



Original article

Scand J Work Environ Health [1997;23\(2\):140-148](#)

doi:10.5271/sjweh.191

Assessing agreement between two job-exposure matrices

by [Hawkes AP](#), [Wilkins JR III](#)

Key terms: [agreement](#); [concordance](#); [job-exposure matrix](#); [occupational exposure](#); [reliability](#)

This article in PubMed: www.ncbi.nlm.nih.gov/pubmed/9167237



This work is licensed under a [Creative Commons Attribution 4.0 International License](#).

Assessing agreement between two job-exposure matrices

by Allison P Hawkes, MD,^{1,2} JR Wilkins III, DrPH¹

Hawkes AP, Wilkins JR III. Assessing agreement between two job-exposure matrices. *Scand J Work Environ Health* 1997;23(2):140—8.

Objectives This study was undertaken to compare the operating characteristics of the Hoar et al job-exposure matrix and the job-exposure matrix of the National Occupational Hazard Survey (NOHS) of the National Institute for Occupational Safety and Health in the United States. Both matrices were applied to occupations in the metal, paper and wood, and chemical, drug and paints industries, and the extent of agreement between the two was estimated.

Methods The agents (exposure entities) selected for the comparison were common to both matrices; altogether 214 agents were included. For the comparison of occupations, the occupational codes of the Hoar et al matrix were converted to those of the NOHS-based matrix. For every occupational group analyzed, each of the 214 exposure entities was classified as to its "presence" or "absence", and a kappa statistic (κ) was calculated.

Results The κ_i ranged from 0.27 to -0.12. In general, most of the agreement between the 2 matrices occurred because they agreed as to the absence of exposure for the exposure entities. Variation in agreement was observed for both the individual exposure entities and for chemical families.

Conclusions These results indicate poor to only fair agreement between the 2 job-exposure matrices, and they underscore the importance of understanding the rules for classifying exposures when any such matrix is used. Investigators also need to be aware of the methods and limitations of the occupational coding systems.

Key terms agreement, concordance, job-exposure matrix, occupational exposures, reliability.

A job-exposure matrix is a classification system linking occupation and industry titles with job-related exposures. Job-exposure matrices were developed to infer specific occupational exposure more accurately on the basis of simple industry or occupation statements. If accurate, one principal advantage of such a matrix would be the reduction in misclassification that occurs when workers are grouped on the basis of exposure rather than job title (1).

Several matrices have been developed (2—14). Two of these were examined in our analyses, the one developed by Sheila Hoar Zahm and her colleagues at Harvard (6) and the one based on data obtained in the National Occupational Hazard Survey (NOHS) conducted by the National Institute for Occupational Safety and Health (NIOSH) (5). Both of these matrices were developed for use in population-based studies and cover a broad range of occupations and agents.

The matrix developed by Hoar et al was designed to study occupational carcinogenesis. In this data base, there

are 334 unique agents in the system, most of which are known or suspected chemical carcinogens, although some agents associated with noncancer health outcomes are also included. The list of exposure entities was compiled from several comprehensive sources in the relevant scientific and engineering literature (ie, no worksite visits were made). Occupations were coded according to the *Dictionary of Occupational Titles* (15); industries were coded according to the Standard Industrial Classification (SIC) (16). Agents were assigned a 4-digit code, and approximately 500 industry-occupation pairs are contained in the matrix. In addition to linking each industry-task combination with the exposure entities, the degree of linkage to each agent was assigned when possible (1 = light, 2 = moderate, 3 = heavy).

The NOHS-based matrix was developed with data obtained in 1972—1974 (17—19). The NOHS was a walk-through survey of more than 4000 United States industries representing over 600 SIC codes. Data on >8000 agents ("hazards") were collected by engineers

¹ Division of Epidemiology and Biometrics, School of Public Health, Columbus, Ohio, United States.

² Currently: Department of Human Services, Maine Bureau of Health, Occupational Health Program, Augusta, Maine, United States.

Reprint requests to: Dr JR Wilkins III, Division of Epidemiology and Biometrics, School of Public Health, B-150 Starling-Loving Hall, 320 West Tenth Avenue, Columbus, OH 43210, United States.

trained to recognize potentially harmful occupational exposures. Although occupations were originally coded using 1967 SIC codes for industries and 1970 census codes for occupations, a version of the job-exposure matrix has been developed using 1980 census codes (20). In this data base, each hazard was given a 5-digit code. In addition, the NOSH-based matrix contains the estimated proportion of workers potentially exposed to each hazard.

