Applied Occupational and Environmental Hygiene

ISSN: 1047-322X (Print) 1521-0898 (Online) Journal homepage: http://www.tandfonline.com/loi/uaoh20

Exposure Assessment for a Study of Workers Exposed to Acrylonitrile. III: Evaluation of Exposure Assessment Methods

Patricia Ann Stewart , John N. Zey , Richard Hornung , Robert F. Herrick , Mustafa Dosemeci , Dennis Zaebst & Linda M. Pottern

To cite this article: Patricia Ann Stewart , John N. Zey , Richard Hornung , Robert F. Herrick , Mustafa Dosemeci , Dennis Zaebst & Linda M. Pottern (1996) Exposure Assessment for a Study of Workers Exposed to Acrylonitrile. III: Evaluation of Exposure Assessment Methods, Applied Occupational and Environmental Hygiene, 11:11, 1312-1321, DOI: 10.1080/1047322X.1996.10389418

To link to this article: http://dx.doi.org/10.1080/1047322X.1996.10389418

Exposure Assessment for a Study of Workers Exposed to Acrylonitrile. III: Evaluation of Exposure Assessment Methods

Patricia Ann Stewart,[A,E] *John N. Zey,*[B] *Richard Hornung,*[B] *Robert F. Herrick,*[B,C]
Mustafa Dosemeci,[A] *Dennis Zaebst,*[B] *and Linda M. Pottern*[A,D]

[A]Environmental Epidemiology Branch, National Cancer Institute, Executive Plaza North, Room 418, Rockville, Maryland 20892; [B]National Institute for Occupational Safety and Health, 4676 Columbia Parkway, Cincinnati, Ohio 45226; [C]Current address: Harvard University, School of Public Health, Department of Environmental Health, 665 Huntington Avenue, Boston, Massachusetts 02115-6069; [D]Current address: National Institutes of Health, Women's Health Initiative, 7550 Wisconsin Avenue, Bethesda, Maryland 20892; [E]To whom correspondence should be addressed

Retrospective exposure assessment in epidemiologic studies is dependent on the availability of historical monitoring results, yet rarely are there sufficient results to rely upon them exclusively. An approach is described that has a formal structure for developing exposure estimates for an epidemiologic study using a variety of methods depending on the information available. The approach identifies criteria for determining what data are needed for each method and the hierarchy for using the methods. The estimation methods include: (1) calculating a mean from the monitoring results of a job; (2) identifying homogeneously exposed jobs and using the mean of the measurements for the jobs as the estimate; (3) applying a ratio of the measurement means of two jobs in one operation to a third job in another operation to estimate a fourth; (4) weighting by time various areas or personal (short or full-shift) measurements representing tasks or locations; (5) taking a deterministic approach that modifies a more recent exposure by estimates of exposure modifiers to reflect how changes in the workplace affected exposures; and (6) using professional judgment. The homogeneous exposure group, ratio, time-weighted, and deterministic approaches were evaluated for bias, precision, accuracy, and correlation. First, a subset of the monitoring results was removed from the entire data set for use as referent values. Estimates were developed using the four methods without the removed data and compared with the referent value. On average, the estimates tended to overestimate the measurements in the ratio method (bias = 77%) and underestimate them in the time-weighted method (−24%). The average difference between the means and the estimates using the homogeneous exposure group method and the deterministic method was zero. The imprecision was threefold to fourfold for the ratio and deterministic methods and 1.5fold for the time-weighted and homogeneous exposure group methods. Correlations between the estimates and the measurement means ranged between 0.60 and 0.75 and were statistically significant. When the methods were evaluated using only the same cells, the homogeneous exposure group method performed best, followed by the deterministic method and then the ratio method. The time-weighted average method could not be evaluated because of the lack of measurement data for these cells. STEWART, P.A.; ZEY, J.N.; HORNUNG, R.; HERRICK, R.F.; DOSEMECI, M.; ZAEBST, D.; POTTERN, L.M.: EXPOSURE ASSESSMENT FOR A STUDY OF WORKERS EXPOSED TO ACRYLONITRILE. III. EVALUATION OF EXPOSURE ASSESSMENT METHODS. APPL. OCCUP. ENVIRON. HYG. 11(11):1312–1321; 1996.

Quantitative retrospective exposure assessment is a difficult task in occupational epidemiologic studies, because rarely are sufficient monitoring results available for all jobs over all the years in the study. Various approaches have been taken to estimate exposures, but only two, to our knowledge, have compared quantitative exposure estimates to measurement data.(1,2) The failure to conduct such an evaluation in other studies is not surprising, however, because if sufficient data were available to fully validate the exposure estimates, the estimates themselves would be unnecessary. Nonetheless, validation of estimates is crucial to the field.

We are engaged in a cohort mortality study of workers exposed to acrylonitrile (AN).(3,4) The cohort has been employed in eight plants that started manufacturing AN, acrylamide, and acrylic fibers or resins between 1952 and 1965. Each of the companies conducted air monitoring since the late 1970s, but only one monitored before this time. Even after 1977, a large number of jobs were not monitored (about 3200) compared with the number of jobs in the study (3500). Thus, it was necessary to develop a large number of estimates both for the unmonitored jobs and the unmonitored time periods to allow the inclusion of all study subjects in the mortality analysis. This report describes the methods used to estimate historical exposures and an evaluation of their performance.

Methods

All companies in the study started conducting personal monitoring in the late 1970s. Accompanying documentation maintained by the plant generally included job title, department, date, duration of the sample, an evaluation of the sample's representativeness of typical exposures, and type of sample (area or personal and time-weighted average or short-term), but not all plants retained the same information. All data were computerized (either by the company or by the study investigators) and put into a standard format.

