Scand J Work Environ Health 2003;29(5):347–353

Use of Monte Carlo simulation to investigate uncertainty in exposure modeling

by Sean E Semple, PhD,[1] Louise A Proud, MSc,[2] John W Cherrie, PhD[1, 2]

Semple SE, Proud LA, Cherrie JW. Use of Monte Carlo simulation to investigate uncertainty in exposure modeling. *Scand J Work Environ Health* 2003;29(5):347–353.

Objectives This study used Monte Carlo (MC) simulation to examine the influence of uncertainty on an exposure model and to determine whether a difference exists between two worker groups in a ceramic fiber manufacturing plant.

Methods Data on work practices and conditions were gathered in interviews with long-serving employees. With the use of previously developed deterministic modeling techniques and likely distributions for model parameters, MC simulations generated exposure profiles for the two job titles.

Results The exposure profiles overlapped considerably, although the average estimated exposure for one job was approximately double that of the other. However, when the correlation between the model parameters in the two jobs was considered, it was concluded that there was a significant difference in the two estimates.

Conclusions Models are increasingly being used to estimate exposure. Different work situations inevitably result in different exposure estimates. However, it is difficult to determine whether such differences in estimated exposure between worker groups are simply the result of uncertainty with respect to the model parameters or whether they reflect real differences between occupational groups. This study demonstrates the value of MC simulation in helping define the uncertainty in deterministic model estimates.

Key terms epidemiology, exposure assessment.

Exposure levels vary between persons doing the same job and from one day to another for the same person. This type of variability results from differences in the way the work is carried out or from changes in the exposure circumstances. Variability in exposure may be the result of simple differences in worker behavior (eg, "the dirty worker effect") or the difference in environmental parameters. Kromhout and his co-workers (1) demonstrated that almost a third of job titles had between-worker variation that spanned an order of magnitude.

Although there have been great advances in the technology available to measure exposure, it is still relatively expensive and time consuming to obtain such data. Therefore, for many current situations, the number of available measurements may be limited, and there may be no reliable data available from the past. This situation has led to the development of several deterministic models that have been used in retrospective epidemiologic studies (2–3) or in regulatory risk assessments (4).

Cherrie and his co-workers (5) have developed a deterministic exposure model in which the coefficients are assigned by the human assessor on the basis of his or her experience. The model has a simple multiplicative structure that is used to provide a single point estimate of exposure for a given worktask and these estimates can then be combined as a time-weighted average to estimate exposure for work within a specific job title. Guidance is provided for the selection of parameters within the model although there is still a great deal of

1 Department of Environmental & Occupational Medicine, University of Aberdeen, Aberdeen, United Kingdom.
2 Institute of Occupational Medicine, Edinburgh, United Kingdom.

Reprint requests to: Dr Sean Semple, Department of Environmental & Occupational Medicine, University of Aberdeen, Aberdeen, AB25 2ZD, United Kingdom. [Email: sean.semple@abdn.ac.uk]

personal judgment required. The technique has been used in several epidemiologic studies, for example, those by Boffetta et al (6) and Dick et al (7), and has been shown to provide reasonably reliable exposure estimates when compared with corresponding measured values (8–9). However, it does not take account of the variability in exposure nor the uncertainty in the assessments from the modeling process.

Monte Carlo (MC) simulation was originally devised to help predict atomic reactions during the development of the atomic bomb. It is a process whereby randomly generated numbers drawn from a given distribution are used to reflect the variability or uncertainty in a deterministic model. The simulation is then repeated many times to produce a distribution of the outcome parameter. With the use of MC simulation it is possible to specify the shape of the distribution for each parameter in the model and to set maximum and minimum values for each variable. Distributions are often specified as one of the following four main types: normal, lognormal, triangular, or uniform.

