AMERICAN JOURNAL OF EPIDEMIOLOGY
Vol. 126, No 2

Printed in U S A

A CASE-CONTROL STUDY OF BLADDER CANCER USING CITY DIRECTORIES AS A SOURCE OF OCCUPATIONAL DATA

KYLE STEENLAND,[1] CAROL BURNETT,[1] AND ANA MARIA OSORIO[2]

Steenland, K. (NIOSH, Cincinnati, OH 45226), C. Burnett, and A. M. Osorio. A case-control study of bladder cancer using city directories as a source of occupational data. *Am J Epidemiol* 1987;126:247–57.

Commercial city directories, currently produced in 1,250 United States cities, potentially provide yearly information on occupation and employer for all city residents over age 18 years. To investigate the usefulness of these occupational data, the authors have conducted a case-control study of male bladder cancer mortality in Hamilton County, Ohio (which includes Cincinnati). A total of 731 bladder cancer cases who died during 1960–1982 were matched on age, sex, race, date-of-death, and residence at death to two controls per case. Risks of bladder cancer death were calculated by occupation, industry, and specific employer, using both city directories (multiple statements) and death certificates (single statement). Four companies showed a significant excess bladder cancer risk when using city directories. Only one would have been identified using death certificates, which ask for usual lifetime type of industry rather than a specific company name. Using city directories, significant positive associations were found between bladder cancer and occupation as an engineer, tailor, carpenter, furnace operator, blending machine operator, chemist, pressing machine operator, house cleaner, or salesman. For industry, the authors found significant positive associations for the textile, chemical, grain mill, foundry, petroleum, building service, entertainment, and advertising industries. A significant increase in risk for those with 20 or more years of employment was seen for those employed as truck drivers and furnace operators, or those employed in the railroad industry. A check of the validity of city directory data indicated that 77 per cent of the listings agreed with Social Security earnings reports for employer in any given year. One limitation of Hamilton County city directory data was the fairly large number of yearly listings without any occupational data (15 per cent for occupation, 36 per cent for employer). While city directory data do provide work history over time, unlike death certificates, such data are available only for years of residence in the city in question.

occupation; bladder cancer; epidemiologic methods

Occupational data from death certificates have been used frequently to study occupationally related disease (1, 2). Death certificate occupational information is limited to "usual lifetime occupation" and "usual lifetime industry," and is provided

Received for publication June 2, 1986, and in final form December 5, 1986.

Abbreviation: ICD, *International Classification of Diseases*.

[1] National Institute for Occupational Safety and Health, Cincinnati, OH.

[2] California State Health Dept., Berkeley, CA.

Reprint requests to Dr. Kyle Steenland, Mailstop R15, National Institute for Occupational Safety and Health, 4676 Columbia Parkway, Cincinnati, OH 45226.

The authors gratefully acknowledge the work of Kim Jenkins and Marcus Wilkins in coding occupational data.

by next-of-kin. This information may not be accurate (3), and fails to reflect the many jobs a person may have held throughout a lifetime.

Commercial city directories are based on a door-to-door yearly census of all residents over age 18 years and include individual data on occupation and employer. They may provide an alternative source for occupational information over the lifetime of a city resident. Such directories have been used only rarely as a source for occupation in an epidemiologic study (4–6). One particular advantage of city directory data is that they identify specific employers. We thought it possible that city directories might offer a better source of occupational data than death certificates for relatively rapid and inexpensive record-based epidemiologic studies which could be conducted in metropolitan areas. City directories are available in most medium-sized cities in the United States.

In order to test the usefulness of city directories, we conducted a case-control study of male bladder cancer deaths in Hamilton County, Ohio, which includes the city of Cincinnati. Occupational information came from both city directories and death certificates. In addition, we also assessed the accuracy and completeness of city directory data by comparing it to work history known from other sources (Social Security and union records).

We chose Hamilton County for our study because it is an industrialized area with high bladder cancer rates (7), because city directories cover approximately 85 per cent of people living in Hamilton County, and because these directories go back in time to the beginning of the century.

Materials and methods

Case-control study methods

Computerized data were obtained from the Ohio Department of Vital Statistics which listed all male bladder cancer deaths (n = 731) and all other male deaths (n = 95,057) in Hamilton County, Ohio from 1960–1982. A nosologist had previously coded all death certificates for cause of death, according to the *International Classification of Diseases* (ICD) revision in use at the time of death.

