Comparison of Occupation and Industry Information from Death Certificates and Interviews

MARY CATHERINE SCHUMACHER, MD, MSPH

Abstract: Information on occupation and industry obtained via an interview prior to death was compared to occupation and industry on the death certificate of 184 colon cancer patients in Utah. The data were coded blindly using a five-digit code. Overall, agreement in the five-digit codes was found for 63 per cent. The industry codes agreed for 67 per cent of the individuals, and the occupation was identical for 68 per cent. Agreement by subjective evaluation of the two data sources, disregarding the five-digit codes, was 73 per cent. There were no differences in agreement of the five-digit codes by age, sex, and county of residence. The number of years worked at the job given by interview was related to agreement. Misclassification occurred in a random manner. It is concluded that the use of death certificates to study the association of occupation and disease is most appropriate for pilot studies. (*Am J Public Health* 1986; 76:635–637.)

Introduction

The use of death certificates to investigate associations of illness and occupation is controversial.[1,2] On one hand, occupation and industry data are available in most states on death certificates. The data are relatively easy to obtain, and inexpensive studies can be done using these data. On the other hand, death certificates contain no exposure information, job descriptions are often nonspecific, the temporal relationship between the job and the illness is not known, and few states routinely code occupation and industry.[3]

Previous studies investigating the reliability of death certificate occupational information have found variable results.[4–8] The information on death certificates is usually compared with another source of information, such as census data or interview data obtained prior to death. The accuracies that are reported range from 51–93 per cent. In general, those authors who rigorously define their criteria for agreement of the two data sources report lower results for accuracy.[4,8]

Because the literature results for accuracy of occupation and industry as listed on death certificates are variable, and because the accuracy may vary from state to state, the current study was performed in Utah. Death certificate data were compared with interview data obtained prior to death, using precise definitions for agreement of the two data sources.

Methods

Patients with colon cancer were identified through the Utah Cancer Registry for two separate case-control studies of diet and cancer between 1977 and 1981.[9]* A total of 475 White, Utah residents over age 40 with colon cancer were interviewed in their homes between one and three months after diagnosis.

As part of the interviews patients were asked:

- During the last 15 years (or prior to age 65) did you have a full-time job outside the home?
- If yes, what type of work did you do most often during this time?
- Type of place work was done?
- Title or position?
- Years worked there?

During the fall of 1984, we identified 114 cases from the first study, and 83 from the second who had died. The interview data were obtained, and the responses to the above questions were abstracted. Occupation and industry, defined as "usual occupation," were then obtained from the death certificates of these people. Additional data collected were age at death, sex, year of death, and county of residence.

Occupation and industry from each source were coded independently and blindly. A five-digit code devised to link occupation to specific exposures was used, the first two digits representing industry, and the next three occupation.[10,11] The data were also coded for "clusters", which are groups of occupations combined because of similar exposures.[11] Up to two occupations were coded for each entry.

The coding system used contains 18 industry categories derived from the US Bureau of the Census *1970 Census of the Population. Alphabetical List of Industries and Occupations.*[12] The occupational titles are derived from the US Department of Labor *Dictionary of Occupational Titles.*[13] This coding system is useful for grouping occupational information from death certificates into broad categories, because most often the occupational information obtained from death certificates is not specific. Also, the codes can be linked to specific exposures. The coding system and occupational cluster analysis is well described by Hsieh, *et al.*[11]

The data were analyzed to see if the five-digit codes obtained via death certificate agreed with those obtained via interview, and also to determine if industry, occupation, and cluster from the two sources agreed. Agreement was also assessed by subjectively comparing each occupation and industry entry from the two sources without regard to the five-digit codes.

The categories for agreement were: 1) agreement; 2) two jobs listed on one source, one of which agrees with the other source; 3) housewife on the death certificate and no job given on the interview; 4) no agreement; 5) job listed on the death certificate and no job from interview; and 6) housewife recorded on the death certificate with a job listed on interview. The first three categories can be considered as agreement, and categories 4 through 6 as disagreement. Agreement was analyzed by sex, age, county of residence, and years worked (obtained via interview).

Results

Both death certificate and interview data were available for 184 individuals (94 per cent). The mean age at death was 67 (range 40–84); the years of death were 1977 to 1984; 112 of

*West DW: Diet and Colon Cancer in Man: The Effects of Fiber. National Cancer Institute Grant # 1 R01 CA25580 03.

Address reprint requests to Mary Catherine Schumacher, MD, MSPH, Department of Family and Community Medicine, University of Utah School of Medicine, 50 North Medical Drive, Salt Lake City, UT 84132. This paper, submitted to the Journal July 2, 1985, was revised and accepted for publication November 8, 1985.

© 1986 American Journal of Public Health 0090-0036/86$1.50

the individuals were male, and 75 per cent were from urban counties. Years worked (obtained from the interview) ranged from 1 to 64 years, with a mean of 24 years.