There were about 18,000 monitoring results. The quantity of monitoring data available on the jobs in the study varied

1047-322X/96/1111-1312$15.00/10 • PII S1047-322X(96)00253-0

considerably over time and among plants. For some of the jobs directly involved in making the product, there were over 100 full-shift, personal monitoring results between 1977 and 1983 (the close of the collection of work history records). In contrast, a few jobs supporting the production operations (e.g., maintenance, quality control, the materials handlers for the AN monomer, environmental control, utilities, engineering, and research and development) had a variable number of results, depending on the job, the plant, and the exposure level. Administrative, non-AN-exposed, and other support operations (e.g., the materials handling department not handling the monomer) were rarely monitored. Area measurements were available in three plants, but in only one of these (a fiber plant) had they been collected prior to 1977. The approximately 4000 area measurements in this plant were short-duration samples and existed back to 1963 (start-up of the plant was in 1958). The area measurements collected in the other two plants included both short-duration and full-shift results.

Prior to determining what estimation procedure should be used, a review of the published exposure assessment literature was conducted. Because so many jobs were missing measurements, calculation of mean exposures for all jobs was not possible (e.g., Reference 5). Grouping jobs into a small number of occupational groups so that mean exposures could be calculated for all groups,(6) derivations of estimates using area measurements weighted by time,(7,8) or use of statistical models(1,8) was not possible prior to 1977 in seven of the plants because processes and controls had changed over the years and measurements for these conditions were nonexistent. In addition, the fiber plant that had collected earlier area measurements was substantially different from the other two fiber operations. It was a continuous wet operation, whereas the other two dried the polymer before making the fiber. Thus, extrapolation to the other two plants was inappropriate. There were also no other parallel chemicals used in the processes that had been measured and could be used to estimate exposure levels.(9) Moreover, no single estimation approach could be identified that would produce estimates for all cells. As a result, it was decided to use a variety of methods, the selection of the method for any particular job/year combination being determined by the available information, by defined criteria, and by how well each method performed in the validation exercise where the predicted exposures were compared with the actual measurements (described below).

The original intent of the investigators was to develop an estimate for each job/department/plant combination for every year. Many of the years, however, had so few monitoring results that it was decided to combine years, where appropriate, to form time period cells. Information was collected from interviews with long-term employees at each of the plants on when major changes occurred that were likely to have affected exposure levels in the workplace. The dates of these changes provided boundaries around time periods, which were made up of years in which exposure levels were likely to have been similar. The periods may have been composed of a single year or may have included several years, and they established the unit of time for which the exposure estimates were developed.

Performing a retrospective exposure assessment usually has two components due to the availability of monitoring data and supporting documentation. First, because it is rare that all the jobs have been monitored in a company, the industrial hygienist must develop at least one exposure estimate (generally the most recent) for each of the jobs in the study. The first step in this study, therefore, was to develop baseline exposure estimates for all cells from monitoring results wherever possible. Cells for which means could not be calculated were evaluated to determine if they could be estimated using the homogeneous exposure group (HEG), ratio, or time-weighted average (TWA) method. For all remaining jobs that still had no estimate, professional judgment was used to assign an estimate for a single cell. The second component in exposure assessment is to develop estimates through time for the remaining empty cells. In this study this was done by modifying the baseline estimates, using estimates of exposure modifiers to complete all remaining empty cells. This was called the deterministic method. All estimates in this study were plant/department/job/time period specific.

Development of Baseline Estimates

Mean of the Monitoring Measurements

The best estimates of the true exposures experienced by the study subjects when cumulative exposure is the primary exposure measure of interest are the arithmetic means of the measurements.(10) Arithmetic means were therefore calculated in this study. Means based on small numbers, however, can be easily affected by a few outliers that weigh more heavily in the calculation of the mean than appropriate. Criteria were therefore needed that ensured, to the extent possible, that the means reflected true exposures. Three variables were evaluated that could have influenced the accuracy of the mean estimates: the duration of the monitoring result, the representativeness of the monitoring result to typical exposures (as indicated in the company records), and the number of monitoring results used in the calculation of the mean.

Sample duration was identified in the documentation of the measurement data in five plants. Documentation from two other plants indicated whether the result represented an 8-hour TWA, a ceiling, or a peak sample. Personnel from the eighth company indicated to the study investigators that generally full-shift samples were taken; therefore, all results obtained from this company were assumed to be full-shift in duration.

Over 80 percent of all measurements were greater than 6 hours in duration. Because the less than 4-hour samples were generally task oriented rather than full-shift oriented, these measurements were excluded. The 4- to 6-hour measurements appeared to be a mixture of both full-shift and task measurements. These results were also excluded, however, because they contributed only slightly to the database (8%). A minimum 6-hour duration was therefore selected to ensure that the measurements were representative of full-shift exposures. (Shorter duration measurements were available using other methods; see below.)

Many of the companies designated the measurements as typical or atypical of normal operating conditions. Exclusion of the atypical measurements was not considered appropriate because they were actual measured exposures. Including them with measurements taken under typical conditions was also troublesome, because in calculating a mean, they might be given more weight than appropriate. For example, if a job was monitored every day during a shutdown operation lasting 2 weeks, and was otherwise monitored monthly, the measurements taken over the 14 days of the shutdown would contribute 54 percent of the total (14/14+12) measurements in a year, which is considerably more than the true contribution of these measurements (14/365 days or 4%) to an annual average.

An examination of the measurements found that fewer than 20 percent of the measurements were taken under unusual conditions. Of these, 60 percent were taken during shutdowns, start-ups, or major repair of the operating units, events that usually occurred once a year or less. The frequency of some of these events could be estimated, so that they were weighted in the calculation of the mean according to their frequency of occurrence over a year (e.g., 4% in the above example). When frequency could not be determined (e.g., spills, upsets, no information), the results were given the same weight as the typical measurements. Although not the perfect solution, this was considered acceptable because these other measurements comprised only about 5 percent of the measurements used in the analysis.

Finally, the minimum number of measurements necessary to develop a cell mean was evaluated by looking at the coefficient of variation (CV). CVs were calculated for all cells with at least five measurements and then pooled to derive an average CV. This calculation was also done for cells with at least six, seven, eight, nine, and ten measurements. The mean CVs changed very little (less than 30% from the lowest to the highest CV) regardless of the number of measurements. Six was therefore arbitrarily selected as the minimum number of measurements required for a cell. (Shorter duration measurements were available for other methods; see below.)