Cases were male bladder cancer deaths (ICD revision 7 code no. 181, ICD revisions 8 and 9 code no. 188). Controls were a representative sample of all male deaths with the exclusion of deaths from urinary tract tumors and pneumonia. Controls were matched to cases on sex, residence in Hamilton County at the time of death, year of death, age at death (within five years), and race. A pool of six matched controls was created for each case.

In this study, two analyses were performed, one based on city directory data and the other based on death certificate data. In the city directory analysis, cases and controls were restricted to individuals who had at least one yearly directory listing with some occupational data. Pertinent identifying information from death certificates (last name, first name, middle initial, address, and spouse's name) was used to search city directories backwards in time from the date of death of the subject until the first listing was found, and then at five-year intervals, until the decedent would have been 20 years old. For this analysis, controls were required to have listings in the city directories which extended backwards in time to at least within five years of the first listing of their matched cases. The first two controls from the pool of six who met this requirement were used. This analysis included 648 cases and 1,275 controls; 627 cases had two controls, while 21 cases had only one control.

A second analysis relied on the single statement of usual lifetime occupation and industry on the death certificate. All 731 bladder cancer deaths were used in this analysis, and all cases had two matched controls. For most cases, the same controls were used in both types of analyses.

Data on occupation and industry were coded with three-digit US Census codes, using the method adopted by the US Bu-

reau of the Census (8). Type of industry was determined from the names of companies listed in city directories for 99.1 per cent of all companies listed. Companies found in the directories were coded numerically in two digit codes assigned by us, whenever three or more cases had worked at that company.

Approximately five per cent of the occupational data (1,278 codes) were re-coded for occupation and industry by a second coder. There was a high degree of reproducibility (93.7 per cent for occupation, 97.1 per cent for industry).

Preliminary analyses consisted of the calculation of odds ratios for bladder cancer for the exposed versus nonexposed, ignoring matching. For associations with unmatched chi-squares of 2.00 or more, odds ratios were calculated taking matching into account. Significance tests and summary odds ratios were calculated utilizing a Mantel-Haenszel procedure, without a continuity correction (9). Significance of differences between odds ratios were tested by a heterogeneity chi-square test (10).

Assessing the accuracy and completeness of city directory data

The validity of occupational data from the city directories was ascertained using two methods. Under the first method, data for employers from the city directory were compared to data from the Social Security Administration Quarterly Earnings reports, available since 1939. We sought release forms to obtain Social Security Earnings reports from 241 spouses of men who had died during 1975–1982. We obtained 102 release forms (42.3 per cent), representing 5.3 per cent of the study subjects in the city directory analysis. We compared the employers on the earnings reports for these men with the employers shown in the city directories, for each city directory listing for each individual. When the employers were the same, that listing was defined as a "match." A "match" was also allowed when the two listings were consistent, although not identical. This would occur, for example, if an individual for a given year had no listing in the city directory and had an employer outside of Hamilton County as shown on quarterly earnings reports. This would indicate that this individual was probably not living in Cincinnati for these years.

Frequently, for a given year, the city directory did not list study subjects at all, or listed them without showing employer. If Social Security gave an employer in Hamilton County for that year, then that listing was determined to be a mismatch.

While Social Security records were available to check the accuracy of city directory listings for employer, we had no comparable source to check the accuracy of occupation in the city directory, since Social Security records do not include occupation. As an alternative, we obtained known occupations from union records for 58 individuals who were not study subjects but who were residents of Hamilton County. The work history of these men was available from another study. We then reviewed the city directory listings at five-year intervals for these men to see if these listings for occupation matched their known occupation. While these 58 men were neither study subjects nor a random sample of the population, this comparison provided some measure of the accuracy of occupation in the directories.

We also reviewed the city directory data for a 15 per cent sample of cases and controls (n = 281). For these men, we first recorded the number of years during which a man was not listed in the city directories at all. When these men were listed, we recorded how many listings lacked any data on occupation or employer.

Agreement between city directories and death certificates

Occupational data on death certificates and in city directories were compared for all study subjects who had occupational information on both sources (1,806 for occupation, 1,727 for industry). We concluded that a match existed when *any* of the city

directory listings had three-digit Census codes for occupation (or industry) which matched the corresponding codes on death certificates for the same individual, on his death certificate. While it would have been useful to determine what the "usual" occupation and industry was for each individual using the city directory data, any definition of "usual" would have been an arbitrary decision based on the number of years during which a man was listed in the directories, which varied for each study subject.

