Risk Analysis, Vol. 6, No. 1, 1986

Computerized Estimates of Potential Occupational Health Risk Due to Chemical Exposure

David H. Pedersen[1] and Richard W. Hornung[1]

Received June 18, 1984; revised January 9, 1985

Estimating the potential health risk encountered by workers due to their exposure to various chemicals is enormously complex, since many chemicals be involved and each may have multiple toxic effects. As an aid to this estimation process, a computer program, or model, which computes index numbers expressing the relative health risk of occupational groups due to their potential chemical exposures was developed at the National Institute for Occupational Safety and Health (NIOSH). This model considers an inventory of the chemicals to which specific occupational groups are potentially exposed, the published information regarding the toxic effects of each chemical, and the conditions of occupational exposure. The system then develops indices of potential occupational group health risk by considering weighted combinations of eight distinct health effects. No direct comparison with external occupational risk indices is currently possible, but internal testing of the model reveals no obvious inconsistencies.

KEY WORDS: chemicals; computerized model; health effects; toxicity; occupational risks.

1. INTRODUCTION

Estimation of potential worker health risk due to chemical exposure in the workplace is a critical factor in occupational health programs, since these estimates are commonly used to prioritize efforts to eliminate or minimize adverse effects on workers. National Institute for Occupational Safety and Health (NIOSH) personnel have developed a computer system, or model, based on published toxicological and occupational survey data, which provides an expression of potential occupational health risk due to chemical exposure. The computer system estimates of occupational health risk are derived in a three-step process: (1) access relative toxicity (i.e., a quantitative judgement of the relative potency) of the chemicals to which workers are potentially exposed, (2) characterize the national incidence and conditions of workers exposure to each of the chemicals in the model, and (3) estimate the potential health risk due to the multiple chemical exposure associated with occupational groups.

Output from the system produces three basic formats, which provide: (1) a listing of 1721 chemicals rank-ordered by relative toxicity and user weighting of health effects, (2) a listing of chemicals rank-ordered on the basis of both toxicological data and data on the conditions and extent of potential worker exposures, and (3) a listing of industries or occupations rank-ordered by potential health risk due to the multiple chemical exposure of each group, and the potential health effects, extent, and conditions of such exposure. Those desiring a comprehensive discussion of the development of the model and the caveats associated with its use are referred to the NIOSH Technical Report[(1)] on the model.

[1] National Institute for Occupational Safety and Health, Robert A. Taft Laboratories, Mail Stop R-19, 4676 Columbia Parkway, Cincinnati, Ohio 45226.

0272-4332/86/0300-0003$05.00/1

2. THE DATA SOURCES

The data used in the modeling system exists primarily in two unique NIOSH data bases, the National Occupational Hazard Survey (NOHS),[2] and the Registry of Toxic Effects of Chemical Substances (RTECS).[3] The RTECS file is a compilation of published toxicological and related data derived from a review of the international professional literature. The January 1981 version contains data on 45,156 different chemicals involving 73,357 toxicity data measurements across a total of 31 different test species and 26 routes of administration. The toxicity data in RTECS consist of the lowest doses producing a specific effect reported for a wide variety of test animals and regimens. We acknowledge that the data available are not all-inclusive, and that RTECS does not evaluate the validity of the cited studies.

The NOHS is a 2-year study conducted by the NIOSH starting in 1972. The survey was designed to describe the health and safety conditions in the American work environment and to determine the extent of worker exposure to chemical and physical agents and the conditions under which such exposure occurs. The sample of businesses in the survey represents all nonagricultural business covered under the Occupational Safety and Health Act of 1970. Exposure data drawn from the NOHS base must be considered "potential," since no environmental level measurements exist in the base for the observations of worker exposure to chemical agents.

3. ASSESSMENT OF RELATIVE CHEMICAL TOXICITY

The first major output of the modeling system is an assessment of the relative toxicity of individual chemicals and a rank-ordering of chemicals on the basis of this calculated relative toxicity. This output is termed the hazard risk index (HRI), and is the end result of a multistage effort involving the categorization, standardization, and normalization of the dose-response data contained in RTECS.

Initial analysis of the RTECS data indicated that all reported toxic endpoints could be contained in eight general categories: (1) acute toxicity (AT), (2) carcinogenicity (CAR), (3) equivocal tumorigenic agent (ETA), (4) mutagenicity (MUT), (5) neoplastigenicity (NEO), (6) primary [skin & eye] irritation (PI), (7) teratogenicity (TER), and (8) other toxic effects (TFX).

