Evaluating Occupation and Industry Separately to Assess Exposures in Case-Control Studies

Mustafa Dosemeci, Patricia A. Stewart, and Aaron Blair
Occupational Studies Section, Environmental Epidemiology Branch, National Cancer Institute, Executive Plaza North Room 418, Bethesda, Maryland 20892

In most occupational case-control studies, assessment of exposures is based on occupation and industry due to the lack of any other exposure information. The traditional approach to assessing exposures to specific substances is to estimate exposure for each occupation/industry combination. A modified procedure that substantially reduces the effort necessary to complete such assessments is described. In this procedure, exposure levels (none, low, medium, and high) are estimated for each occupation and industry separately. Numeric weights are assigned to these exposures levels. Final estimates for each occupation/industry combination are derived using exposure weights for occupations and industries. This method was applied to a case-control data set to estimate levels of silica exposure for the work histories of subject. The traditional method required 13,445 individual evaluations while the new approach required only 2,000 estimates. Results from the new approach were compared with those from the traditional method for a sample of 210 occupation/industry combinations. The kappa statistic for the agreement of the methods was found to be 0.82 with an exact agreement by level for 89 percent of the occupation/industry combinations. The markedly smaller amount of time and the high correlation with the more time-intensive procedure suggest that this new procedure may be an efficient approach to exposure assessment in case-control studies. Dosemeci, M.; Stewart, P.A.; Blair, A.: Evaluating Occupation and Industry Separately to Assess Exposures in Case-Control Studies. Appl. Ind. Hyg. 4:256–259; 1989.

Introduction

Estimating exposure is a crucial component of occupational epidemiologic studies. In most occupational case-control studies, information on exposures is limited to titles of occupation and industry only. In many studies, workplace risk factors are evaluated by classifying cases and controls into either occupational and/or industry groups.[1–17]

In some case-control studies,[18–26] however, risk factors are evaluated by using both occupation and industry titles. This latter approach allows specification of exposures experienced by subjects in a particular occupation in a particular industry. If the characterization of exposure is carried out using occupational titles only, the same occupation in different industries would be considered to have the same exposure. For example, machine operators in the construction, rubber, communication, or textile industries would be assigned the same exposure levels despite the fact that these machine operators are likely to have different exposures. Similarly, if the exposure characterization is carried out using industry titles only, different occupations in the same industry would be considered to have the same exposure. For instance, laborer, engineer, painter, or foreman in the construction industry would all be assigned to the same exposure level although they may have very different exposure levels. In the National Bladder Cancer study,[27] Zahm *et al* evaluated the risk of bladder cancer by characterizing the exposure with industry, occupation, and occupation/industry combination. The accuracy of this approach was much higher than the assessment of exposure by occupation or industry alone, but the number of specific job titles for which estimates were required was greater due to the increased number of occupation/industry combinations. A large number of occupation/industry combinations may also affect consistency because as the number of estimates required increases, it is easier to make mistakes or forget the decisions made earlier. This is especially true when estimating exposure levels for the same occupation in similar industries. In order to overcome problems associated with this, an estimation procedure has been developed which reduces the number of estimates required in the occupation/industry combination approach, but it increases the likelihood of consistency by using standard occupational and industrial classification systems. The estimation procedure of the new approach and its correlation with the previous approach is described here.

Method

Data from six lung cancer case-control studies conducted by the National Cancer Institute (NCI) were used to demonstrate the proposed exposure estimating procedure. There were 2,251 unique occupations and 1,130 unique industries. The total number of occupation/industry combinations was 13,445. Assignment of a semiquantitative (none, low, medium, and high) exposure level to silica dust was made by two industrial hygienists reaching a consensus agreement. In the first step, each unique industrial title held by the study subjects was coded according to the *Standard Industrial Classification (SIC) Manual*.[28] For some industries, two-digit SIC codes were found to be specific enough to characterize the exposure, e.g., "banking" sector (SIC code 60). For some other industries, three- or four-digit SIC codes were assigned depending on their ability to characterize the exposure. For instance, the three-digit SIC code 753 was assigned to the less specific "general auto repair shop," while the four-digit SIC code 7535 was assigned to the more specific "auto paint shop." In general, four-digit SIC codes were found to be specific enough to characterize the exposure. In some cases, however, fifth digits were added to the existing four-digit SIC codes to create more specific and descriptive industry classes. For instance, the less specific "highway and street construction" industry (SIC code 1611) was broken down into the following groups to differentiate their exposure potentials: 1) asphalt paving (SIC code 16111); 2) concrete road construction (SIC code 16112); 3) curb construction (SIC 16113); 4) highway sign installation (SIC code 16114); 5) street maintenance or repair (SIC code 16115); 6) trail building (SIC code 16116); and 7) other highway and street construction (SIC code 16119). Similar procedures were applied to the occupations using the *Standard Occupational Classification (SOC) Manual*.[29] Unlike the industry classification, most four-digit SOC codes were not found to be specific enough to characterize the exposure for the occupations, particularly for construction operations, repairers, and production operations. Therefore, an additional fifth digit was generally necessary to make the SOC system useful for the exposure estimation procedures. For example, "packaging and filling machine operators and tenders" (SOC code 7662) was broken down into "bagger" (SOC code 76621), "wrapper layer" (SOC code 76622), "packaging machine operator" (SOC code 76623), "loading machine operator" (SOC code 76624), "barrel filler" (SOC code 76625), "cotton baler" (SOC code 76626), "bottle packer" (SOC code 76627), "tobacco packer" (SOC code 76628), and "other packaging and filling machine operators" (SOC code 76629).