Results

Associations of bladder cancer with employer

Using the city directories, seven of 56 employers were positively associated with bladder cancer in the preliminary significance testing. For four employers, the association was statistically significant at the five per cent level or less in a matched analysis.

Although death certificates are supposed to list usual *type* of industry, many death certificates do list company name (3). We have therefore also included the findings for specific employers (companies) from death certificates in table 1.

Associations of bladder cancer with occupation and industry

For occupations in which the subjects had ever been employed, the city directory data in table 2 showed significant positive associations for engineers ($p = 0.01$), tailors ($p < 0.01$), carpenters ($p < 0.01$), furnace operators ($p = 0.03$), blending machine operators ($p = 0.01$), chemists ($p = 0.03$), salesmen ($p = 0.03$), pressing machine operators ($p = 0.03$), and housecleaners ($p = 0.03$). Other positive associations (table 2) showed a large test statistic, but small numbers preclude the assumption that the test statistic was distributed as a chi-square. Given the large number of comparisons ($n = 318$), 16 of the 23 associations would have been expected to occur by chance at the five per cent level of significance. A significant increase in the odds ratio with increased duration of employment was observed (table 3) for truck drivers ($p = 0.01$) and furnace operators ($p = 0.05$), while the increase for printers ($p = 0.06$) and shoe repairers ($p = 0.07$) was close to significance.

For industry ("ever employed"), significant positive associations were seen (table 4) for textiles ($p = 0.01$), chemicals ($p = 0.01$), grain mills ($p = 0.03$), advertising ($p = 0.01$), foundries ($p = 0.03$), petroleum ($p = 0.03$), building services ($p < 0.01$), and entertainment ($p = 0.04$) (202 comparisons). A significant increase in the odds ratio with 20 or more years of employment was noted (table 5) for the railroad industry ($p < 0.01$), while a similar trend just short of significance was seen for laundries ($p = 0.08$).

Accuracy and completeness of city directory data, and correspondence to death certificate data

Comparisons between employers known from Social Security records and employers as listed in the city directories (at five-year intervals) were done for 102 study subjects. For 77.1 per cent of the 653 comparisons, the two sources of employer data agreed; for 12.6 per cent, the two sources did not agree; while for 10.3 per cent, Social Security showed an employer while the city directory listing showed none. It should be noted that the percentage of agreements would have increased if the employer on Social Security records for a given person matched an employer in *any* city directory listing for that person, rather than requiring that the same employer be shown for the same year.

For occupation, we compared the known histories of occupation for 58 union members who were residents of Cincinnati (but who were not study subjects) to occupations shown in their city directory listings, at five-year intervals. We reviewed 237 city directory listings, and found that the directory occupation matched known occupation for 76.4 per cent of the listings. We found

TABLE 1

Matched analysis, with exposure defined as "ever employed" by a given employer, from city directories and from death certificates for 1960–1982: case-control study of bladder cancer, Hamilton County, Ohio

Employer	Death certificates			City directories		
	Exposed cases	Exposed controls	Odds ratio (χ^2)*	Exposed cases	Exposed controls	Odds ratio (χ^2)*
Breweries (all)	5	5	1.73 (0.78)	11	11	1.96 (2.49)
Chemical plant 1	4	3	2.50 (1.42)	7	2	7.00 (8.00)
Chemical plant 2	3	0	– (6.00)	3	0	– (6.00)
Foundry	1	1	2.00 (0.25)	4	2	4.00 (3.00)
Machinery plant	3	1	5.50 (2.45)	4	1	7.50 (4.12)
Printing company	4	6	1.15 (0.05)	5	4	2.50 (2.00)
Railroad	3	13	0.46 (1.53)	3	15	0.38 (2.49)
Valve company	7	4	3.00 (3.60)	11	5	4.40 (9.03)

* All employers with a test statistic of 2.00 or greater are presented here. If fewer than five cases were exposed for a given employer, the test statistic may not be distributed as a chi-square. Odds ratios were not calculated when there were either no cases or no controls exposed. Breweries were grouped and included here because the US Census codes do not provide a category for breweries, which then could not have been included in table 3.