A major task in preparing the RTECS data for use in the computer system was the standardization of the dose units reported. Accordingly, the dose units from all inhalation studies were converted to ppm notation, and all other dose units were converted to millimole/kg. The millimole notiation was selected because it corresponded to the number of mulecules producing the observed effect and was, therefore, considered more indicative of the toxicological activity of specific chemicals than a mg/kg notation.

In addition, certain data were empirically rejected. In the interests of report standardization, only those tests performed on rabbits and using the Draize procedure were utilized under the (PI) category, and only those mutagenic tests utilizing mammels or cells of mamallian origin were accepted. Due to the wide variety of dose units, test species or cell types, and toxic effect notations encountered in RTECS, a small number were found to be incompatible with the test class or dose unit standardization procedures. Complete documentation on these exclusion decisions and on the procedures to convert dose amounts to the standard units described here are available from the senior author on request, and/or are contained in ref. (1). A brief description of the algorithm used to derive the HRI can be found in the appendix to this paper.

An example listing of a single chemical as produced in HRI printout is shown is Fig. 1. The data display format is in several parts, as follows:

1. The first line of the heading at the top of each page identifies the index, the date of the RTECS data used in its preparation, the date the index was produced, and the sequential page number.
2. The second and third lines of the heading identify the values set by the user in calculating the HRI. They also identify the use or nonuse of the carcinogen option and neoplastigen option.
3. The entry for each chemical is in three sections. The first line gives the rank-order number of the chemical (out of the 1721 common to RTECS and NOHS), identifies it as a chemical common to both NOHS and RTECS ("H"), cites the RTECS accession number, and gives the RTECS primary name. The

```
HRI Calculations based on RTECS data as of 01/81                DATE:   08/17/82    page 1

Multipliers for:   AT=1   CAR=1   ETA=1   MUT=1   NEO=1   PI=1   TER=1   TFX=1   NEO OPTION= NO

Constants for:     AT=0   CAR=0   ETA=0   MUT=0   NEO=0   PI=0   TER=0   TFX=0   CAR OPTION= NO

SEQ# 1       H     AB1925000  Acetaldehyde

  AT=.186   CAR=.000   ETA=.000   MUT=.440   NEO=.000   PI=.120   TER=.131   TFX=.393   HRI=1.271
```

ROUTE	SPECIES	TEST	OUTCOME	# OF TESTS	DOSE VALUES
ANY	ANY	ANY	TER	300	.132
ANY	ANY	DND	MUT	189	.128
ANY	ANY	SCE	MUT	88	.192
ANY	ANY	VIV	MUT	142	1.000
EYE	RBT	SSSS	PI	1454	.228
IHL	HMN	TCLO	TFX	89	.394
IHL	RAT	LCLO	AT	424	.340
IPR	RAT	LDLO	AT	865	.163
IVN	MUS	LD50	AT	10478	.165
ORL	RAT	LD50	AT	5118	.148
SCU	MUS	LD50	AT	3581	.168
SCU	RAT	LD50	AT	755	.133
SKN	RBT	SSSS	PI	1265	.014

Fig. 1. Example Hazard Risk Index (HRI).

second line type gives the calculated value for each health effect for which there is data and the final HRI number. The third section presents the RTECS toxicity data for the chemical; lines of this type are entered in whatever number is needed to display all RTECS data utilized in calculation of a relative toxicity value for the chemical. Each test class is identified in terms of route, species, test, and outcome using RTECS abbreviations. This is followed by a numeric entry showing the number of chemicals for which there are data within that test class. Finally, there is a number showing the normalized dose value within each test class (to three decimals). Printing of the third section, which displays all input data, is an option in the computer program, and is referred to as the "long form."

4. ESTIMATES OF CHEMICAL SPECIFIC EFFECTS ON WORKER HEALTH

The second major output of the modeling system is an estimate of the relative health effect due to specific chemical agents through simultaneous consideration of toxicity and conditions or occupational exposure. This output is termed the adjusted hazard risk index (AHRI). The alórithm considers five factors these are:

1. The HRI value for the specific chemical which has been calculated in accordance with the health effects weights set by the user.
2. The number of workers observed exposed to the chemical during the NOHS survey.
3. The number of workers observed exposed to *any* NOHS chemical during the survey.
4. The percentage of observed exposures considered to be controlled under NOHS protocol.
5. The percentage of observed exposures considered to be full time under NOHS protocol (more than 4 hr per working day).