Since job-exposure matrices are used to infer exposures and ultimately to evaluate the association between these exposures and disease risk, it should be informative to examine how well the matrices agree with regard to the job-exposure links contained therein. Previous comparisons of the 2 matrices appear limited to the work of Linet et al (21), who coded the occupations of cases and referents in a study of chronic lymphocytic leukemia and assessed concordance for 7 selected agents by computing the kappa statistic κ , a widely used quantitative measure of agreement (22). Based on the guidelines suggested by Landis & Koch (23) for the interpretation of κ values, slight to fair agreement was found for 5 of the 7 exposures (κ range = 0.01 to 0.28), while the κ values were somewhat better for asbestos and butadiene, ranging from 0.41 to 0.60, indicating "moderate" agreement.

The present study was undertaken to compare the operating characteristics of the Hoar et al matrix and the NOHS-based matrix more comprehensively by estimating the extent of agreement on the exposures for occupations in the following 3 industries: metal (SIC code 06), paper and wood (SIC code 04), and chemicals, drugs and paints (SIC code 11). (Limited resources precluded the examination of more industries.)

Material and methods

Differences in the occupational coding systems and in the number of agents contained in each matrix necessitated completion of a 2-step process prior to comparing the matrices of interest. First, agents had to be selected that were common to both matrices, and, second, the occupational codes of 1 system had to be converted to equivalent occupational codes of the other.

Selection of agents (step 1)

Since the Hoar et al matrix contains many fewer agents than the NOHS-based matrix, the agent list for the comparison was based on the Hoar et al matrix. Agents were selected according to the following rules.

1. For each exposure entity (a single agent or group of agents in the Hoar et al matrix), the identical agent or group was sought in the NOHS-based matrix. If the Hoar

et al matrix specified only a single agent such as "benzene", only the compound benzene was included in the comparison even though the NOHS-based matrix contained benzene along with many benzene-containing compounds. On the other hand, if the Hoar et al matrix specifically indicated that the entry was "X and derivatives" or "X and compounds", then all the derivatives (or compounds) in the NOHS-based matrix were included. In the analysis, the exposure was classified as present if the job was positive for any of the separate compounds in the NOHS-based matrix.

For groups of compounds in the Hoar et al matrix, such as "pigments", the identical group was sought in the NOHS-based matrix. Thus, for the Hoar et al entry "pigments", the entry used in the NOHS-based matrix was "pigments", and individual pigments were not included in the comparison.

2. If an entry in the Hoar et al matrix included 2 compounds entered as X/Y, both compounds had to be found in the NOHS-based matrix to be used in the comparison. For example, "sodium hydroxide/potassium hydroxide", although 2 compounds, represent 1 entry in the Hoar et al matrix and has the 4-digit code 6825. Therefore, both sodium hydroxide and potassium hydroxide had to be located in the NOHS-based matrix in order for entry 6825 to be included in the comparison. However, sodium hydroxide and potassium hydroxide did not have to be in the combined form in the NOHS-based matrix. For "potassium cyanide/ferrocyanides", code 6800, only potassium cyanide was contained in the NOHS-based matrix, so that the entity with code 6800 was not included in the comparison.

3. If the identical entry could not be found, then an agent or group of agents was selected when acceptably similar.

A total of 214 agents were included in the comparison (64% of the total number of entries in the Hoar et al matrix). One hundred and sixty-six of the 214 were entries for which 1 entry in the Hoar et al matrix corresponded to 1 identical entry in the NOHS-based matrix. Thirty-five were entries for which 1 entry in the Hoar et al matrix corresponded to more than 1 entry in the NOHS-based matrix (the majority of these were "metal and compounds") and 13 were a similar match, but not identical.

Conversion of occupational codes (step 2)

In the Hoar et al matrix, occupations are coded according to the *Dictionary of Occupational Titles*, which groups jobs together on the basis of similar tasks and assigns each of these groups a number. The occupational codes in the Hoar et al matrix thus represent *groups* of jobs. On the other hand, the NOHS-based matrix codes each *indi-*

vidual occupation with a 1980 Industry and Occupation Census Bureau code.

The occupational codes of the Hoar et al matrix were converted to equivalent codes in the NOHS-based matrix by assigning a 1980 Industry and Occupation Census Bureau code to each of the individual jobs in a given occupational group in the *Dictionary of Occupational Titles*. The third edition of the *Dictionary of Occupational Titles* was used in the analysis since this was the one used for the development of the Hoar et al matrix. Only occupations identified as base titles in the *Dictionary of Occupational Titles* were used. If a census code for the base title could not be located, then the synonym for the base title was used to try to find the census code. Jobs noted as occurring in "any industry" in the *Dictionary of Occupational Titles* were not coded.

Two occupational groups, 899 (miscellaneous structural work occupations, not elsewhere classified) and 929 (packaging and materials handling occupations, not elsewhere classified), contained jobs from diverse industries. For our analysis, only jobs pertaining to the 3 industries under study were chosen for analysis.