TABLE 2

Matched analysis, with exposure defined as "ever employed" in a given occupation, from city directories and from death certificates for 1960–1982: case-control study of bladder cancer, Hamilton County, Ohio

Occupation (code)	Death certificates			City directories		
	Exposed cases	Exposed controls	Odds ratio (χ^2)*	Exposed cases	Exposed controls	Odds ratio (χ^2)*
Engineers (53, 59)	5	9	1.11 (0.36)	12	8	3.00 (6.40)
Chemists (73)	0	0	–	5	2	5.00 (4.57)
Social workers (174)	1	0	– (2.00)	3	0	– (4.84)
Editors, reporters (195)	1	1	2.00 (0.25)	4	1	8.00 (4.90)
Sales, ads (256)	1	0	– (2.00)	5	2	5.00 (4.57)
Sales, hardware (268)	4	2	4.00 (3.00)	5	2	4.95 (4.57)
News vendors (278)	4	3	2.67 (1.79)	2	0	– (4.00)
Bookkeepers (337)	1	7	0.29 (1.76)	7	36	0.39 (5.63)
Rate clerks (343)	1	1	2.00 (0.25)	0	9	– (4.50)
Firefighters (417)	1	7	0.29 (1.76)	0	9	– (4.50)
Cook supervisors (433)	1	0	– (2.00)	4	0	– (8.00)
Bartenders (434)	8	5	3.20 (4.65)	13	13	1.99 (3.25)
Housecleaners (449)	1	2	1.00 (0.00)	5	2	5.00 (4.57)
Animal caretakers (487)	1	0	– (2.00)	4	1	8.00 (4.90)
Carpenters (567)	17	15	2.46 (6.45)	27	23	2.36 (9.30)
Construction workers, nec† (599)	2	3	1.33 (0.10)	3	0	– (6.00)
Machinists (637)	25	67	0.73 (1.70)	45	124	0.69 (3.96)
Sheet metal workers (653)	0	15	– (7.50)	2	24	0.17 (7.69)
Tailors (667, 676)	10	15	1.33 (0.50)	21	17	2.56 (8.68)
Grinding machine operators (709)	3	3	2.00 (0.75)	11	11	2.00 (2.75)
Pressing machine operators (747)	2	0	– (4.00)	5	2	5.00 (4.57)
Blending machine operators (756)	5	1	10.00 (6.75)	8	4	4.00 (6.00)
Furnace operators (766)	6	2	6.00 (6.25)	13	10	2.50 (4.84)
Drivers, nec (814)	0	5	– (2.50)	26	84	0.63 (4.39)
Soldiers (905)	0	0	– –	26	81	0.61 (4.47)

* All occupations associated with bladder cancer with a test statistic of 3.84 or greater are included in this table. Bartenders and grinding machine operators were included due to a priori interest. If fewer than five cases were exposed for a given occupation, the test statistic may not be distributed as a chi-square. Odds ratios were not calculated when there were either no cases or no controls exposed.

† nec, not elsewhere classified.

TABLE 3

Matched analysis, by duration of employment in a given occupation, using data from city directories. case-control study of bladder cancer, Hamilton County, Ohio

Occupation (code)	Duration of employment (years)					
	<20			≥20		
	Exposed cases	Exposed controls	Odds ratio (χ^2)*	Exposed cases	Exposed controls	Odds ratio (χ^2)*
Engineers (53, 59)	11	6	3.67 (7.53)	1	2	1.00 (0.00)
Sales supervisors (243)	42	85	0.99 (0.83)	6	34	0.31 (7.31)
Police (418)	3	3	1.83 (0.51)	7	5	3.13 (3.28)
Bartenders (434)	11	12	1.83 (2.17)	2	1	4.00 (1.50)
Janitors (453)	57	86	1.33 (2.39)	12	9	2.51 (4.47)
Carpenters (567)	18	15	2.40 (6.68)	9	8	2.19 (2.63)
Machinists (637)	30	83	0.68 (3.12)	16	32	0.88 (0.17)
Tailors (667, 676)	14	10	2.75 (6.35)	7	6	2.33 (2.67)
Shoe repairers (669)	6	17	0.66 (0.77)	4	2	4.00 (3.00)
Printers (734–7)	7	31	0.44 (4.12)	6	7	1.69 (0.84)
Grinding machine operators (709)	11	11	2.00 (2.75)	0	0	–
Furnace operators (766)	9	10	1.70 (1.27)	4	0	– (8.00)
Machine operators, nec† (777)	16	55	0.58 (3.94)	4	3	2.67 (1.34)
Drivers, nec (814)	27	70	0.73 (2.01)	1	13	0.15 (4.68)
Laborers, nec (889)	87	133	1.31 (2.79)	7	12	0.86 (0.77)
Truck drivers (804–5)	26	49	1.06 (0.45)	6	1	12.00 (8.64)

