응용통계연구 제 6권 2호
1993년 9월 pp.201-215

Statistical analyses in an occupational health study

Domyung Paek[1], Jung-Keun Choi, Mi-A Son.

1. Abstract

The health status of workers in a foundry was analyzed in a study which consisted of evaluations of respiratory health together with environmental measurements.

The results from environmental measurements showed values exceeding permissible exposure limits. A t-test was done with log transformed and untransformed data to examine the statistical significance for the noncompliance with exposure standards. For the analysis of categorical health outcomes, χ-square test with 2×2 tables and logistic regression analysis were employed. For continuous variables, multiple linear regression was done against assessed risk factors.

Pros and cons of different parameters in the compliance (or noncompliance) testing were presented. Respiratory function did not show any relation with occupational exposures, which may be due to the healthy worker effects. Strategies for controlling time dependent covariates were discussed in relation to the healthy worker effect.

The scope of statistical analysis in occupational health studies is still limited in Korea without a suitable external comparison group such as credible vital statistics for the whole nation. Internal comparisons between different exposure status often result in unstable estimates of effect, and proportional morbidity study is discussed as an alternative potential research tool.

2. Introduction

A study was conducted in a steel roll manufacturing foundry where cases of silicosis had been previously reported 15 years prior to this report (Straub and Rostrand 1976). The specific objectives of the study were 1) to assess the exposure status of workers in relation to promulgated standards; and 2) to examine the respiratory health status of workers and to determine the causes of their health impairment.

Based on the presentation and comments received during the Symposium of Health Statistics, this paper reports statistical analysis strategies in answering specific questions posed in the study. The problems inherent in the analysis of occupational health data and specific limitations of analysis in Korea are also presented.

1) (110-460) 서울시 종로구 연건동 28 서울대학교 보건대학원 교수

3. Methods

The study consisted of a self administered questionnaire, spirometry, and chest x-ray examination as well as environmental measurements.

The questionnaire was adopted from the British MRC (Medical Research Council) questionnaire and IUAT (International Union Against Tuberculosis) questionnaire. Respiratory symptoms including cough, phlegm, and wheezing were investigated along with questions about previous medical and family histories and occupational exposures. Smoking habits, hobbies and other personal and residential risk factors for respiratory illness were also asked.

Lung function testing was done according to 1987 American Thoracic Society Guidelines. Workers performed the tests at least five times on an Ohio Medical Model 827 dry seal spirometer attached to an HF4 dedicated computer. The flow-volume curve data were stored on magnetic tape and later retrieved to calculate Forced Vital Capacity (FVC), Forced Expiratory Volume in 1 sec (FEV1), ratio of FEV1 over FVC (FEV1/FVC), Forced Expiratory Flow at 25-75% of Vital Capacity (FEF 25-75%), and Peak Expiratory Flow Rate (PEFR).

Chest x-rays were taken on a full-size (11x4 inch) film and read independently by two NIOSH-certified B-readers for any presence of reticulo-nodular changes using the 1980 International Labor Organization (ILO) classification system. All the films which were categorized as having any pneumoconiotic changes by either of the two B-readers were re-read by a third B-reader.

Environmental samples were collected during typical working shifts (day, evening, night) in the foundry in an attempt to evaluate the workers' exposures to total particulate, respirable silica, metals, isopropyl alcohol, carbon monoxide and sulfur dioxide. Personal breathing zone samples were obtained on a number of foundry occupations, including molding, furnace tending, and maintenance. Area samples were also placed throughout the foundry near jobs thought to have excessive exposures.

4. Analysis

1. Compliance determination

All air samples were analyzed for respirable dusts and total respirable crystalline free silica (alpha quartz, tridymite and cristobalite). Respirable dust content was analyzed gravimetrically according to NIOSH Method 0500, and respirable crystalline silica dust content was analyzed by NIOSH Method 7500 using X-ray diffraction. Samples for the estimation of sulfur dioxide were collected using Dräger long-term diffusion indicator tubes.

The range of measurements of each airborne contaminant (f) was examined to determine the compliance with workplace exposure standard (Std) after considering the measurement errors (CVt: coefficient of variation in repeated measurements) as below.