Ratio Method

In this method, the assumption was made that similar jobs in similar environments in different plants are likely to have similar exposure levels relative to other jobs in the same environment. This assumption appears to have validity based on work by Eisen *et al.*[11] In their report on granite workers they presented exposure data from seven granite sheds. For each shed, the exposure estimates of six jobs were presented. Although the particular environmental conditions were not presented in the report, it appears from those data that the relationship of exposure levels between two jobs in one shed approximated that of the same jobs in another shed, despite the wide range of exposure levels among sheds. That study provides the underlying assumption that

$$E_{RTOi'j'} = \frac{(E_{MONij'})(E_{MONi'j})}{E_{MONij}} \tag{1}$$

where:

$E_{RTOi'j'}$ = the exposure estimate derived from the ratio method for job i'j'
E_{MON} = the estimate derived from the mean of the monitoring measurements method
i = jobs
j = operations or plants

The criteria for selecting the jobs for this method were as follows: (1) the tasks being performed by job_i in plants j and j' were similar, as were those by $\text{job}_{i'}$; (2) the amounts of time spent by job_i performing those tasks were similar in plants j and j', as were those of $\text{job}_{i'}$; (3) the sources of exposure for the two jobs (i and i') in plants j and j' were similar; and (4) the controls influencing those sources were similar. Because of the requirement of similar controls, the estimate for job_i and $\text{job}_{i'}$ in plant_j were from the same time period, and those of job_i and $\text{job}_{i'}$ in $\text{plant}_{j'}$ were from the same time period, but the estimates for plants j and j' did not have to be from the same time period.

HEG Method

Another approach used to estimate exposure was based on the concept of HEGs.[12] Jobs identified as being similar were grouped, and the measurements for those jobs were used to calculate a mean value. The criteria for grouping these jobs were that: (1) they had to have the same sources of exposure (i.e., they were in the same operating unit and plant); (2) the tasks being performed were similar; (3) the time spent performing those tasks was similar; (4) the controls influencing those sources were the same (resulting in all the measurements being in the same time period); and (5) there had to be at least six measurements of at least 6 hours in duration on the jobs within the group.

Time-Weighted Average (TWA)

In this method, a TWA was calculated from short-duration or area measurements or from other job estimates and weighted by the time spent at those concentrations by

$$E_{TWA} = \frac{C_1t_1 + C_2t_2 \ldots + C_nt_n}{t_1 + t_2 \ldots t_n} \tag{2}$$

where:

E_{TWA} = the estimate derived from the TWA method
C_i = the concentration at location_i or performing task_i
t_i = the minutes the job spent at location_i or task_i

Three different sources of exposure concentrations (C_i) were available: (1) a concentration entered by the industrial hygienist based on professional judgment, which was usually 0.00, designating the exposure level experienced when working in a nonexposed office; (2) a mean of area samples or a mean based on less than six personal samples or on samples that were less than 6 hours in duration (These data had been excluded from the mean of the monitoring measurements method); and (3) a previously developed estimate for another job or a mean of other jobs' estimates. For example, an engineer who spent 1 hour a day in an operating unit was given the same concentration for that hour as an engineer who spent 8 hours a day in the unit. The time estimates were either specifically identified from interviews or they were estimated by the study investigators based on an understanding of the tasks.

Estimates derived from these four methods were considered to be of higher confidence, and these methods were used to develop the estimates wherever possible. For the jobs to which these methods could not be applied, a single estimate was developed, usually the most recent, using one of two professional judgment techniques:

TABLE 1. Examples of Significant Changes and Estimates of Their Size and Effect

Plant Type	Change	Engr Estimate[A]		IH Estimate[B]	
		Size[C]	Effect[D]	Size	Effect
Fiber	Installed vents on reactors	M	1–1.5	30	1.3
Fiber	Quality control sample points enclosed in ventilated sample boxes	M	10–100	30	55
Fiber	Pilot plant closed; quality control samples no longer analyzed	L	0.9–0.9	30	0.9
Monomer	Began closed dome loading of tank cars	M	2–10	30	6
Monomer	Single seals replaced by double seals/tandem seals	M/L	2–5	40	3.5
Monomer	Floating roofs installed at tank farm	S	1–1.1	5	1.1
Monomer	Oil added to settling ponds to reduce evaporation	S	1.1–3	5	2
Resin	Increased preventative maintenance	M	1–1.3	20	1.1
All	AN continuous analyzer installed	M	1.1–2	10	1.5
All	New OSHA standard became effective	M	2–5	10	3.5
All	Hoods added to laboratory	M/L	2–10	20	6
All	Air conditioning added to control room	M	1–5	30	3

[A] These represent the estimates developed by the engineer of the study. They were used as guidelines for the industrial hygienist.

[B] These represent the estimates assigned by the industrial hygienist of the study. They generally represented the midpoint of the engineer's estimate.

[C] Percent of a job's total exposure that the source being affected contributed to prior to the change: S = small (1 to 10%); M = medium (11 to 49%); L = large (50 to 100%).

[D] A correction factor designating the reduction or increase that the change had on the emissions being released from the source. For example, 1 would indicate no effect and 1.5 would indicate that emissions from a source were 1.5 times higher before the change was in place.

Department-Wide Method

This method allowed the industrial hygienist to enter one exposure estimate for all jobs in a department. It was generally used for nonexposed departments, such as administration, for example, where 0.00 was entered.

Professional Judgment

In this method the industrial hygienist developed an estimate based on: (1) a mean of personal measurements where fewer than six existed or where the duration of the samples was less than 6 hours; (2) a mean of area measurements; (3) a mean of measurements after 1983; and/or (4) anecdotal information, such as frequency of health effects, odor, task descriptions, etc.