* Included in this table were all occupations with a test statistic of 2.00 or greater in the city directory analysis of "ever employed," and with at least 10 cases or controls exposed. Truck drivers, shoe repairers, and printers were included for a priori interest. If fewer than five cases were exposed for a given test, the test statistic may not be distributed as a chi-square. Odds ratios were not calculated when there were either no cases or no controls exposed.

† nec, not elsewhere classified.

TABLE 4

Matched analysis, with exposure defined as "ever employed" in a given industry, from city directories and from death certificates for 1960–1982. case-control study of bladder cancer, Hamilton County, Ohio

Industry (code)	Death certificates			City directories		
	Exposed cases	Exposed controls	Odds ratio (χ^2)*	Exposed cases	Exposed controls	Odds ratio (χ^2)*
Grain mills (110)	1	1	2.00 (0.25)	6	3	4.00 (4.50)
Textile (150, 151)	11	21	1.05 (0.16)	19	16	2.58 (7.33)
Chemical (192)	26	29	1.82 (4.90)	30	33	1.87 (5.98)
Petroleum (200–201)	4	3	2.50 (1.42)	6	3	4.00 (4.50)
Cement (251)	2	0	– (4.00)	4	1	7.91 (4.90)
Foundries (271)	3	4	1.50 (0.29)	7	4	3.50 (4.54)
Appliances manufacturing (340)	3	6	1.00 (0.00)	1	13	0.15 (4.31)
Miscellaneous manufacturing (391)	8	12	1.31 (0.40)	8	35	0.45 (4.41)
Wholesale trade (532)	3	2	3.00 (1.60)	4	1	7.50 (4.90)
Advertising (721)	5	5	2.00 (1.25)	13	8	3.06 (7.25)
Building services (722)	4	2	4.01 (3.00)	8	2	8.00 (9.80)
Management consulting (732)	0	1	– (0.50)	3	0	– (6.00)
Bowling alleys (801)	1	1	2.00 (0.25)	3	0	– (6.00)
Entertainment (802)	6	1	12.00 (8.64)	5	1	10.00 (4.20)
Physicians' offices (812)	0	5	– (2.50)	0	9	– (4.50)

* All industries associated with bladder cancer with a test statistic of 3.84 or greater are presented in this table. If fewer than five cases were exposed for a given industry, the test statistic may not be distributed as a chi-square. Odds ratios were not calculated when there were either no cases or no controls exposed.

TABLE 5

Matched analysis, by duration of exposure in a given industry, using data from city directories: case-control study of bladder cancer, Hamilton County, Ohio

Industry (code)	Duration of exposure (years)					
	<20			≥20		
	Exposed cases	Exposed controls	Odds ratio (χ^2)*	Exposed cases	Exposed controls	Odds ratio (χ^2)*
Textile (151–2)	14	13	2.25 (4.32)	5	2	4.75 (3.95)
Printing (171–2)	21	37	1.13 (0.18)	12	19	1.25 (0.34)
Chemical (192)	21	25	1.68 (3.02)	9	7	2.57 (3.78)
Leather (201–2)	12	17	1.33 (0.62)	2	9	0.37 (1.60)
Railroad (400)	10	42	0.42 (5.70)	22	19	2.21 (6.67)
Advertising (721)	10	5	4.00 (7.50)	3	3	1.71 (0.51)
Laundry (771)	4	11	0.73 (0.30)	5	3	3.33 (3.06)

* Included in this table were all industries with a test statistic of 2.00 or greater in the city directory analysis of "ever employed" in which at least 10 cases and controls were exposed. Printing, leather, railroad, and laundry were included for a priori interest. If fewer than five cases were exposed for a given test, the test statistic may not be distributed as a chi-square. Odds ratios were not calculated when there were either no cases or no controls exposed.

a mismatch for 15.6 per cent, while 8.0 per cent of the directory listings showed no occupational information.

