Taylor & Francis
Taylor & Francis Group

American Industrial Hygiene Association Journal

ISSN: 0002-8894 (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/aiha20

Application of a Pilot Study to the Development of an Industrial Hygiene Sampling Strategy

MARTIN R. PETERSEN , WAYNE T. SANDERSON & STEVEN W. LENHART

To cite this article: MARTIN R. PETERSEN , WAYNE T. SANDERSON & STEVEN W. LENHART (1986) Application of a Pilot Study to the Development of an Industrial Hygiene Sampling Strategy, American Industrial Hygiene Association Journal, 47:10, 655-658, DOI: 10.1080/15298668691390395

To link to this article: https://doi.org/10.1080/15298668691390395

Published online: 04 Jun 2010.

Am. Ind. Hyg. Assoc. J. 47(10):655-658 (1986)

Application of a Pilot Study to the Development of an Industrial Hygiene Sampling Strategy

MARTIN R. PETERSEN, WAYNE T. SANDERSON and STEVEN W. LENHART*
National Institute for Occupational Safety and Health, 944 Chestnut Ridge Road, Morgantown, WV 26505-2888

An industrial hygiene pilot study was conducted to estimate the concentrations of respirable dust likely to be encountered during the personal sampling phase of a large-scale morbidity study of the portland cement industry. An analysis of the pilot study data showed little variability in exposure for subjects working in the same job in the same area of the same plant. Thus, one could estimate mean exposure by sampling several subjects rather than sampling the same subject several times. It was concluded that for statistical considerations, the best approach would be to sample four jobs per area and six subjects per job. Practical considerations required one more often to select six jobs with two subjects per job, however. Overall, the collection of fewer samples was required during the morbidity study than was anticipated originally. In turn, the reduction in the number of samples to be collected resulted in a savings of time and resources.

Introduction

An important phase of any large-scale study designed to investigate the exposure-health effect relationships of a worker population is the collection of environmental samples for the accurate estimation of daily exposure. Certainly, the most accurate approach for obtaining exposure values is to collect personal samples on every worker in the industry of interest on several randomly selected occasions over a specified time period. Since the practical aspects of a study, such as feasibility and fiscal resources, must be considered, however, the collection of samples on every worker would not be possible. Realistically, enough samples need to be collected to answer questions regarding exposure-health effect relationships, yet avoid wasting time and money over-sampling a worker population.

This report describes the use of a pilot study to ascertain the number of environmental samples needed for characterizing the exposures experienced by workers in the U.S. portland cement industry, as part of the National Institute for Occupational Safety and Health (NIOSH) Cement Workers Morbidity Study.[1] The idea for developing a pilot study came out of necessity, because data from previous industrial hygiene surveys of portland cement plants were inadequate for developing a large-scale sampling strategy. Data from a previous study of the portland cement industry[2] was too old to be of value, since production methods and sampling techniques have changed. Personal sampling data collected by compliance officers of the Mine Safety and Health Administration were reviewed, but generally consisted only of those jobs encountering the highest exposures. Sampling data was sought from the management of several portland cement plants with the support and assistance of the Portland Cement Association. The data they had were not suitable for determining sources of variability for respirable dust levels, however. The industrial hygiene sampling methods of this large-scale study were intended to be comprehensive enough to detect all potential respiratory hazards (*i.e.*, those which might affect lung function, chest symptoms or chest radiographs) and measure exposure accurately, but be selective enough to exclude unnecessary tests and measurements. The purpose of this pilot study was to calculate the number of jobs and the number of subjects per job for which respirable dust samples needed to be collected in the main cement study, as well as to calculate the number of replications needed per subject. The methods are not new, but the proper interpretation of the statistics has not been well defined. This paper illustrates one method of interpretation.

Methods

The process of manufacturing portland cement begins with the grinding of carefully proportioned raw materials into a powder. This raw mix then is subjected to pyroprocessing in a rotary kiln in either dry form or as a slurry of raw mix and water. The intense heat of the rotary kiln changes the chemical properties of the mix, and round, marble-size, glass-hard balls called clinker are produced. After cooling, clinker is combined with gypsum, and a final grinding process results in the production of the finished cement.