It was decided that the best approach to describing the interrelationship between chemical specific toxicity and .the extent of worker exposure was to multiply the hazard risk index number (HRIN) by

```
AHRI Calculation based on HRIN Trial 02 Run 01     RTECS Date 01/81     NOHS Date 07/80
                                                    Date:  08/31/82      page 1

Multipliers for:    AT=1   CAR=1   ETA=1   MUT=1   NEO=1   PI=1   TER=1   TFX=1   NEO OPTION= YES

Constants for:      AT=0   CAR=0   ETA=0   MUT=0   NEO=0   PI=0   TER=0   TFX=0   CAR OPTION=  NO

SEQ#      8     LP8925000/33640          Formaldehyde

    HRI SEQ=  21   HRI=2.223   PES=10907   PEN=545569   PC=48   PFT=08   AHRI=136.337
```

Fig. 2. Example Adjusted Hazard Risk Index (AHRI).

those factors describing the extent of worker exposure. The algorithm for the adjusted hazard risk index number (AHRIN) is as follows:

$$AHRIN = (HRIN)(PES/PEN)(1.0 - 0.9\ PC)(0.5 + 0.5\ PFT)(K)$$

where:

HRIN = Hazard risk index number.
PES = Number of workers potentially exposed to the hazard in the NOHS.
PEN = Number of workers observed potentially exposed to *any* NOHS hazard. (Therefore, the PEN factor is a constant 545,569 in all AHRIN calculations.)
PC = Percentage of exposures controlled.
PFT = Percentage of full-time exposures.
K = A constant used to remove leading zeroes.

In this algorithm the HRINs come directly from the HRI, the PES/PEN ratio indicates the proportion of all workers observed in the NOHS survey who were potentially exposed to the hazard, and the (1.0–0.9 PC) factor reduces the significance of controlled exposures to 10%[2] of that represented by uncontrolled exposure.

The result is an index number that has a direct relationship to the estimated occupational health risk due to exposure to the chemical.

The higher the computed value, the higher the potential health risk to the workforce as a whole.

A sample of the AHRI output is shown in Fig. 2. The data display format for the AHRI is in several parts, as follows:

1. The first three lines are similar to those of the HRI output, except that the first line of the heading at the top of each page also identifies the AHRI in terms of the HRIN trial and run[3] input.
2. The data entry for each chemical consists of two lines. The first line gives the sequence (rank-order) number for the chemical, its RTECS and NOHS identification codes (in RTECS/NOHS) format), and the prime RTECS name for the chemical. The second line gives the sequence (rank-order) number (out of 1721) for the chemical in the source HRI, the source HRI value used, the number of people in the NOHS sample potentially exposed to the chemical (PES), the number of people observed exposed to any NOHS chemical (PEN), the percent of observed potential exposures controlled in some way according to NOHS protocol, the percent of potential exposures that were more than 4 hr per working day, and the final AHRI calculated.

5. ESTIMATES OF POTENTIAL OCCUPATIONAL GROUP HEALTH RISK

The third major output of the model is an estimate of the relative health risk associated with occupational groups defined by two-, three-, or four-digit Standard Industrial Classification (SIC) or Bureau of the Census occupational code. These outputs are termed the industrial risk index (IRI), or the

[2] The 10% factor for controlled exposures corresponds to the minimum acceptable effect produced by use of a respirator, and is considered to be indicative of the order of magnitude protection that can be expected from the use of exposure controls.

[3] The words "trial" and "run" and attached reference numbers are internal tracking systems references that document the source HRI and the health effects weights used to produce it.

```
            IRI Calculation based on HRI Trial 02 Run 01     RTECS Date 01/81  NOHS Date 07/80
                                                              CENSUS OPTION=YES  Date: 09/01/82  page 1

Multipliers for:    AT=1   CAR=1   ETA=1   MUT=1   NEO=1   PI=1   TER=1   TFX=1   NEO OPTION= YES

Constants for:      AT=0   CAR=0   ETA=0   MUT=0   NEO=0   PI=0   TER=0   TFX=0   CAR OPTION=  NO

SEQ#      5      SIC 28        CHEMICALS AND ALLIED PRODUCTS        IRI =    224.829

                               NUMBER OF CHEMICALS =    1162                +/- 1.529
```

Fig. 3. Example Industrial Risk Index (IRI).

occupational risk index (ORI). The system user may optionally utilize census population data[4,5] specific to each occupational group in order to consider the total population at risk. Use of this option results in an expression of potential health risk for the occupational group as a whole rather than an expression of potential health risk to an individual worker employed in such a group. Listing of this option is documented in the output as Census Option = Yes, or Census Option = No.

An example of the IRI is shown in Fig. 3. The description of the format shown applies to both the IRI and ORI, and is in several parts, as follows:

1. The first three lines are similar to the AHRI output, except that the use or nonuse of the census option is also documented.
2. The first line of the index entry for each SIC or occupational group consists of a sequence number, SIC or occupational numerical code and title, and the calculated IRI or ORI. The second line of the entry for each SIC or occupational group consists of the number of linked chemicals (out of 1721) to which workers in the industrial or occupational group are potentially exposed. The approximate 95% confidence limits of the IRI (based on an analysis of the number of worker observations considered) are also displayed on this line. Printing of the second line is a program option.