The second step of the exposure assessment procedure was to estimate the relative level of exposure to silica and the degree of certainty of this estimate for each industry and occupation code independently (Table I and Table II, respectively). For the assignments of semiquantitative (none, low, medium, and high) exposure levels, intensity of exposure, frequency of exposure, type of contact with the agent (direct or indirect), and potential of other relevant exposure parameters (e.g., peaks) were taken into account, depending on the type of the agent under consideration. For example, if the agent is silica, full-time, highly intensive, direct/indirect exposures were considered to be high level; part-time, highly intensive or full-time, moderately intensive, direct/indirect exposures were medium level; and occasionally highly intensive or part-time, moderately intensive direct/indirect exposure or full-time, less intensive, indirect exposures were low level. If the agent is carbon monoxide, more emphasis is given to the peak exposure.

TABLE I. Assigning Levels of Exposure to Silica (E_i) and Certainty (C_i) For Industry (SIC) Codes

SIC	SIC Name	(E_i)	(C_i)
1423	Crushed Stone	High	High
3731	Painting, Decorating	High	Low
326	Pottery Products	Med	High
3321	Gray Iron Foundries	High	High
1661	Highway Construction	Med	Low
16611	Asphalt Paving	Low	High
48	Communication	None	High
721	Laundry and Garment	None	High

TABLE II. Assigning Levels of Exposure to Silica (E_o) and Certainty (C_o) For SOC Codes

SOC	SOC Name	(E_o)	(C_o)
75491	Sandblast Operator	High	High
463	General Clerk	None	High
16	Engineer	Low	Low
1624	Mining Engineer	Low	High
32	Writer	None	High
7662	Packing Mach. Oper.	High	Low
8461	General Worker	High	Low
711	Prod. Supervisor	Med	Med

In addition to exposure levels, degree of certainty of the exposure estimates were also assigned semiquantitatively (low, moderate, and high) to each industry and occupation code. The degree of certainty characterized the confidence of the industrial hygienists in the estimates of exposure levels based on the specificity of the potential of exposure in that industrial and occupational group. For example, "stone, clay, and glass products" (SIC code 32), which includes nonsilica-exposed industries (manmade minerals, asbestos products) together with the exposed ones (pottery, glass products), would have a lower degree of certainty for silica exposure than "cut stone" (SIC code 32811) since SIC code 32811 excludes the "asbestos products" (SIC code 3292), "lime" (SIC code 3274), and "gypsum products" (SIC code 3275). Similar certainty assignments were carried out for occupations, considering the specificity of the duties inferred from a particular job title. For example, in the assessment of exposure to silica, a high level of exposure and a high degree of certainty were assigned for "sandblast operator" (SOC code 75491). For "general clerk" (SOC code 463), a none level of exposure with high degree of certainty was assigned. For the less specific title of "engineer" (SOC code 16), a low level of exposure and a low degree of certainty were assigned, but for the more specific title of "mining engineer" (SOC code 1624), a high degree of certainty with low level of exposure was assigned.

In the third step of the assessment, numeric weights were assigned to each level of exposure (0 for none, 1 for low, 2 for medium, and 3 for high) by occupation (E_o) and industry (E_i). The selection of these scores was based on the studies of the other investigators.[30,31] In order to derive the final estimates for each occupation/industry combination ($E_{o/i}$), several formulas are presented in Table III.

