

## Comparison of Death Certificate Occupation and Industry Data With Lifetime Occupational Histories Obtained by Interview: Variations in the Accuracy of Death Certificate Entries

William J. Schade, MPH, and G. Marie Swanson, PhD, MPH

---

This study compares usual and recent occupation and industry data from lifetime work histories obtained by interview with death certificate entries for occupation and industry for 2,435 persons diagnosed with cancer. Match rates are calculated as the percent of death certificate occupation and industry entries that were confirmed by interview data and are compared for exact 3-digit 1980 U.S. Census Bureau occupation and industry codes and for groups of these codes. The overall match rate for individual usual occupation codes was 47.9% and for exact usual industry codes it was 61.8%. Significant differences between the interview data for usual occupation or industry and the death certificate entry were observed by race and gender, marital status, number of years worked, and occupation and industry groups and by age for industry. Misclassification or overreporting of occupation and industry data on the death certificate ranged from 30 to 50% in this study. Our results suggest that the utility of death certificate data for investigations into the occupational risk factors for cancer may be quite limited.

**Key words:** occupational epidemiology, neoplasms, methods

---

### INTRODUCTION

Increasing interest in the relationship between occupation and disease has prompted researchers to compare the quality of occupational data obtained from various sources. Most studies in occupational cancer epidemiology have utilized death certificate data to obtain the subjects' occupation and industry. Questions have been raised as to whether occupational data obtained from death certificates adequately represent decedents' occupational histories [Illis et al., 1987]. In cancer etiology, the best summary measure of occupation and industry is thought to be the usual occupation and industry, defined as the occupation and industry held for the longest number of years of the total work history. Investigations of the quality of death certificate data began in the mid-1950s, evaluating both the completeness of reporting and accuracy of occupation and industry data [Buechley et al., 1956;

Division of Epidemiology, Michigan Cancer Foundation, Detroit.

Address reprint requests to Dr. G.M. Swanson, Division of Epidemiology, Michigan Cancer Foundation, 110 E. Warren Avenue, Detroit, MI 48201.

Accepted for publication November 30, 1987.

Alderson, 1972; Wegman and Peters, 1978; Frazier and Wegman, 1979; Rosenberg et al., 1979; Rousch et al., 1980; Swanson et al., 1984; Balarajan, 1985; Gute and Fulton, 1985; Steenland and Beaumont, 1984; Schumacher, 1986; Turner et al., 1987], as well as the codability of these entries [Kaminski et al., 1981]. Many of these studies have compared one or more occupations and industries obtained by interview with the death certificate entries [Buechley et al., 1956; Alderson, 1972; Wegman and Peters, 1978; Swanson et al., 1984; Schumacher, 1986; Turner et al., 1987], while others have compared death certificate data with cancer registry data [Swanson et al., 1984; Balarajan, 1985], city directory information [Rousch et al., 1980], company or union work records [Steenland and Beaumont, 1984], or both interview and city directory data [Gute and Fulton, 1985]. Most studies have utilized usual occupation from a lifetime work history as the indicator against which death certificate entries were compared [Alderson, 1972; Wegman and Peters, 1978; Swanson et al., 1984; Steenland and Beaumont, 1984; Rousch et al., 1980; Gute and Fulton, 1985], while two studies utilized both usual and last occupation [Buechley et al., 1956; Turner et al., 1987]; another very limited study combined occupation and industry and matched the combination with either of two "jobs" from the most recent 15 years of the occupational history [Schumacher, 1986]; three studies also included usual industry from interview or industry records [Swanson et al., 1984; Steenland and Beaumont, 1984; Gute and Fulton, 1985]; one utilized last industry from city directory data [Gute and Fulton, 1985]; and one very small study (87 cases) obtained both usual and recent industry from interview data [Turner et al., 1987].

The coding methods utilized to determine whether the death certificate information matched that on the comparison source varied across these studies. The most specific and detailed approach utilized the U.S. Census Bureau's 3-digit occupation and industry codes [U.S. Department of Commerce, 1982]. This method also is the most widely utilized throughout published reports of occupational epidemiology [Dubrow and Wegman, 1983]. Specific 3-digit Census Bureau codes were utilized in five studies [Buechley et al., 1956; Wegman and Peters, 1978; Swanson et al., 1984; Steenland and Beaumont, 1984; Gute and Fulton, 1985], while one study combined the specific Census Bureau codes into grouped categories [Turner et al., 1987]. Two studies conducted in the United Kingdom utilized the United Kingdom's Registrar General's Classification of Occupations [Alderson, 1972; Balarajan, 1985]; a study conducted in Utah utilized a code developed by Hoar [Hoar et al., 1980] that combines Census Bureau industry codes with Labor Department occupation codes and adds an exposure code [Schumacher, 1986]; and a third study coded the work histories in terms of whether there had been exposure to nickel, cutting oils, or wood dust [Rousch et al., 1980]. The accuracy of agreement with usual occupation reported in these studies ranges from 45 to 95%; for industry the match rates range from 62 to 72%. The majority of these studies included 300 or more subjects [Buechley et al., 1956; Alderson, 1972; Rousch et al., 1980; Swanson et al., 1984; Steenland and Beaumont, 1984; Gute and Fulton, 1985; Balarajan, 1985], although three studies included fewer than 200 subjects [Wegman and Peters, 1978; Schumacher, 1986; Turner et al., 1987].