Combining MC simulation with deterministic exposure modeling offers the opportunity to explore the possible variability and uncertainty of exposure estimates. In the present investigation, we chose to use it for the latter purpose. In particular, we investigated whether the reconstructed exposure for two jobs carried out in the same work area are significantly different given the uncertainty in the exposure estimates. We did not consider any uncertainty that might have arisen from the particular model selected or the impact of variation in exposure from between- and within-worker variability. For any given job intra- and interworker variability will have a random effect on exposure levels. In epidemiologic studies in which we are concerned with average exposures over a long period of time, this random effect on estimated exposures tends towards zero, some workers having lower exposures one day and higher exposures on others. In this study we were solely concerned with the uncertainty in the value of the parameters selected by the assessor for the model and whether, given this uncertainty, it is sensible to consider exposure estimates for the jobs as distinct.

Subjects and methods

The objective of this study was to examine the uncertainty of generated exposure estimates using previously validated exposure reconstruction methodology. With the use of detailed descriptions of two job titles, it was possible to estimate ranges for model parameters for each job and for each job in different time periods. Using MC iteration techniques, we then calculated the range and distribution of likely exposures within each job. Finally, by contrasting and comparing these exposure distributions, we were able to identify whether or not the jobs were distinct.

This work was carried out as part of a study to investigate the health of workers in the man-made mineral fiber (MMMF) manufacturing industry (10). It was carried out within a vacuum-forming process in which MMMF were processed to form complex shapes. Information about the historical development of the process, and more generally for the whole manufacturing plant, was obtained using questionnaires and structured interviews with long-service employees. Additional information was provided from company records and plant management. The reconstruction of respirable fiber exposure levels was carried out for two jobs in this area (serviceman and molding operator). Exposure was estimated for eight distinct time periods, during which process changes, such as the introduction of local exhaust ventilation or changed output of finished products, were identified.

Exposure reconstruction methodology

The strategy for the exposure reconstruction has been described earlier by Cherrie et al (5) and Cherrie (11). This exposure model is based on the characterization of emissions from sources in the work environment and the way in which workers interact with the dispersed pollutant. The first part of the model comprises three components: the *intrinsic emission* (ε_i) of the pollutant, the method of *handling* (h) or processing involved at the source, and the effects of any *local controls* (η_{lv}), such as local ventilation. It was assumed that these three factors are all independent and that they act in a multiplicative way, their product being the *active emission* from the source.

Exposure can also arise from *passive emission* (ε_p) or fugitive sources. *Total emission* from a source is considered to be the sum of passive and active emission.

Two other factors are important in determining the impact of a source: the *time that the source is actively emitting* (t_a) and the use of *personal protective equipment* (η_{ppe}). It was assumed that these two terms also affect exposure level in a multiplicative way, with a reduction in the time the source is active producing a pro-rata reduction in cumulative exposure.

Sources can arise in the following two zones: the *near field* or *far field*. The near field is defined as a cube of 2 meters side, centered around the breathing zone of the worker. Sources within the near field act directly to cause exposure. Sources in the far field will have their impact reduced because of dilution or *general ventilation* (d_{gv}).

Exposure sources in the worker's near field (NF) and far field (FF) are then assessed individually according to the following equation:

$$C=[(\varepsilon_i \times h \times (1-\eta_{lv}) \times t_a)+\varepsilon_p](1-\eta_{ppe}) \times d_{gv}.$$

The reconstruction of exposure levels for a job title can be carried out by subdividing the work into component tasks or operations and dealing with each separately. It is assumed that individual tasks and the exposure arising from these tasks are independent of each other. By reconstructing exposure levels for both NF and FF sources for each task (i) in a job title using the time fraction worked at each task (Δi), a time-weighted average exposure level can be estimated (C_T) as shown in the next equation:

$$C_T = \Sigma_{i=1}^{n} (C_{NFi}+C_{FFi}) \times \Delta i.$$

The descriptive information gathered from questionnaires and employee interviews was summarized, and the time between the start of the vacuum-forming operations and 1995 was divided into a number of time periods for assessment. Two of these periods were chosen to coincide with existing sets of cross-sectional occupational hygiene measurements (12–13).

