Pergamon

Ann. occup. Hyg., Vol. 41, No. 5, pp. 591-604, 1997
© 1997 British Occupational Hygiene Society
Published by Elsevier Science Ltd. All rights reserved
Printed in Great Britain
0003-4878/97 $17.00+0.00

PII: S0003-4878(97)00009-4

QUANTIFICATION OF HISTORICAL DUST EXPOSURES IN THE DIATOMACEOUS EARTH INDUSTRY

Noah S. Seixas, Nicholas J. Heyer, Esther A. E. Welp and Harvey Checkoway
Department of Environmental Health, University of Washington, Seattle, WA 98195, U.S.A.

(*Received in final form* 30 *January* 1997)

Abstract—Quantitative estimates of dust exposure in a diatomaceous earth (DE) mining and milling operation have been derived based on air sampling records for the period 1948–1988. A total of 6395 records was included in the analysis. Conversion of results obtained by particle counting, expressed as millions of particles per cubic feet (mppcf) or gravimetrically from a filter cassette and expressed as mg m^{-3} total, were converted to mg m^{-3} respirable dust using a conversion factor derived from data obtained during the same periods at the plant. Conversion factors were calculated as the average difference of means on the log scale in order to provide stable and consistent conversions and as a ratio of arithmetic means so that the results could be compared with similar studies. After converting the available data to mg m^{-3} respirable dust, geometric mean (geometric standard deviation) concentrations were 0.37 (2.43) during the 1950s and 0.17 (2.35) during later periods. Exposures were estimated using two linear models, one estimating the changes in concentration over time, and the other providing job-specific mean exposures during the more recent period. Extrapolation of the estimates to periods prior to the availability of any data was done using a subjectively-determined scaling factor. The average estimated respirable dust concentrations for 135 jobs were 3.55 (±1.25), 1.37 (±0.48), 0.47 (±0.16) and 0.29 (±0.10) mg m^{-3} prior to 1949, 1949–1953, 1954–1973 and 1974–1988, respectively. Despite the limitations of the available data, the estimation procedures used are expected to provide reasonable quantitative estimates of silica-containing dust exposure for subsequent exposure–response analyses. © 1997 British Occupational Hygiene Society. Published by Elsevier Science Ltd.

INTRODUCTION

Strong associations between exposure to crystalline silica in a cohort of diatomaceous earth (DE) industry workers and mortality due to non-malignant respiratory disease and lung cancer have previously been demonstrated (Checkoway *et al.*, 1993). In that study, exposure was represented by an index derived from a set of qualitative factors representing dustiness in different jobs, temporal changes in exposures, use of respiratory protection and the percentage crystalline silica (primarily in the form of cristobalite) in various DE products. Despite the availability of a substantial quantity of industrial hygiene data from the plant, dust exposure was estimated subjectively using the scalars 0, 1, 3 and 6 for jobs considered unexposed, low, medium and highly exposed, respectively. This approach was adopted because of limitations of the available data, including sparse data for certain jobs or time periods, changes in sampling technique and analytical method over time, the use of area samples and poorly specified methods such as worst-case sampling strategies. Such limitations are typical impediments to accurate exposure quantification and may result in a significant degree of measurement error in the exposure–response analysis. In addition, the job-specific concentration scalars were not expressed in units that could be directly interpreted for standard-setting or risk assessment. The exposure assessment described here was designed to consider these limitations by modelling the available data and deriving quantitative estimates of dust exposure at the facility where the majority of the cohort worked.

BACKGROUND

DE is comprised of the skeletal remains of cell walls of microscopic marine algae (diatoms) that were deposited in large quantities in a few locations worldwide (ILO, 1989). Mining and development of products from DE deposits began around the turn of the 20th century and developed rapidly during the subsequent 50 years. DE was originally cut directly from the quarries for use as refractory bricks and later used as a constituent of formed insulation bricks. Currently, DE products are used extensively for a wide range of products, including filtration media, insulation, as carriers for catalysts, pesticides and fertilizers, and as fillers in paints and plastics.

In its natural form, DE is an amorphous silica that can be dried at temperatures up to about 400°C. When calcined by heating to about 1000°C, natural DE converts to 10–30% crystalline silica, mostly in the form of cristobalite. If a fluxing agent such as sodium carbonate, is added during the calcining process, the product contains 40–60% cristobalite. Cristobalite differs from α-quartz in its crystalline structure, has larger number of reactive surface sites and is considered to be more fibrogenic (Davis, 1996). The American Conference of Governmental Industrial Hygienists (ACGIH, 1980, 1995) and the National Institute for Occupational Safety and Health NIOSH (1974) recommended exposure limit for an 8-h time weighted average is 0.05 mg m^{-3}, which is half the recommended limit for quartz.