If maximum of $f/Std < 1 - 1.645\ CVt$ then compliant,
otherwise noncompliant. -------- (Strategy 1)

Environmental measurements are variable over time (t) and area(a). When multiple measurements are made, probability of its maximum being greater than a certain fixed value increase with repeated measurements and the above logic is not practical (Rappaport 1984). More realistic approach will be determining probabilities (P) as follows.

If P($f_{t,a}$/Std > 1) < 0.05 then compliant,
otherwise noncompliant. -------- (Strategy 2)

The distribution of environmental measurements was then examined with histogram and goodness of fit test was done. Two parameters, median and mean, were examined with t-test to determine the statistical significance of difference from exposure limits.

2. Health outcome quantification

Chronic bronchitis and asthma were defined based on the answers to the symptom questionnaire: chronic bronchitis if chronic cough and/or chronic phlegm production have been present for at least 3 months a year for 2 or more years; asthma if any symptoms of wheezing, attacks of shortness of breath, chest tightness, or nocturnal cough/breathlessness are present together with an indication of repeated or continuous breathing trouble as opposed to never or rarely having trouble.

Predicted lung function values based on race, age and height for a normal unexposed population were calculated (Knudson et al. 1983). For each lung function test, the 95 percentile lower limit of normal is the value below which less than 5% of normal subjects show results. The lower limit of normal values from Knudson et al. (1983) were applied to determine abnormal lung functions as follows: restrictive changes if the FVC is below the lower limit of normal, but the ratio of FEV_1/FVC is within normal; obstructive changes if the ratio of FEV_1/FVC is below normal, but the FVC is within normal; combined changes if both FVC and the ratio of FEV_1/FVC are below normal limits.

According to the 1980 ILO classification, the profusion of pneumoconiotic changes of the lung is graded on a 12 point scale, from 0/- to 3/3. The presence of pneumoconiosis on chest x-ray was determined if two or more readers out of three reported the presence of reticulo-nodular opacities with profusion 1/0 or greater on a 12-point scale. Report of reticulo-nodular changes with profusion 0/1 was considered as minimal pneumoconiotic change. The minimal pneumoconiotic changes were included in the analysis if reported by two or more B-readers.

3. Risk factor quantification

Tobacco use was quantified as pack-years. This was calculated for each worker by multiplying the average packs of cigarettes smoked per day by the number of years of active smoking. For non-smokers, zero pack-years was assigned.

The mean of environmental measurements was calculated for each job, and assigned as a measured exposure level for each worker. Cumulative occupational exposure was estimated from the tenure of employment.

The exposure level of air contaminants varies with repeated measurements and the measured value retains measurement errors. In this study, environmental measurement and health outcome assessment were carried independently, and the measurement errors should be random. To avoid the diluting effect of random measurement errors on the exposure outcome relationship, every job title was classified into high or low exposure categories for both silica dust and SO_2 gas by industrial hygienists based on their knowledge of work process at this particular study site and published findings of exposures in typical foundries (Karava et al. 1976, NIOSH 1985, Landrigan et al. 1986). These a priori classifications were used as exposure status when measured exposure level showed significant relationships with health outcomes.

4. Statistical analysis

Lung function test results were separately analyzed using multiple linear regression to test the association with age, height, smoking, tenure, and measured exposure level (or a priori occupational exposure status) of silica and SO_2 .

Categorical health outcomes (respiratory symptoms, abnormal lung function patterns and the presence of pneumoconiotic changes on chest x-ray) were analyzed for their relation to the measured exposure level and tenure using t-test. Other categorical risk factors were examined in 2×2 tables, and the chi-square statistic was calculated. Other risk factors examined include smoking status, history of allergy in the worker or his family, history of other respiratory diseases, such as tuberculosis or pneumonia, in the worker or his family, type of residential heating source (classified into vented versus unvented), and presence of any dusty or hazardous hobbies. Those statistically important (p value less than 0.05) risk factors were further examined in stratified tables, and a Mantel -Haenszel test was done if necessary.

Step-wise logistic regression analysis was also done to assess the effects of risk factors, including a priori occupational exposure status (high versus low exposure to silica and SO_2), on the categorical health outcomes while controlling the effects of competing risk factors simultaneously. The pack-years of smoking, tenure and age were examined together with other risk factors in this logistic regression analysis. The results of both 2×2 table and logistic regression analyses for risk factors gave similar results and they are discussed together.