With the use of MC methods with the @Risk™ software package, model parameters were assigned distributions on the basis of data gained from the questionnaires and interviews. @Risk™ is an add-in for the Microsoft Excel spreadsheet program that allows "distributions" to be entered into individual cells in a spreadsheet rather than single numbers. When calculations are made, a number is randomly generated from the defined distribution in each cell. Because the point-estimate reconstructions were made within a spreadsheet, incorporating the MC simulation was convenient. It is also possible to simultaneously model two or more jobs that have linked parameters by pasting these data into dynamically linked spreadsheet pages. The distributions used for each job are shown in table 1.

Either a triangular or uniform distribution was used for each parameter. For the triangular distributions the model allows the minimum, midpoint, and maximum values to be specified; in the table this specification is shown as 0.2, 0.3, and 0.4. For the uniform distributions, only the minimum and maximum values are specified. For example, for the emission parameter for the molding operator's tasks, a triangular distribution with minimum and maximum values of 0.2 and 0.4, respectively, was employed. This distribution was assigned on the basis of the subjective judgment of the user, and the spread indicated the likely range of values within which the user feels the true value lies. Local ventilation and respiratory protection were not incorporated into the MC simulation because neither was used at the time relevant to this simulation.

Where the same numerical parameter value was used in several tasks in the single point estimate (eg, intrinsic emission for wet fiber) the same random number was pasted into the spreadsheet cells in each case rather than a different random number being chosen. Because both jobs were carried out in the same area (also for the cement production and the other molding job) the near-field exposure contribution of one job was taken as the basis for the far-field contribution to the other, and vice versa.

Using MC iteration, the model for both the molding operator and serviceman jobs was run using the parameter distributions. Five thousand iterations were run for each job title and the output exposure distributions produced. An analysis of the distributions was carried out to provide values for the mean and 5–95th percentile exposure levels.

Table 1. Model parameter values and distributions.

	Molding operator		Serviceman	
Model parameter [a]	Preparation and cleaning	Forming	Unloading	Clean-up
Emission	Triangular (0.2,0.3,0.4)		Triangular (0.7,1,1.3)	
Handling	Triangular (2,3,4)	Triangular (2,3,4)	Triangular (0.07,0.1,0.13)	100 times the unloading value
Local controls	1	1	1	1
Time source active	Uniform (0.1, 0.3)	Uniform (0.22,0.42)	Uniform (0.3,0.7)	Uniform (0.9,1.0)
Passive	0.03	0.03	0.03	0.03
General ventilation	Triangular (0.7,1,1.3)		Triangular (0.7,1,1.3)	
Personal protection	1	1	1	1
Time fraction on task	Uniform (0,0.26)	1 minus the preparation and cleaning value	Uniform [b] (0.2,0.6)	Uniform[A] (0.02,0.06)

[a] These values refer to the near-field parameter values; far-field exposures were generated for the molding operator from the serviceman's values and vice versa.
[b] The serviceman's "time fraction on task" does not sum to 1, as other tasks not involving exposure to man-made mineral fibers were also carried out.

Description of the work activities

The process involves the formation of an aqueous mix containing MMMF, colloidal silica, and modified potato starch. The mix was formed in a tank and then transferred into forming tanks located directly underneath the mixing tank. Using the mix, the operators used vacuum extraction molds to form shapes, which were then dried in an oven.

The serviceman spent approximately half of his workshift preparing mixes. The remainder of the workshift was spent unloading the drying ovens and carrying out quality control checks of the product. The serviceman typically added MMMF to the mixing tanks for 90 minutes during each shift. The mixing platform was cleaned at the end of the shift by means of brushing. This activity took approximately 20 minutes to complete. The serviceman was also responsible for unloading carts containing the MMMF forms from the drying ovens. He also prepared the used boxes, which were usually recycled and contained residual MMMF dust.

The molders worked at the forming tanks. At the beginning of each shift, the mold was cleaned by a high-pressure water jet. The mold was then fitted onto a vacuum line and lowered into the fiber suspension. The vacuum was switched on, and the fiber suspension was drawn onto the mold. The mold was then lifted from the forming tank by a hoist, the vacuum was switched off, and air was used to blow the form off the mold.