The facility where this study was conducted is located in Lompoc, California and is a major producer of DE products in the U.S. Operations began at the plant in 1902. The company was bought by the Johns Manville Corporation in 1928 and it remained with the company until 1992 when it was purchased by the Celite Co., which is the current operator. Crude diatomite was quarried in both bulk and solid brick form, transported to the mill areas, dried as crude DE products or crushed in mills and sent through large rotary calcining kilns with or without a carbonate flux. Powdered DE could be further processed by mixing with other materials before being bagged and shipped. A large quantity of airborne dust was generated in each production step during early years and efforts to reduce the amount of dust released to the atmosphere have been made since the 1930s. Significant dust suppression efforts at the plant included the change to paper bags from burlap for shipping product around 1940, and the introduction of cyclone dust collectors during the 1940s, and baghouses for air cleaning beginning around 1945. Of particular importance was the extensive study of pneumoconiosis and dust exposure in the DE industry, including this plant, by the California State Department of Public Health and the U.S. Public Health Service (USPHS) from 1952 to 1954 (Cooper and Cralley, 1958). The study resulted in a recommended exposure limit for crude DE (with $<5\%$ crystalline silica) of 20 million particles per cubic foot (mppcf), and for calcined DE containing cristobalite of 5 mppcf. In addition, this study resulted in major new efforts to reduce exposures at the source, and in an on-going programme of quantitative exposure monitoring.

MATERIALS AND METHODS

Quantitative air monitoring data available in computer files were provided by the Manville Corporation for the original study (Checkoway *et al.*, 1993). These data included 5714 records of samples taken over the period 1962–1988 and formed the basis for the initial exposure characterization. These data were coded according to the year in which the sample was obtained and the units in which it was measured: mppcf (millions of particles per cubic feet determined by light microscopy from an impinger sample), $mg\,m^{-3}$ total (milligrams per cubic metre determined gravimetrically from open face filter cassette sample), or $mg\,m^{-3}$ respirable (milligrams per cubic metre determined gravimetrically from a filter cassette sample using a cyclone pre-selector to obtain the respirable fraction). The location and type of sample was determined by sampling station codes which specified the department in which the station was located and whether the sample was an area or personal sample. Two of the major departments in which much of the exposure data were derived (Baghouse and Powder Mill) were further classified according to area within the department, reflecting the major process or type of material handled. Maps provided by the company were used to identify the location of the sampling station and assign the area codes. The department and area codes are described in Table 1.

Early in this study, it was reported to us that routine air monitoring had begun during the late 1940s to early 1950s. The substantial monitoring effort conducted during the USPHS study in the early 1950s added support to this contention. Because these data could not be located through direct contacts, a search was conducted at the Manville Archives which maintains all records concerning health and the workplace environment from the Manville Corporation. Among the records of the Manville Archives numerous reports were found detailing a wide variety of quantitative air sampling activities for the years 1948 to 1962 (no later records were obtained, because data since 1962 had been computerized).

Data obtained from 41 documents covering the period 1948 through 1962 were extracted, coded and entered into a database and summarized as arithmetic mean values by year and department. Because these data were poorly described in many of the reports, it was not consistently possible to determine if they were personal or stationary samples; thus, they were defined as unknown sample type. All these results were reported in units of mppcf. Some reports which used qualitative descriptors of dust conditions such as ‘heavy’ or ‘extremely heavy’ were not be used in the analysis. Reports concerning measurements of ambient releases to the environment, or reporting tests of specific pieces of equipment such as tractor cab enclosures were also not included. Thus, the reports utilized in this analysis include only those specifically measuring dust concentrations in defined work areas. The sampling station codes were inconsistently reported in these data and were coded by the department in which they occurred, without additional location or job specificity. Unweighted mean concentrations were used in order to give equal weight to each report because the number of samples used to generate the value was frequently missing from the reports. Duplicate records including 25 records that were redundant with the company's dataset (in 1962) were excluded.

Exposure data were assigned to a set of 135 job codes which had been defined previously from the original company job titles. Matching exposure data to

Table 1. Definitions of locations within the DE mine and mill

Department	Area	Description
Powder mill		Main powder processing areas including crude crushers, drying and calcining kilns and bag packing stations
	Packing station	Bag packing stations
	Dry end	Processing after drying on mills 3, 4, 5, 6
	Wet end	Crushing and feeding crude materials prior to drying, mills 3, 4, 5, 6
	Natural 11	Mill 11 in which no calcined material was processed
	Mill 11 Unspecified	Processing or baghouse for mill 11
	Dry end 7	Mill 7, prior to drying
	Wet end 7	Mill 7 after drying
	Other	Any other areas in powder mill
	Unspecified	Unknown
Baghouse		Dust control and maintenance operations
	Dry end	Baghouse operations located at the dry end of the powder mills 3, 4, 5, 6
	Natural 11	Dedicated to natural product in mill 11
	Dry end 7	Baghouse for mill 7, near dry end of mills
	Natural 7	Baghouse used only for natural DE
	Other	All other baghouse areas
	Unspecified	Unknown
Speciality products		Mixing and packing various mineral and chemical powdered products with lesser amounts of DE
Quarry		Quarrying and transport of natural product to mill
Brick plant		Production and drying of DE bricks
Maintenance		Building and maintenance operations
Quality Control		Quality control jobs throughout plant
Shipping/warehouse		Shipping and warehouse operations
Other		All other departments