6. INTERNAL TESTING

Having developed a model for the identification of high risk industrial and occupational groups, the obvious question is "how well does it work?" The best way of evaluating the model's credibility would be to match its rankings of occupational groups to results from the many epidemiological studies in recent years. Although this approach was initially considered, it soon became obvious that there was no practical and objective way of ranking the results of the widely diversified types of occupational epidemiology studies. This is because no other system has been discovered by the authors that provides rank-ordered risk assessments solely on the basis of chemical exposure. Instead, an attempt was made to anticipate problem areas and examine some assumptions made in developing the algorithms. Several internal analyses were then performed to shed some light on the utility and validity of the model.

Since the RTECS data were divided into eight categories of health effects, the degree of correlation of pairs of these categories across all chemicals was of obvious interest. Table I presents the nonparametric (Kendall's Tau b)[6] correlation coefficients for each possible pair of health categories. These values could range from -1.0 to 1.0, and measure the degree between pairs of health effect categories.

Individual indices equal to zero were excluded because they were indeterminate, in that a zero could mean either no data on a particular chemical for a health effect category or no known effect.

Inspection of the coefficients in Table I reveals no major surprises. All coefficients are positive, which one would expect since each of the eight categories is a measure of an adverse health condition. Also, as expected, the correlation of "neoplastigenicity" with "carcinogenicity" was relatively high ($r = .449$). However, somewhat suprisingly, the highest correlation ($r = .533$) was between "neoplastigenicity" and "other toxic effects." The fact that "carcinogenicity" and "mutagenicity" did not have a high correlation ($r = .224$) may seem questionable at first glance. However, RTECS mutagen data in the model was limited to only whole animal data or data form tests performed on mammalian cells, which excluded test procedures of the Ames type.

Table I. Kendall's Tau *b* Correlation Coefficients (Number of Chemicals)

	AT	CAR	ETA	MUT	NEO	PI	TER	TFX
AT	1.00000 (1594)	0.33063 (92)	0.32956 (171)	0.20672 (154)	0.15830 (91)	0.24415 (453)	0.40601 (129)	0.43610 (173)
CAR		1.00000 (108)	0.41682 (58)	0.22353 (35)	0.44949 (31)	0.12184 (30)	0.24638 (24)	0.16364 (11)
ETA			1.00000 (236)	0.13273 (62)	0.35994 (53)	0.22843 (51)	0.36572 (39)	0.11385 (26)
MUT				1.00000 (162)	0.21270 (36)	0.12262 (44)	0.08780 (48)	0.02834 (39)
NEO					1.00000 (109)	0.07407 (26)	0.34000 (25)	0.53333 (16)
PI						1.00000 (481)	0.23645 (29)	0.13037 (74)
TER							1.00000 (147)	0.36036 (37)
TFX								1.00000 (187)

One assumption made in developing the AHRI was that the use of controls reduced the extent of chemical exposure by an average of 90%. Since this might be the subject of considerable debate among industrial hygienists, alternatives were considered. A sensitivity analysis was employed to test the effects on the relative rankings of chemicals within the AHRI by changing the 90% parameter to various estimates between 80% and 98%. Results indicated that the size of this parameter had little effect on the AHRI chemical ranking. For example, with the parameter set at the extreme of 80% or 98%, 19 of the first 20 chemicals in the resulting ordered lists are identical when all other settings are held constant.

An assumption made in developing the occupational risk indicates (IRI and ORI) is that the net health effect for a group exposed to many chemicals is the cumulative sum of the risk associated with each individual chemical without regard to the possible synergistic or antagonistic effects that may be encountered. However, the overall health effect of exposure to a mixture of contaminants is of considerable debate among toxicologists and epidemiologists. If the actual health effect of multiple chemical exposure is nonadditive, then the rank orderings produced by the model would be somewhat changed. A related problem is that two equivalent IRI values could be associated with either a large number of low toxicity exposures or a small number of highly toxic exposures.