The pilot study data consisted of personal respirable dust measurements which were made at five portland cement plants. Selection of a plant for the pilot study depended upon work force stability, low probability of closing down in the near future, and expected management-worker cooperation. Plants were not selected randomly for the pilot study, but care was taken not to make the sample too restrictive. For example, care was taken to select both "wet" and "dry" process plants, plants of various ages, and plants located in various geographic regions.

*Present address: North Carolina Department of Labor, 214 West Jones Street, Raleigh, NC 27603.

The workers in each plant of the pilot study were divided into four exposure categories based on the predominant type of dust in their work areas, *i.e.*, raw material, clinker, finished product, and mixed exposures. Within each of these areas, two to five job categories were selected randomly from all possible job categories, and within each job category one to four test subjects were selected randomly from all possible subjects. Most test subjects were sampled on two successive days, but some were sampled on only one day.

Selected workers were requested to wear a respirable dust sampler. For respirable dust sampling, air was drawn at a flow rate of 1.7 L/min through a 10 mm nylon cyclone and collected on a 37 mm polyvinyl chloride filter. Gravimetric analyses to the nearest 0.01 mg were used in calculating TWA exposures in mg/m^3.

Statistical Analysis

The statistical view of a respirable dust measurement is the sum of an overall mean for the particular plant and area, a number (either positive or negative) corresponding to the job for which the dust measurement was taken (called the effect of job), another number (either positive or negative) corresponding to the subject who was wearing the respirable dust sampler (called the effect of subject), and a number (either positive or negative) which is called random error and results from unknown sources. The mean is a constant (for a particular plant and area), but is unknown. Different dust measurements arise from varying the last three numbers. Two dust measurements on the same subject (and thus the same job) generally will be different because of two different values of the random error term. Two different subjects working the same job generally will have different measurements because the corresponding values of the latter two terms will be different. Subjects working different jobs within the same area of the same plant will have different measurements because the last three terms may vary. For subjects in different plants or areas, the mean may change also. Because these last three terms may vary, they have a variance, and these three variances are called variance components. (If a term does not vary, its variance and thus variance component will be zero.)

Estimated variance components for job ($\hat{\sigma}^2_{job}$), subject ($\hat{\sigma}^2_{sub}$), and error ($\hat{\sigma}^2_e$) were obtained by Henderson's Method 1.[3] Briefly, the method requires obtaining observed and expected sums of squares for each variable, and solving for components of variance by three equations in three unknowns (see Table I). Usually the method is not appropriate when some factors are fixed (*e.g.*, plant and area) while other factors are random (*e.g.*, job, subject and error). For the present data, however, because all of the random effects are nested within all of the fixed effects, it is possible to obtain variance components for the random effects. Nesting of jobs within a particular plant area was preserved even though jobs within other plant areas were sometimes quite similar. This was done to simplify the analysis.

Once variance components were obtained, they were used to calculate the approximate half-widths of the 95% confidence intervals on the mean dust level of an area for varying numbers of jobs, subjects and measurements. The intervals are approximate because the data are described best by a log-normal distribution, but logarithms were not taken so that the industrial hygienist could make an easier judgment as to desired half-width.

Results

Table I contains the observed and expected sums of squares as well as the estimated variance components. All three of the *true* variance components being estimated, σ^2_{job}, σ^2_{sub}, and σ^2_e, must be non-negative, but the *estimate* of the subject to subject variance component, $\hat{\sigma}^2_{sub}$, is negative. Thus, it is reasonable to assume that the true subject to subject variance component is close to zero, because if this true variance component were much larger than zero, it would be unlikely to obtain a negative estimate of it ($\hat{\sigma}^2_{sub}$ is an unbiased estimator of σ^2_{sub}). Therefore, with respect to the variance of an area or job mean, it should make little difference whether one selects only one person in a certain job and measures that person's exposure several times, or selects several persons and measures each one's exposure only once. To avoid placing a burden on a few subjects and to obtain the necessary data in as few days as possible in the large-scale morbidity study, several people were to be selected and measured only once.