Approach to performing comparison

The analysis was performed by comparing the exposure entities associated with an occupational group in the Hoar et al matrix with those associated with the corresponding census codes in the NOHS-based matrix. Each of 214 exposure entities was classified as to its "pres-

ence" or "absence" in a matrix, and a kappa statistic (κ) was calculated. Table 1 details the results of comparing the 2 data bases with respect to occupational group 819 in the metal industry. As can be seen, there were 7 exposure entities (3.3% of all 214 hazards) present in both matrices (acetylene, copper and compounds, silver and compounds, titanium and compounds, zinc and compounds, carbon dioxide, and ultraviolet radiation). The NOHS-based matrix indicated 17 entities as present (7.9% of total) that the Hoar et al did not, and the Hoar et al matrix indicated 22 entities as present (10.3% of total) that the NOHS-based JEM indicated as absent. One hundred and sixty-eight agents were not present in either matrix (78.5% of the total). Kappa here equaled 0.16, indicative of "slight" agreement (23).

This procedure was followed for each of the occupational groups considered. Note that the Hoar et al exposure-intensity measures were ignored in the comparisons of the 2 data bases, as was the percentage of workers exposed to a particular hazard estimated by the NOHS-based matrix.

To complement the κ values calculated for each occupational group, the G index was estimated (24, 25). G is equal to the proportion of agreeing responses minus the proportion of nonagreeing responses and can be calculated even when a cell entry is 0 and κ is undefined (26).

In addition to the assessment of the possible variation in agreement between the 2 data bases by occupation and industry, differences between the matrices for single ex-

Table 1. Comparison of the Hoar et al job-exposure matrix and job-exposure matrix based on the National Occupational Hazard Survey (NOHS) for exposure entities in occupational group 819 in the metal industry (welders, flame cutters, and related occupations)

Agreement		Disagreement	
Present in both matrices (+/+)	Present in neither matrix (-/-)	Present in NOHS-based matrix only (+/-)	Present in Hoar et al matrix only (-/+)
Acetylene	(numerous)	Paints	Benzene
Copper and compounds		Toluene	Epoxides
Silver and compounds		Methylene chloride	Acetamide
Titanium and compounds		1,1,1-trichloroethane	Acetic acid
Zinc and compounds		Ethanol	Diethylene glycol
Carbon dioxide		Acetone	Rosin
Ultraviolet radiation		Lubricants	Beryllium and compounds
		Mineral oil	Arsine
		Surfactants	Bismuth and compounds
		Aluminum/aluminum oxides	Brass
		Phosphorus	Cadmium and compounds
		Boric acid	Chromium and compounds
		Enamels	Iron
		Graphite	Lead
		Sulfuric acid	Manganese
		Abrasives	Magnesium and compounds
		Glass fibers	Tin and compounds
			Phosphine
			Zinc chloride
			Carbon monoxide
			Nitrogen oxides
			Ozone
Number	7	17	22
Percentage of 214	3.3	7.9	10.3

Table 2. Comparison of the job-exposure matrix based on the National Occupational Hazard Survey (NOHS) and the Hoar et al job-exposure matrix for inferred exposure to aldehydes, ketones, and ethers in the chemicals, drugs and paints industry.^a

Agents in family	Occupational groups from the <i>Dictionary of Occupational Titles</i>														Frequency of agents			
	550	551	552	553	554	555	556	557	558	559	737	849	899	929	+/+	-/-	+/-	-/+
Acetaldehyde	+/+	+/+	-/+	+/+	-/+	-/+	-/+	-/+	+/+	+/+	+/-	-/-	-/-	-/-	5	3	1	5
Acetone	+/-	+/-	+/-	+/-	+/-	-/+	+/-	+/-	+/-	+/-	+/-	+/-	-/-	+/-	2	2	10	0
Acrolein	-/+	-/+	-/+	-/+	-/+	-/+	-/+	-/+	-/+	-/+	-/+	-/+	-/+	-/+	0	4	0	10
Anesthetics	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	-/-	0	14	0	0
Ethyl ether	+/+	+/+	+/+	+/+	+/+	-/+	-/+	-/+	+/+	+/+	+/+	+/-	-/-	-/-	8	2	1	3
Formalin	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/-	-/-	+/-	10	2	2	0
Ketones, not otherwise specified	+/+	-/+	-/+	+/+	-/+	-/+	-/+	-/+	+/+	+/+	-/+	-/+	-/+	-/-	4	1	0	9
Oil of clove	-/-	-/+	-/-	+/-	+/-	-/+	+/-	-/+	+/-	+/-	-/+	-/+	-/+	-/-	0	9	5	0
Total number															29	37	19	27
Percentage of total (rounded)															26	33	18	23