Two other activities were performed in the vacuum-forming plant that would have influenced the MMMF exposure levels. During 1985, another molding process was situated in the vacuum-forming plant. Circular saws were used to trim the modules manufactured in this process, and the dust created was transmitted into the vacuum-forming area. In addition, a small pilot plant for cement production was present in the area for 2–3 years in the late 1980s. Bales of bulk mineral fiber were cut open with a knife and added to a vat with other chemicals.

Table 2. Summary of process changes in the plant.

Period	Years	Comments	Other processes
1	1973–1975	Pilot plant	
2	1976–1978	Early plant development	
3	1979–1984	Slow increase in tonnage produced	
4	1985	Disposable respirators available	Other molding
5	1986	Local ventilation installed on mixing tank	Cement
6	1987	First occupational hygiene survey	Cement
7	1988–1990	Increased growth in tonnage produced	
8	1991–1993	Rapid increase in tonnage produced, general ventilation improved	
9	1994–1995	Second occupational hygiene survey	

Historical development of the plant

Table 2 briefly summarizes the historical development of the vacuum-forming operations. Nine time periods were identified between the introduction of the pilot plant in 1973 and 1995. The first period encompassed only the pilot plant, and, because none of those interviewed had any knowledge of these operations, it was not possible to make any exposure reconstruction for this period. The remaining periods ranged from 1 to 6 years.

Results

The measured respirable fiber concentration data for 1987 were known to the assessor (JWC), and the exposure estimates for this period were therefore adjusted to correspond approximately to the measured values. The exposure levels for the remaining time periods were then adjusted by multiplying by the figure for the 1987 measurement divided by the 1987 estimate. These adjusted estimates are shown in figure 1.

The results from the MC simulation for period 2 (1976–1978) showed that the molding operator and serviceman had considerable overlap in the distribution of the estimated exposures. Figure 2 shows the distribution with the 5th and 95th percentile values for the estimated exposures of both jobs.

The serviceman's mean estimated exposure level (0.44 fibers/ml) was approximately twice that of the molding operators (0.25 fibers/ml). In addition, the molding operators had a much narrower distribution of estimated exposures, 90% of the estimates being between 0.18 and 0.35 fibers/ml. This situation is in contrast to that of the serviceman, who had a much wider distribution of estimated exposures, 90% of the data being between 0.25 to 0.72 fibers/ml. All of the distribution of molding operator exposure was within the serviceman's distribution.

As noted earlier, the models for these jobs were not independent because of common parameters (eg, intrinsic emission) and because they were undertaken in the same building; plus we have used the near field from one for the far field of the other. Using the MC simulations across all periods, we found that the correlation between the estimates for the two jobs was 0.7. We further investigated the effect of the correlation on the estimated exposures by taking the ratio of each simulation output for the two job titles for period 2 (figure 3). It became apparent that the serviceman had a greater estimated exposure concentration in over 98% of all cases.

We also carried out a sensitivity analysis of the MC simulations on the basis of the rank order correlation coefficient of each of the model input variables with the

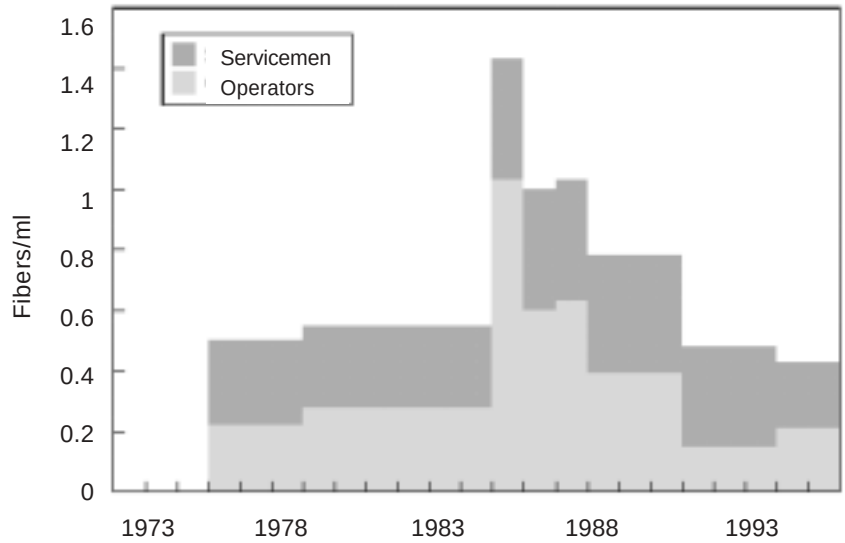

Figure 1. Reconstructed respirable fiber exposure levels for molding operators and servicemen.