Only four of the nine studies reviewed included females [Swanson et al., 1984; Steenland and Beaumont, 1984; Gute and Fulton, 1985; Schumacher, 1986]. Variation in the accuracy of death certificate occupational data by race was reported by only one study [Steenland and Beaumont, 1984]; one evaluated the death

certificate entries by education [Turner et al., 1987]; one assessed death certificate occupational data by occupation and industry group [Steenland and Beaumont, 1984]; and none has analyzed these data by marital status or cancer site. The purpose of this report is to examine the accuracy of occupation and industry information from death certificates for a sample of incident cancer cases. Death certificate entries for occupation and industry are compared with lifetime work histories obtained by interview. Comparisons are made by race and gender, marital status, length of employment, and specific occupational and industry groups.

## MATERIALS AND METHODS

The Occupational Cancer Incidence Surveillance System (OCISS) is a population-based study of occupational factors in the etiology of cancer in the Detroit metropolitan area. OCISS is integrated with the Michigan Cancer Foundation (MCF) Division of Epidemiology's Metropolitan Detroit Cancer Surveillance System (MDCSS), a founding participant in the National Cancer Institute's SEER (surveillance, epidemiology, and end results) program [Swanson and Brennan, 1981]. Case-finding is conducted through the MDCSS rapid reporting system, which identifies cases within 2 to 4 weeks after diagnosis. Incident cases of cancers of the esophagus, lung, colon, rectum, urinary bladder, liver, eye, and salivary gland, and mesotheliomas occurring in black and white males and females between the ages of 40 and 84 are included in OCISS. Telephone interviews are conducted with the cases or with a surrogate (spouse or other close relative of the case) to obtain complete lifetime occupational and smoking histories, demographic data, health history, and residential history. The response rate to date for this study is 94.1%. There were 2,519 cases selected for analysis diagnosed between November of 1984 and June of 1986. These cases include all decedents in the study between November of 1984 and January of 1987.

The lifetime work history obtained includes the occupation and industry titles of all jobs ever held, a complete description of the duties performed, the dates each job began and ended, and whether the job was full- or part-time. Death certificate data were obtained from microfilm copies of the full Michigan death certificate. The Michigan death certificate form includes an entry for usual occupation and industry. The occupation and industry data from both the interviews and death certificates were coded according to the 3-digit codes of the 1980 U.S. Census Bureau classification [U.S. Department of Commerce, 1982]. Staff of the MCF Division of Epidemiology performing this coding have been trained by National Institute of Occupational Safety and Health staff and the coding supervisor has attended U.S. Census Bureau training courses. Thus, comparisons made between the death certificates and interviews are based upon codes assigned by the same coding staff. Coding for both the interviews and death certificates was conducted separately and without identifiers; thus assignment of codes was blinded in terms of the diagnosis and cause of death. MCF coders code death certificate information regarding occupation and industry for all cancer deaths each year; thus cases in this study were selected from the computer files of all deaths for the years 1984–1987. This further blinded the coding process, since it was performed as a routine function, rather than specifically for this analysis. Usual occupation and industry from the interview data are defined by summing the total number of months a person was employed in a specific industry or occupation over

the entire work history and then selecting the occupation and industry for which the person had accumulated the largest number of months of employment. Recent occupation and industry are defined as the job held immediately prior to retirement or before leaving work due to illness.

Of the 2,519 cases selected for analysis, occupation or industry information was missing from either the death certificate or the interview for 3.3% of cases, which were excluded from the analysis. The remaining 2,435 cases were then analyzed in terms of four specific categories of matching between the interview and death certificate data: 1—an exact match between the death certificate 3-digit code for occupation or industry with the 3-digit code for usual occupation or usual industry from the interview; 2—a match between death certificate data and interview data for 16 groups of codes for usual industry and 17 groups of codes for usual occupation (Appendix A); 3—an exact match of the death certificate 3-digit codes with the 3-digit code for recent occupation or recent industry from the interview; or 4—a match between the 16 industry groups and 17 occupation groups (Appendix A) between the death certificate data and the recent occupation or industry from the interview. Through these procedures, match rates define the extent to which death certificate occupation and industry codes were confirmed by comparison with usual or recent occupation or industry obtained by interview. In Tables I–IV, percentages were calculated first by selecting all entries on the death certificate for occupation or industry that were confirmed by the exact 3-digit code from the interview for usual occupation or industry; then, cases that did not match were compared to the broad group of usual industry or occupation data from the interview; next, of those cases with no match in either of the first two categories, matches between the death certificate 3-digit entry and the 3-digit code of the interview entry for recent occupation or industry were obtained; and finally, cases that remained with no match in any of the first three categories were matched with the broad groups of recent occupation and industry from the interview. This procedure was utilized to emphasize the match rates for usual occupation and industry, since the death certificate form requests usual employment data and since studies in cancer epidemiology require information about usual, rather than recent, occupation and industry in order to encompass an adequate latency period. The variation in proportions across the four matching categories was tested utilizing the Pearson chi-square [Dixon et al., 1985].