^a Key:
 +/+ = Both matrices noted agent present for occupational group.
 -/- = Both matrices noted agent absent for occupational group.
 +/- = Agent noted as present in NOHS-based matrix, but absent in Hoar et al matrix.
 -/+ = Agent noted as absent in NOHS-based matrix, but present in Hoar et al matrix.
 550 = Mixing and blending occupations.
 551 = Filtering, straining, and separating occupations.
 552 = Distilling occupations.
 553 = Heating, baking, drying, seasoning, melting, and heat-treating occupations.
 554 = Coating, calendaring, laminating, and finishing occupations.
 555 = Grinding and crushing occupations.
 556 = Casting and molding occupations, not elsewhere classified.
 557 = Extruding occupations.
 558 = Reacting occupations, not otherwise classified.
 559 = Occupations in processing of chemicals, plastics, synthetics, rubber, paint & related products, not otherwise classified.
 737 = Occupations in fabrication of ammunition, fireworks, explosives, and related products.
 849 = Painting, plastering, waterproofing, cementing, and related occupations, not otherwise classified.
 899 = Miscellaneous structural work occupations, not otherwise classified.
 929 = Packaging and materials handling occupations, not otherwise classified.

posure entities and for chemical families by industry were also of interest. The approach taken is illustrated in table 2, where group-specific concordance-discordance results of the *Dictionary of Occupational Titles* are shown for each exposure entity in family 12 (aldehydes, ketones, ethers) in the chemical, drugs and paints industry. The tabulated concordance-discordance percentages (25.9%, 33.0%, 17.9%, 23.2%) were derived on an industry-specific basis by summing the frequency counts across the individual hazards in the family. These concordance-discordance percentages were also plotted in bar chart form by industry and family.

Results

Proportion of occupations coded and located in the NOHS-based job-exposure matrix

Table 3 shows the number of occupational groups selected for analysis in the industries under study. These groups represent all the occupational groups in the Hoar et al matrix which were not managerial, professional, or secretarial. The number of jobs selected from each group varied considerably with an average of 25 jobs per group in the metal industry, 21 in the paper and wood industry, and 56 in the chemicals, drugs and paints industry.

Table 3. Proportion of occupations coded with a 1980 Census Bureau code and located in a job-exposure matrix based on the National Occupational Hazard Survey (NOHS) for each industry.

Industry as denoted in Hoar et al job-exposure matrix	Number of occupational groups selected for analysis from the <i>Dictionary of Occupational Diseases</i>	Total number of jobs selected for the comparison, excluding those noted as occurring in "any industry"	Jobs in each occupational group		1st step of code conversion		2nd step of code conversion	
			Mean	Range	Jobs in an occupational group, for which a 1980 census code was found		Jobs in an occupational group, located in NOHS-based job-exposure matrix	
					Mean (%)	Range (%)	Mean (%)	Range (%)
Metal	22	543	25	3—79	71.1	55—100	44.8	9—83
Paper and wood	17	356	21	3—59	76.1	58—100	36.8	8—58
Chemicals, drugs and paints	15	845	56	2—251	65.8	31—100	33.7	0—52

Census codes could not be found for all the jobs selected in the occupational groups. An average of 71.1% of the jobs in each group of the metal industry could be coded, 76.1% in the paper and wood industry, and 65.8% in the chemicals, drugs and paints industry.

Once the jobs were coded, the NOHS-based matrix was searched for these census codes. Not all jobs that were coded could be located in the NOHS-based matrix, since the National Occupational Hazard Survey was a sample of industries and occupations. In general, less than half of all jobs in each occupational group were found in the NOHS-based matrix. For the metal industry, the mean percentage located was 44.8%. In the paper and wood industry, this percentage was 36.8%, and for the chemicals, drugs and paints industry 33.7% were found, on the average. For 1 group in the chemicals, drugs and paints industry (number 955), none of the census codes were found in the NOHS-based matrix, and therefore this group was eliminated.

Agreement between the matrices for presence or absence of agents in each occupational group

It was possible to calculate kappa for 20 of the 22 occupational groups in the metal industry. Kappa could not be calculated in occupational group 503 because none of the agents identified in the NOHS-based matrix were used in the comparison, and in group 899 none of the agents were present in both matrices. The highest κ of 0.27 occurred in group 510 (mixing occupations in ore refining). The lowest κ of 0.02 was found for group 809 (structural work occupations). Despite the low kappas, many were statistically significant in conventional terms ($P \leq .05$). In general, the NOHS-based matrix indicated

more agents present for each occupational group than did the Hoar et al matrix.