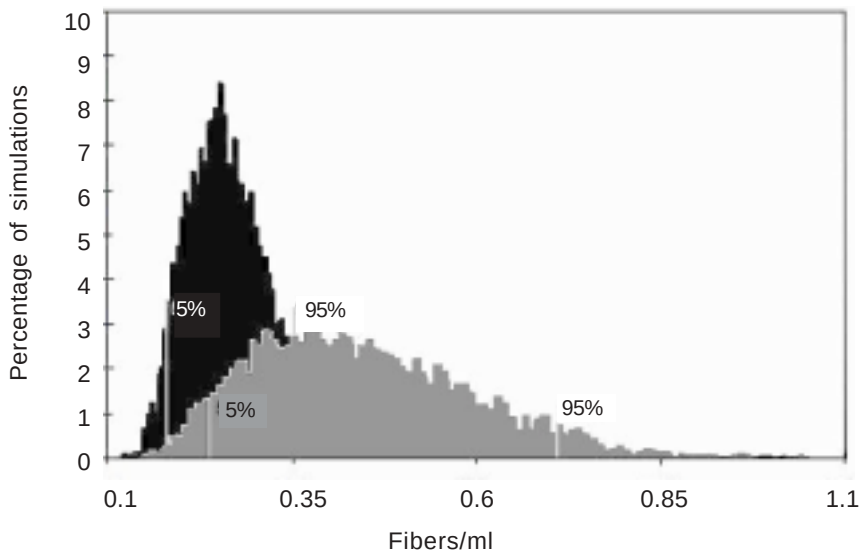

Figure 2. Distribution of modeled exposure levels for the jobs of molding operator and serviceman for period 2.

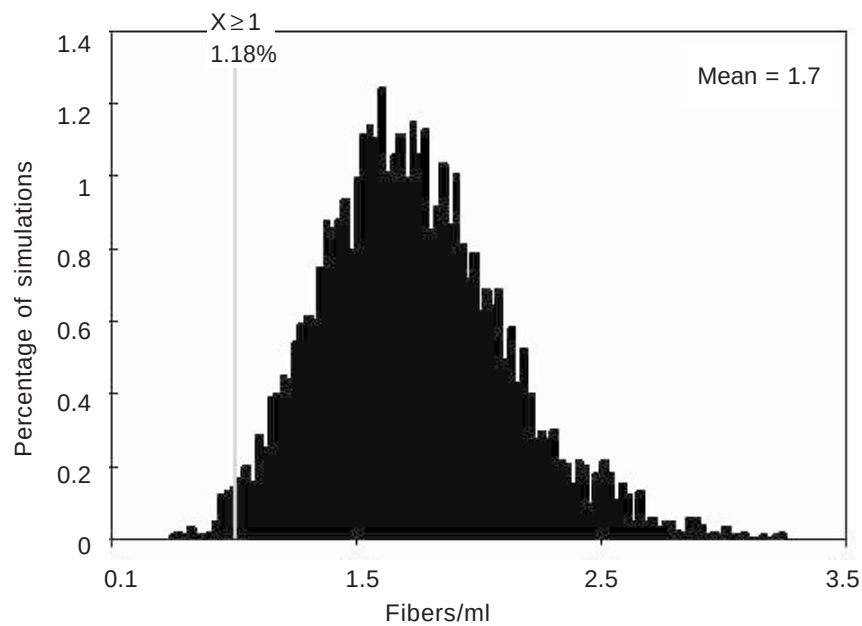

Figure 3. Distribution of ratio of serviceman to molding operator exposure levels for period 2.

estimated exposure for both the serviceman and molding operator job titles. This analysis showed that the three most important input parameters for molding operator exposure were the time spent by the serviceman in the clean-up task (r = 0.60), the general ventilation parameter (r = 0.48), and the time source active for the molding operator's forming task (r = 0.26). The three most important input parameters for the serviceman's exposure were the time spent by the serviceman in the clean-up task (r = 0.79), the general ventilation parameter (r = 0.34), and the emission factor during unloading by the serviceman (r = 0.32).