If an industry's ranking were more a function of the total number of potential exposures rather than of the average toxicity of these exposures, then one would expect to see a decreasing trend in the number of chemical exposures from the top to the risk rankings to the bottom. In order to test this hypothesis, the 389 4-digit SIC in the model were ranked from highest to lowest IRI. These values resulted from a trial of the model which weighed all health effects equally. The ordered list of 389 industries was then divided into ten roughly equal groups. The average number of chemicals and average risk per chemical were calculated for each group. The results (Table II)

Table II. Mean of Number of Chemicals and Average Toxicity for 10 Equal Intervals of 389 Sorted IRIN values

Group[a]	Mean number of chemicals	Average toxicity
1 (highest (IRI's)	162.0	0.0298
2	134.8	0.0133
3	131.5	0.0116
4	132.8	0.0103
5	101.7	0.0087
6	107.9	0.0087
7	92.7	0.0075
8	74.5	0.0058
9	60.4	0.0048
10 (lowest IRI's)	23.9	0.0057

[a]All groups contain 39 IRI values except Group 1 which contains 38.

Table III. Number of Industries in Each of 10 Groupings Produced by Cluster Analysis of IRI Values

Cluster	Number of industries within cluster
1 (highest)	4
2	1
3	3
4	1
5	5
6	8
7	17
8	74
9	185
10 (lowest)	91

show a clearly decreasing trend in number of chemicals from group 1 to group 10. The result for average risk per chemical also shows a decreasing trend, but to a lesser degree. This would seem to indicate that the industrial ranking produced by the model is more influenced by the number of chemicals associated with each industry than by the average toxicity of those chemicals.

The real effect of exposure to a mixture of many chemicals probably lies somewhere between the average risk per chemical and the cumulative risk for all chemicals. Since the true effect is unknown, a cluster analysis[7] of the industrial risk indices resulting from one specific trial of the system was employed in order to divide the industries into groups on the basis of a statistical procedure which considers the magnitude of the calculated IRI values. This approach considers those industries within each cluster to be virtually equivalent in potential health risk. Results are shown in Table III. The authors feel that clusters of this kind may represent a more realistic assessment of the potential health risk associated with industries or occupations than a literal interpretation of the rank-ordered lists produced by the model. Accordingly, an optional procedure to compute and display the aver-

```
IRI Calculation based on HRI Trial 02 Run 01  RTECS Date 01/81  NOHS Date 07/80  Census Option = YES
                                                                  Date:  09/01/82  page 1

Multipliers for    AT=1  CAR=1  ETA=1  MUT=1  NEO=1  PI=1  TER=1  TFX=1  NEO OPTION= YES

Constants for      AT=0  CAR=0  ETA=0  MUT=0  NEO=0  PI=0  TER=0  TFX=0  CAR OPTION= NO

SEQ#    1     SIC 2893      PRINTING INK                AVG TOX=0.040               IRI=5.985
        AVG TOX SEQ#= 2     Number of chemicals=148     TOX Range=0.001-0.832       +/-   .048

SEQ#    2     SIC 2834      PHARMACEUTICAL PREPARATIONS AVG TOX=0.016               IRI=5.659
        AVG TOX SEQ#= 3     Number of chemicals=360     TOX Range=0.003-0.968       +/-   .062

SEQ#    3     SIC 3496      COLLAPSIBLE TUBES           AVG TOX=0.046               IRI=5.626
        AVG TOX SEQ#= 1     Number of chemicals=123     TOX Range=0.001-0.828       +/-   .053

SEQ#    4     SIC 2851      PAINTS AND ALLIED PRODUCTS  AVG TOX=0.015               IRI=5.618
        AVG TOX SEQ#= 4     Number of chemicals=361     TOX Range=0.005-0.827       +/-   .036
```

Fig. 4. Example Industrial Risk Index (optional format).

Table IV. Industrial Risk List (Accute Toxicity Effect Only)

		% Workers potentially exposed		% Potential exposures controlled		% Potential exposures more than 4 hr/day	
SIC[a]	#Chemicals[b]	(avg.)	(range)	(avg.)	(range)	(age.)	(range)
28	1079	1.2	0.0–29.2	47.0	0.0–100.0	0.7	0.0–85.0
80	634	1.3	0.0–53.4	9.4	0.0–100.0	0.3	0.0–100.0
55	184	3.5	0.1–31.1	3.7	0.0–100.0	0.0	0.0–0.0
72	182	3.1	0.1–18.5	1.8	0.0–67.0	0.3	0.0–46.0
73	316	1.9	0.1–29.1	34.4	0.0–100.0	0.0	0.0–0.0
65	113	3.7	0.1–33.9	1.4	0.0–80.0	0.0	0.0–0.0
79	224	2.0	0.0–20.1	3.1	0.0–100.0	0.0	0.0–1.0
45	315	1.5	0.0–25.5	10.6	0.0–100.0	0.2	0.0–36.0

[a]See Table VIII for definition of SIC codes.
[b]There are 1594 chemicals in the model, which animal studies indicate are acutely toxic.