For example, when the degree of certainty for an occupation was high (line 1 in Table III), the overall exposure was calculated using the occupational information only (E_o). In other words, if the occupational information was specific enough to characterize the overall exposure, then industry information was not taken into account for the calculation of the final exposure. This occurred for occupational groups which have very specific tasks related to the exposure. For example, sand blasters were assigned a high level of exposure to silica regardless of the industry because the tasks association with sand blasting operations are specific enough to characterize exposure to silica. As a result, a sand blaster in the gray iron foundry industry or in TV manufacturing were assigned the same level of exposure to silica despite the exposure differences between these industries. On the other

TABLE III. Formulae in Calculating Final Exposure to Silica for Occupation/Industry Combinations

No.	Certainty Combinations (C_o)	(C_i)	Equations for ($E_{o/i}$)
1	H	L,M,H	$(E_o)^2$
2	M	L	$(E_o^2 \cdot E_i)^{2/3}$
3	M	M	$(E_o \cdot E_i)$
4	M	H	$(E_o \cdot E_i^2)^{2/3}$
5	L	L	$(E_o \cdot E_i)$
6	L	M	$(E_o \cdot E_i^2)^{2/3}$
7	L	H	$(E_i)^2$

where:
C_o = the certainty category of exposure in the occupational group
C_i = the certainty category of exposure in the industrial group
E_o = the estimate of exposure level for occupational group
E_i = the estimate of exposure level for industrial group
$E_{o/i}$ = the exposure level for occupational/industry combination

hand, physicians or cooks in the mining industry were unlikely to have silica exposure despite the fact that the mining industry overall has high level silica exposure. When the degree of certainty for occupation was medium or low, the exposure level of the occupation was adjusted by the exposure level of the industry using the equations in Table III, which take into account the degree of certainty for both the occupation and the industry. If the degree of certainty for a given occupation is higher than the degree of certainty for the industry, more weight is placed on the exposure level of occupation. On line 7 in Table III, where the degree of certainty is low for the occupation and high for the industry, overall exposure ($E_{O/I}$) is calculated by using only the exposure level of the industry independent of the exposure level of the occupation. This last formulation is used for occupations with nonspecific jobs in specific industries. For example, the silica exposure level of a cutting machine operator in a stone quarry or in the meat processing industry depends solely on the exposure level of the industry even though they both are doing the same cutting operation.

The rationale for the equations in Table III was to obtain scores between 0 to 9 by giving a special emphasis to the exposure levels for occupation (E_O) or for industry (E_I) depending on their degree of certainty. These scores are then put into four final exposure categories (none for 0, for 0.1–3.0, medium 3.1–5.9, and high for 6.0–9.0) (Table IV).

For comparison with the estimates of silica exposure obtained from the new procedure, semiquantitative exposure levels were assigned directly to a stratified random sample of 210 occupation/industry combinations. The same industrial hygienists who followed the first approach also made direct semiquantitative estimates (none, low, medium, and high) for each of these occupation/industry combinations. Agreement (kappa statistics[32] and sensitivity and specificity[33]) of the usual and the new approach were measured.[33]

Results

The distribution of the estimates of exposure levels is cross tabulated in Table V. The estimates for 89 percent of the assignments from the two methods were in the same exposure level category. The two methods disagreed by more than one rank for less than 1 percent of the comparisons. The kappa statistic for the agreement between the approaches was high (0.82) and the Spearman correlation coefficient was 0.83. The sensitivity and specificity of the new approach were found to be 0.91 and 0.87, respectively.

Discussion

The traditional approach to the assessment of exposure in case-control studies is direct assignment of estimated exposure levels for each occupation/industry combination held by the study subjects. Due to the large number of occupation/industry combinations, this approach is time-consuming and may result in inconsistencies. This article describes another approach for exposure assessment in an attempt to reduce the time and minimize the consistency problems of the combination approach. In the application of this new approach to the 13,445 occupation/industry combinations of the NCI lung cancer case-control study, only 2,000 exposure assessment decisions were made for all combinations, much lower than what would have been required.

The new approach has several advantages. Apart from having a substantial decrease in the number of decisions, it is more likely to increase the consistency among the estimates due to the reduced number of estimates. Another advantage of this approach is its ease of transferability to other studies. Once the estimates of exposure to the agent under consideration are developed for the standard occupations and standard industries, this *a priori* job/industry exposure matrix may be used in any other epidemiologic study on the same agent. This approach also allows the establishment of a common exposure criteria for different occupational studies. This feature of the new approach may enhance the comparability of risk estimates obtained from different studies. This may be particularly important when evaluating studies with contradictory results.