Since the effect of inter-subject variability (σ^2_{sub}) is estimated to be zero, it will be assumed to be zero, and thus ignored when determining a sampling strategy. Therefore, the model for individual dust measurements within a particular plant-area category can be written as:

TABLE I
Mean Squares and Estimated Variance Components

Variance Component	Sums of Squares: Observed	Sums of Squares: Expected
Job	153.82772	$43\sigma^2_e + 77.10833\sigma^2_{sub} + 152.808080\sigma^2_{job}$
Subject	81.42787	$72\sigma^2_e + 115.33333\sigma^2_{sub}$
Error	137.25500	$93\sigma^2_e$

Estimates obtained by equating observed and expected sums of squares:

Variance *Component*	*Calculation*[A]
Error	$\hat{\sigma}^2_e = \frac{137.25500}{93} = 1.47586$
Subject	$\hat{\sigma}^2_{sub} = \frac{81.42787 - 72\hat{\sigma}^2_e}{115.333} = \frac{81.42787 - 72(1.47586)}{115.3333} = -0.21532$
Job	$\hat{\sigma}^2_{job} = \frac{153.8277 - 43\hat{\sigma}^2_e - 77.10833\hat{\sigma}^2_{sub}}{152.80808} = \frac{153.8277 - 43(1.47586) - 77.10833(-0.21532)}{152.80808} = 0.70002$

[A]When the observed sums of squares are equated with their expected sum of squares, the true components (without hats, *e.g.*, σ^2_e) are replaced by their estimates (with hats, *e.g.*, $\hat{\sigma}^2_e$).

TABLE II
Half-Widths of Approximate 95% Confidence Intervals on Mean Plant-Area Respirable Dust

No. of Jobs	Number of Subjects (or Replicates) Per Job 1	2	3	4	5	6	7	8	9	10
1	2.95	2.40	2.18	2.07	2.00	1.95	1.91	1.88	1.86	1.84
2	2.09	1.70	1.54	1.46	1.41	1.38	1.35	1.33	1.31	1.30
3	1.70	1.38	1.26	1.19	1.15	1.12	1.10	1.09	1.07	1.06
4	1.48	1.20	1.09	1.03	1.00	0.97	0.95	0.94	0.93	0.92
5	1.32	1.07	0.98	0.92	0.89	0.87	0.85	0.84	0.83	0.82
6	1.20	0.98	0.89	0.84	0.81	0.79	0.78	0.77	0.76	0.75
7	1.12	0.91	0.83	0.78	0.75	0.74	0.72	0.71	0.70	0.70
8	1.04	0.85	0.77	0.73	0.71	0.69	0.67	0.67	0.66	0.65
9	0.98	0.80	0.73	0.69	0.67	0.65	0.64	0.63	0.62	0.61
10	0.93	0.76	0.69	0.65	0.63	0.62	0.60	0.59	0.59	0.58

$$Y_{ijk} = \mu_i + J_{j(i)} + \epsilon_{k(ij)}$$

where Y_{ijk} is a respirable dust measurement,

μ_i is the true mean of the ith plant-area combination,

$J_{j(i)}$ is the effect of the jth job in the ith plant-area combination and

$\epsilon_{k(ij)}$ is the random error associated with repeated measurements within the jth job within the ith plant-area combination.

Now we can consider how precisely the mean dust exposure for a particular job category can be estimated. Ideally, we would take each job, make a certain number of exposure measurements for that job, and assign the mean of these measurements to each subject in that job category. The number of measurements, (r), to be taken would be chosen such that the variance of this mean, $\hat{\sigma}_e^2/r = 1.47586/r$, was acceptably small. For the Cement Workers Morbidity Study, however, it was not possible to sample all jobs because this would require more staff and equipment or more time than was available. The approach taken was to estimate the mean for a plant-area combination with an acceptable accuracy (namely, ± 1.0 mg/m^3 with approximately 95% confidence), although there was no intention of assigning this mean to anyone. This value is the mean of the dust exposures for all jobs in that area.