For both city directory and death certificate data, missing information was not uncommon. Among our potential cases, 10.8 per cent had to be excluded from the city directory analysis because they could not be located in the city directories, or because their listings gave no occupational data. Persons accepted for the study had an average of seven listings. However, for 36 per cent of these persons, listings lacked employer, while 15 per cent either lacked occupation or had uncodable data (e.g., "foreman" or "laborer") (table 6).

Finally, we compared city directory and death certificate information for each study subject. Results indicated a match for occupation based on the three-digit US census code between at least one city directory listing and the occupation on the death certificate for 68.3 per cent of study subjects. A match was found for industry for 80.0 per cent of study subjects.

TABLE 6

Missing or uncodable occupational data in city directories and death certificates case-control study of bladder cancer, Hamilton County, Ohio*

Type of missing data	Death certificates (%)	City directories (%)
Occupation	6.0	15.0
Industry	12.0	NA
Employer	NA†	35.5

* The city directory data considered here include all 1986 listings for the 281 study subjects who died in 1961, 1971, and 1981, for the years when these subjects were actually listed in the directories (n = 1,986). The death certificate data covers all 2,193 deaths used in the study. Type of industry is available on death certificates but not in city directories, while employer is available in city directories but not on death certificates. For death certificates, the query is for "usual lifetime" occupation and industry, while for city directory data, the query is for current occupation and employer for any given year.

† NA, not applicable.

DISCUSSION

Associations of bladder cancer and employers, occupations, and industries

One advantage of city directories versus death certificates, in "surveillance" studies such as this one, is that analysis for specific employers is possible. Death certificates, in contrast, call for "usual lifetime industry" rather than employer.

By using city directories, we were able to identify four companies with a significant excess risk for bladder cancer. In contrast, only one of these plants would have been identified using death certificate data.

The National Institute for Occupational

Safety and Health has conducted an initial investigation at six plants (table 1) which showed the strongest positive associations with bladder cancer. Three of these plants were no longer in operation. Table 7 indicates the potential bladder carcinogens which were used at these plants. Chemical plant 1 is worthy of special mention because it had produced benzidine and was known to have been the site of a bladder cancer epidemic from a prior study (11). The ability of city directory data to identify this plant was reassuring. In contrast, death certificate analysis for this plant yielded a nonsignificant odds ratio.

Regarding analyses by occupation and industry, most of the occupational exposures associated with bladder cancer in this study have been noted in previous studies. Among the more interesting findings was the significantly increased risk for those employed for more than 20 years as truck drivers or in the railroad industry, in view of recent suggestions that exposure to diesel fumes may increase bladder cancer risk (12, 13). Also of interest was the finding of an excess of bladder cancer among workers in foundries and among furnace operators, in light of recent work indicating that polycyclic aromatic hydrocarbons in potrooms of aluminum workers may be bladder carcinogens (14). The elevated risk for janitors ($p = 0.06$) has not, to our knowledge, been previously cited in the literature. Finally, while machinists have been identified as a high-risk group in other studies (12), presumably due to cutting oils, here they were at decreased risk. On the other hand, machinists did show an increased risk with increased duration of employment.

As can be seen from tables 3 and 5, the direction and size of the odds ratios are similar for the data from death certificates and from city directories, but the number of exposed individuals is almost always greater for the data from city directories. The larger numbers of exposed in the city directories were expected, since the city directories list a lifetime work history with many possible jobs.

Given the large number of comparisons in this study, a number of the associations seen in tables 3–6 may be due to chance. Various confounders may have also caused false positives. Given that the relative risk of bladder cancer due to smoking is a moderate one (about 2–3), it is unlikely that differences in smoking habits between cases and controls could account for most of the relative risks observed in this study

TABLE 7

Suspected bladder carcinogens at six plants identified as sites of high risk: case-control study of bladder cancer, Hamilton County, Ohio

Site	Suspected agent	Comment (reference no.(s))
Chemical plant 1	Benzidine	Human bladder carcinogen (19)
Chemical plant 2	Dicholorobenzidine, toluidine, dianisidine	Animal bladder carcinogens, suspect human carcinogens (20)
Foundry	Polycyclic aromatic hydrocarbons	Suspect human carcinogen (15)
Machine manufacturer	Cutting oils with nitrosamines	Suspect human carcinogen (12, 16)
Printing ink company	Printing inks with aromatic amines	Suspect human carcinogen (20)
Valve manufacturer	Cutting oils with nitrosamines	Suspect human carcinogen (12, 16)

(16, 17). A review of the data to assess confounding by other hazardous jobs found no such confounding.