5. Result

Study Subjects

Among a total of 183 eligible workers, 129 (70.5%) participated. When company records were examined, there were no significant differences between participants and non-participants in age, tenure, and race (Table 1). Mean age of the participants and non-participants was 42.6 and 42.7 in years, tenure was 16.6 and 16.7 years respectively, and 92% of the participants and 94% of the non-participants were white. All the workers were male. Among the participants, there were 32 (25%) non-smokers, 41 (32%) ex-smokers, and 55 (43%) current smokers. Current smokers had an average of 31 pack-years of smoking, ranging up to 96 pack-years.

Table 1.
Characteristics of Participants and Non-Participants

	Participants	Non-Participants
Number	129 (70.5%)	54 (29.5%) (6: injured 5: vacation 15: laid off)
Age	42.6±8.0	42.7±8.4
Tenure	16.6±5.8	16.7±6.9
Race White Black	 119 (92%) 10 (8%)	 51 (94%) 3 (6%)

Exposure status and compliance assessment

The environmental results indicated that the primary health hazard to workers at the foundry was exposure to free crystalline silica. To summarize, 13 (23%) of all (personal and area) samples collected exceeded the Permissible Exposure Limit (PEL) of 100 $\mu g/m^3$ for respirable quartz. The overall mean value of quartz in the personal samples was 106 $\mu g/m^3$ with maximum of 280.1 $\mu g/m^3$. Three samples of sulfur dioxide exceeded PEL of 2 ppm Time Weighted Average (TWA) exposure. Concentrations of sulfur dioxide ranged from 0.40 to slightly above 7 ppm.

The overall precision as expressed by the coefficient of variation in sampling and analysis of quartz in a typical laboratory is less than 10%. The maximum quartz concentration of 280.1 $\mu g/m^3$ divided by the exposure limit of 100 $\mu g/m^3$ is over (1-1.645*0.1), and non-compliance could be declared following the strategy 1. For sulfur dioxide, the maximum is also far over the exposure limit.

The distributions of quartz measurements and their log-transformed values were examined with Shapiro-Wilk statistic for normality. Each distribution was significantly different from the normal distribution ($p < 0.05$). When area and personal measurements were treated separately, only the distribution of personal measurements was not significantly different from normal distribution. When the mean value of personal exposure measurements was compared to the exposure limit with t-test, the mean was not significantly lower than the permissible exposure limit, and non-compliance could be declared also following the strategy 2.

Respiratory health and its relation to risk factors

In the multiple linear regression analysis of lung function tests, age, height and pack-years of smoking explained most of the findings (Table 3). Tenure and current occupational exposure levels were not significantly associated with the lung function test results in this analysis.

Table 2.
Exposure Measurements

	Sample Type	No of Measurements	Range	Mean ± SD
Respirable Dust	Personal Area	18 38	0.1-1.3mg/m3 0.01-2.8mg/m3	0.46±0.34 0.30±0.21
Quartz	Personal Area	18 38	ND-280μg/m3 ND-199μg/m3	106±79 37±49
Sulfur Dioxide	Personal Area	5 23	1.1-3.6ppm 0.3-7.1ppm	1.6±1.1 1.3±1.4
Isopropyl Alcohol	Personal Area	1 6	708ppm 18-136ppm	 61.2±40
Aluminum	Personal Area	10 25	0-0.07mg/m3 0-0.07mg/m3	0.02±0.02 0.01±0.01
Calcium	Personal Area	10 25	0-0.21mg/m3 0-0.08mg/m3	0.07±0.06 0.02±0.02
Chromium	Personal Area	10 25	0-0.04mg/m3 0-0.02mg/m3	0.01±0.01 0.002
Copper	Personal Area	10 25	0-0.01mg/m3 0-0.01mg/m3	0.003 0.0005
Iron	Personal Area	10 25	0.1-2.15mg/m3 0-1.56mg/m3	0.69±0.62 0.24±0.35
Magnesium	Personal Area	10 25	0-0.08mg/m3 0-0.21mg/m3	0.03±0.03 0.03±0.05
Magnanese	Personal Area	10 25	0-0.12mg/m3 0-0.04mg/m3	0.03±0.04 0.007
Molybedenum	Personal Area	10 25	0-0.02mg/m3 0-0.15mg/m3	0.005 0.007
Sodium	Personal Area	10 25	0-0.09mg/m3 0-0.06mg/m3	0.03±0.03 0.009
Nickel	Personal Area	10 25	0-0.05mg/m3 0-0.04mg/m3	0.01±0.01 0.006
Phosphorus	Personal Area	10 25	0-0.02mg/m3 0	0.002
Lead	Personal Area	10 25	0-0.02mg/m3 0-0.21mg/m3	0.007 0.01±0.04

Table 3.
Multiple Linear Regression Analysis
of the Lung Function Test Results.