Based on the model, the variance of the sample mean for a plant-area combination is $\sigma_{job}^2/m + \sigma_e^2/mr$, where m is the number of jobs selected, and r is the number of measurements per job. The criterion used to decide how many jobs and replications to choose was based upon twice the standard error of the mean for a plant-area combination,[4] *i.e.*, $2(\sigma_{job}^2/m + \sigma_e^2/mr)^{0.5}$. This represents one-half of the approximate 95% confidence interval on the true mean, μ_i (half-width).

Table II contains these half-widths, namely $2(0.70002/m + 1.47586/mr)^{0.5}$, for various values of m and r. Based on a subjective judgment of feasibility and the results of Table II, a half-width of 1.0 mg/m^3 was selected as being acceptable. Also, because in the morbidity study it was decided to sample a subject only once, replications, r, also refer to the number of subjects. From Table II, one can see that the following choices satisfy the desired criterion: nine jobs with one subject per job, six jobs with two subjects per job, five jobs with three subjects per job, or four jobs with five subjects per job. Because the main purpose of measuring the dust was to assign the mean level for a job to all subjects doing that job, it seemed reasonable to control the variability of this mean also. This half-width would be $2(1.47586/r)^{0.5}$, which is less than 2 for r = 2 and less than 1 for r = 6. Possibly the best approach would be to take four jobs and six subjects per job, because this would keep the half-widths of both plant-area and job at 1.0. There were not always six subjects available for some jobs, however. Therefore, in practice six jobs with two subjects per job often were used, provided six jobs per category were available. Also, this sampling strategy required twelve samples instead of twenty-four, and that was an important consideration.

Discussion

There are many ways to interpret the estimated variance components. Essentially, after the pilot study data was collected, we decided to ignore the component if the estimate was less than or equal to zero. If the estimate was positive, we considered the component to be important. Had we decided to ignore the component if the estimate was less than or equal to one, the component for job would have been ignored also, and the sampling would have been simplified just to taking observations in an area and assigning the mean to all subjects working in the area. In this case, one would have the choice of sampling one subject multiple times, a few subjects from several jobs several times, or several subjects from several jobs once. One possible way to decide where to make the cutoff would be to prepare Table II as was done and see if the entries vary as one looks down a column, which would indicate that the job component was important. In the present Table II, the entries do vary, and so it seems reasonable to take job into account.

The pilot study proved to be a valuable effort in the design of the overall Cement Workers Morbidity Study. It was found that only the components for job and random error were important, and that subject to subject variability could be ignored. Based on the results of this pilot study, it was evident that fewer samples (twelve: six jobs with two observations per job) were needed than previously was believed necessary. This of course resulted in a considerable savings in time and resources. Although there is no single best way to interpret this type of data, it seems reasonable that the statistical analysis of a pilot study of this type could be advantageous for many studies in which dust sampling will be performed.

Acknowledgment

The authors gratefully acknowledge the assistance provided by Mrs. Linda DeVor in the preparation of this manuscript.

References

1. **Abrons, H.L., W.T. Sanderson and M.R. Petersen:** Respiratory Effects of Portland Cement Dust. (DHHS/NIOSH) Morgantown, W. Va.: 1985.
2. **Thompson, L.R., D.K. Brundage, A.E. Russell and J.J. Bloomfield:** The Health of Workers in Dusty Trades. I. Health of Workers in a Portland Cement Plant (Public Health Bulletin No. 176). Washington, D.C.: Government Printing Office, 1928.
3. **Henderson, C.R.:** Estimation of Variance and Covariance Components. *Biometrics 9*:226-252 (1953).
4. **Snedecor, G.W. and W.G. Cochran:** *Statistical Methods.* 6th ed. Ames, Iowa: Iowa State University Press, 1967. pp. 279-282.

4 November 1985; Revised 15 April 1986