The five-digit code agreed exactly for 75 of the 184 individuals (41 per cent). Including those with two jobs listed on one or the other source, one of which agreed with the other source, increased the number that agreed to 88 (48 per cent). Finally, assuming that housewife on the death certificate and no occupation on the interview represented agreement, yielded an overall agreement of 63 per cent. Of the 37 per cent for whom the two sources disagreed, 10 per cent were those for whom housewife was listed on the death certificate, and a job was listed on the interview. In a subjective evaluation of the data, the overall agreement increased to 73 per cent because in some cases visual examination concluded that the two entries were similar, even though the five-digit codes were different.

The data were analyzed for occupational "clusters", as described by Hsieh, *et al.*[11] These are groups of occupations combined because of similar exposures. If misclassification of occupation occurred on the death certificate, causing an incorrect five-digit code, but if that misclassification placed the individual into the correct exposure group, then it should be less important.

The results for agreement of death certificate cluster with that obtained via interview (65 per cent) were similar to those obtained for agreement of the five-digit code (63 per cent). In other words, the agreement by "cluster" was no better than agreement for the five-digit code, and the misclassification did place the individual into a different exposure group.

The agreement for industry alone was 67 per cent, and agreement for occupation alone was 68 per cent. The similarity in agreement between the five-digit code, industry alone, and occupation alone indicates that most often, when the two sources differed, *both* the occupation and the industry were different.

There were no differences in overall per cent agreement when the data were analyzed by age, sex, and urban-rural county of residence. If an occupation was listed on a woman's death certificate, it was accurate 76 per cent of the time, compared to 67 per cent for men. However, 40 per cent of those women listed on the death certificates as housewives had some other job listed on interview. There was no difference in per cent agreement between those dying at ages less than 65 years, and those dying at greater than or equal to 65 years.

Table 1 shows that the per cent agreement increased with years worked. Those with greater than 35 years of employment had a 82 per cent agreement, and those with 1–14 years agreed only 40 per cent of the time. However, most of the discrepancy for those in the 1–14 year group occurred because housewife was listed on the death certificate and another job was listed on the interview. Excluding housewives from the analysis improved the per cent agreement for the group working 1–14 years to 67 per cent.

The frequencies of the 18 industry categories and of the 30 occupational clusters obtained from death certificates and interviews were compared, and the results were similar. The low-exposure industry category (99) contained 30 per cent of those occupations identified by death certificate, and 35 per cent of those from interview. Similarly, 40 per cent of occupations taken from death certificates were classified as having low exposures by the occupational cluster (clusters 0 and 1), and 41 per cent from interview were so classified.[11] There was no major discrepancy in the frequencies of the other industry or cluster groups, comparing interview with death certificate data.

TABLE 1—Agreement of Death Certificate Occupation and Industry with Those Obtained via Interview, Analyzed by Years Worked, Using Five-Digit Occupation-Industry Codes*

	Years Worked					
	1–14		15–34		35+	
Agreement Code	N	%	N	%	N	%
Codes identical	10	29	38	49	25	76
Two jobs	4	11	6	8	2	6
Housewife/no job*						
Total agreement	14	40	44	57	27	82
No agreement	7	20	28	36	6	18
No job*						
Housewife/job	14	40	5	7	0	
Total disagreement	21	60	33	43	6	18
Total	35	100	77	100	33	100

*There were three individuals for whom the "years worked" was missing, 27 individuals with "housewife" on death certificate and no job on interview, and nine individuals with a job on the death certificate and no job on the interview.

In addition, the percentage of agreement for each industry group, and for each cluster group, were similar. These results indicated that the misclassification of occupation and industry on death certificates occurred in a random manner.

Discussion

The results of this study are similar to those from previous studies where the methods used to determine agreement are described. Buechley, *et al*, compared interview data obtained prior to death to death certificate data for 518 male lung cancer patients, aged 35–64.[4] They used three-digit codes from the US Bureau of Census Alphabetical List of Occupations and Industries[14] and found that usual occupation agreed only 51 per cent of the time, and last occupation agreed with the death certificate 70 per cent of the time. Agreement improved somewhat if two- and one-digit codes were used. Steenland and Beaumont compared NIOSH records for 2,198 long-term employees of several industries with the death certificate records, using as criteria for agreement a match in the three-digit Bureau of Census codes, and found occupation agreed only 64 per cent of the time, and industry 70 per cent.[8] Results were lower for women and non-Whites, and higher for White males.

Other often quoted studies list occupation and industry as being about 75 per cent accurate but the methods to determine agreement are not given.[5–7] Subjective evaluation of the two data sources in this study showed agreement for 73 per cent, similar to the 75 per cent accuracy reported in the literature. These similarities indicate that death certificate accuracy for occupation and industry may at least be consistent from state to state.