This proposed approach should not be taken as the only alternative to existing retrospective exposure assessment procedures. Other approaches may be more desirable depending upon study objectives and the availability of exposure information. Certainly, further studies are necessary to test the validity of the proposed approach, especially the external validity, by having exposure estimates from different groups of industrial hygienists. The reliability of the results of the proposed approach could be compared with the results of the quantitative environmental mea-

TABLE IV. Final Exposures for Occupation/Industry Combinations

Occupation	Industry	E_o	E_i	C_o	C_i	$E_{o/i}$
Timekeeper	Bldg. Construction	0	1	H	M	0 (none)
Bricklayer	Bldg. Construction	1	1	H	M	1 (Low)
Laborer	Bldg. Construction	3	1	L	M	3 (Low)
Sandblaster	Car Factory	3	3	H	L	9 (High)
Engineer	Car Factory	1	3	M	L	2 (Low)
Janitor	Car Factory	1	3	L	L	3 (Low)
Timekeeper	TV Factory	0	0	H	M	0 (None)
Engineer	TV Factory	1	0	M	M	0 (None)
Sandblaster	TV Factory	3	0	H	M	9 (High)
Laborer	Granite Quarry	3	3	L	H	9 (High)
Engineer	Granite Quarry	1	3	M	H	3 (Low)
Foreman	Granite Quarry	2	3	M	H	6 (Medium)
Timekeeper	Grain Elevator	0	1	H	L	0 (None)
Foreman	Grain Elevator	2	1	M	L	3 (Low)
Janitor	Grain Elevator	1	1	L	L	1 (Low)
Timekeeper	School	0	0	H	H	0 (None)
Janitor	School	1	0	L	H	0 (None)
Bricklayer	School	1	0	H	H	1 (Low)

TABLE V. Comparison of the Estimates from New Approach and The Combination Approach

New	Usual None	Low	Medium	High	Total
None	104	4	0	0	108
Low	7	48	4	0	59
Medium	0	6	28	0	34
High	0	1	2	6	9
Total	111	59	34	6	210

Kappa Value for the Agreement:	0.82
Spearman Correlation Coefficient:	0.83
Sensitivity:	0.91
Specificity:	0.87

surements obtained for the same occupation/industry combinations. The authors plan to apply the new approach to a well-established, exposure–effect relationship such as asbestos and lung cancer. A significant trend of lung cancer risk with the estimated asbestos exposure levels obtained by the new approach may indicate the level of its reliability.

These practical advantages and the high correlation with the traditional approach suggest that this proposed procedure may be an efficient approach to exposure assessment in retrospective occupational case-control studies.

