation and industry as reported on the death certificate were evaluated against usual lifetime occupation and industry as reported by the respondent, or a relative of the respondent, during an interview. Both studies utilized proxy interviews with the patient's closest relative when the patient was too ill for interview or was deceased by the time of contact for interview. There were 171 (48.6 per cent) of the

TABLE 2—Cancer Occupational Reporting* on Initial Case Abstracts Compared with Death Certificates: Distribution by Age for Patients Diagnosed in 1979 Who Died Before January 1, 1981

Age (years)	N	% Occupation Reported on Abstract	% Occupation Reported on Death Certificate
≤24	37	64.9	73.0
25–34	23	95.7	95.7
35–44	94	67.0	95.7
45–54	399	60.4	96.0
55–64	931	54.9	95.2
65–74	1216	33.4	93.9
75–84	1076	26.5	94.3
85+	525	23.6	93.5
TOTAL	4301	39.0	94.3
Age (years)	**N**	**% Industry Reported on Abstract**	**% Industry Reported on Death Certificate**
≤24	37	89.2	94.6
25–34	23	82.6	91.3
35–44	94	78.7	90.4
45–54	399	70.9	92.0
55–64	931	71.0	91.1
65–74	1216	59.0	93.6
75–84	1076	57.4	95.7
85+	525	61.7	93.7
TOTAL	4301	63.5	93.4

*Includes US Census Bureau coded occupations and housewife, student, disabled, and unemployed.

TABLE 3—Occupational Reporting from Initial Cancer Case Abstracts Combined with Death Certificates: An Estimate of the Total Data Available within One Year after Diagnosis for Cases Diagnosed in 1980 and 1981, by Age

Age (years)	N	% Estimates of Total Occupational Data Available*	% Cases with No Information
≤24	773	47.2	52.8
25–34	1391	51.9	48.1
35–44	1747	58.2	41.8
45–54	3884	63.6	36.4
55–64	7688	57.2	42.8
65–74	7846	45.7	54.3
75–84	5263	49.7	50.3
85+	1597	65.4	34.6
All Ages	30,194‡	53.7	46.3

*Includes US Census Bureau coded occupations and housewife, disabled, and unemployed.
‡Includes five cases with missing ages.

interviews conducted with the cancer patient, while 181 (51.4 per cent) were proxy interviews with relatives of the patient.

Overall, 76.1 per cent of the death certificates reported the same usual occupation and industry as provided by interview (Table 5). An additional 11.4 per cent had housewife reported by one source and some other occupation by the other. Only 12.5 per cent of the death certificates had occupation and industry, other than housewife, that differed on the two reporting sources. Of the 44 in this latter "no match" category, 20 of the death certificates had the same occupation and industry reported as found on the interview for most recent occupation and industry, rather than usual. If one accepts all three groups as providing some accurate data on the death certificate—actual agreement for usual occupation and industry, differing only for housewives, or a match with current occupation—then the proportion of valid data from death certificates is 93.2 per cent. Excluding matches with current occupation, the total valid information is 87.5 per cent.

Only minor differences in accuracy were observed for proxy interviews compared with patient interviews, as Table 5 indicates. Complete agreement between sources regarding usual occupation and industry was obtained for 74.3 per cent of case interviews, compared with 77.9 per cent of proxy interviews.

Discussion

In terms of population-based incidence reporting of occupation for cancer patients, data adequate for analysis are clearly not available from the initial case abstract of the hospital medical record.

For mortality studies, however, it appears that death certificate data on occupation and industry are very complete and match well with interview data describing usual occupation and industry. Our results are comparable to those of two other studies. An NCHS-NIOSH (National Center for Health Statistics-National Institute for Occupational Safety and Health) study of the codability of occupation and industry data for a sample of 5,000 US death certificates found that 91 per cent had a codable entry for occupation, as did 81 per cent for industry. For both items, percentages of codable entries were somewhat higher for females than for males and for Whites compared with Blacks. For persons age 65 and older, 92 per cent of the entries for occupation were codable, compared with 88 per cent for those under age 65. In previous analyses of US death certificates, for persons age 25–64, occupation was codable for 92.8 per cent of 1950 death certificates reviewed, for 86.9 per cent of 1973 death certificates reviewed, and for 86.1 per cent of 1975 death certificates studied.[7]