Results for the paper and wood industry showed higher G scores (compared with the metal industry results) and thus indicated more agreeing responses; however the majority of the κ_i were less than 0.10, and some of the κ_i were negative and therefore indicated poor agreement beyond what one would expect by chance alone. Kappa could not be calculated for 3 groups because there were either no agents present in both matrices, or there was no discordance for 1 of the matrices. The highest κ of 0.24 occurred in this industry in group 659 (printing occupations), and the lowest of -0.07 was seen in group 561 (preserving occupations in wood processing).

Somewhat different results were obtained for the chemicals, drugs and paints industry. For the occupational groups, the percentage of agreeing responses for reported exposures tended to be lower and the majority of the κ_i were less than 0.10, several being less than 0. The highest κ of this industry, 0.14, occurred for group 849 (painters), and the lowest of -0.12 for grinding and crushing occupations. In contrast to results for the metal and paper and wood industries, the Hoar et al matrix usually indicated more agents present in an occupational group than did the NOHS-based matrix.

Agreement between matrices for the presence or absence of agents within a chemical family

Figures 1 and 2 graphically illustrate the concordance-discordance results by industry and chemical family. For example, in the metal industry, the best agreement (99%) was seen for family 6 (alicyclic halogens), followed by family 10 (aliphatic nitro compounds), family 2 (ar-

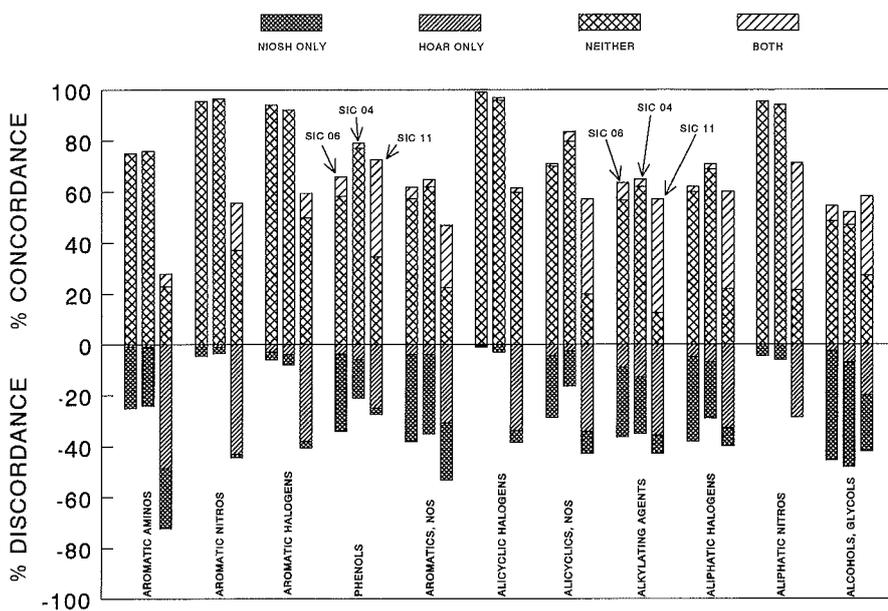


Figure 1. Concordance-discordance results by industry for hazard families (NIOSH = National Institute for Occupational Safety and Health in the United States, ie, based on the National Occupational Hazard Survey; SIC = standard industrial classification; NOS = not otherwise specified) (part 1).

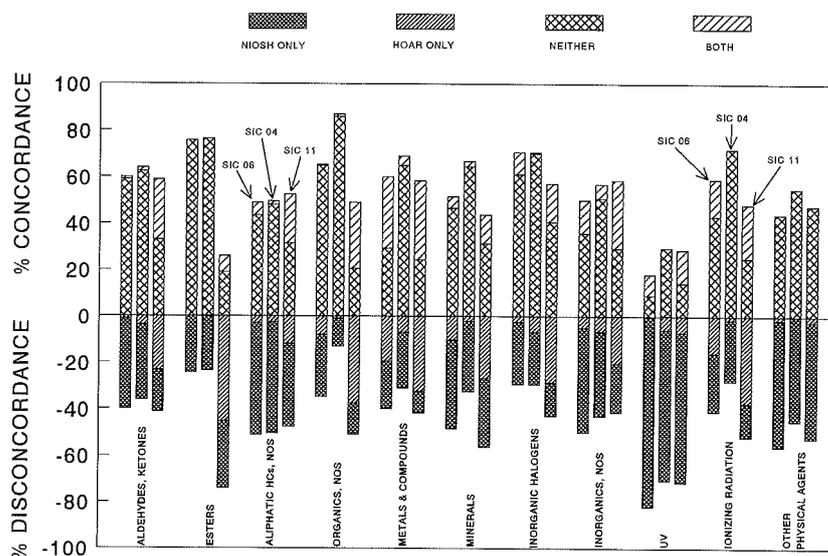


Figure 2. Concordance-discordance results by industry, for hazard families (NIOSH = National Institute for Occupational Safety and Health in the United States, ie, based on the National Occupational Hazard Survey); SIC = standard industrial classification; NOS = not otherwise specified; HCs = hydrocarbons, UV = ultraviolet radiation) (part 2).