References

1. Malker, H.R.; McLaughlin, J.K.; et al: Occupational Risks for Bladder Cancer among Men in Sweden. Cancer Res. 47: 6763 (1987).
2. Kjuus, H.; Skjaerven, R.; Langard, S.; et al: A Case-referent Study of Lung Cancer, Occupational Exposure and Smoking. Scand. J. Work. Environ. Health 12:193 (1986).
3. Roush, G.C.; Meigs, J.W.; Kelly, J.; et al: Sinonasal Cancer and Occupation: A Case-control Study. Am. J. Epid. 111(2):183 (1980).
4. Musicco, M.; Filippini, G.; Bordo, B.; et al: Gliomas and Occupational Exposure to Carcinogens: Case-control Study. Am. J. Epidemiol. 11(5):782 (1982).
5. McLaughlin, J.K.; Malker, H.S.R.; Ericsson, J.L.E.; et al: Occupational Risks for Meningiomas of the CNS in Sweden. J. Occup. Med. 29(1):66 (1987).
6. Thomas, T.L.; Fontham, E.T.H.; Norman, S.A.; et al: Occupational Risk Factors for Brain Tumors. Scand. J. Work Environ. Health 12:121 (1986).
7. Hayes, R.B.; Gerin, M.; Raatgever, J.W.; et al: Wood-related Occupations, Wood Dust Exposure, and Sinonsal Cancer. Am. J. Epidemiol. 124(4):569 (1986).
8. Brinton, L.A.; Blot, W.J.; Becker, J.A.; et al: A Case-control Study of Cancer of the Nasal Cavity and Paranasal Sinuses. Am. J. Epidemiol. 119(6):896 (1984).
9. Zagraniski, R.T.; Kelsey, L.J.; Walter, S.D.: Occupational Risk Factors for Laryngeal Carcinoma: Connecticut, 1975–1980. Am. J. Epidemiol. 124(1):67 (1986).
10. Flanders, D.W.; Cann, C.I.; Rothman, K.J.; et al: Work-related Risk Factors for Laryngeal Cancer. Am. J. Epidemiol. 119(1):23 (1984).
11. McLaughlin, J.K.; Malker, H.S.R.; Malker, B.K.; et al: Registry-based Analysis of Occupational Risks for Primary Liver Cancer in Sweden. Cancer Res. 47:287 (1987).
12. Stemhagen, A.; Slade, J.; Altman, R.; et al: Occupational Risk Factors and Liver Cancer. Am. J. Epidemiol. 117(4):443 (1983).
13. McLaughlin, J.K.; Malker, H.S.R.; Stone, B.J.; et al: Occupational Risks for Renal Cancer in Sweden. Br. J. Ind. Med.44:119 (1987).
14. Olsen, J.H.; Jensen, O.M.: Occupation and Risk of Cancer in Denmark. Scand. J. Work. Environ. Health 13(Suppl):3 (1987).
15. Wilkins, J.R.; Sinks, T.H.: Occupational Exposures Among Fathers of Children with Wilms' Tumor. J. Occup. Med. 26(6):427 (1984).
16. Morrison, A.S.; Ahlbom, A.; Verhoek, W.G.; et al: Occupation and Bladder Cancer in Boston, USA, Manchester, UK, and Nagoya, Japan. J. Epidemiol. Community Health 39:294: (1985).
17. Hayes, R.B.; Raatgever, J.W.; Bruyn, A.; et al: Cancer of the Nasal Cavity and Paranasal Sinuses, and Formaldehyde Exposure. Int. J. Cancer 37:487 (1986).
18. Vaughan, T.L.; Strader, C.; Davis, S.; et al: Formaldehyde and Cancer of the Pharynx, Sinus, and Nasal Cavity: I. Occupational Exposures. Int. J. Cancer 38:677 (1986).
19. Thomas, T.; Stewart, P.A., Stemhagen, A.; et al: Risk of Astrocytic Brain Tumors Associated with Occupational Chemical Exposures. Scand. J. Work Environ. Health 13:417 (1987).
20. Coggon, D.; Pannett, B.; Acheson, E.D.: Screening for New Occupational Hazards of Cancer in Young Persons. Ann. Occup. Hyg. 28(1):145 (1984).
21. Hinds, W.M.; Kolonel, L.N.; Lee, J.: Application of a Job–Exposure Matrix to a Case-Control Study of Lung Cancer. J. Nat. Cancer Inst. 75(2):193 (1985).
22. Coggon, D.; Pannett, B.; Acheson, E.D.: Use of Job-Exposure Matrix in an Occupational Analysis of Lung and Bladder Cancers on the Basis of Death Certificates. J. Nat. Cancer Inst. 72(1):61 (1984).
23. Macaluso, M.; Vineis, P.; et al.: Job Exposure Matrices; Experience in Italy. In: Proceedings of a Conference Held in April 1982 at the University of Southampton, Southampton, U.K., pp. 22–30. Hobbs, The Printers of Southampton (1983).
24. Decoufle, P.; Stanislawczyk, K.: A Retrospective Survey of Cancer in Relation to Occupation. NIOSH Research Report. DHEW (NIOSH) Pub. 77-178 (1977).
25. Brown, L.M.; Mason, T.J.; Pickle, L.W.; et al: Occupational Risk Factors for Laryngeal Cancer on the Texas Gulf Coast. Cancer Res. 48:1960 (1988).
26. Spirtas, R.; Hoar, S.K.; Kaminski, R.; et al: Aplastic Anemia Mortality and Occupational Exposures. In: Proceedings of the 1985 Public Health Conference on Records and Statistics. DHHS (PHS) Pub. No. 86-1214 (1986).
27. Zahm, S.H.; Hartge, P.; Hoover, R.: The National Bladder Cancer Study: Employment in the Chemical Industry. J. Nat. Cancer Inst. 79(2):217 (1987).
28. Standard Industry Classification Manual. Executive Office to the U.S. President, Office of Management and Budget. U.S. Government Printing Office, Washington, DC (1972).
29. Standard Occupational Classification Manual. Executive Office to the U.S. President, Office of Management and Budget, U.S. Government Printing Office, Washington, DC (1972).
30. Siemiatycki, J.; et al: Association Between Smoking Habits and Levels of Workplace Exposures. Am. J. Ind. Med. 13:59 (1988).
31. Kromhout, H.; et al: Agreement Between Qualitative Exposure Estimates, and Quantitative Exposure Measurements. Am. J. Ind. Med. 12:551 (1987).
32. Fleiss, J.L.: Statistical Methods for Rates and Proportion, 2nd ed., pp. 212–236. John Wiley & Sons, Inc., New York (1981).
33. Fleiss, J.L.: Statistical Methods for Rates and Proportion, 2nd ed., p. 6. John Wiley & Sons, Inc.; New York (1981).

Received 11/28/88; review decision 4/21/89; revision 6/2/89; accepted 6/26/89