Other studies found 75 per cent agreement between the death certificate and lifetime occupation interview data,[8] and strongly support inclusion of death certificate information as an element of occupational health surveillance.[9] An earlier study conducted in the mid-1950s by the State of California Health Department found 51.9 per cent agreement between interview data and death certificate entries for usual occupation and 69.7 per cent agreement between these two sources for current or last occupation.[10]

A recent analysis of the availability of coded death certificate information for occupation and industry reported 11 states that routinely code occupation, seven that routinely code industry, and six that code one or the other on a limited basis.[18] Thus, most states will have uncoded information from death certificates.

Our study and these previous reports suggest that population-based reporting of occupation and industry for mortality studies of cancer and other diseases is feasible and likely to produce useful information. It also seems feasible to obtain usual occupation and industry, rather than current classifications. Due to the long latency period of most cancers, usual occupation and industry are essential starting points for the assessment of potential risk. Attaining an equally satisfactory level of reporting from existing records for incidence studies will be more complex.

TABLE 4—Frequency of Matches between Initial Cancer Case Abstract and Death Certificate Occupation by Age for Cases with Occupation and Industry Reported from Both Sources among Those Diagnosed in 1979 Who Died Prior to January 1, 1981

Matching Occupation*		
Age (years)	N	%
≤24	24	100.0
25–34	22	77.3
35–44	63	82.5
45–54	241	82.2
55–64	511	84.3
65–74	406	77.8
75–84	285	83.9
85+	124	94.4
All Ages	1676	83.2
Matching Industry‡		
Age (years)	N	%
≤24	33	100.0
25–34	18	88.9
35–44	69	98.6
45–54	271	95.9
55–64	630	97.1
65–74	694	94.8
75–84	608	97.4
85+	320	99.4
TOTAL	2643	96.7

*Includes US Census coded occupations and housewife, student, disabled, and unemployed.
‡Includes US Census coded industries and housewife and student.

TABLE 5—Comparison of Usual Occupation and Industry Data Obtained from Death Certificates with Interview Data

	N	%
Interviews with Cases		
Sources agreed for both occupation and industry	127	74.3
Sources did not agree for occupation and industry	27	15.8
One source reported housewife, alternative source reported other occupation	17	9.9
TOTAL	171	100.0
Interviews with Relatives of Cases		
Sources agreed for both occupation and industry	141	77.9
Sources did not agree for occupation and industry	17	9.4
One source reported housewife, alternative source reported other occupation	23	12.7
TOTAL	181	100.0
All Interviews		
Sources agreed for both occupation and industry	268	76.1
Sources did not agree for occupation and industry	44	12.5
One source reported housewife, alternative source reported other occupation	40	11.4
TOTAL	352	100.0

The issue of readily available data regarding usual occupation and industry is important for epidemiologic study of many serious illnesses, in addition to cancer research. Our experience is limited to a single geographic area, therefore it is possible that other population-based systems for cancer incidence reporting may be able to achieve higher rates of occupational reporting from the hospital medical records than we have reported. Mortality data are readily available and should be coded and used for analysis of cancer deaths by occupation and industry. Incidence monitoring is essential for complete and accurate occupational and industrial cancer surveillance. Therefore, cost-effective and accurate methods for incidence surveillance must be developed and tested.