Coefficients of Significant ($p<0.05$) Explanatory Variables

	COEFFICIENTS					
	Age (/year)	Height (/inch)	Pack-Year (/pack-year)	Current Smoking	Intercept	R*
FVC	-0.0215	0.0199	-0.0060		-7.6535	0.75
FEV1	-0.0219	0.1345	-0.0110		-4.2647	0.71
FEV1/FVC			-0.1655		79.3663	0.46
FEF 25-75%	-0.0259	0.0746	-0.0255		-0.1667	0.54
PEFR		0.2267		-0.8249		0.43

* R value of linear regression

As mentioned, the 2×2 table and logistic regression analyses of health outcomes and risk factors showed the same result. Risk factors with statistical significance are summarized in Table 4. Silica dust exposure categories significantly explained chronic phlegm, chronic bronchitis and especially hemoptysis. The current smoking status explained significantly the occurrence of chronic cough and chronic phlegm. While current smokers had a lower risk of wheezing, increase in tenure was associated with more wheezing symptoms. Asthma was associated with past medical history of allergy, suggesting the importance of atopic status in association with asthmatic symptoms. Neither silica nor SO2 exposure categories significantly explained the presence of asthmatic symptoms as defined in this study. The occurrence of obstructive lung function changes was most significantly explained by pack-years of smoking, but occupational exposures, including tenure, were not significant. When the presence of suggestive pneumoconiosis on x-rays including minimal pneumoconiotic changes was examined in the logistic regression analyses for the association with risk factors including age, tenure, exposure level to silica dust and SO_2 gas, smoking status, and pack-years, only job tenure was significantly associated with the radiologic changes. Other risk factors were not significant. Among workers in this study, every 10 years of tenure at the foundry was associated with a 1.8 fold and 2.7 fold increased risk of having wheezing symptoms and pneumoconiotic changes on x-rays respectively. The prevalence of suggestive pneumoconiotic findings was plotted for different tenure categories to further examine the relationship with tenure (Figure 1).

Table 4.
Logistic Regression Analysis
for the Occurrence of Respiratory Symptoms,
Abnormal Lung Function Patterns, and X-ray Findings

Respiratory Symptoms	Risk Factors*	Odds Ratio
chronic cough for more than 2 years	current smoker	4.5
chronic phlegm for more than 2 years	current smoker	4.5
	silica dust	3.1
	family history of resp dis	3.1
chrònic bronchitis	current smoker	4.4
	silica dust	2.4
hemoptysis	silica dust	13.4
	family history of resp dis	4.8
chest tightness or difficult breathing	family history of resp dis	2.7
wheezing	current smoker	0.3
	tenure	#
(#: parameter estimate in logistic regression - 0.0635/year)		
asthma	past history of allergy	2.7

Abnormal Lung Function Patterns	Risk Factors*	
obstructive lung function	pack-year of smoking	#
(#: parameter estimate in logistic regression - 0.0245/pack-year)		

X-ray findings	Risk Factors*	
Pneumoconiotic changes	tenure	#
(#: parameter estimate in logistic regression - 0.1003/year)		