Table V. Industrial Risk List (Skin and Eye Irritation Effects Only)

		% Workers potentially exposed		% Potential exposures controlled		% Potential exposures more than 4 hr/day	
SIC[a]	# Chemicals[b]	(avg.)	(range)	(avg.)	(range)	(avg.)	(range)
80	184	2.6	0.0–53.4	10.9	0.0–100.0	0.7	0.0–100.0
55	106	4.2	0.1–31.1	2.7	0.0–59.0	0.0	0.0–0.0
28	362	2.0	0.0–29.2	47.6	0.0–100.0	1.0	0.0–85.0
72	90	3.9	0.1–18.5	2.6	0.0–67.0	0.5	0.0–46.0
79	109	2.8	0.0–17.2	4.5	0.0–100.0	0.0	0.0–0.0
65	65	4.3	0.1–33.9	1.2	0.0–80.0	0.0	0.0–0.0

[a]See Table VIII for definition of SIC codes.
[b]There are 483 chemicals in the model, which animal studies indicate are skin and eye irritants.

Table VI. Industrial Risk List (Mutagenic/Teratogenic Effects Only)

		% Workers potentially exposed		% Potential exposures controlled		% Potential exposures more than 4 hr/day	
SIC[a]	# Chemicals[b]	(avg.)	(range)	(avg.)	(range)	(avg.)	(range)
28	182	1.2	0.0–29.2	49.7	0.0–100.0	1.2	0.0–60.0
80	119	1.3	0.0–46.7	14.8	0.0–100.0	0.0	0.0–1.0
72	30	2.7	0.1–12.1	2.5	0.0–50.0	0.0	0.0–0.0
73	63	1.2	0.0–11.8	30.7	0.0–100.0	0.0	0.0–0.0
62	7	14.4	0.0–29.0	0.0	0.0–0.0	0.0	0.0–0.0
79	44	1.6	0.0–14.9	1.6	0.0–50.0	0.0	0.0–0.0
75	6	4.8	0.5–22.6	0.0	0.0–0.0	0.0	0.0–0.0

[a]See Table VIII for definition of SIC codes.
[b]There are 261 chemicals in the model, which animal studies indicate are mutagenic and/or teratogenic.

Table VII. Industrial Risk List (Carcinogen/Equivocal Tumorigen/Neoplastigenic Effects Only)

		% Workers potentially exposed		% Potential emposures controlled		% Potential exposures more than 4 hr/day	
SIC[a]	# Chemicals[b]	(avg.)	(range)	(avg.)	(range)	(avg.)	(range)
28	228	1.2	0.0–29.2	41.3	0.0–100.0	1.0	0.0–85.0
89	13	10.2	0.8–32.7	2.4	0.0–31.0	0.0	0.0–0.0
81	2	65.0	65.0–65.0	0.0	0.0–0.0	0.0	0.0–0.0
62	13	8.4	0.4–51.5	0.0	0.0–0.0	0.0	0.0–0.0
80	115	1.3	0.0–46.7	12.1	0.0–100.0	0.0	0.0–0.0
79	51	1.9	0.0–14.9	3.4	0.0–100.0	0.0	0.0–0.0
55	42	2.5	0.1–11.5	0.8	0.0–35.0	0.0	0.0–0.0
45	80	1.5	0.0–15.9	11.9	0.0–100.0	0.0	0.0–1.0
13	58	1.8	0.0–16.6	26.2	0.0–100.0	0.0	0.0–50.0
32	93	1.2	0.0–20.6	22.1	0.0–100.0	1.6	0.0–61.0

[a]See Table VIII for definition of SIC codes.
[b]There are 333 chemicals in the model, which animal studies indicate are carcinogens, equivocal tumorigens, and/or neoplastigens.

presented in Tables IV–VII. Each of these tables represents data form a separate trial of the model resulting in rank-ordered listings of two-digit SICs.

As seen in Tables IV–VII, industrial rank orderings from the model based on selected health effects seems to both confirm "conventional wisdom" and provide some surprises, such as the relatively high ranking of miscellaneous business services (SIC 73). In order to help the reader evaluate the date presented in Tables IV–VII, a brief explanation of the