individual jobs was done in the previous study by interviewing knowledgeable plant personnel and by walk-throughs of the facility (Checkoway *et al.*, 1993). Individual sampling stations were assigned to jobs which were conducted in the same area. Any particular job code could have several sampling stations assigned to it, and some sampling stations were not assigned to any specific job.

The first step in the quantitative analysis was conversion of all data to a uniform measure. Respirable dust in mg m^{-3} was selected as the measure to be used in the analyses because the recent data were all measured in these units, respirable dust is the most biologically relevant exposure fraction for silicosis or lung cancer, and any likely recommendations derived from the study would most appropriately be expressed in these units. Therefore, data collected as mppcf or mg m^{-3} total were converted to mg m^{-3} respirable. Data for years in which samples were obtained using both mg m^{-3} respirable and either mg m^{-3} total or mppcf were selected for calculation of conversion factors. Means (and mean of the logs) were calculated by sampling station. If means were available in two units for the same sampling station, the difference in the mean of the logs, and the ratio of the means were calculated. These factors were then averaged across appropriate categories of year, area or department. The relationship between the two units was thus expressed as:

$$\Delta_s = \overline{\ln(x)}_{s(resp)} - \overline{\ln(x)}_{s(mppcf)} \tag{1a}$$

or

$$P_s = \frac{\overline{x}_{s(resp)}}{\overline{x}_{s(mppcf)}} \tag{1b}$$

where Δ_s is the conversion factor for sampling station s based on the log scale, P_s is the ratio factor based on the natural scale. $x_{s(resp)}$ is the station-specific mean of the dust concentration, expressed in the units specified within parentheses. The sampling station-specific factors were then averaged across appropriate categories. The average value of the factors, Δ, was selected as the appropriate conversion factor because the data were generally lognormal and it helped minimize the large influence of a few high samples in the calculation of sampling station-specific means. Additionally, P_s was calculated in order to provide a comparison to previously cited conversion factors (Jacobsen and Tomb, 1967; Montgomery *et al.*, 1991; Tomb and Haney, 1988). After selecting a final conversion factor, the data originally measured in mppcf or mg m^{-3} total were converted to mg m^{-3} respirable by the use of the factor on each sample. For instance,

$$\hat{x}_{i(resp)} = \exp(\ln(x_{i(mppcf)}) + \Delta) \tag{2}$$

where $\hat{x}_{i(resp)}$ is the estimated concentration expressed in mg m^{-3} respirable dust and $\ln(x_{i(mppcf)}\hat{x}_{i(resp)}$ is the actual sample result in mppcf. The data originally collected as $x_{i(resp)}$, and the estimated concentrations, $\hat{x}_{i(resp)}$, were then used in all subsequent analyses.

The company data from 1962–1988 and the Manville Archives data from 1948 to 1962 were combined into a single dataset, and analysed using a linear regression model to describe the location-specific exposures and changes over time (Model 1):

$$\hat{x}_{i(resp)} = \alpha + \sum \beta_L L + \sum \beta_T T + \sum \beta_S S + \varepsilon \tag{3}$$

in which L represents indicator variables for location using department or department and area codes (see Table 1), T represent time periods and S represent indicators for sample type (area, personal or unknown). The definition of the time periods was made examining average dust concentration changes by year. Model selection was conducted by considering its parsimony and fit using the model R-squared. The change in concentration over time from this model, β_T, was used to estimate concentrations for each earlier period as described below.

There was no simple method by which the jobs could be linked directly to the model-derived exposure estimates because each specific job may have been associated with only a portion of the process described by its department, some jobs spanned multiple area assignments and samples from any particular station could be assigned to multiple jobs working in a given area. In addition, the archives data could not be matched completely to specific jobs because the location frequently was inadequately described. An alternative approach was therefore used to derive job-specific exposures.