This study reports variations by demographic and occupational variables regarding the extent to which death certificate entries for “usual” occupation and industry actually indicate the usual occupation and industry of incident cancer cases who died during the first 27 months of the study.

## RESULTS

Of the 2,435 cases included in the analysis, 66.7% were incident cases of lung cancer, 15.7% were colon cancer, 4.9% were esophageal cancers, 4.5% were cancers of the urinary bladder, 3.9% were cancers of the rectum, 3.6% were cancers of the liver, and less than 1% were mesotheliomas, salivary gland cancers, or cancers of the eye. Death certificate information for cause of death or any underlying cause of death specified cancer for 83.2% of cases. No significant differences were observed for the match rate between death certificate occupation or industry data and the interview

**TABLE I. Proportion of Matches Between Death Certificate Occupation and Interview Usual and Recent Occupation by Race and Gender, Age at Diagnosis, Education, Marital Status, and Years Worked**

Subgroup	Total no. from death certificate subgroups (denomi- nators)	Usual occupation		Recent occupation	
		1. % exact <sup>a</sup> match	2. % grouped <sup>b</sup> match (for those not matched in column 1)	3. % exact <sup>a</sup> match (for those not matched in columns 1 or 2)	4. % grouped <sup>b</sup> match (for [denominators] matched in column 1, 2, or 3)
1. Race and gender					
White males	1,238	47.7	12.9	5.3	2.2
White females	633	60.0	4.4	6.8	2.1
Black males	369	28.5	13.0	6.8	2.2
Black females	173	49.1	2.3	11.6	1.2
p < .0001					
2. Age at diagnosis					
40-49	109	35.8	11.9	9.2	3.7
50-59	440	47.0	9.8	6.6	2.5
60-69	881	47.3	11.3	6.5	2.2
70-79	759	49.7	8.8	6.5	2.1
80-84	246	51.6	8.9	3.3	0.8
p < .3188					
3. Years of education					
1-9	607	45.6	11.4	6.4	2.0
10-12	1,170	48.7	9.2	7.3	2.4
13-16	390	48.7	10.8	5.9	2.3
17+	47	66.0	19.1	0.0	2.1
p < .0574					
4. Marital status					
Single	128	32.0	14.1	5.5	2.3
Divorced	306	37.3	13.4	5.0	2.3
Widowed	551	50.5	7.1	5.6	2.5
Married	1,445	50.6	10.0	6.9	1.9
p < .0001					
5. Years worked in usual or recent occupation					
1-9	52	30.8	13.5	7.7	3.8
10-19	234	26.5	9.8	11.1	3.4
20-29	583	37.2	11.5	10.8	3.6
30-39	655	45.6	12.4	6.1	2.3
40-49	503	56.3	8.3	3.2	0.8
50+	376	73.1	4.5	1.1	0.5
Total	2,435	47.9	10.0	6.3	2.1
p < .0001					

<sup>a</sup>Match on 3-digit Census Bureau codes.<sup>b</sup>See Appendix A for occupation groups.

usual or recent occupation and industry data by cancer site or by source of interview (case or surrogate respondent; 67% of interviews were conducted with cases).

In Tables I and II the exact and grouped (see Appendix A) match proportions between the death certificate entries for occupation and industry and the interview information for usual and recent occupation and industry are presented. These data show the match proportions for each matching category among cases that did not

**TABLE II. Proportion of Matches Between Death Certificate Industry and Interview Usual and Recent Industry by Race and Gender, Age at Diagnosis, Education, Marital Status, and Years Worked**

Subgroup	Total no. from death certificate subgroups (denomi- nators)	Usual industry		Recent industry	
		1. % exact <sup>a</sup> match	2. % grouped <sup>b</sup> match (for those not matched in column 1)	3. % exact <sup>a</sup> match (for those not matched in column 1 or 2)	4. % grouped <sup>b</sup> match (for [denominators] matched in column 1, 2, or 3)
1. Race and gender					
White males	1,238	62.0	13.6	5.0	1.4
White females	633	63.7	5.1	8.8	1.7
Black males	369	60.4	6.0	8.4	0.8
Black females	173	57.8	0.6	14.5	0.6
p < .0001					
2. Age at diagnosis					
40-49	109	45.9	11.0	13.8	0.9
50-59	440	62.3	9.8	8.4	0.5
60-69	881	59.7	9.6	7.3	1.5
70-79	759	65.1	9.0	6.6	2.1
80-84	246	65.4	8.9	4.5	0.4
p < .0055					
3. Years of education					
1-9	607	63.4	8.1	6.6	1.6
10-12	1,170	60.9	9.9	8.3	1.4
13-16	390	60.0	8.5	7.2	1.0
17+	47	68.1	17.0	2.1	2.1
p < .2834					
4. Marital status					
Single	128	47.7	10.9	6.2	3.9
Divorced	306	49.3	12.0	9.3	1.3
Widowed	551	61.5	6.0	8.3	1.5
Married	1,445	65.7	9.7	6.6	1.1
p < .0001					
5. Years worked in usual or recent industry					
1-9	41	34.1	4.9	7.3	0.0
10-19	221	36.2	10.4	12.2	1.4
20-29	514	49.2	13.2	15.0	3.7
30-39	650	64.9	8.6	7.4	0.9
40-49	563	70.0	8.3	3.4	0.7
50+	414	76.7	6.3	0.7	0.2
Total	2,435	61.8	9.2	7.3	1.4
p < .0001					