matic nitro compounds), and family 3 (aromatic halogens). Two families demonstrated more disagreement than agreement: family 22 (other physical agents) and family 20 (nonionizing radiation/ultraviolet radiation). In general, most concordance occurred because the matrices agreed on the absence of exposure to the agents in a family. Only 2 families — metals, metalloids and compounds and nonionizing radiation — showed a substantial portion of the overall agreement to be explained by both matrices agreeing on the presence of exposures. Consistent with the data obtained for the occupational groups, much of the disagreement within a family occurred because the NOHS-based matrix indicated that the agents were present in the industry when the Hoar et al matrix did not.

Table 4 summarizes the family-specific concordance-discordance findings for each industry examined. The summary statistics indicate that, on the average, there was more concordance than discordance, particularly for the metal and paper and wood industries (total average concordance 65.5% and 71.3%, respectively). Except for chemicals, drugs and paints, the “neither” component of concordance (in the 60–70% range) generally far outweighed the “both” component (in the 2–5% range).

It should be stressed that the proportions illustrated in figures 1 and 2 represent an average of all the agents in a chemical family across all the occupational groups in the industry. It is possible that, even though a family shows only 50% agreement, for a single agent, there may be a high proportion of agreement between the matrices. For example, in table 2, the overall proportion of agreement for the chemical family aldehydes, ketones and ethers in chemicals, drugs and paints was 59%. However, 2 separate agents, formalin and anesthetics, showed a high pro-

portion of agreement. This pattern was not unique to this chemical family or this industry.

Discussion

The magnitude of the κ values for the 3 industries studied suggest, on the whole, poor to only slight or fair

Table 4. Summary statistics for the family-specific concordance-discordance results, by industry. (NOHS = National Occupational Hazard Survey)

Industry	Mean (%) (N = 22)	SD	Median (%)	Range
Metal				
Concordance				
Both (+/+)	5.4	7.4	3.2	0–30.8
Neither (-/-)	60.1	23.0	58.7	9–98.8
Discordance				
NOHS-based only (+/-)	29.9	18.7	27.1	0.6–82.0
Hoar et al only (-/+)	4.5	5.2	3.0	0–19.4
Paper and wood				
Concordance				
Both (+/+)	1.6	1.8	1.1	0–6.1
Neither (-/-)	69.7	17.5	69.5	29.4–96.5
Discordance				
NOHS-based only (+/-)	24.9	15.8	23.2	2.4–64.7
Hoar et al only (-/+)	3.8	3.4	3.2	0–13.2
Chemicals, drugs, and paints				
Concordance				
Both (+/+)	23.3	14.0	23.9	0–50.0
Neither (-/-)	29.1	11.4	26.1	12.5–55.4
Discordance				
NOHS-based only (+/-)	18.4	16.1	14.3	0–64.3
Hoar et al only (-/+)	29.2	11.7	31.6	2.4–48.6

agreement between the two matrices, at least when the interpretive guide suggested by Landis & Koch is used (23). On the other hand, Schmidt et al (26) found that kappa weights conservatively for concordance on "no exposure" as compared with other measures of agreement. For example, for the paper and wood industry, the G index (the proportion of agreeing responses minus the proportion of disagreeing responses) tended to be higher than for the metal industry because there was a high proportion of agreeing responses due to agreement on "no exposure". The κ values for the paper and wood industry were lower, however, because of the way in which κ was calculated. Kappa was defined as $(p_o - p_e)/(1 - p_e)$, where p_o = observed probability of concordance between the 2 surveys and p_e = expected probability of concordance between the 2 surveys. In addition, $p_e = (a_i b_i)$, where a_i and b_i are the marginal probabilities for the i th category in the $c \times c$ contingency table relating the 2 surveys (27). Consequently, if there is a high proportion of agreeing responses due to "no exposure", as occurred for jobs in the paper and wood industry, the p_e is larger and κ smaller.

While other authors have noted that time and place are important considerations in using a matrix to infer occupational exposure (2, 3, 28), the 2 matrices examined in this study were developed in the United States using data from approximately the same time period so that a difference in place or time probably did not significantly contribute to the observed differences. In addition, there could be some bias due to miscoding of the jobs, but it is difficult to estimate the magnitude and direction of such bias.