Sampling stations were assigned to each job. All samples derived from those stations were used to calculate the job-specific means (on the log scale). Because these data included both area and personal data, and spanned the years 1962–1988, the means were calculated stratified by sample type and time period. These means were then used in a linear model (Model 2), to estimate exposure on the basis of personal samples for the most recent time period.

$$\overline{[\ln(x)]}_{STJ} = \alpha + \sum \beta_S S + \sum \beta_T T + \sum \beta_J J + \varepsilon \qquad (4)$$

where $[\ln(x)]_{STJ}$ is the mean of the log concentrations for a specific sample type, S, time period, T and job, J. Because the data used in this analysis were not independent (the same sample result was assigned to multiple jobs and the dependent variable represented the means of multiple samples), this model violates the assumption of independent observations. Regression with correlated (non-indepedent) observations can usually result in inflated variances, but valid point estimates; thus, Model 2 was used to provide estimates of job-specific mean personal exposures in the later time period using all of the relevant data.

Final job and time specific dust exposure estimates were calculated by adjusting for historical period and converting to the arithmetic mean:

$$\hat{x}_{STJ} = \exp\overline{[\ln(x)]}_{STJ} + \beta_T + \frac{1}{2}\hat{\sigma}^2 \qquad (5)$$

where $\overline{[\ln(x)]}_{STJ}$ is the mean of the logs for personal samples in the most recent time period for job J in the most recent time period estimated from Model 2, and β_T is the factor associated with historical time periods, and $\hat{\sigma}^2$ is the variance of the logs of the exposure data estimated from Model 1.

The final job- and time-specific dust estimates were compared to the qualitative ratings of high, medium and low given each job and the factors associated with time used in the previous analysis (Checkoway *et al.*, 1993), in visual plots and by analysis of variance. These respirable dust cocnentration estimates will subsequently be modified to estimate respirable crystalline silica concentrations, based on percentages of crystalline silica measured in various product mixes and historical production records (Checkoway *et al.*, 1993).

RESULTS

There were 140 station-specific matched means obtained during 1979 and from 1983 to 1988 with which measurements in mppcf could be compared to measurements in mg m^{-3} respirable units (Table 2). Likewise, there were 160 matched means from 1980–1983 with which measurements in mg m^{-3} total could be compared to measurements in mg m^{-3} respirable units. The average difference in mean of logs (Δ) was -2.67 for mppcf and -1.48 for mg m^{-3} total. When calculated within major departments, there were no clear or substantial differences. The Δs were not substantially different between the baghouse (-2.69) and all other departments (-2.67), although the ratio of arithmetic means (P) differed, especially comparing mppcf to mg m^{-3} respirable (0.18 vs 0.09). However, this difference was due to a limited number of samples within a single year and was not generally representative of the data. As a result of this analysis, a single conversion factor for

Table 2. Factors calculated to convert measurements of total dust in mg m^{-3} or mppcf into respirable dust in mg m^{-3}

Classification	mppcf to mg m^{-3} respirable n_m*	Δ†	P‡	mg m^{-3} Total to mg m^{-3} respirable n_m	Δ	P
All	140	−2.67	0.11	160	−1.48	0.29
By department						
Powder Mill	47	−2.62	0.10	47	−1.73	0.20
Baghouse	20	−2.69	0.18	18	−1.29	0.38
Specialty Products	12	−2.77	0.09	12	−1.24	0.39
Quarry	16	−2.90	0.07	21	−1.54	0.25
Shipping/Ware house	12	−2.54	0.08	14	−1.76	0.20
Maintenance	10	−2.57	0.14	20	−1.15	0.40
Other	23	−2.67	0.09	30	−1.35	0.32
Non−Baghouse	120	−2.67	0.09	142	−1.51	0.28

*Number of means matched by sampling station containing data in paired sets of units (for example, mppcf and respirable mg m^{-3}).

†Difference in the mean of the log values by sampling station defined in Equation (1A).

‡Ratio of arithmetic means by sampling station defined in Equation (1B).

total mg m^{-3} ($\Delta = -1.48$) and one for mppcf ($\Delta = -2.67$) was used to convert the data into consistent units of mg m^{-3} respirable.

Of the 5714 samples in the company dataset, five samples were very high (maximum, 52 mg m^{-3}) and appeared as outliers compared the remaining data. These five samples were removed from all subsequent analyses, leaving 5709 samples. These data were combined with those derived from the Manville archives ($n=686$) resulting in a master dataset used for analysis containing 6395 air concentrations over the years 1948 to 1988 (Table 3). After conversion to mg m^{-3} respirable units, these data were well described by the lognormal distribution. Reasonableness of the lognormality assumption was determined by visual inspection of probability plots and calculation of skewness for the whole dataset, and within sampling location strata. The measured concentrations were generally low, with mean values well under 1.0 mg m^{-3} and an overall geometric mean of 0.18 mg m^{-3}. About half the dataset ($n=3268$) were personal samples and about 1000 samples, including 686 from the archives data, were of unknown sample type.