REFERENCES

1. Wong O, Hurley FL: A biostatistical and epidemiological perspective of an occupational health record system. J Am Med Rec Assoc 1981; 52:56–62.
2. Walker B: Occupational Cancer. J Environ Health 1982; 44:176–179.
3. Spirtas R, Sundin DS, Sestito JP, Behrens VJ, French J: A Conceptual Framework for Occupational Health Surveillance. DHEW (NIOSH) Pub. No. 78–135. Cincinnati: NIOSH, 1978.
4. Ashford NP: Crisis In the Workplace. Cambridge, MA: MIT Press, 1976.
5. Schilling RSF (ed): Occupational Health Practice, 2nd Ed. London: Butterworths, 1981.
6. Baumgarten M, Oseasohn R: Studies on occupational health: a critique. JOM 1980; 22:171–176.
7. Rosenberg HM, Burnham D, Spirtas R, Valdisera V: Information from the death certificate: assessment of the completeness of reporting. *In:* DelBene L, Scheuren F (eds): Statistical Uses of Administrative Records with Emphasis on Mortality and Disability Research. Washington DC: Social Security Administration, Office of Research and Statistics, 1979.
8. Wegman DH, Peters JM: Oat cell lung cancer in selected occupations. JOM 1978; 20:793–796.
9. Frazier TM, Wegman DH: Exploring the use of death certificates as a component of an occupational health surveillance system. Am J Public Health 1979; 69:718–720.
10. Buechley R, Dunn JE, Linden G, Breslow L: Death certificate statement of occupation: its usefulness in comparing mortalities. Public Health Rep 1956; 71:1105–1111.
11. Alderson A: Some sources of error in British occupational mortality data. Br J Indust Med 1972; 29:245–254.
12. Logan WPD: Cancer Mortality by Occupation and Social Class, 1851–1971. IARC Scientific Pub. No. 36. Lyons: International Agency for Cancer Research, 1982.
13. Guarlnick L: Mortality by Occupation and Industry among Men 20 to 64 Years of Age, United States 1950. USDHEW, Vital Statistics, Special Report 53. Washington, DC, Govt Printing Office, 1962.
14. Milham S: Occupational Mortality In Washington State, Vol. 1. NIOSH Research Report. Cincinnati: NIOSH, 1976.
15. Young JL Jr, Percy CL, Asire AJ (eds): Surveillance, Epidemiology, and End Results: Incidence and Mortality Data, 1973–1977. National Institutes of Health, National Cancer Institute Monograph 57. Washington, DC: Govt Printing Office, 1981.
16. Swanson GM, Brennan MJ: Cancer incidence and mortality in metropolitan Detroit, 1973–1977. *In:* Young JL, Percy CL, and Asire AJ (eds): Surveillance, Epidemiology, and End Results: Incidence and Mortality Data, 1973–1977. National Institutes of Health, National Cancer Institute Monograph 57. Washington, DC: Govt Printing Office, 1981.
17. US Bureau of the Census, 1980. Alphabetical Index of Industries and Occupations. Washington, DC: US Govt Printing Office, 1981.
18. Kaminski R, Brockert J, Sestito J, Frazier T: Occupational information on death certificates: a survey of state practices. Am J Public Health 1981; 7:525–526.

ACKNOWLEDGMENTS

This research was supported by the National Cancer Institute Contract No. NO1 CP-61028 and by the United Foundation. The authors wish to express their gratitude to Ruth Knapp and her staff of the Cancer Surveillance Section, Division of Epidemiology, Michigan Cancer Foundation, for their careful coding of occupation and industry for 1980 and 1981 diagnoses; to Dr. Debra Silverman, Biometry Branch, National Cancer Institute, for providing a computerized list of usual occupation and industry from interview data; and to Dr. William Satariano and Steven Belle, Division of Epidemiology, Michigan Cancer Foundation, for their thoughtful review of the manuscript.

Conference on Genetic Counseling in Perinatal Period Scheduled

"Strategies in Genetic Counseling: Issues in Perinatal Care," the fourth annual Professional Education Conference of the National Society of Genetic Counselors, will be held June 15–16, 1984 at the Regency Hotel, Denver, Colorado. As in previous years, this Conference precedes the March of Dimes Birth Defects Foundation Conference.

The purpose of the conference is to promote the education of and communication among genetic counselors and related professionals. Through workshops, lectures, and panel discussions, the impact of genetic and congenital problems in the perinatal period—and appropriate intervention strategies—will be addressed.

For more information contact Helen Travers, MS, Conference Chairperson, Mailman Center for Child Development (D-820), PO Box 016820, Miami, FL 33101.