Table VIII. Selected Industry Titles and Activities Contained In the Model

SIC Code	Title
13	*Oil and Gas Extraction*—Activities contained within this code are crude petroleum and natural gas, oil and well drilling and field services.
28	*Chemicals and Allied Products*—Activities contained within this code are the production of cyclic intermediate and crudes, pigments, organic/inorganic chemicals, plastics, synthetic rubbers, cellulosic fibers, pharmaceuticals, soaps, polishes, surface active agents, sanitation goods, paints, inks, and agricultural chemicals.
32	*Stone, Clay, and Glass Products*—Activities contained within this code are the production of flat glass, pressed or blown glassware, hydraulic cement, structural clay products, brick, pottery, concrete, gypsum, abrasive products, and asbestos products.
45	*Transportation By Air*—Activities contained within this code are certificated and noncertificated transport of passengers or freight, the operation and maintenance of flying fields, and handling of air freight or passengers at airports.
55	*Automobile Dealers and Service Stations*—Activities contained within this code are new and used vehicle dealers, including on-site repair and maintenance facilities, and gasoline service stations, including auto repair operations.
62	*Security, Commodity Brokers, and Services*—Activity contained within this code is the operation of brokerage offices, including related clerical operations.
65	*Real Estate*—Activities contained within this code are the rental, buying, and management of real estate. Included are the operation and maintenance of residential and commercial buildings.
72	*Personal Services*—Activities contained within this code are laundries and dry cleaning facilities, including clothing and fabric dyeing.
73	*Miscellaneous Business Services*—Activities contained within this code include advertising agencies, janitorial, maintenance, and housekeeping services, research and development laboratories, and miscellaneous business services such as business and statistical consultants.
75	*Auto Repair, Services, and Garage*—Activities contained within this code include the repair, rental, and storage of automobiles owned by the general public.
79	*Amusement and Recreation Services, Not Elsewhere Classified*—Activities contained within this code include the operation of bowling alleys, amusement parks, golf and country clubs, and miscellaneous operations such as athletic clubs and carnivals.
80	*Medical and Other Health Services*—Activities contained within this code include the operation of physicians' offices, hospitals, sanatoria, and rest homes. Also included are miscellaneous services such as blood banks and therapists.
81	*Legal Services*—Activity contained within this code includes the operation associated with legal offices and their clerical work.
89	*Miscellaneous Services*—Activities contained within this code include architectural and engineering services, engineering research, and accounting and bookkeeping services.

industrial activity represented by each two-digit SIC code (1967 Standard Industrial Classification) for the 14 SIC codes displayed is presented in Table VIII. Note that the model contains 60 groups at the two-age risk per chemical within an industry has been developed to further aid the researcher in his assessment of the relative health risk of each industry. For example, a researcher examining the data for one individual cluster (as shown in Fig. 4) might conclude that the workers in SIC 2893 and SIC 3496 are at higher risk than those in the other two industries, since their average chemical toxicity is approximately three times as high. The optional data illustrated in Fig. 4 includes average toxicity, a sequence or rank-order number based on average toxicity, and the range of chemical specific toxicity values associated with the group.

7. EXAMPLES

Output from this computer model is heavily dependent upon user decisions in setting health-effects weights, selection of tumorigenic options, and use of census data. Accordingly, no single listing of chemicals or occupational groups rank-ordered by descending risk index can be veiwed as totally accurate, nor (as discussed under internal testing of the model) should the precise ordering of any rank-ordered list be too literally interpreted.

However, to demonstrate the potential utility and flexibility of the model, some data examples are digit SIC level, 269 groups at the three-digit level, 389 groups at the four-digit level, and 262 occupation codes.

8. CONCLUSION

While a user of this modeling system should recognize its limitations (primarily the correlation of animal test outcome with similar effects in the human and limited chemical toxicity and exposure data), the several possible outputs from the model make it a research tool of great versatility.

Although originally conceived as an aid to the prioritization of occupational health research, post-development use of the model has indicated its considerable strengths in the identification of possible occupational cohorts for detailed study and as a reference tool in the documentation of the potential chemical exposure characteristics of occupational and industrial groups. However, risk estimates provided by the model should be regarded only as preliminary indications of the potential for adverse health effects which can be verified only through field research.

Due to the design of the model, it can be utilized by researchers with widely disparate concerns. The most obvious examples of this are the simultaneous production of average and cumulative toxicity values for the chemical exposure of occupational groups and optional utilization of variable health-effects weights. Any researcher exercising the weighting options should be fully aware that model output is heavily dependent upon these settings.

Continuing development of this model, to include at least the updating of both the chemical toxicity and exposure data, is expected to further enhance its value to the occupational health community.

APPENDIX: HAZARD RISK INDEX ALGORITHM CONSIDERATIONS

All input RTECS dose data were normalized. This was accomplished by working with each test class individually, and expressing individual doses as a function of the range of doses within a test class. After expressing all test doses as the natural logarithm of the dose (and adding a constant of 30.00 to assure positive values), the normalized dose was calculated as follows:

$$d_n = \frac{d_{max} - d_i}{d_{max} - d_{min}}$$

where

d_n = normalized dose,
d_i = observed or cited dose,
d_{max} = maximum observed dose in the test class,
d_{min} = minimum dose in the test class.