* Risk factors of statistical significance with p value less than 0.05

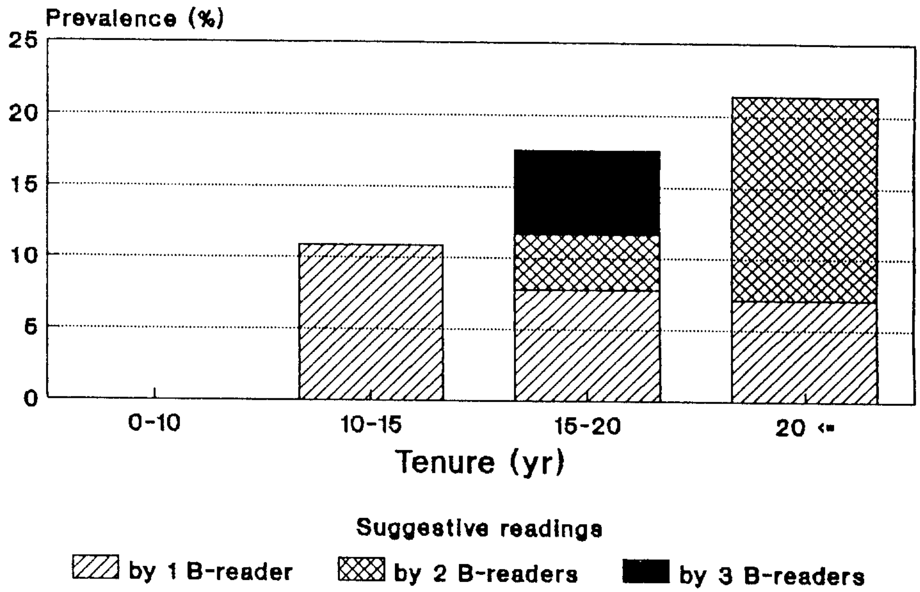

Figure 1.
Prevalence of Suggestive Pneumoconiotic Changes
on X-ray for Different Tenures

6. Discussion

Distribution of exposure determination

Compliance determination is one of the major area of occupational health study. However, few, often only one, measurements are made in Korea and either the mean or maximum value is directly compared to the exposure limits without considering variabilities inherent in environmental measurements. The decision logic can be expressed as follows.

If maximum or mean of f/Std < 1 then compliant,
otherwise noncompliant.

Variabilities in measurements stem from two major sources, measurement errors, including those in the analysis of samples, and changes in air contaminant concentration with time and space. When measurement and analysis errors are considered (CVt, coefficient of variation in repeated measurements), the above expression becomes the one in strategy 1.

If maximum or mean of f/Std < 1 - 1.645 CVt then
compliant, otherwise noncompliant.

The magnitude of error in measurements as expressed by the coefficient of variation in repeated measurements is usually less than 10% (Leidel et al. 1977). When compared to the variability with time and space which may be more than 10 folds, the measurement errors are small and can be neglected as in the following expression of strategy 2.

If P(f/Std > 1) < 0.05 then compliant,
otherwise noncompliant.

To determine the probability in the above expression, one should know the distribution of the underlying population. Environmental measurements are often described as log-normally distributed (First 1989). They should be greater than 0, but the maximum is not bounded by any condition. The log-normal distribution has been demonstrated when multiple measurements are made in a day for one worker, when daily averages are determined over a period for one worker, or when daily average is determined for group of workers.

In this study, the distribution of environmental measurements conformed to neither normal nor log-normal distribution. Small sample size and non-random sampling such as selective sampling of potentially dusty jobs and areas may have contributed the problem, which is a common practice of industrial hygiene measurements. Non-parametric test, such as signed rank test, can be employed to determine the probability, but low efficiency will require a larger sample size than parametric tests (Hawkins et al. 1991).

Under the log-normal distribution, one way of determining the probability is to determine whether the geometric mean is significantly lower than the exposure limit. Even though this approach has an intuitive strength of being statistically appropriate, geometric mean of exposures may not bear any biological (or toxicological) relevances.

When the exposure response relationship can be expressed as a linear function, the parameter that should be used to indicate the response probability is the mean, not the median of exposure measurements. This is especially true for chronic diseases such as pneumoconiosis as in this study, for which exposure variability dampens inside the body and should be integrated over time to reflect the cumulating response (Roach 1966, Rappaport and Spear 1988). In this regard, another approach is to test whether the mean, not the geometric mean, is significantly lower than the exposure limit as in the following expressions (Rappaport and Selvin 1987).