<sup>a</sup>Match on 3-digit Census Bureau codes.<sup>b</sup>See Appendix A for occupation groups.

match in the previous column. Thus, the main criterion for matching is presented in the first column—an exact match between the 3-digit code from both sources for usual occupation or industry—followed by a grouped match for usual occupation or industry. The last two columns present exact matches between death certificates and interview 3-digit codes for recent occupation or industry and then grouped matches for recent occupation or industry. The overall rate of matching for exact usual

occupation was 47.9%, and for exact usual industry it was 61.8%. Matches with either exact or grouped usual occupation rose to 57.9%, while for industry this rate was 71.0%. The total matching proportion across all four categories of matches was 66.3% for occupation and 79.7% for industry. Matches with exact code for recent occupation increase the overall proportion matching by a range of 1 to 11%; for industry this proportion ranges from 2 to 14%. For every category of each variable for which matching was evaluated, the proportion of matches with exact usual industry codes was higher than that for exact usual occupation code (Tables I and II). The match rates varied significantly by race, gender, marital status, years worked for both occupation and industry, and age at diagnosis for industry. White females have the highest exact match proportions for usual occupation as well as for usual industry, black males have the lowest exact match proportion for usual occupation, and black females have the lowest exact match proportion for usual industry. Considering those matching usual occupation for either the exact or grouped categories, white females have the highest proportion (64.4%) and black males have the lowest (41.5%); for industry matches with either exact or grouped usual industry, white males have the highest proportion (75.6%), while black females have the lowest (58.4%; Tables I and II). When one excludes housewives from the matching criteria for women, the proportion matching exact usual occupation decreases from 57.7% for all females including housewives to 30.6% for all females excluding housewives; including matches with either exact or grouped usual occupation increases this proportion to 38.8%. This pattern is identical for black and white females.

Matches with exact usual occupation increase with age at diagnosis, ranging from 35.8% at 40–49 years to 51.6% at 80–84 years. For industry, the age differences are significant, ranging from 45.9% at 40–49 to 65.4% at 80–84. Although the differences in matching by education are not statistically significant, persons with 17 or more years of education have the highest proportion matching exact 3-digit codes for both usual occupation and usual industry, as well as for matches with either exact or grouped usual occupation or industry. The variation in matches by marital status is significant, with married persons having the highest matching proportions with exact usual occupation (50.6%), exact or grouped usual occupation (60.6%), exact usual industry (65.7%), and exact or grouped usual industry (75.4%).

The proportion of matches also varies significantly by the number of years worked in the usual occupation or industry, with a tendency for the proportion matching usual occupation or industry to increase with increasing number of years employed. The highest proportion matching the exact 3-digit code for usual occupation or industry occurs among persons who accumulated 50 or more years worked. Among those who worked for 50 years or more, the match proportion for exact plus grouped usual occupation was 77.6%; for exact plus grouped usual industry, it was 84.3% (Tables I, II).

Analysis of matching between death certificate entries and interview data by occupational groups shows a range from a 14.0% match on exact usual occupation for laborers to 63.3% match for transportation workers and material movers among males (Table III). Among females, the highest matches with exact usual occupation are for housewives (82.6%) and for private household workers (50.0%), while the lowest are for laborers (0.0%), technicians (25.0%), and sales workers (25.6%). Matches with either exact or grouped usual occupation is 70% or greater for only two occupational

**TABLE III. Proportion of Matches Between Death Certificate Occupation and Interview Usual and Recent Occupation by Occupational Group and Gender**