Time and location variables were grouped in order to provide stable estimates and a parsimonious model. The crude average concentration by year (Fig. 1) was somewhat variable from year to year, but three reasonably distinct periods could be

Table 3. Air sampling data available by data source, sample method and sample type

	n Total (mppcf)	n Total (mg m^{-3})	n Respirable (mg m^{-3})	Converted to respirable (mg m^{-3}) n	AM	GM	GSD
Archived data (1948–1962)	686	0	0	686	0.68	0.37	2.43
Company data (1962–1988)	2931	973	1805	5709	0.25	0.17	2.35
Sample type							
Area	1066	337	689	2092	0.19	0.13	2.38
Personal	1574	625	1069	3268	0.27	0.20	2.23
Unknown	291	11	47	349	0.30	0.20	2.25
Complete data	3617	973	1805	6395	0.30	0.18	2.44

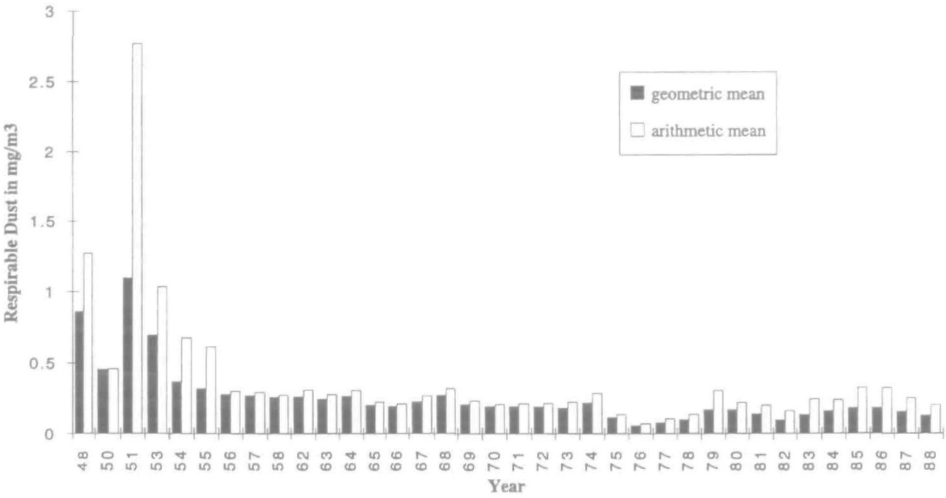

Fig. 1. Crude arithmetic and geometric mean respirable dust concentrations by year.

identified. Prior to 1954, the average concentration was clearly higher. Through the 1950s and 1960s, there were no clear changes in concentrations, but after about 1974, average concentrations were somewhat lower, but more variable. Thus, time was grouped into three categories (<1954, 1954–1973, $\geqslant 1974$). These particular boundaries were also chosen to be consistent with the earlier analysis of these data (Checkoway *et al.*, 1993). A four-level grouping was also considered, by dividing the middle period into 1954–1962 and 1963–1973, representing the change in the source of the exposure data. Grouping the data by sampling station was rejected because of the large number of stations (419), and because sampling station was not consistently available for the pre-1962 years. Thus, location was considered on the basis of department (nine levels) or department/area (22 levels), as defined in Table 1.

Linear models [Model 1, Equation (3)] were developed to compare alternative levels of grouping for time and location, while also considering the difference between area and personal samples. The dependent variable was the log of the converted dust concentration ($\ln(\mathrm{mg\ m^{-3}})$ respirable). Alternative models are compared in Table 4. Area and personal samples were significantly different and explained about 10% of the total variability. Year explained 20% of the variability; however, grouping year into the three distinct periods explained 13%. There was no substantial improvement in the model ($R^2=0.14$) when the years were divided into four periods representing the change in source of data in 1962, suggesting that the data were consistent across this time period. Department, surprisingly, explained only 4% of the variability; this was somewhat improved by sub-dividing the major dusty departments, baghouse and powder mill, into distinct areas. With the combination of sample type, time in three groups and department/area, the model explained almost 22% of the variability and this model was selected to provide estimates of time-related changes in dust concentration. The complete Model 1 results are presented in Table 5.

The available exposure data from the company (1962–1988), in which the

Table 4. Alternative models (Model 1) for describing exposure by time, location and sample type variables*

Variable	DF	SS†	P	Model r^2
Area/personal	2	534	<0.001	0.105
Year	35	1049	<0.001	0.206
Year Group 4	3	725	<0.001	0.142
Year Group 3	2	669	<0.001	0.132
Department	8	209	<0.001	0.041
Department/area	21	412	<0.001	0.081
Area/personal	2	403	<0.001	0.184
Year Group 3	2	268	<0.001	
Area/personal	2	294	<0.001	0.208
Year Group 3	2	363	<0.001	
Department	8	119	<0.001	
Area/personal	2	263	<0.001	0.218
Year Group 3	2	367	<0.001	
Department/area	21	172	<0.001	

*Dependent variable, log (respirable dust in mg m^{-3}).