$$\overline{x}_1 = \sum \ln(x_i)/n\ , \qquad s_1^2 = \sum(\ln(x_i) - \overline{x}_1)^2/(n-1)$$

$$\overline{x}_c = \exp[\ \overline{x}_1 + 0.5 s_1^2]\ , \qquad s^2_{\overline{x}_c} = (PEL)^2 (s_1^2 + 0.5 s_1^4)/(n-2)$$

$$p(t|t \geq (\ \overline{x}_1 - \ln(PEL))/s_1)$$

$$p(t|t \geq (\ \overline{x}_C - PEL)/s_{\overline{x}_c})$$

Exposure measurement errors and strategies for the control

In the exposure measurements, a group of errors can arise from calibration of sample collecting device, measurement of collection time, sample analysis and calculations among others. On the other hand, the level of air contaminants fluctuates with time and place, and exposure level measured in the breathing zone of a worker varies accordingly. These variations together result in errors in the estimation of true exposure and consequently dilution of exposure response relationships.

Different approaches have been utilized to combat these diluting effects. One of the approaches, which was adopted in this study, is to group continuous exposure values into several categories, employing a priori categorical classification of exposure levels which should not be biased by the exposure measurement itself. If the a priori categories are free from bias, the errors in the measurements are random and the average value for each category should be close to the true mean. The dose response relationships based on these a priori categories will be steeper than those based on all observed measurements, but the number of observations becomes smaller resulting in much less efficiency.

Another approach is using the analysis of variance technique to separate the within subject from the between subject variance and to evaluate the potential bias in calculating the slope of association with the observed exposure measurements. If more than one measurement are available per subject, the relative magnitude of intraindividual and interindividual variability can be estimated (Liu et al. 1978). The intraindividual variance reflects the error variance in a subject and interindividual variance reflects the true variance between subjects, and the variance ratio (λ) can be used to assess the magnitude of the bias in regression and correlation coefficients, with the following equations:

$$b = B\ (1 + \lambda\ )^{-1}$$
$$r = R\ (1 + \lambda\ )^{-0.5}$$

where b and B are observed and true regression coefficients and r and R are observed and true correlation coefficients, respectively.

Confounding factors in occupational health studies and strategies for the control

Occupational health studies which compare workers with the general population often show results of better health status among workers even under hazardous exposures. On other occasions, workers with poor health are often removed to unexposed jobs, while workers with good health are continuously exposed, resulting in poorer health status in lower (cumulative) exposure groups. 'Healthy worker effect' has been named to encompass such study results as above.

In this study, when the lung function of nonsmoking foundry workers was compared to the expected values, there was a significant difference. However, in the internal comparisons among the foundry workers, we did not find any association with different jobs. Neither the current exposures to dust or gases nor tenure was significant. The same finding of no association was reported in other studies of foundry workers. (Karava et al.

1976, Landrigan et al. 1986, Myers et al. 1987). This may be related to the cross-sectional design of the studies and the fact that only the healthiest workers can work continuously in adverse conditions, so called 'healthy worker effect'. It has been reported that workers who left a foundry had more health problems and more bronchitis and diagnosed lung diseases than current workers (Karava et al. 1976).

If exposure is a risk factor for poor health, and change of employment status due to poor health is a risk factor for further poor health, then change of employment status is an intermediate risk factor. Because change of employment status is associated with further exposure, it is also a confounder on the exposure-disease relationship.

The traditional approach of stratifying with the confounding variable will result in lowering the risk of poor health from exposure in the case of healthy worker effect mediated through employment status, because it will also result in controlling the intermediate variable even though we are interested in the overall risk of those who were exposed compared to never-exposed. Analysis strategy of controlling time-dependent covariate which is simultaneously a confounder and an intermediate variable on the causal pathway from exposure to disease has been proposed for the longitudinal studies in which data on exposure, covariate, and disease outcome are obtained at several points in time (Robins 1987). The prerequisite for this analysis is that the time interval between exposure measurements has to be sufficiently short that the number of subjects who change their exposure status between measurements is negligible, which may not be satisfied in many occupational health studies.