Subgroup	Total no. from death certificate subgroups (denomi- nators)	Usual occupation		Recent occupation	
		1. % exact <sup>a</sup> match	2. % grouped <sup>b</sup> match (for those not matched in column 1)	3. % exact <sup>a</sup> match (for those not matched in column 1 or 2)	4. % grouped <sup>b</sup> match (for [denominators] matched in column 1, 2, or 3)
<b>1. Males</b>					
Administration and managers	136	46.3	8.1	6.6	2.2
Professionals	104	51.9	13.5	1.0	1.9
Technicians	21	33.3	9.5	4.8	9.5
Sales workers	93	48.4	15.1	5.4	6.5
Administrative support	80	28.7	18.7	6.2	3.7
Protective services	28	42.9	21.4	17.9	7.1
Other services	100	38.0	7.0	15.0	3.0
Farming and fishing	8	37.5	12.5	37.5	0.0
Mechanics	108	50.0	13.9	5.6	2.8
Construction workers	118	56.8	8.5	4.2	0.8
Precision production workers	220	46.8	12.3	7.7	1.4
Machine operators, assemblers, and inspectors	214	50.5	23.8	3.7	2.8
Transportation workers and material movers	139	63.3	12.2	4.3	1.4
Laborers	207	14.0	10.6	1.4	0.5
Students and retirees	44	15.9	0.0	2.3	0.0
Total	1,620	43.1	13.0	5.5	2.3
p < .0001					
<b>2. Females</b>					
Administration and managers	23	30.4	4.3	4.3	0.0
Professionals	48	41.7	2.1	4.2	0.0
Technicians	8	25.0	0.0	0.0	0.0
Sales workers	43	25.6	4.7	11.6	11.6
Administrative support	82	35.4	17.1	7.3	3.7
Private household services	16	50.0	0.0	0.0	0.0
Other services	68	32.4	11.8	11.8	4.4
Precision production workers	10	40.0	10.0	20.0	0.0

(continued)

**TABLE III. Proportion of Matches Between Death Certificate Occupation and Interview Usual and Recent Occupation by Occupational Group and Gender (Continued)**

Subgroup	Total no. from death certificate subgroups (denomi- nators)	Usual occupation		Recent occupation	
		1. % exact <sup>a</sup> match	2. % grouped <sup>b</sup> match (for those not matched in column 1)	3. % exact <sup>a</sup> match (for those not matched in column 1 or 2)	4. % grouped <sup>b</sup> match (for [denominators] matched in column 1, 2, or 3)
2. Females (continued)					
Machine operators, assemblers, and inspectors	43	32.6	11.6	4.7	7.0
Laborers	23	0.0	0.0	0.0	4.3
Students and retirees	18	0.0	0.0	0.0	0.0
Housewives	420	82.6	0.0	8.6	0.0
Total	802	57.5	4.0	7.8	1.9
p < .0001					

<sup>a</sup>Match on 3-digit Census Bureau codes.

<sup>b</sup>See Appendix A for occupation groups.

groups among males—machine operators, assemblers, and inspectors; and transportation workers and material movers. For occupations other than housewife among women, only three groups achieve at least a 50% match with exact or grouped usual occupation—administrative support workers, private household workers, and precision production workers. Matches with exact recent occupation increase the overall proportion of matches by a range of 1 to 37% for males and by 0 to 20% for females. Utilizing matches with any of the four categories of usual or recent occupation increases matching to 80% or greater for four of the 15 groups for men and to 70% for two of 11 groups other than housewife among women (Table III).

Matches on exact usual industry range from rates of over 70% among manufacturing of nondurable products and transportation, communication, and public utilities to less than 30% for wholesale sales of durable and nondurable products and nonworkers among men (Table IV). Among women in industries other than housewife, matches of 50% or greater are observed for three groups: finance, insurance, and real estate; personal services; and transportation, communications, and public utilities; and matches of 20% or less are observed among manufacturing of nondurable products, public administration, and business and repair services. Adding matches with grouped usual industry to those with exact usual industry increases matching to 75% or greater for three groups among males—manufacturing of durable products; manufacturing of nondurable products; and transportation, communication, and public utilities—and to 55% or greater among three groups among females—transportation, communications, and public utilities; finance, insurance, and real estate; and personal services. Matches with exact recent industry increase the overall proportion of matches by a range of 2 to 23% for males and by 0 to 20% for females. Matches with any of the four categories of usual or recent industry increase matching to 80% or greater for only five of the 15 groups for men and to 70% or greater for one of the nine groups other than housewife for women.

TABLE IV. Proportion of Matches Between Death Certificate Industry and Interview Usual and Recent Industry by Industry Group and Gender

Subgroup	Total no. from death certificate subgroups (denomi- nators)	Usual industry		Recent industry	
		1. % exact <sup>a</sup> match	2. % grouped <sup>b</sup> match (for those not matched in column 1)	3. % exact <sup>a</sup> match (for those not matched in column 1 or 2)	4. % grouped <sup>b</sup> match (for [denominators] matched in column 1, 2, or 3)
1. Males					
Agriculture, forestry and fishing	6	33.3	33.3	16.7	0.0
Construction	169	68.6	0.0	4.1	0.0
Manufacturing, nondurable products	98	74.5	2.0	4.1	2.0
Manufacturing, durable products	780	67.3	18.6	3.8	1.5
Transportation, communications and public utilities	133	70.7	6.8	6.0	0.8
Wholesale trade, nondurable products	18	27.8	0.0	5.6	0.0
Wholesale trade, durable products	27	29.6	14.8	0.0	3.7
Retail trade	103	52.4	12.6	7.8	0.0
Finance, insurance, and real estate	34	52.9	2.9	23.5	2.9
Business and repair services	53	34.0	1.9	9.4	0.0
Personal services	23	60.9	0.0	8.7	0.0
Entertainment	13	46.2	7.7	23.1	0.0
Professional services	80	46.2	6.2	21.2	2.5
Public administration	53	37.7	13.2	1.9	1.9
Retiree, student, housewife	33	24.2	3.0	3.0	0.0
Total	1,623	61.5	11.8	5.9	1.2
p < .0001					
2. Females					
Manufacturing, nondurable products	25	20.0	16.0	12.0	0.0
Manufacturing, durable products	67	38.8	9.0	10.4	6.0
Transportation, communications and public utilities	17	52.9	5.9	5.9	0.0
Retail trade	68	35.3	13.2	8.8	5.9