Development of Estimates Over Time

After all jobs had at least one baseline estimate, all remaining cells were estimated using the deterministic method, which modified baseline estimates based on estimates of the effect that the major changes that occurred in the workplace, or exposure modifiers,(13) had on exposure levels. Three variables were identified as having an important effect on AN exposure levels: changes in production rates,(14) changes in the frequency of exposure, and changes in the operation and in engineering controls. Annual production rates were received from the companies in the study, but because data were missing for several years, production rates were not included in the model. The frequency of exposure was obtained from the interviews or estimated. For example, production employees usually were given daily exposure, whereas an engineer may have been exposed only 1 day a week.

Estimating the effect of the changes in the operation or in engineering controls was a more difficult task. For many changes there were no monitoring data available prior to and after implementation of the change. A chemical engineer with experience in these types of plants, therefore, compiled monitoring data, information from the interviews, information from other plants in the study, and data from the published literature. Using these sources and his own experience, he estimated the effect the changes had on exposure levels for two variables: size and effect. Size was defined as the percent of the job's total exposure that the source being affected contributed to prior to the change, and was indicated as small (1 to 10%), medium (11 to 49%), or large (50 to 100%). Effect was a correction factor indicating the reduction or increase that the change had on the emissions being released from the source, and was generally indicated as a range. Unless information was available that suggested otherwise, the midpoint of the ranges for both size and effect was used. Table 1 identifies examples of several changes and their estimated size and effect.

To develop an estimate, the baseline value for an adjacent (generally more recent) time period was used to derive an estimate for an earlier time period as follows:

$$E_{DET(y-1)} = [B_y/(1 - \Sigma\, S_n + \Sigma\, S_n/CF_n)](F_{(y-1)}/F_y) \qquad (3)$$

where:

E_{DET} = the estimate being derived from the deterministic method
B = the baseline estimate
S = the size of the source being affected by $change_i$
CF_n = the correction factor or effect of $change_i$
F = the frequency of exposure (1.0 = 4 to 5 days/week, 0.4 = 2 to 3 days/week, and 0.1 = <2 days/week)
y = the time period for the baseline value
$y - 1$ = the time period for the estimate

(See Appendix for a derivation of the equation.) Thus, suppose in 1977 the exposure estimate was 2 ppm, and two changes took place that reduced the exposure between 1977 and 1976. The size and effect of these two changes were 10 and 3.5 for one change and 20 and 1.2 for the other, respectively. The frequency of exposure was daily in both years:

$$E_{DET} = [2/(1 - 0.10 + 0.10/3.5 - 0.20 + 0.20/1.2)] \cdot (1/1)$$
$$= (2/0.90)(1)$$
$$= 2.22 \text{ ppm}$$

The E_{DET} then became the baseline estimate for the next adjacent time period. In addition to this estimate, maximum and minimum estimates were calculated using the extremes of the size and effect values developed by the chemical engineer. These additional estimates provided an estimate of uncertainty about the estimates.

Other Estimates

The estimates described thus far were estimates of air concentrations. These are typically the principal estimates used in epidemiologic analyses. In some instances, however, these may not be the best estimates to evaluate disease risks. For example, in some jobs respirator use was mandatory, suggesting that an air concentration may not be the best exposure estimate. Exposure estimates were therefore derived that accounted for respiratory protection when respirator use was mandatory for 8 hours a day (E_{RES}). For these jobs, a protection factor (PF) was applied to the air concentration estimate (E) based on the type of respirator worn (half-mask: PF = 10; full-facepiece: PF = 50; supplied air: PF = 2000) and multiplied by 0.65(15) to allow for imperfect protection, that is,

$$E_{RES} = E/PF(0.65) \qquad (4)$$

Because the amount of an airborne chemical received by the body depends on the amount inhaled, a third estimate was derived that took into account respiratory rates based on low, medium, and high levels of physical activity. These levels were defined to correspond to ventilation volumes of 750, 1450, and 2150 cm^3/breath, respectively, assuming 15 breaths per minute at all levels.(16) An adjusted exposure estimate for a job was therefore calculated by multiplying the estimate of the air concentration derived for the job by the appropriate ventilation volume and the number of breaths taken in 8 hours.

Finally, AN can be absorbed dermally.(17) Because the absorption rate in humans was not found in the literature, a dermal exposure score was calculated by multiplying the frequency of dermal exposure, arbitrarily selected as 5 for frequent (more than once a day) and 1 for infrequent (less than once a day) by the concentration of AN in the liquid.

Method Evaluation

Evaluation as to how well the estimation methods performed is crucial for the interpretation of an epidemiologic study. In this study, however, there were no full-shift, personal monitoring data prior to 1977 with which to evaluate how well the assessment methods performed during the earlier years of interest. There were, however, monitoring data available after 1983, the last year for which estimates were developed. In some plants, up to 10 years of monitoring data were available. It was decided to test the estimation methods using these data where possible.

Four methods could be tested: the ratio, the HEG, the TWA, and the deterministic methods. To test these methods, a subset of the cells for which there were monitoring results was removed from the estimation process. The assessment methods were then applied to the remaining cells to develop estimates, which were compared with the corresponding measurement means. All possible job/time period cells were used for all estimates combined and for type of operation (fiber, monomer, and resin).

In the ratio method all sets of four jobs in the ratio that met the criteria listed for this method were identified. One of the four jobs was selected randomly to be the estimated job. The means of the measurements for the remaining three jobs were used to calculate the estimate for the fourth, using Equation 1. The estimate was compared with the mean of the measurements of that job. All possible combinations of four jobs were used (i.e., a single job may have been used in comparison with several different jobs).

In the HEG method, at least three jobs, each with measurements, were identified as being in an HEG using the criteria described earlier for this method. One job was randomly selected from the jobs in the HEG and removed. The mean of the remaining measurements was the estimate and the mean value of the measurements of the removed job was the referent value.

For the TWA method, jobs with mean of the monitoring measurements estimates were identified. Wherever possible, estimates were developed for these cells using area measurements, short-duration measurements, or estimates from other jobs, weighted by time. These estimates were compared with the mean of the monitoring measurements values.