Kappa is a measure of reliability and all that can be concluded from the present study is that the matrices are only partly reliable when compared with each other for reported exposures in an occupational group. The interpretation of the difference between the 2 matrices is difficult, however, because in the Hoar et al matrix, "exposure" is not defined, and the degree of linkage was excluded from our comparison. We also excluded the estimated percentage of employees potentially exposed to any given agent in the NOHS-based job-exposure matrix. It is possible that, had the analysis been restricted to comparing only agents for which there was a high (or low) degree of exposure in the Hoar et al matrix with agents that had a high (or low) proportion of employees potentially exposed in the NOHS-based matrix, the observed agreement might have been greater.

One limitation of both of these matrices is that neither of them was constructed in a way that allows for the identification of differences in exposure to workers with the same job title or in the same occupational group. (The NOHS-based matrix identifies the percentage of workers potentially exposed to a given agent for an industry-occupation combination, but it does not quantify

the level of exposure.) This type of distinction is important in industries such as chemical manufacturing where the highest exposures to chemicals can occur, but only for a small fraction of the workers in an occupational group. If either of these matrices were applied to workers in such an industry, some of the workers would be misclassified as to exposure status.

Unfortunately, there is no universally accepted "gold standard" by which to judge the sensitivity and specificity of these 2 matrices. Therefore, it is not possible to assess whether 1 of the matrices misclassifies workers or whether the matrices are simply classifying workers for different things.

Hinds et al (29) previously suggested that the Hoar et al matrix may have a low sensitivity. When they used it in a case-referent study of lung cancer, only 2 of 7 known or suspected lung carcinogens produced a statistically significant increase in the risk of lung cancer, as well as a dose-response effect. The authors did indicate, however, that there were potential biases in their data that may have influenced their results, and they concluded that the low sensitivity might "indicate a general shortcoming or one peculiar to [their] data set" (29).

Linet et al (21) assessed the sensitivity and specificity of the NOHS-based matrix and the Hoar et al matrix by assuming self-reported exposure to benzene and asbestos as the "gold standard". When using this method, they found the sensitivity to be low, but the specificity to be greater than 80% for both. The high specificity was attributed, however, to the low proportion determined to be exposed. Kromhout et al (30) used self-reported exposure to asbestos (as well as to several other agents) to assess the sensitivity and specificity of the Hoar et al matrix. They found results similar to those of Linet et al for asbestos, but a somewhat higher sensitivity for exposure to wood dust and pesticides (greater than 0.50).

Ciccione et al (31) also evaluated the NOHS-based matrix. They used occupational data from patients with mesothelioma (obtained from a tumor register) to compare assignment of exposure to asbestos by the NOHS-based matrix with the authors' own assignment based on job title and found only a weak correlation between the 2.

Our results suggest that the 2 matrices we compared have different sensitivities and specificities and that when applying these matrices to the same population, researchers could reach different conclusions regarding clues as to the etiology of disease. In addition, if the sensitivity and specificity of the 2 matrices are different, the positive predictive value of each will be affected.

Concluding remarks

The results of the present comparison underscore the importance of classification rules in developing and using

matrices. Our results show that the NOHS-based matrix and the Hoar et al matrix generally differ and that, when applied to the same population, different conclusions are reached. The reasons for these differences are not clear, as differences in exposures were noted for both individual agents and families of agents and they also appear to be somewhat dependent on industry. It is important to note, however, that there was also some agreement between the 2 matrices. The variation in the assignment of exposure by these 2 matrices, as well as the limitations in the occupational coding systems, should be considered when either of these matrices are used.

Acknowledgments

The authors wish to acknowledge Dr Mac Crawford for his assistance in computing the kappa statistics, Dr Mel Moeschberger and the staff of the Division's Biometrics Laboratory for additional data management and analysis support, and Ms Patricia Price for her preparation of the manuscript.