†SS, Sum of Squares for Model. Total SS for each model = 5082.

sampling station information was complete, were matched by station to 135 distinct jobs and means (of log concentrations) were calculated for each job by sample type and time period. Estimates of the mean personal exposure in the post 1973 time period were calculated using Model 2 [Equation (4)] with the means (of logs) as the dependent variable. The model-derived job-specific exposures (on the log scale for personal samples in the post-1973 period) were then extrapolated to the earlier time periods using Equation (5). The estimates of the change in dust concentrations over time (β_T), derived from Model 1, were 0.48 for the 1954–1973 and 1.54 for periods prior to 1954. No data were available prior to 1948; however, evidence from the company's history suggested that substantially higher concentrations were present in the 1930s and early 1940s. In about 1944 the conversion from paper to burlap bags was completed and other substantial dust control efforts were reported during this period. To be consistent with the previous analysis, we chose 1944 for the boundary and assigned a factor of 2.5 (log scale) to extrapolate the dust concentrations to jobs prior to this time.

In order to convert to arithmetic means, an estimate of the variance was required. Because the variance was highly unstable, and is even more sensitive to biases in the data than the mean, we chose to use the average variance observed in the later time period derived from Model 1 (Table 5). The estimated variability represented by σ was 0.909 which represents as geometric standard deviation of 2.48. This variance estimate was applied using Equation (5) to estimate arithmetic mean concentrations for all jobs and time periods. The final estimated mean dust concentrations are presented in Table 6.

In the earlier analyses (Checkoway *et al.*, 1993), each job was assigned a concentration represented on a qualitative scale of high, medium or low dustiness. A

Table 5. Final model estimates used to describe exposure over time

Variable	Level	β	SE(β)
Intercept		−1.81	0.05
Sample type	Area	−0.44	0.04
	Personal	0.12	0.04
	Unknown*	0.0	—
Year Group	< 1954	1.54	0.08
	1954–1973	0.48	0.02
	⩾ 1974*	0.00	—
Department/area	Powder mill: unspecified	0.04	0.08
	Powder mill: packing station	−0.15	0.04
	Powder mill: dry end	0.13	0.06
	Powder mill: wet end	0.26	0.09
	Powder mill: natural 11	−0.33	0.30
	Powder mill: dry end 7	−0.35	0.08
	Powder mill: wet end 7	−0.004	0.20
	Powder mill: other	−0.03	0.06
	Baghouse: unspecified	0.59	0.09
	Baghouse: dry end	0.25	0.06
	Baghouse: natural 11	0.33	0.09
	Baghouse: dry end 7	0.38	0.12
	Baghouse: natural 7	0.09	0.10
	Baghouse: other	−0.25	0.21
	Specialty products	−0.49	0.06
	Mill 11 unspecified	−0.03	0.06
	Quarry	−0.08	0.05
	Brick plant	0.13	0.09
	Maintenance	0.01	0.05
	Quality control	0.33	0.15
	Other	−0.07	0.04
	Shipping/warehouse*	0.00	

Independent variable, log [dust (mg m^{-3})]. Model $r^2 = 0.21$.
*Baseline level, coefficient is 0 by definition.

Table 6. Estimated arithmetic mean exposures by time period for 135 jobs (mg m^{-3} respirable dust)

Period	Mean	SD	Minimum	Maximum
< 1948	3.55	1.25	0.97	7.14
1949–1953	1.37	0.48	0.38	2.75
1954–1973	0.47	0.16	0.13	0.95
1974–1988	0.29	0.10	0.08	0.59

comparison of those subjective assignments to our quantitative estimates is presented in Fig. 2. While there is some trend toward higher concentrations among the jobs classified as medium or high, there is a large degree of overlap between each of the categories indicating the large potential for exposure misclassification. Analysis of variance of the estimates between the three levels indicated a statistically significant difference ($P = 0.02$), although the average estimate in the high category (0.29 mg m^{-3}) was slightly lower than that for the medium category (0.31 mg m^{-3}).

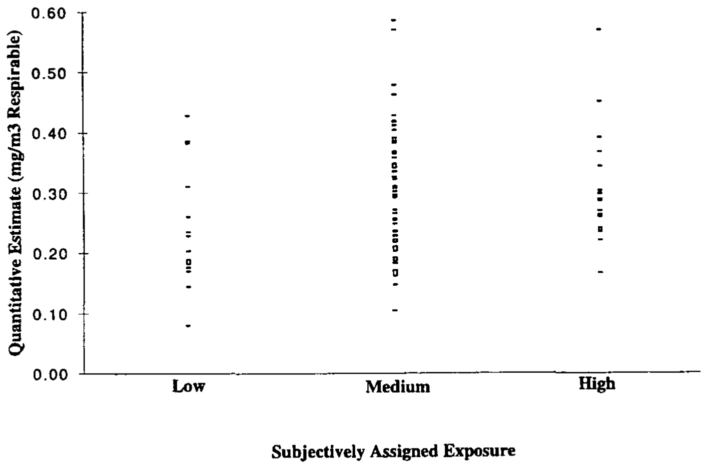

Fig. 2. Estimated job-specific means for the post-1973 period in comparison with the previously used subjective exposure assignments.