In the deterministic method, all jobs with at least two estimates calculated from the mean of the monitoring measurements method were identified. Frequency of exposure was not relevant here because it designated number of days exposed per week, which is not relevant to a full-shift measurement. The estimate of the job in the most recent time period (e.g., 1987) was modified using the estimates of size and effect generated by the chemical engineer to derive an estimate for each time period back to 1977. The estimates were compared with the referent values (i.e., the means of the measurements) when they existed.

Statistical Methods

Bias, precision, and accuracy were calculated for each method.(18) Bias is the average difference between the measurement means and the estimates. Precision (a more intuitive word would be imprecision, and therefore this term is used here) is the standard deviation of those differences. Accuracy is the square root of the sum of the bias squared and the imprecision squared. The closer these three values are to zero, the better the estimates are. Relative bias, imprecision, and accuracy were calculated by dividing each of these values by the measurement means. To determine how well the ranking of the estimates compared with the ranking of the referent values, Spearman rank correlation coefficients were calculated. All possible estimates for each of the methods were included in the analysis to see how well the methods performed overall. To

TABLE 2. Bias, Imprecision, Accuracy, and Correlation of Three Exposure Assessment Methods

	n	$\bar{x}_m$[A] (ppm)	SE[B] (ppm)	$\bar{x}_e$[C] (ppm)	Bias (ppm)	Rel Bias (%)	Imp (ppm)	Rel Imp (%)	Acc (ppm)	Rel Acc (%)	r[D]
Ratio	51	0.93	0.14	1.64	0.71	77	3.82	412	3.88	419	0.76[E]
HEG	86	0.72	0.14	0.73	<0.01	0	1.08	149	1.08	149	0.65[E]
TWA	32	0.62	0.02	0.47	−0.15	−24	1.03	166	1.04	168	0.59[F]
Deterministic	177	1.34	0.25	1.33	0.01	1	3.17	236	3.17	236	0.58[E]

[A] Arithmetic mean of the measurements.
[B] Standard error.
[C] Mean of the estimates.
[D] Spearman rank correlation coefficient.
[E] $p < 0.001$.
[F] $p < 0.005$.

establish which of the three methods was preferable when the data allowed selection of more than one method, only cells for which estimates were derived for the ratio, HEG, and deterministic methods were included in the calculation of the statistics. None of these cells could be derived from the TWA method, so this method was not evaluated.

Results

A random sample of one job from each of 51 sets of four jobs was used for evaluating the ratio method (Table 2). The overall mean of the cells based on the mean of monitoring results was 0.93 ppm, whereas the overall mean of the estimates for those same cells was 1.64 ppm. The estimates therefore overestimated the measurements by 0.71 ppm or about 80 percent. The relative imprecision was about 400 percent. Overall, the estimates were accurate to within about four times the measurement means. The Spearman correlation was 0.76, which was statistically significant ($p < 0.001$). The average of the mean of the measurements in the 86 cells in the HEG method was 0.72 ppm, with the mean of the estimates being 0.73 ppm. The relative bias was zero. The imprecision was 150 percent and overall accuracy was 150 percent. The Spearman coefficient was statistically significant at 0.65 ($p < 0.001$). The TWA method had overall bias of 25 percent, the relative imprecision was about 165 percent, and the relative accuracy was about 170 percent. The correlation was 0.59 ($p < 0.001$). The deterministic method had low bias (mean of the measurements = 1.34 ppm; mean of the estimates = 1.33 ppm; relative bias = 1%). The imprecision and overall accuracy were both about 240 percent and the correlation was 0.58 ($p < 0.001$).

The methods were examined by type of operation. The mean of the estimates using the ratio method (Table 3) was higher than the mean of the monitoring results (2.96 versus 1.48 ppm, respectively) in the fiber operations. This resulted in a relative bias of 100 percent and a relative imprecision of 380 percent. Overall accuracy was about 400 percent and the correlation coefficient was 0.87 ($p < 0.001$). The estimates for the monomer operations were only slightly higher than the monitoring results (means = 0.56 and 0.48 ppm, respectively). The relative bias was 18 percent, relative imprecision 139 percent, and relative accuracy 140 percent. The correlation was lower than for the fiber operations (monomer $r = 0.48$), although still significant ($p < 0.01$). Estimates could not be developed using this method for the resin operation because there was no comparable operation to use in the ratio.

The results of the HEG method by type of operation are found in Table 4. The means of the estimates for the fiber, monomer, and resin plants were fairly close to the means of the measurements (relative biases of −9, 25, and 12%, respectively). The imprecision and accuracy were 220 percent or less. The correlation between the measurements and the estimates was excellent to moderate for the fiber and monomer plants ($r = 0.80$ and 0.47, $p < 0.01$, respectively), but poor for the resin plant (0.31, $p > 0.05$). This low correlation may be due to the narrow range of exposures in this plant (mean = 0.23, SE = 0.06).

The TWA method underestimated the measurements by 20 and by 56 percent when it was evaluated for the fiber and monomer operations, respectively (Table 5). There were too few measurements ($n = 2$) for the resin plant to evaluate the method. Relative imprecision and accuracy for the two operations was about 75 and 125 percent for the two operations, respectively. The correlations were 0.57 ($p < 0.01$) and 0.43 ($p > 0.05$).

TABLE 3. Bias, Imprecision, and Accuracy of the Ratio Method by Type of Occupation

	n	$\bar{x}_m$ (ppm)	SE (ppm)	$\bar{x}_e$ (ppm)	Bias (ppm)	Rel Bias (%)	Imp (ppm)	Rel Imp (%)	Acc (ppm)	Rel Acc (%)	r
Fiber	23	1.48	0.26	2.96	1.48	100	5.61	380	5.80	393	0.87[A]
Monomer	28	0.48	0.07	0.56	0.09	18	0.66	139	0.67	140	0.48[B]

Abbreviations are explained in Table 2.
[A] $p < 0.001$.
[B] $p < 0.01$.