References

1. Hoar SK. Job-exposure matrix methodology. *J Toxicol Clin Toxicol* 1983—1984;21(1+2):9—26.
2. Kauppinen T, Partanen T. Use of plant- and period-specific job-exposure matrices in studies on occupational cancer. *Scand J Work Environ Health* 1988;14:161—7.
3. Macaluso M, Vineis P, Continenza D, Ferrario F, Pisani P, Andisio R. Job exposure matrices: experience in Italy. In Acheson ED, editor. *Job-exposure matrices*. Southampton (England): University of Southampton, Medical Research Council, 1983:22—30. Conference report.
4. Bonassi S, Merlo F, Pearce N, Puntoni R. Bladder cancer and occupational exposure to polycyclic aromatic hydrocarbons. *Int J Cancer* 1989;44:648—51.
5. Sieber WK, Sundin DS, Frazier TM, Robinson CF. Development, use, and availability of a job-exposure matrix based on National Occupational Hazard Survey data. *Am J Ind Med* 1991;20:163—74.
6. Hoar SK, Morrison AS, Cole P, Silverman D. An occupation and exposure linkage system for the study of occupational carcinogenesis. *J Occup Med* 1980;22:722—6.
7. Pannett B, Coggon D, Acheson ED. A job-exposure matrix for use in population based studies in England and Wales. *Br J Ind Med* 1985;42:777—83.
8. Steineck G, Plato N, Alfredsson L, Lorell S. Industry related urothelial carcinogens: application of a job-exposure matrix to census data. *Am J Ind Med* 1989;16:209—24.
9. Vaughan TL, Strader C, Davis S, Daling JR. Formaldehyde and cancers of the pharynx, sinus, and nasal cavity: I. occupational exposures. *Int J Cancer* 1986;38:677—83.
10. Vineis P, Magnani C. Occupation and bladder cancer in males: a case-control study. *Int J Cancer* 1985;35:599—606.
11. Gamble J, Spirtas R. Job classification and utilization of complete work histories in occupational epidemiology. *J Occup Med* 1976;18:399—404.
12. Coggon D, Pannett B, Acheson ED. Use of a job-exposure matrix in an occupational analysis of lung and bladder cancers on the basis of death certificates. *JNCI* 1984;72:61—5.
13. Siemiatycki J, Day NE, Fabry J, Cooper JA. Discovering carcinogens in the occupational environment: a novel epidemiologic approach. *JNCI* 1981;66(2):217—25.
14. Gerin M, Siemiatycki J, Kemper H, Begin D. Obtaining occupational exposure histories in epidemiologic case-control studies. *J Occup Med* 1985;27:420—6.
15. US Department of Labor. *Dictionary of occupational titles; vol I (Definition of titles) & vol II (Occupational classification)*. Washington (DC): United States Government Printing Office, 1965.
16. Office of Management and Budget. *Standard industrial classification manual*. Washington (DC): United States Government Printing Office, 1972. Stock no 041-001-0066-6.
17. Seta JA, Sundin DS, Pedersen DH. *National occupational exposure survey field guidelines*. Cincinnati (OH): United States Department of Health and Human Services, Public Health Service, Centers for Disease Control, National Institute for Occupational Safety and Health, Division of Surveillance, Hazard Evaluations and Field Studies, 1988.
18. Sieber WK. *National occupational exposure survey sampling methodology*. Cincinnati (OH): United States Department of Health and Human Services, Public Health Service, Centers for Disease Control, National Institute for Occupational Safety and Health, Division of Surveillance, Hazard Evaluations and Field Studies; 1990.
19. Pedersen DH, Sieber WK. *National occupational exposure survey analysis of management interview responses*. Cincinnati (OH): United States Department of Health and Human Service, Public Health Service, Centers for Disease Control, National Institute for Occupational Safety and Health, Division of Surveillance, Hazard Evaluations and Field Studies, 1988.
20. United States Bureau of the Census. *1980 census of population: alphabetical index of industries and occupations*. Washington (DC): United States Government Printing Office, 1980.
21. Linet MS, Stewart WF, VanNatta ML, McCaffrey LD, Szklo M. Comparison of methods for determining occupational exposure in a case-control interview study of chronic lymphocytic leukemia. *J Occup Med* 1987;29:136—11.
22. Fleiss JL. *Statistical methods for rates and proportions*. 2nd edition. New York (NY): John Wiley and Sons Inc, 1981: 212—35.
23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159—74.
24. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20(1):37—46.
25. Holley JW, Guilford JP. A note on the G index of agreement. *Educ Psychol Meas* 1964;24(4):749—53.
26. Schmidt S, Lemasters G, Rice C, Succop P. A comparison of reliability estimates when industrial hygienists determine workplace exposures (abstract). *Am J Epidemiol* 1992;136(8): 952.
27. Rosner B. *Fundamentals of biostatistics*. 2nd edition. Boston (MA): PWS Publishers, 1986:425.
28. Acheson ED. What are job exposure matrices? In: Acheson ED, editor. *Job exposure matrices*. Southampton (England): University of Southampton Medical Research Council, 1983:1—4. Conference report.
29. Hinds MW, Kolonel LN, Lee J. Application of a job-exposure

- matrix to a case-control study of lung cancer. *JNCI* 1985;75: 193—7.
30. Kromhout H, Heederik D, Dalderup LM, Kromhout D. Performance of two general job-exposure matrices in a study of lung cancer morbidity in the Zutphen cohort. *Am J Epidemiol* 1992;136(6):698—711.
 31. Ciccione C, London SJ, Garabrant DH, Bernstein L, Phillips K, Peters JM. Occupational asbestos exposure and mesothelioma risk in Los Angeles county: application of an occupational hazard survey job-exposure matrix. *Am J Ind Med* 1991;20: 371—9.

Received for publication: 10 October 1995