DISCUSSION

Retrospective quantitative exposure assessment is by its nature replete with potential errors. Although the availability of extensive industrial hygiene measurements greatly improves the potential for accurately estimating exposure levels, there are numerous sources of error which may still hamper the effort. The errors include sparse data for certain individuals, jobs or time periods, inappropriate, inaccurate or changing sampling and analytical methods, use of area samples to represent personal exposures and biased sampling strategies such as task-specific sampling or 'worst case' sampling. All of these sources of potential error are present to a substantial degree in the history of the cohort examined here, and inherently limit the accuracy of the final results. Nevertheless, by considering the problems, and by the application of statistical modelling, potential measurement error bias may be minimized, and the advantages of using quantitative data may be exploited.

The data available for this study were collected using three substantially different methods: by impinger and quantified by particle counting under a light microscope and expressed as mppcf up to 1983, by gravimetric analysis of a filter cassette sample for total dust, and finally, respirable dust beginning in 1979. Quantitative relations between these different measures are dependent on the precision of the sampling and analysis method, which especially for the impinger method is highly dependent on the skill of the analyst, and the particle size distribution which may change substantially as the product, process machinery and control technologies are altered over time.

In support of the earlier study of this plant, an experimental investigation was conducted to compare the various methods used at this plant (Montgomery *et al.*, 1991). That investigation demonstrated differences between the baghouse and powder mill exposures, presumably because of the smaller particle size distribution of the 'fines' found released from the baghouse operations. However, this experimental approach is limited for historical reconstruction because current exposure conditions may be substantially different from conditions during the historical periods of interest. Also, measurement skills and technique undoubtedly vary significantly between the experimental scientist and the routine practitioner, especially for microscopic particle counting. As a result, we chose to use the data available from the plant during periods in which two comparable measurement methods were being used to derive our conversion factors. Because we did not have control over data quality, nor complete information on past sampling strategies, we chose to calculate a conversion factor based on the log concentrations, as this should minimize large effects that non-representative worst case samples might create. Nevertheless, the more standard ratio of the arithmetic means was also calculated in order to compare with previously cited conversion factors. As demonstrated in Table 7, the conversion factors calculated from our data lie between those derived experimentally in the DE industry by Montgomery *et al.* (1991) and those generally observed in other industries (Tomb and Haney, 1988), or calculated based on theoretical grounds (ACGIH, 1980, 1995). Given that the factors calculated for this study were derived from the actual data used in the study, and are within the same range as those observed in this industry and others, the resulting estimated concentrations appear reasonable.

Limited information was available for this study on sampling strategy, and in particular, the likelihood that much of the data were obtained by a worst case strategy. Although this problem cannot be completely avoided when using historical data, we attempted to control its influence in the analysis. By conducting the analysis on the log scale, the potential influence of particularly high worst case or non-representative samples is dampened. Conversion back to the appropriate dose estimator, the arithmetic mean, was conducted using a single estimate of the variability (GSD = 2.48) derived from Model 1 for the post 1973 period. By using this single model-derived estimate of variability for all jobs and time periods, we

Table 7. Comparison of conversion factors with those found in previous studies

Study	Data source	Conversion factor*
Current	DE mining and milling company sampling using both methods	0.11
	Baghouse 'fines'	0.18
	Non-baghouse	0.09
Montgomery *et al.*, 1991	DE mining and milling, side by side experimental	
	Baghouse fines	0.064
	Powdermill	0.018
Tomb and Haney, 1988	Experimental samples in various industries	
	DE—non-calcined	0.046
	Graphite	0.12
	Perlite	0.18
	Talc	0.19
	Mica	0.15
Jacobsen and Tomb, 1967	Coal mine dust	0.18
ACGIH, 1980	Theoretical, assuming silica particles with mass median diameter of 1.5 μm	0.16

*Conversion factor is mg m^{-3} respirable/mppcf.

reduced the error associated with estimation of the geometric standard deviation, which is very high in most situations (Buringh and Lanting, 1991). This estimate based on the later time period probably represents data least likely to be affected by severe sampling and analytic error. Some additional assurance that this was a reasonable approach comes from the suggestion of Buringh and Lanting (1991), that when an accurate estimate of the GSD cannot be derived from the data, a value of 2.7 might be reasonably assumed.