TABLE 4. Bias, Imprecision, Accuracy, and Correlation of the HEG Method by Type of Operation

Plant	n	$\bar{x}_m$ (ppm)	SE (ppm)	$\bar{x}_e$ (ppm)	Bias (ppm)	Rel Bias (%)	Imp (ppm)	Rel Imp (%)	Acc (ppm)	Rel Acc (%)	r
Fiber	36	1.22	0.30	1.11	−0.11	−9	1.38	114	1.39	114	0.80^{A}
Monomer	33	0.44	0.12	0.55	0.11	25	0.97	219	0.97	221	0.47^{B}
Resin	17	0.23	0.06	0.26	0.03	12	0.30	128	0.30	129	0.31

Abbreviations are explained in Table 2.
$^{A}p < 0.001$.
$^{B}p < 0.01$.

The estimates for the deterministic method were, on average, higher than the measurements in the fiber and resin operations, whereas they were lower than the measurements for the monomer plants (Table 6). The magnitude of the relative bias, on the other hand, was similar for the fiber and monomer plants (18 and −32%, respectively), but large for the resin plant (170%). Relative imprecision and relative accuracy were over twice as high for the monomer and resin plants as compared with the fiber plants. The correlation in all operations was moderate ($r = 0.38$ to 0.58).

To evaluate what hierarchy should be used when more than one method was possible, only the cells for which estimates were developed using all methods were evaluated (Table 7). There were 32 estimates that the ratio, the HEG, and the deterministic methods had in common, but none of these could be developed using the TWA method. The ratio method developed estimates that were much higher than the measurements (2.70 versus 1.57 ppm, respectively). The relative bias was 71 percent, the relative imprecision was 322 percent, and the relative accuracy was 330 percent. The correlation coefficient was 0.63 ($p < 0.001$). The HEG and the deterministic methods also overestimated the measurements, but by much less, with relative biases of 6 and 32 percent, respectively. The relative imprecisions and accuracies of these two methods were about and 90 and 120 percent, respectively, and the correlations were moderate (0.65, $p < 0.001$, and 0.54, $p < 0.01$, respectively).

Discussion

This report describes the exposure assessment methods used in an epidemiologic study to assess historical exposures when measurement data were not available. Because measurement data varied in quantity by job and by year, several methods were developed to allow the best use of the data. The methods were then evaluated to ensure that they developed reasonable estimates.

Most of the methods performed acceptably. Relative bias was 77 percent for the ratio method, 24 percent for the TWA method, and almost zero for the HEG and the deterministic methods. Relative imprecision was reasonable for the latter three methods (150, 165, and 240%, respectively), but high for the ratio method (400%). Correlation between the measurements and the estimates was moderate (for the HEG, TWA, and deterministic methods, $r = 0.58$ to 0.65) to excellent (ratio method, $r = 0.76$). These statistics were generated using every possible cell for each method.

In many cases, however, the study investigators were able to develop estimates using more than one method, so that the methods were also evaluated using the same subset of job/time periods. (The TWA method could not be evaluated because no estimates could be developed for any of the cells that were developed for the other three methods.) Bias and imprecision were found to be lowest for the HEG method (6 and 88%, respectively), and it had a correlation equal to or higher than the other two methods ($r = 0.65$). The HEG method, therefore, was the method of choice when the criteria described earlier for using means of the measurements were not met. The deterministic method performed better than the ratio method, so it was the next preferable method.

The results of the evaluation were examined by type of operation to determine where the estimation methods performed poorly. Of the ten method/operation evaluations, six estimates overestimated the measurements and four underestimated the measurements. There was no consistency by type of operation, although the estimates averaged higher than the measurements for both sets of resin operation comparisons. The monomer operation generally had a lower bias, imprecision, and accuracy than did the fiber operations, but the average differences were slight. The methods did much worse for the resin operation. The fiber operation had, on average, higher correlations than the monomer or resin operations, but all were moderate.

TABLE 5. Bias, Imprecision, Accuracy, and Correlation of the TWA Method by Type of Operation

	n	$\bar{x}_m$ (ppm)	SE (ppm)	$\bar{x}_e$ (ppm)	Bias (ppm)	Rel Bias (%)	Imp (ppm)	Rel Imp (%)	Acc (ppm)	Rel Acc (%)	r
Fiber	22	0.80	0.65	0.64	−0.16	−20	0.58	73	0.61	76	0.57*
Monomer	8	0.27	0.25	0.13	−0.15	−56	0.34	126	0.37	137	0.43

Abbreviations are explained in Table 2. Resin plant results are not presented because $n = 2$.
*$p > 0.01$.

TABLE 6. Bias, Imprecision, Accuracy, and Correlation of the Deterministic Method by Type of Operation

Plant	n	$\bar{x}_m$ (ppm)	SE (ppm)	$\bar{x}_e$ (ppm)	Bias (ppm)	Rel Bias (%)	Imp (ppm)	Rel Imp (%)	Acc (ppm)	Rel Acc (%)	r
Fiber	74	1.48	0.18	1.75	0.26	18	1.81	122	1.83	123	0.38*
Monomer	69	1.72	0.61	1.18	−0.54	−32	4.66	271	4.69	273	0.55*
Resin	34	0.28	0.04	0.75	0.47	170	0.83	301	0.95	346	0.58*

Abbreviations are explained in Table 2.
*$p < 0.001$.

Where large differences occurred between the estimates and the measurements, almost all (88%, n = 14) occurred when the measurement data appeared to be unusually low or high when compared with means in other years or with means of other similar jobs in the same year. For example, in one job the mean of the measurements in 1978 was 1.83 ppm; in 1979, 8.04 ppm; in 1980, 0.41 ppm; and in 1981, 0.56 ppm. (All these means were based on more than ten measurements.) The three methods for which estimates were developed for this job underestimated the 8.04 ppm measurement (the ratio estimate was 1.61 ppm; the HEG estimate, 4.06 ppm; and the deterministic estimate, 2.28 ppm). These values are more in line with the measurement means of the other years. In one sense this is reassuring, in that the methods produced estimates that appeared reasonable, even when the referent value appeared unreasonable. In another sense, however, this difference is troublesome, because it points out the problem of using measurement data as the referent value.