Validity and completeness of city directory and death certificate data

It must be stressed that city directory data are only available in medium-sized US cities. Polk Directories, the largest producer of city directories, currently publishes directories for 1,250 US cities. All contain data on occupation. Large cities such as New York, Chicago, Philadelphia, Los Angeles have not had directories since the 1930s, while cities such as Baltimore and Washington, DC were dropped by Polk in the 1960s. The largest cities for which Polk directories continue to be produced are Houston, San Francisco, and Dallas. However, for most middle-sized cities (such as Cincinnati, Sacramento, San Diego, and Indianapolis) there are city directories dating from the present back to the turn of the century (personal communication, R. D. Monforton, R. L. Polk Directories, 1986).

In our review of the listings for 58 union members with known work history, occupation in the city directories was accurate 76 per cent of the time. City directory listings for employer for 102 study subjects compared to Social Security data were accurate 77 per cent of the time. Although it is difficult to compare city directory data (multiple listings) with death certificate data (single listing), the level of accuracy of our directory data compared rather well with what is known about death certificates. One recent study of death certificates of 2,200 long-term US workers (3) indicated that the death certificate listing for usual lifetime occupation was accurate 64 per cent of the time. Death certificates were accurate for usual lifetime industry about 70 per cent of the time.

While the validity of the city directory data may be fairly good, the major limitation of the city directories is missing data.

Some residents may be missed by the city directories altogether. A study done by the city of Cincinnati indicated that in 1978 the Cincinnati city directories included 92 per cent of the individuals enumerated by the 1980 US Census (personal communication, Dick Moran, city planner, City of Cincinnati, 1983). Discrepancies occurred primarily in poor neighborhoods.

Among our potential study subjects, 10.8 per cent of the cases had to be excluded from the city directory analysis because they could not be located at all, or because they had no listing with any occupational data. Some of these men may have moved to Cincinnati just before death, or may have lived in small towns which were not covered by the city directories in Hamilton County. On the other hand, some may have simply been missed by city directory enumerators. Among our actual study subjects, who had at least one listing in the city directories, 25 per cent of the years checked showed no listing (going back to age 20 years). Most of the years with no listings occurred prior to the first listing for an individual, which may indicate that the man had not yet moved to Hamilton County. When we reviewed those years for which a man *was* listed in the directories, 15 per cent of these listings were missing occupation, and 36 per cent were missing employer.

Even with this level of missing data, given the large number of listings per person, in our view it was possible to construct a fairly complete work history for most individuals during those years in which they were living in Hamilton County.

Three previous studies (4, 5, 18) have compared occupational data on death certificates and city directories, although these studies did not assess the accuracy of city directories versus another source known to be valid. Gute and Fulton (18) analyzed the correspondence between city directories in Rhode Island and death certificates. They found directory listings for about 65 per cent of their subjects. They also found approximately a 60 per cent correspondence for occupation and industry between directory and death certificate data, after deter-

mining a subject's "usual" occupation and industry from the multiple directory listings. Roush et al. (4, 5) used city directories and death certificates in Connecticut in two case-control studies. In both studies, about 65 per cent of study subjects had listings in the city directories, and at least one city directory listing matched occupation on the death certificate for 90–95 per cent of the subjects.

Conclusion

If we assume that the Cincinnati directories are representative of other directories in the United States, then our conclusion is that city directories, despite their limitations, provide a useful data source which has been little used by occupational epidemiologists. They are particularly useful for identifying high-risk employers in specific geographic areas. However, this conclusion holds only for men. For women, the change from maiden to married name may mean that they cannot be traced in some city directories.

City directories are most useful in certain types of occupational case-control studies. When a hypothesis has already been established, a more expensive interview study may be warranted. However, case-control studies have relatively low power to detect elevated occupational risks, and the large sample sizes necessary may preclude more expensive interview studies (13). Furthermore, there are many situations in which a hypothesis-generating study based on records alone may be in order.