Discussion

Previous work has employed MC simulations for risk assessment purposes (14) or to gauge variability in exposure estimates. However, a recent study by Nicas & Jayjock (15) utilized MC simulations to determine uncertainty in exposure estimates produced by modeling as opposed to monitoring. An analysis of variability in monitoring results led the authors to suggest that in circumstances in which the geometric standard deviation of the exposure distribution is greater than 2.3, mathematical modeling is a more cost-effective and reliable method of estimating the mean exposure level when three or less exposure measurements are available.

The current study also uses MC analysis to examine the uncertainty of exposure estimates. An initial analysis of the available data relating to the molding operator and serviceman indicated that these two jobs were closely related. It was unclear whether there was sufficient difference to warrant their continued separation for the purposes of any epidemiologic study. Using our exposure model to reconstruct exposure levels suggested that the serviceman had exposures that were about double those of the molding operator throughout the period under study. However, there was uncertainty in many of the model parameters, and, therefore, it was difficult to know if the differences in the estimated exposure level were sustainable.

The importance of accurate exposure classification in epidemiologic studies is well understood. Misclassification of exposure generally gives rise to bias so that any relationship between exposure and response may be underestimated, perhaps even obscured entirely (16). It is therefore important to maintain differences between estimated exposures when justified, and we believe that the method described in this report can be used to identify when real differences exist.

The interlinked nature of the two jobs and the fact that the MMMF exposures generated by one group of workers were in the far field of the other group posed particular problems for the model. By dynamically linking the worksheets of the spreadsheet for each group, we were able to incorporate the influence of the other workers' far-field exposure in the model simulations.

This type of correlation between model parameters is more likely to reflect real-life situations than simplistic assumptions of independent actions. The relationships of model parameters should be fully explored when MC simulations are carried out to examine uncertainty. Many examples of interconnectivity within models exist. These examples may take the form of physical or behavioral influences. For example, physical relationships clearly apply to time periods—a longer time spent on one task will necessarily reduce the time available for another task in a workshift. Behavioral influences can include the increasing likelihood of using control factors when longer tasks are performed or when concentrations reach levels that can be detected by smell or trigger symptoms. In general, these physical and behavioral interrelationships act to reduce the uncertainty of the generated exposure levels, and their identification and incorporation into models can only improve the reliability of the produced estimates.

Models are also used in regulatory risk assessments for occupational exposure. MC simulations have been used to help determine the risks to health from benzene (17) and ethylene glycol ethers (18). Cox (17) combined MC uncertainty analysis and a physiologically based pharmacokinetic (PBPK) model to demonstrate that benzene exposures of less than 1 ppm do not increase the risk of tumor development. Within the European Existing Substances program the EASE model, developed by the UK Health and Safety Executive in the United Kingdom, is commonly used for risk assessment (4). It provides some indication of the uncertainty of the estimated exposure by expressing the output as a range of exposure levels. However, these ranges are arbitrarily defined without any reference to the uncertainties inherent in the model or the likely levels of variability in the factors employed to predict concentrations. MC simulation could provide a much better way of taking account of uncertainty in this arena, although, as currently formulated, the EASE model would not be amenable to this procedure, and we are not aware of other occupational exposure models incorporating MC simulation.

Our study shows how MC simulation can be successfully used to incorporate uncertainty in model parameters and to identify differences when they exist between occupational groups. There is a need for more work to characterize both the interdependency of used parameters in probabilistic models and the typical variability of these parameters. The parameter values used in the generation of exposure levels were assigned subjectively by an expert using data from worker interviews. More-validated information is needed on the

parameter values, and detailed guidance or a structured approach for assigning parameter distributions should be developed.