Measurements are generally considered to represent truth, but because they are generally few in number, they can be influenced by a few unrepresentative values. An attempt was made to reduce such an influence by requiring at least six 6-hour or more measurements, with nonrepresentative measurements generally weighing less in the calculation of the mean than the sample results taken under typical conditions. It may be that this approach was not totally successful, however. One explanation for these outlying data is that the measurements may have been taken on days that were typical, but on the high end of the typical scale (i.e., day-to-day variability). There was, however, no way to determine this possibility from the data. The data could also represent undocumented atypical conditions.

There are few studies evaluating historical exposure assessment methods with which to compare these results. Hornung *et al.*(1) compared the results of a regression model with measurement data in a study of ethylene oxide workers and found a relative bias of 32 percent, a relative precision of 105 percent, and a relative accuracy of 109 percent (Hornung, R., personal communication). Investigators of a man-made mineral fiber study simulated historical environmental conditions and found that their estimates were about fourfold the measurements taken under simulation.(2) Two studies evaluated assessments made of current exposures using professional judgment. One found approximate relative biases of 200 and −4 percent (calculated from published data) when the raters estimated exposures without and with monitoring data, respectively.(19) A second study found correlations of 0.67 to 0.73 between exposure estimates and measurements of methylene chloride and 0.12 to 0.29 for styrene.(20) Overall, our results compare favorably with these other studies.

In conclusion, methods were described that were used to estimate historical data when measurements were nonexistent. Four methods were evaluated by comparing estimates derived from these methods with measurements. All performed satisfactorily, although three (the HEG, TWA, and deterministic methods) gave more accurate results than the fourth (the ratio method).

Recommendations

When developing historical exposures, the estimation procedures described in this report should be considered when measurements are nonexistent or limited and more rigorous methods, such as regression models, are not possible. These procedures require fewer monitoring data than more rigorous methods and can be used with descriptive information. Attempts should be made, however, to evaluate their performance.

Acknowledgment

The authors thank Dave Chesnut, Mary Ann Heyer, and Winner Ricker of IMS. The advice of Dr. Roy Shore, Dr.

TABLE 7. Bias, Imprecision, Accuracy, and Correlation of the Ratio, HEG, and Deterministic Methods Using the Same Estimates

Method	n	$\bar{x}_m$ (ppm)	SE (ppm)	$\bar{x}_e$ (ppm)	Bias (ppm)	Rel Bias (%)	Imp (ppm)	Rel Imp (%)	Acc (ppm)	Rel Acc (%)	r
Ratio	32	1.57	0.29	2.70	1.13	71	5.07	322	5.19	330	0.63^A^
HEG	32	1.57	0.29	1.67	0.10	6	1.39	88	1.39	89	0.65^A^
Deterministic	32	1.57	0.29	2.08	0.51	32	1.88	119	1.94	123	0.54^B^

Abbreviations are explained in Table 2.
$^{A}p < 0.001$.
$^{B}p < 0.01$.

Richard Monson, Dr. Robert Harris, Dr. Carol Rice, and Dr. Katharine Hammond is also gratefully acknowledged.

References

1. Hornung, R.W.; Greife, A.L.; Stayner, L.T.; et al.: Statistical Model for Predicting Retrospective Exposure to Ethylene Oxide in an Occupational Mortality Study. Am. J. Ind. Med. 25:825–836 (1994).
2. Dodgson, J.; Cherrie, J.; Groat, S.: Estimates of Past Exposure to Respirable Man-made Mineral Fibers in the European Insulation Wool Industry. Ann. Occup. Hyg. 31:567–582 (1987).
3. Stewart, P.A.; Lemanski, D.; White, D.; et al.: Exposure Assessment for a Study of Workers Exposed to Acrylonitrile. I. Job Exposure Profiles: A Computerized Data Management System. Appl. Occup. Environ. Hyg. 7:820–825 (1992).
4. Stewart, P.A.; Triolo, H.; Zey, J.; et al.: Exposure Assessment for a Study of Workers Exposed to Acrylonitrile. II. A Computerized Exposure Assessment Program. Appl. Occup. Environ. Hyg. 10:698–706 (1995).
5. Smith, T.J.; Hammond, S.K.; Laidlaw, F.; Fine, S.: Respiratory Exposures Associated with Silicon Carbide Production: Estimation of Cumulative Exposures for an Epidemiological Study. Br. J. Ind. Med. 41:100–108 (1984).
6. Seixas, N.S.; Moulton, L.H.; Robins, T.G.; et al.: Estimation of Cumulative Exposures for the National Study of Coal Workers' Pneumoconiosis. Appl. Occup. Environ. Hyg. 6:1032–1041 (1991).
7. Rice, C.; Harris, R.L.; Lumsden, J.C.; Symons, M.J.: Reconstruction of Silica Exposures in the North Carolina Dusty Trades. Am. Ind. Hyg. Assoc. J. 45:689–696 (1984).
8. Dement, J.M.; Harris, R.L.; Symons, M.J.; Shy, C.M.: Exposures and Mortality among Chrysotile Asbestos Workers. Part I: Exposure Estimates. Am. J. Ind. Med. 4:399–419 (1983).
9. Armstrong, B.G.; Tremblay, C.G.; Cyr, D.; Theriault, G.P.: Estimating the Relationship Between Exposure to Tar Volatiles and the Incidence of Bladder Cancer in Aluminum Smelter Workers. Scand. J. Work Environ. Health 12:486–493 (1986).
10. Seixas, N.S.; Tobins, T.G.; Mouton, L.H.: The Use of Geometric and Arithmetic Mean Exposures in Occupational Epidemiology. Am. J. Ind. Med. 14:465–477 (1988).
11. Eisen, E.A.; Smith, T.J.; Wegman, D.H.; et al.: Estimation of Long Term Dust Exposures in the Vermont Granite Sheds. Am. Ind. Hyg. Assoc. J. 45:89–94 (1984).
12. Corn, M.; Esmen, N.A.: Workplace Exposure Zones for Classification of Employee Exposures to Physical and Chemical Agents. Am. Ind. Hyg. Assoc. J. 40:47–57 (1979).
13. Schneider, T.; Olsen, I.; Jorgensen, O.; Lauersen, B.: Evaluation of Exposure Information. Appl. Occup. Environ. Hyg. 6:475–481 (1991).
14. Esmen, N.: Retrospective Industrial Hygiene Surveys. Am. Ind. Hyg. Assoc. J. 40:58–65 (1979).
15. Toney, C.R.; Barnhart, W.L.: Performance Evaluation of Respiratory Protective Equipment Used in Paint Spraying Operations. U.S. Department of Health, Education, and Welfare, Cincinnati, OH (1976).
16. Task Group on Lung Dynamics: Deposition and Retention Models for Internal Dosimetry of the Human Respiratory Tract. Health Physics 12:173–207 (1966).
17. U.S. Environmental Protection Agency: Health Assessment Document for Acrylonitrile. EPA—600/8-82-007. USEPA, Washington, DC (1982).
18. Hornung, R.: Statistical Evaluation of Exposure Assessment Strategies. Appl. Occup. Environ. Hyg. 6:516–520 (1991).
19. Hawkins, N.C.; Evans, J.S.: Subjective Estimation of Toluene Exposures: A Calibration Study of Industrial Hygienists. Appl. Ind. Hyg. J. 4:61–68 (1989).
20. Post, W.; Kromhout, H.; Heederik, D.; et al.: Semiquantitative Estimates of Exposure to Methylene Chloride and Styrene: The Influence of Quantitative Exposure Data. Appl. Occup. Environ. Hyg. 6:197–204 (1991).