(continued)

**TABLE IV. Proportion of Matches Between Death Certificate Industry and Interview Usual and Recent Industry by Industry Group and Gender (Continued)**

Subgroup	Total no. from death certificate subgroups (denomi- nators)	Usual industry		Recent industry	
		1. % exact <sup>a</sup> match	2. % grouped <sup>b</sup> match (for those not matched in column 1)	3. % exact <sup>a</sup> match (for those not matched in column 1 or 2)	4. % grouped <sup>b</sup> match (for [denominators] matched in column 1, 2, or 3)
2. Females (continued)					
Finance, insurance, and real estate	13	61.5	7.7	7.7	0.0
Business and repair services	15	13.3	6.7	20.0	0.0
Personal services	33	54.5	3.0	0.0	3.0
Professional services	115	42.6	6.1	18.3	0.9
Public administration	15	20.0	13.3	6.7	13.3
Retiree, student, housewife	430	83.3	0.0	8.8	0.0
Total	798	62.9	4.0	10.2	1.5
p < .0001					

<sup>a</sup>Match on 3-digit Census Bureau codes.

<sup>b</sup>See Appendix A for industry groups.

Another way of looking at the accuracy of death certificate data is to assess the proportion of matches either with exact 3-digit code for usual occupation or with most recent occupation (Table V). The variation by race and gender for this level of matching is considerable. Black males have the highest proportion of no matches (64.0%), while white females have the lowest (33.2%). The pattern is somewhat different for industry, with the proportion of no matches about the same for males of either race and for females of both races. Overall exact matches of death certificate data with either recent occupation or recent industry are higher than those for usual occupation or usual industry for each group except white females. Combining usual occupation and usual industry, black males again have the lowest proportion matching both exact 3-digit codes (23.6%), while white females have the highest proportion of matches (55.0%; Table V).

Finally, analysis of the shifts among five major occupational groups for males and females shows a trend toward increasing the proportion of males in white collar or service occupations on death certificates when compared to the interview usual occupation as well as a 50% increase in the proportion of females in white collar occupations on the death certificates, compared to the interview data (Table VI). For males, we see that the least accurate group is farmers, who were most often shifted to blue-collar occupations on the death certificate. The largest movement into white-collar occupations on the death certificate was from blue-collar workers, while movement into service occupations came from each of the other four groups. Among females, the largest shifts occur among women who had the usual occupation of housewife on the interview and moved into white-collar occupations and blue-collar occupations on the death certificate.

**TABLE V. Exact Agreement\* Between Usual or Most Recent Occupation, Industry, and Occupation and Industry Combined From Interviews With Death Certificate Data**

Matching category	White		Black	
	Males (%)	Females (%)	Males (%)	Females (%)
<b>A. Occupation</b>				
1. Match on usual occupations	47.6	60.0	28.5	49.1
2. Match on most recent occupation	49.6	55.9	31.2	52.0
3. No match on either usual or most recent occupation	45.5	33.2	64.0	38.2
<b>B. Industry</b>				
1. Match on usual industry	61.9	63.7	60.4	57.8
2. Match on most recent industry	64.6	61.8	64.2	60.2
3. No match on either usual or most recent occupation	30.9	27.3	30.4	27.7
<b>C. Occupation and industry</b>				
1. Match on both usual occupation and industry	35.1	55.0	23.6	47.4
2. No match on either usual occupation or industry	25.4	31.3	34.7	40.5

\*Exact match on 3-digit U.S. Census Bureau code.

**TABLE VI. Misclassification Shifts in Broad Occupational Categories for Males and Females**

Death certificate occupation	Total from death certificate	Usual occupation from interview				
		White collar workers	Service workers	Farm workers	Blue collar workers	Retirees
<b>Males</b>						
White-collar workers	434	401	—	—	29	4
Service workers	128	10	106	6	5	1
Farm workers	8	2	—	6	—	—
Blue-collar workers	1,006	—	—	14	992	—
Retirees	44	—	—	2	20	22
	Total from interview	413	106	28	1,046	27
<b>Females</b>						
Usual occupation from interview						
Death certificate occupation	Total from death certificate	White collar workers	Service workers	Blue collar workers	Housewives	Retirees
White-collar workers	204	129	9	4	61	1
Service workers	84	—	75	3	6	—
Blue-collar workers	76	—	—	61	15	—
Housewives	420	—	—	—	420	—
Retirees	18	—	1	3	9	5
	Total from interview	129	85	71	511	6