This resulted in the creation of an index number ranging from 1.000 for the most toxic dose in a test class to 0.000 (arbitrarily set at 0.001) for the least toxic dose. The use of logarithmic scale also resulted in the capability to recognize small differences at the highly toxic end of the scale. This calculation was performed for all doses contained in the 98 test classes in the computer system. The third section of Fig. 1 illustrates the test class specific values associated with the reported tests on a single chemical for thirteen test classes as well as reporting the number of other chemicals in the computer system that were tested within the same test class parameters.

In designing the algorithm to calculate the HRI number, several factors were considered. We wanted to retain the ability to compare the relative toxicity of chemicals within as well as across health effects. In addition, we had to allow for the facts that the RTECS data are not all-inclusive, and that there may be no data for a specific health effect.

Therefore, two decisions were made which fundamentally affected the form of the algorithm: (1) It was decided to average the normalized values within each individual health effect as the best approximation of potency of the chemical for that outcome. (2) The nonzero values calculated for each individual health effect would be summed as the expression of the overall relative toxicity of the chemical. While acknowledging that this summary procedure equates a lack of data with no effect, we believe that it is the best solution, given the comprehensive data content of RTECS. In addition, a procedure involving the averaging of nonzero health-effect values was noted to produce values biased toward chemicals with a single, highly toxic test outcome while assigning relatively low values to well-researched chemicals with broad spectrum toxicological properties.

Finally, in order to retain the ability to compare chemical toxicity within health effects, and to allow maximum user flexibility, the algorithm had to permit the optional weighing of individual health-effect values. This permits users to produce a rank-ordered listing (by descending relative toxicity) of chemicals for a single health effect, or any combination of health effects weighed according to user priorities.

In the case of tumorigenic data, two additional options are available. Should the user elect to consider carcinogens a metastisizing subset of neoplastigens, neoplastic data can be produced for a chemical for which only a carcinogenic response is reported in RTECS (the Neoplastigen Option). Conversely, the model can create a carcinogenic citation for those chemicals reported only as neoplastigenic in RTECS (the Carcinogen Option).

On the basis of all these considerations, the final HRI algorithm is expressed as:

$$\begin{aligned}\mathrm{HRIN} = &(a\mathrm{AT} \oplus b) + (c\mathrm{CAR} \oplus d)\\ &+ (e\mathrm{ETA} \oplus f)(g\mathrm{MUT} \oplus h) + (i\mathrm{NEO} \oplus j)\\ &+ (k\mathrm{PI} \oplus l) + (m\mathrm{TER} \oplus n) + (o\mathrm{TFX} \oplus p)\end{aligned}$$

where

AT = The average of all acute toxicity-normalized doses.
CAR = The average of all carcinogenic-normalized doses.
ETA = The average of all equivocal tumorigenic agent-normalized doses.
MUT = The average of all mutagenic-normalized doses.
NEO = The average of all neoplastigenic-normalized doses.
PI = The average of all primary irriration-normalized doses.
TER = The average of all teratogenix-normalized doses.
TFX = The average of all other toxic effect-normalized doses.

Lower case letters $a-p$ indicate variable numerical factors (values 0–9) to be set by the system user. And $\oplus$ = an addition that is performed only if the associated values are not equal to zero.

REFERENCES

1. David H, Pedersen, Randy O. Young, and David S. Sundin, *A Model for the Identification of High Risk Occupational Groupes Using RTECS/NOHS Data* (DHHS, NIOSH Technical Report, Publication 83-117, Cincinnati, Ohio, September, 1983).
2. *National Occupational Hazard Survey* (National Institute for Occupational Safety and Health, Cincinnati, Ohio, DHEW Publications No. (NIOSH) 74-127, May, 1974; (NIOSH 77-213; July, 1977; and (NIOSH) 78-114; December, 1977).
3. *Registry of Toxic Effects of Chemical Substances*, 1980 edition, (National Institute for Occupational Safety and Health, DHHS (NIOSH) Publication No. 81-116, Cincinnati, Ohio, February, 1982).
4. U.S. Summary, *County Business Patterns* (Census Bureau, Publication CBP73-1, Table 221, 1973).
5. U.S. Summary, *County Business Patterns* (Census Bureau, Publication CBP-73-1, Table 1B, 1973).
6. William H. Kruskal, "Ordinal Measures of Association," *Journal of the American Statistical Association*, **53**, 814–861 (1958).
7. Stephen C. Johnson, "Hierarchial Clustering Schemes," *Psychometrika*, **XXXII**, 241–254 (1967).