Appendix

The deterministic model was based on the assumption that two important modifiers of exposure levels are changes in the process or in engineering controls or work practices, and changes in frequency of exposure.

In most of the situations in which the deterministic method will be used, measurement data will be available in a more recent time period and the more historical time period is the period being estimated. It is easier to conceptualize the derivation, however, taking the opposite case (i.e., assuming measurement data exist for an earlier time period, and that it is the exposure of the current time period being estimated). The historical exposure is designated as the baseline estimate ($B_{(y-1)}$) (i.e., the estimate that will be modified), and the exposure for the current year (y) is designated as the estimate derived from this method (E_{DETy}).

An exposure (whether a mean (M) or an estimate (E)) is a result of numerous sources generating emissions, and can be measured by summing the individual sources of emissions (e_i). Thus,

$$B_{(y-1)} = \Sigma e_{n(y-1)} \tag{1}$$

If the total exposure (E or M) is equal to the sum of the emissions of all the sources, size, as a percent, can be used to quantify the contribution of each source. The percent of emissions contributed by any source (S_i) is equal to the amount of emissions divided by the total emissions or the mean:

$$S_{i(y-1)} = e_{i(y-1)}/B_{(y-1)} \tag{2}$$

and because the source sizes are percentages, they must equal 1:

$$\Sigma S_{n(y-1)} = 1 \tag{3}$$

If a change affects the level of emissions from a source, the amount of emissions will decrease by a quantity or a correction factor (CF). This term is defined as

$$CF_i = e_{i(y-1)}/e_{iy} \tag{4}$$

Substituting $e_{i(y-1)}$ in Equation 4 with Equation 2, we have

$$CF_i = (S_{i(y-1)})(B_{(y-1)})/e_{iy} \tag{5}$$

and

$$e_{iy} = (S_{i(y-1)})(B_{(y-1)})/CF_i \tag{6}$$

and the estimate is therefore equal to the sum of the new emissions:

$$E_{DETy} = \Sigma e_{iy} \tag{7}$$

Where no change occurs at a particular source, CF = 1. Thus, the estimate for the more recent time period equals the sum of the emissions that were changed (e_{iyc}) and the sum of the emissions that were unchanged ($e_{i(y-1)u}$):

$$E_{DETy} = \Sigma e_{iyc} + \Sigma e_{i(y-1)u} \tag{8}$$

Replacing these two terms with Equation 6, and because CF = 1 where no changes took place,

$$E_{DETy} = [(\Sigma S_{i(y-1)c})(B_{(y-1)})/CF_i] + [(\Sigma S_{i(y-1)u})(B_{(y-1)})] \tag{9}$$

but

$$\Sigma S_{i(y-1)u} = 1 - \Sigma S_{i(y-1)c} \tag{10}$$

Replacing $S_{i(y-1)u}$ in Equation 9, and rearranging the terms:

$$\begin{aligned} E_{DETy} = {} & [(B_{(y-1)})(\Sigma S_{i(y-1)c})/CF_i] + [(B_{(y-1)}) \\ & \cdot (1 - \Sigma S_{i(y-1)c})] \end{aligned} \tag{11}$$

or

$$E_{DETy} = (B_{(y-1)})(1 - \Sigma S_{i(y-1)c} + \Sigma S_{i(y-1)c}/CF_i) \tag{12}$$

If S_n is designated as the size of only those sources changed, the equation becomes

$$E_{DETy} = B_{(y-1)}(1 - \Sigma S_n + \Sigma S_n/CF_n) \tag{13}$$

However, because the mean is generally for the most recent time period and the estimate is for the earlier time period, the inverse occurs, so that the equation becomes

$$E_{DET(y-1)} = B_{y'}(1 - \Sigma S_n + \Sigma S_n/CF_n) \tag{14}$$

Frequency of exposure is directly proportional to exposure level, so that

$$E_{DET(y-1)} = B_{y'}(1 - \Sigma S_n + \Sigma S_n/CF_n)(F_{(y-1)}/F_y) \tag{15}$$