## DISCUSSION

The relative ease and low cost of obtaining death certificates have made them an attractive source of information for occupational disease studies. However, this study demonstrates considerable overreporting in death certificate entries for occupation and industry. Our results are similar to those of other studies which have compared death certificate data with lifetime work histories from interview data [Alderson, 1972; Buechley et al., 1956; Gute and Fulton, 1985]. Steenland and Beaumont [1984] found higher proportions of exact matches with usual occupation (64.7%) and usual industry (70.1%) from industry records. However, their sample included only long-term workers (persons working at least 10 years in these jobs). As this study also indicates, match rates tend to increase with increasing duration of employment. Studies that found higher proportions of death certificates matching usual occupation or industry tended to have smaller sample sizes [Wegman and Peters, 1978; Swanson et al., 1984; Schumacher, 1986], used unusual coding schemes [Rousch et al., 1980; Schumacher, 1986], or included only married whites 60 and older [Gute and Fulton, 1985].

Our results support those of other studies showing that death certificate occupational data more often provide usual occupation or industry for females than males [Gute and Fulton, 1985], for whites compared to nonwhites [Steenland and Beaumont, 1984], and for persons with 10 or more years of employment [Steenland and Beaumont, 1984]. Our study is unique in assessing death certificate occupational data for a population-based sample according to marital status and occupational groups. We demonstrated significant variation in the proportion of matches with usual occupation and industry for both of these variables. These results clearly indicate that the accuracy and utility of death certificate information regarding occupation and industry will depend upon the type of population under consideration. A study sample consisting of predominantly white, married persons who have worked at the same job for 10 or more years will allow a better estimate of their distribution of usual occupations and industries from death certificates than will groups with other characteristics. When housewives are excluded from analyses, the usual occupations and industries of males will be better represented by death certificate data than will females, as demonstrated by this study.

The accuracy and utility of death certificate occupational data also vary by occupational group, with a higher match rate shown by these results for males for the two occupational groups of (1) machine operators, assemblers, and inspectors and (2) transportation workers and material movers, as well as for the groups including manufacturing industries and transportation, communications, and public utility industries. For females, none of the occupations other than housewife achieves even a 60% match rate, while only one industry group approaches 70%—that of finance, insurance, and real estate. There seems to be little question from this analysis that death certificate occupational data should not be utilized to study illness effects of workplace exposures outside the home among females. It also is important to note that, for most occupational groups, recent occupation or industry added little to the overall proportion of cases of death certificates that matched either exact or grouped usual occupation or industry. For studies of cancer and other chronic diseases, the critical summary measure of occupation is usual occupation, because of the long latency period before disease is observed.

These data represent the employment distributions among residents of the Detroit metropolitan area and may differ considerably from those of other areas. Thus, one can expect that there will be geographic variation in the extent to which death certificate data are useful for studies of occupationally induced diseases. Practices in completing the occupational entries on death certificates also vary by state [Dubrow et al., 1987].

This study indicates that it may be useful to group certain occupations or industries into broader categories than the individual 3-digit Census Bureau codes. This strategy enables one to increase the number of cases available for analysis of specific job groups, which is useful if they share known or suspected exposures. However, if such grouping is performed too liberally, then one risks losing the detail required to approximate exposures. The extremely low proportion of death certificates that matched both usual occupation and usual industry from interviews was particularly disappointing, since accurate information about both occupation and industry is essential to estimating exposure from employment data.

Occupational data on death certificates may incorrectly report usual employment for many reasons. As our results suggest, next of kin may promote decedents to higher-status jobs. Further, the person completing the death certificate may not specifically ask the respondent to report the decedent's usual occupation and industry. Although some women may have worked as housewives longer than they worked in other jobs, their relatives may still record the occupation they held either before or after the time they spent as housewives. Finally, death certificates report only titles of occupations and industries, which may not be precise enough for coding exact or grouped categories of usual employment.

## CONCLUSIONS

Our results raise serious questions about the utility of death certificates for studies of occupationally induced disease, especially those such as cancer which require usual employment as the minimum level of accuracy. We have shown that misclassification could easily occur in the range of 30 to 50%, which would compromise results of such studies. Recent discussions call for increased utilization of death certificate occupational data for hypothesis generation and for routine illness surveillance of occupations [Dubrow et al., 1987]. More complete analyses of the accuracy of such data in other geographic areas should precede extensive utilization of these data. Resources for occupational epidemiology are quite limited. If the majority of new hypotheses being developed are derived from a data source that misclassifies or overreports between one-third and one-half of the exposure categories, a large proportion of these resources could be utilized to follow "leads" that are misleading.

## ACKNOWLEDGMENTS

This research is supported in part by National Institute of Occupational Safety and Health grant OH02067, National Cancer Institute contract No. N01-CN-55423, and by the United Foundation of Detroit. The authors wish to thank Ms. Verna Lee Dennert for her careful coding of the occupational and industry data, Ms. Enid Satariano for editing and data management support, and Ms. Valerie Wade for her

careful preparation of the manuscript. The authors are grateful to Dr. William A. Satariano and Mr. William R. Illis for their thoughtful review of the manuscript.