In this study, the occupation and industry listed on the death certificate was the "usual" occupation and industry, and the interview asked about employment in the last 15 years or prior to age 65. This may have resulted in some discrepancies between the two data sources.

In determining death certificate accuracy, this study assumed that the interview data were correct, and that discrepancies between the two sources were attributable to death certificate inaccuracies. There could be errors in the

interviews as well, either in recording of the answers, or in the patient's understanding of the question. Also, it is entirely possible that both sources are correct, in that often people have more than one job during their lives.

The misclassification of occupation and industry on death certificates appeared to occur in a random manner, because there was no association of misclassification with any particular industry or "cluster," and because the overall frequencies of industries and "clusters" were similar for the two data sources. Random misclassification causes a bias towards the null.[15] The degree of this effect depends on the magnitude of the "true" odds ratio, and the proportion of controls with the occupation of interest.

Some of the problems with death certificate data could be corrected if occupation and industry data were collected more specifically. Minor changes in phrasing can cause coding changes. Sometimes the occupation is not listed, or is listed nonspecifically, such as "steelworker". Other times the industry is not listed. Still, if death certificates are available and relatively inexpensive, their use for pilot studies seems appropriate, especially when the occupation of interest is prevalent in the population.

ACKNOWLEDGMENT

This research was presented at the Prevention '85 meeting in Atlanta, Georgia, March 30, 1985.

The author would like to acknowledge Dee West, PhD, and Joseph L. Lyon, MD, MPH, for their help with this study, and for their permission to use the data from the two previous studies. The first study of colon cancer and diet was supported by NCI Grant# 1 R01 CA21007 01, and the second by NCI Grant# 1 R01 CA25580 03. The author would also like to acknowledge the editorial assistance of Nancy B. Johnson, BS.

REFERENCES

1. Milham S: Quality of death certificate occupation information (letter). JNCI 1985; 74:539.
2. Dubrow R, Wegman DH: Setting priorities for occupational cancer research and control: Synthesis of the results of occupational disease surveillance studies. JNCI 1983; 71:1123–1142.
3. Kaminski R, Brockert J, Sestito J, Frazier T: Occupational information on death certificates: a survey of state practices. Am J Public Health 1981; 71:525–526.
4. 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.
5. Petersen GR, Milham S: Hodgkin's disease mortality and occupational exposure to wood. JNCI 1974; 53:957–958.
6. Wegman DH, Peters JM: Oat cell lung cancer in selected occupations. JOM 1978; 20:793–796.
7. 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 Health 1984; 74:464–467.
8. Steenland K, Beaumont J: The accuracy of occupation and industry data on death certificates. JOM 1984; 26:288–296.
9. Sorenson AW: Methodology and strategies for nutritional epidemiology studies using a diet and colon cancer model. *In:* Cairns J, Lyon JL, Skolnik M (eds): Banbury Report 4: Cancer Incidence in Defined Populations. Cold Spring Harbor Laboratory, 1980.
10. Hoar SK, Morrison AS, Cole P, Silverman DT: An occupation and exposure linkage system for the study of occupational carcinogenesis. JOM 1980; 22:722–726.
11. Hsieh CC, Walker AM, Hoar SK: Grouping occupations according to carcinogenic potential: occupation clusters from an exposure linkage system. Am J Epidemiol 1983; 117:575–589.
12. US Bureau of the Census: 1970 Census of population. Alphabetical index of industries and occupations. Washington, DC: Govt Printing Office, 1971.
13. US Department of Labor: Dictionary of Occupational Titles, 4th Ed. Washington, DC: Govt Printing Office, 1977.
14. US Bureau of the Census: Alphabetical index of occupations and industries. Washington, DC, US Govt Printing Office, 1950.
15. Kleinbaum DG, Kupper LL, Morgenstern H: Epidemiologic Research. Belmont, CA: Lifetime Learning Publications, 1982; 221–240.

Better Care of Persons in Middle Life (circa 1913)

We may now summarize our discussion of the factors which enter into the possible reduction of mortality at the middle ages as follows:

1. We must place even greater emphasis upon the municipal control of the communicable diseases of early life in order to reduce the instances of heart and kidney impairments which often result therefrom.

2. We must encourage the movements directed against the spread of venereal disease as well as against the intemperate use of alcoholic beverages.

3. We must further all efforts for the improvement of adequate labor legislation and promote better understanding between employers and employees. This programme will include the improvement of factory sanitation, the medical examination of employees and the instruction of both employers and employees in industrial hygiene.

4. It will be necessary to supplement labor legislation with the careful examination of death certificates to see that in every instance those who are responsible for preventable deaths are properly prosecuted.

5. Finally, we must heartily encourage the movement for public education on all topics connected with personal hygiene that there may be better cooperation between physicians and their patients and that there may be no unnecessary losses sustained through neglect of symptoms pointing to serious organic diseases.

—Dublin LI: Possibilities of reducing mortality at higher age groups. Am J Public Health 1913; 3:1262–1271.