Acknowledgments

We are grateful to our colleagues at the Institute of Occupational Medicine for their helpful comments.

References

1. Kromhout H, Symanski E, Rappaport SM. A comprehensive evaluation of within- and between-worker components of occupational exposure to chemicals. Ann Occup Hyg 1993;37: 253–70.
2. Daniell W, Claypole K, Checkoway H, Smith-Weller T, Dager S, Townes B, et al. Neuropsychological function in retired workers with previous long term occupational exposure to solvents. Occup Environ Med 1999;56:93–105.
3. Glass D, Adams G, Manuell R, Bisby J. Retrospective exposure assessment for benzene in the Australian petroleum industry. Ann Occup Hyg 2000;44:301–20.
4. Friar J. The assessment of workplace exposure to substances hazardous to health: the EASE model. Bootle (UK): Health and Safety Executive; 1996. Internal report.
5. Cherrie JW, Schneider T, Spankie S, Quinn M. A new method for structured, subjective assessments of past concentrations. Occup Hyg 1996;3:75–83.
6. Boffetta P, Saracci R, Andersen A, Bertazzi PA, Chang-Claude J, Cherrie J, et al. Cancer mortality among man-made vitreous fiber production workers. Epidemiology 1997;3:259–68.
7. Dick F, Semple S, Osborne A, Soutar A, Seaton A, Cherrie JW, et al. Organic solvent exposure, genes, and risk of neuropsychological impairment. Q J Med 2002;95:379–87.
8. Cherrie JW, Schneider T. Validation of a new method for structured subjective assessments of past concentrations. Ann Occup Hyg 1999;43:235–45.
9. Semple SE, Proud LA, Tannahill SN, Tindall ME, Cherrie JW. A training exercise in subjectively estimating inhalation exposures. Scand J Work Environ Health 2001;27:395–401.
10. Cowie HA, Wild P, Beck J, Auburtin G, Piekarski C, Massin N, et al. An epidemiological study of the respiratory health of workers in the European refractory ceramic fibre industry. Occup Environ Med 2001;58:800–10.
11. Cherrie JW. The effect of room size and general ventilation on the relationship between near and far-field concentrations. App Environ Occup Hyg 1999;14:539–46.
12. Cherrie JW, Bodsworth PL, Cowie HA, Groat SA, Pettie S, Dodgson J. A report on the environmental conditions at seven European ceramic fibre plants. Edinburgh: Institute of Occupational Medicine; 1989. IOM report no TM.89/07.
13. Groat S, Kauffer E, Lovett M, Miller BG, Kidd MW, Davies L, et al. Epidemiological research in the European ceramic fibre industry 1994–1998; vol. 1 (Workplace concentrations of airbome dust and fibres). Edinburgh: Institute of Occupational Medicine; 1999. IOM report no TM.99/01.
14. Shade WD, Jayjock MA. Monte Carlo uncertainty analysis of a diffusion model for the assessment of halogen gas exposure during dosing of brominators. Am Ind Hyg Assoc J 1997;58: 418–24.
15. Nicas M, Jaycock M. Uncertainty in exposure estimates made by modeling versus monitoring. Am Ind Hyg Assoc J 2002;63: 275–83.
16. Checkoway H, Savitz D, Heyer N. Assessing the effects of nondifferential misclassification of exposures in occupational studies. Appl Occup Environ Hyg 1991;6:528–33.
17. Cox LA. Reassessing benzene risks using internal doses and Monte-Carlo uncertainty analysis. Environ Health Perspect 1996;104 Suppl 6:1413–29.
18. Sweeney LM, Tyler TR, Kirman CR, Corley RA, Reitz RH, Paustenbach DJ, et al. Proposed occupational exposure limits for select ethylene glycol ethers using PBPK models and Monte Carlo simulations. Toxicol Sci 2001;62:124–39.

Received for publication: 22 November 2002