In the current study, one researcher spent about six months abstracting data from the city directories for 1,900 study subjects. An additional six months was required to code and analyze these data. The cost of the city directory analysis, considering salary and computer costs, was approximately $40,000. A comparable interview study of 1,900 subjects would probably have cost ten times that amount.

For mortality studies, city directory data are often superior to the traditional source of occupational data, death certificates. For incident studies, where the investigator seeks to use records rather than expensive interviews to ascertain occupation, city directories may provide the only such records available.

Although the focus of this study has been to compare city directory data and death certificate data, often it will be useful to use the data in combination. Death certificates reflect the usual lifetime occupation, which may be the most relevant information in an occupational study. City directories, with many short-term jobs and multiple jobs per person, may yield a number of false positive associations. On the other hand, short-term jobs will be missing from death certificates, and such jobs may involve critical exposures. Death certificates may provide useful occupational data for individuals who had only recently moved to the area where they died, and who would not have been included in the local city directory. Death certificates will also provide occupational data for individuals in areas without city directories.

References

1. Milham S. Occupational mortality in Washington State 1950–79. Cincinnati, OH: NIOSH, 1983. (NIOSH publication no. 83-116).
2. Dubrow R, Wegman DH. Occupational characteristics of cancer victims in Massachusetts. Cincinnati, OH: NIOSH, 1984. (NIOSH publication no. 84-109).
3. Steenland K, Beaumont JJ. The accuracy of occupation and industry data on death certificates. J Occup Med 1984;26:288–96.
4. Roush GC, Meigs JW, Kelly J, et al. Sinonasal cancer and occupation: a case-control study. Am J Epidemiol 1980;111:183–93.
5. Roush GC, Kelly JA, Meigs JW, et al. Scrotal carcinoma in Connecticut metalworkers: sequel to a study of sinonasal cancer. Am J Epidemiol 1982;116:76–85.
6. Burnett CA Bladder cancer and its association with occupational exposure in Hamilton County. Unpublished MS thesis. University of Cincinnati, Cincinnati, OH, 1982.
7. Riggan WB, Bruggen JV, Acquavella JF, et al. US cancer mortality rates and trends 1950–78. Research Triangle Park, NC: Health Effects Research Laboratory, US Environmental Protection Agency, unpublished, 1983.
8. US Bureau of the Census. Alphabetical index of industry and occupation. Washington, DC: US GPO, 1981.
9. Mantel N, Haenszel W. Statistical aspects of the

analysis of data from retrospective studies of disease. JNCI 1959;22:719–48.
10. Rothman KJ, Boice JD. Epidemiologic analysis with a programmable calculator. Washington, DC: US GPO, 1979. (NIH publication no. 79-1649).
11. Zavon MR, Hoegg U, Bingham E. Benzidine exposure as a cause of bladder tumors. Arch Environ Health 1973;27:1–7.
12. Silverman DT, Hoover RN, Albert S, et al. Occupation and cancer of the lower urinary tract in Detroit. JNCI 1984;70:237–45.
13. Hoar SK, Hartge P, Hoover RN. The National Bladder Cancer Study: employment in the chemical industry. Paper presented at the Society for Epidemiologic Research 18th Annual Meeting, Chapel Hill, NC, June 18–21, 1985.
14. Theriault G, Tremblay C, Cordier S, et al. Bladder cancer in the aluminum industry. Lancet 1984;1:947–50.
15. Vineis P, Magnani C. Occupation and bladder cancer in males, a case-control study. Int J Cancer 1985;35:599–600.
16. Axelson O. Aspects on confounding in occupational health epidemiology. Scand J Work Environ Health 1978;4:85–9.
17. Blair A, Hoar SK, Walrath J. Comparison of crude and smoking-adjusted standardized mortality ratios. J Occup Med 1986;27:881–4.
18. Gute DM, Fulton JP. Agreement of occupation and industry data on Rhode Island death certificates with two alternative sources of information. Public Health Rep 1985;100:65–72.
19. Case RA, Hosker ME, McDonald D, et al. Tumours of the urinary bladder. Br J Ind Med 1954;11:75–104.
20. Boeniger MF. The carcinogenicity and metabolism of azo dyes. Cincinnati, OH: NIOSH, 1980. (NIOSH technical report no. 80-119).