## REFERENCES

- Alderson A (1972): Some sources of error in British occupational mortality data. *Br J Ind Med* 29:245-254.
- Balarajan R (1985): Comparison of occupations recorded at cancer registration and death. *Public Health (Lond)* 99:169-173.
- Buechley R, Dunn JE, Linden G, Breslow L (1956): Death certificate statement of occupation: Its usefulness in comparing mortalities. *Public Health Rep* 71:1105-1111.
- Dixon WJ, Brown MB, Engleman L, Frane JW, Hill MA, Jennrich RI, Toparek JD (1985): "BMDP Statistical Software." Berkeley: University of California.
- Dubrow R, Wegman DH (1983): Setting priorities for occupational research and control: Synthesis of the results of occupational disease surveillance studies. *J Natl Cancer Inst* 71:1123-1142.
- Dubrow R, Sestito JP, Lalich NR, Burnett CA, Salg JA (1987): Death certificate-based occupational mortality surveillance in the United States. *Am J Ind Med* 11:329-342.
- Frazier TM, Wegman DH (1979): Exploring the use of death certificates as a component of an occupational health surveillance system. *Am J Public Health* 69:718-720.
- Gute DM, Fulton JP (1985): Agreement of occupation and industry data on Rhode Island death certificates with two alternate sources of information. *Public Health Rep* 100:65-72.
- Hoar SK, Morrison AS, Cole P, Silverman DT (1980): An occupation and exposure linkage system for the study of occupational carcinogenesis. *J Occup Med* 22:722-726.
- Illis WR, Swanson GM, Satariano E, Schwartz AG (1987): Summary measures of occupational history: A comparison of usual occupation with latest occupation and industry. *Am J Public Health* 77:1532-1534.
- Kaminski R, Brockert J, Sestito J, Frazier T (1981): Occupational information on death certificates: A survey of state practices. *Am J Public Health* 7:525-526.
- Rosenberg HM, Burnham D, Spirtas R, Valdisera V (1979): 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, pp. 83-89.
- Rousch GC, Meigs JW, Kelley J, Flannery JT, Burdo H (1980): Sinonasal cancer and occupation: A case control study. *Am J Epidemiol* 111:183-193.
- Schumacher MC (1986): Comparison of occupation and industry information from death certificates and interviews. *Am J Public Health* 76:635-637.
- Steenland K, Beaumont J (1984): The accuracy of occupation and industry on death certificates. *J Occup Med* 26:288-296.
- Swanson GM, Brennan MJ (1981): Cancer incidence and mortality in Metropolitan Detroit, 1973-77. In Young JL, Percy CL, Asire AJ (eds): "Surveillance, Epidemiology, and End Results: Incidence and Mortality Data, 1973-77." Washington, DC: National Cancer Institute Monograph 57, pp. 219-221.
- Swanson GM, Schwartz AG, Burrows RW (1984): An assessment of occupation and industry data from death certificates and hospital records for population-based cancer surveillance. *Am J Public Health* 74:464-467.
- Turner DW, Schumacher MC, West DW (1987): Comparison of occupational interview data to death certificate data in Utah. *Am J Ind Med* 12:145-151.
- Wegman DH, Peters JM (1978): Oat cell lung cancer in selected occupations. *J Occup Med* 20:793-796.
- U.S. Department of Commerce, Bureau of the Census (1982): "1980 Census of Population: Classified Index of Industries and Occupations, Final Edition." Washington, DC: U.S. Government Printing Office.

**APPENDIX A. List of Occupation and Industry Group Codes**

Groups	3-digit 1980 U.S. Census Bureau Codes
<b>A. Industry groups</b>	
1. Agriculture, forestry, and fishing	010-031
2. Mining	040-050
3. Construction	060
4. Manufacturing, nondurable goods	100-222
5. Manufacturing, durable goods	230-392
6. Transportation, communication, and other public utilities	400-472
7. Wholesale trade, durable goods	500-532
8. Wholesale trade, nondurable goods	540-571
9. Retail trade	580-691
10. Finance, insurance, and real estate	700-712
11. Business and repair services	721-760
12. Personal services	761-791
13. Entertainment and recreation services	800-802
14. Professional and related services	812-892
15. Public administration	900-932
16. Volunteer services, housewife, student, never worked	961
<b>B. Occupational groups</b>	
1. Executive, administrative, and managerial	003-037
2. Professional specialty occupations	043-199
3. Technicians and related support	203-235
4. Sales occupations	243-285
5. Administrative support, including clerical	303-389
6. Private household occupations	403-406
7. Protective service occupations	413-427
8. Service occupations, except protective service and household	433-469
9. Farming, fishing, and forestry	473-499
10. Mechanics and repairers	503-549
11. Construction trades	553-599
12. Precision production occupations	613-699
13. Machine operators, assemblers, and inspectors	703-799
14. Transportation and material moving	803-859
15. Handlers, equipment cleaners, helpers, and laborers	863-889
16. Students, retired, unemployed, volunteers	913, 915-918
17. Housewives	914