Finally, the absence of quantitative data is perhaps the severe limitation of most retrospective exposure assessments. No data were available prior to 1948 and estimates were required for this cohort much earlier. We used a somewhat arbitrary multiplier of 2.5 for the pre-1944 period. This multiplier was derived simply as an extension of the data-derived values we had for the later periods of 0, 0.48 and 1.52. An additional step of 1 unit (on the natural log scale) seemed reasonable, and gave us estimated average concentrations within the range that plausibly could have existed: 1–7 mg m^{-3} respirable dust.

The high degree of apparent misclassification shown in Fig. 2 comparing the previously used subjective exposure categories and the currently estimated dust concentrations is worrisome. However, it is plausible that either the subjective estimates include a high degree of error, as demonstrated by the overlap of the distributions, or that the quantitative estimates are highly inaccurate, as seen in the wide variability within an exposure category and the lack of a difference in means between the two higher categories. In the former case, the previously estimated exposure–response relationship would be biased toward the null, and in the latter case, the newly estimated exposure–response relationship will likely be biased toward the null. In reality, there is probably some truth in both sets of estimates. Methods combining two types of exposure estimates to exploit the strengths of each could help control the random error in each set of exposure estimates, and prove useful in estimating a more accurate exposure–response relationship.

The data available for this analysis inherently limit the estimates in terms of random error (large variation in environment, over years and between jobs, and small numbers of samples in any single category) and bias (worst case and area sampling strategy, sampling and analytic method changes, etcetera). Nevertheless, the methods adopted for this assessment were designed to minimize these limitations and provide valid quantitative exposure estimates for the exposure–response analyses which will follow. Although the high degree of variability in exposures from day to day, and between individuals within a job category is well known (Kromhout *et al.*, 1993), the averaging of the variability through grouping and modeling actually helps to reduce the effect of random error on the exposure–response models (Seixas and Sheppard, 1996). Although biased data cannot be discounted, the effects of these biases have been addressed to the degree possible in the analysis. Ultimately, we will gain insight into the accuracy of the exposure estimates by their performance in the exposure–response analyses and by comparison with other exposure metrics for DE.

Acknowledgements—The authors thank Sylvie Adam, MS, for her rigorous efforts in extracting and coding the historical industrial hygiene data. This work was supported by grant RO1 OH03126 from the U.S. National Institute for Occupational Safety and Health.

REFERENCES

ACGIH (1980, 1995) *Threshold Limit Values for Chemical Substances and physical agents and biological exposure indices.* American Conference of Governmental Industrial Hygienists, Akron, OH.

Buringh, E. and Lanting, R. (1991) Exposure variability in the workplace: its implications for the assessment of compliance. *American Industrial Hygiene Association Journal* **52,** 6–13.

Checkoway, H., Heyer, N. J., Demers, P. A. and Breslow, N. E. (1993) Mortality among workers in the diatomaceous earth industry. *British Journal of Industrial Medicine* **50,** 586–597.

Cooper, W. C. and Cralley, L. J. (1958) *Pneumoconiosis in diatomite mining and processing.* US DHEW, PHS.

Davis, G. S. (1996) Silica. In *Occupational and Environmental Respiratory Disease,* eds P. Harber, M. B. Schenker and J. R. Balmes, pp. 373–399. Mosby-Year Book, St Louis.

ILO (1989). *Encyclopedia of Occupational Safety and Health.* International Labour Office, Geneva.

Jacobsen, M. and Tomb, T. F. (1967) Relationship between gravimetric respirable dust concentration and midget impinger number concentration. *American Industrial Hygiene Association Journal* **28,** 554–556.

Kromhout, H., Symanski, E. and Rappaport, S. M. (1993) A comprehensive evaluation of within and between worker components of occupational exposure to chemical agents. *Annals of Occupational Hygiene* **37,** 253–270.

Montgomery, J. S., Horstman, S., Breslow, N., Heyer, N., Stebbins, A. and Checkoway, H. (1991) A comparison of air sampling methods for airborne silica in the diatomaceous earth industry. *Applied Occupational Environmental Hygiene* **6,** 696–702.

NIOSH (1974) *Criteria for a recommended standard: Occupational Exposure to crystalline Silica.* NIOSH Publication No. 75-120. US Government Printing Office, Washington, DC.

Seixas, N. S. and Sheppard, L. (1996) Maximizing accuracy and precision using individual and grouped exposure assessment. *Scandinavian Journal of Work Environmental Health* **22,** 94–101.

Tomb, T. and Haney, R. (1988) Comparison of number and respirable mass concentration determinations. In *Advances in Air Sampling,* American Conference of Governmental Industrial Hygienists. Lewis Publishers, Chelsea, MI.