Health outcome - Standardized mortality ratio (SMR) and proportionate mortality ratio (PMR) analysis

When disease incidence rate is known for the control population, the rate of disease among study subjects can be compared after standardizing the stratum specific characteristics of control and study population. The standardized mortality ratio is the ratio of observed number of diseases over expected number that is calculated by applying the stratum specific rate among general population to the number of subjects in the specific stratum of study population. When the control population is large as in the national vital statistics, the expected number is very stable. Assuming the denominator to be a constant, an approximate confidence interval can be estimated for the SMR. The formula for the lower and upper limits of an SMR can be given as below (Checkoway et al. 1989).

$$\underline{SMR} = Obs[1 - 1/(9Obs) - (Z/3)(\sqrt{1/Obs})]^3/Exp$$

$$\overline{SMR} = (Obs+1)[1 - 1/9(Obs+1) + (Z/3)(\sqrt{1/(Obs+1)})]^3/Exp$$

The above SMR calculation has been utilized in many occupational studies to explore the potential risks of occupational exposures by characterizing the disease occurrence in specific occupations. However, it is impossible without having a large stable base population statistics, which is not yet available in Korea. On the other hand, information is often lacking about the characteristics of the base population from which the study population originate, and the denominator for the rate may not be available as in the data of compensated occupational cases or cancer registry. Under these circumstances, only the proportion of specific disease of the study group can be compared to the proportion of corresponding disease of the total population.

In Korea, workers' health examination program routinely collects data on easily identifiable diseases such as viral hepatitis, diabetes, and hypertension among general population, as well as specific occupational diseases such as pneumoconiosis, noise-induced hearing loss, and lead poisoning among specifically exposed population. Even in specific medical examinations, all workers exposed to occupational hazards are included, and total population covered entails wide spectrum of exposures, from dangerously high level to negligibly safe level. Therefore, occurrence of a certain disease or laboratory test results in a specific occupation or workplace can be assessed by computing the ratio of proportion of outcome of interest among study subjects over the one among total covered subjects. When the observed and expected numbers of diseases are greater than five, then Mantel-Haenszel χ^2 test can be employed to test the significance. Otherwise statistical tests based on Poisson distribution can be employed with approximation that the proportion of disease of interest is small in each stratum (O'Brien et al. 1991). If this approximation is inappropriate, p-values and confidence limits can be calculated using the noncentral hypergeometric distribution.

Reference

[1] Chekoway H., Pearce N. E., Crawford-Brown D. J. (1989) *Research methods in occupational epidemiology.* Oxford University Press, Oxford.

[2] First M. W. (1989), "Air sampling and analysis for contaminants: an overview." in Hering S. V. Ed. *Air sampling instruments* American Conference of Governmental Industrial Hygienists, Ohio.

[3] Hawkins N. C, Norwood S. K, Rock J. C. (1991), *A strategy for occupational exposure assessment.* American Industrial Hygiene Association, Ohio.

[4] Johnson A, Chan-Yeung M, Maclean L, Atkins E, Dybuncio A, Cheng F, Enarson D. (1985), "Respiratory Abnormalities among Workers in an Iron and Steel Foundry.," *British Journal of Industrial Medicine* 42(2), 94-100.

[5] Karava R, Hernberg S, Koskela R, Luoma K. T. I. (1976), "Prevalence of Pneumoconiosis and Chronic Bronchitis in Foundry Workers," *Scandinavian Journal of Work, Environment and Health* 2(1), 64-72.

[6] Knudson R. J, Lebowitz M. D, Holberg C. J, Burrow B, (1983), "Changes in the Normal Maximal Expiratory Flow-Volume Curve with Growth and Aging," *American Review of Respiratory Disease,* 127, 725-34.

[7] Landrigan P. J, Cherniack M. G, Lewis F. A, Catlett L. R, Hornung R. W. (1986), "Silicosis in a Grey Iron Foundry. The Persistence of an Ancient Disease," *Scandinavian Journal of Work, Environment and Health* 12, 32-39.

[8] Leidel N. A, Bush K. A, Lynch J. R. (1977), *Occupational exposure sampling strategy manual,* National Institute for Occupational Safety and Health, (DHEW (NIOSH) Publication No 77-173), Ohio.

[9] Liu K, Stamler J, Dyer A, McKeever J, McKeever P. (1978), "Statistical methods to

assess and minimize the role of intra-individual variability in obscuring the relationship between dietary lipids and serum cholesterol," *Journal of Chronic Disease,* **31**, 399-418.

[10] Myers J. E, Garisch D, Myers H. S, Cornell J. E. (1987), "A Respiratory Epidemiological Survey of Workers in a Small South African Foundry," *American Journal of Industrial Medicine,* 12, 1-9.

[11] National Institute for Occupational Safety and Health. (1985), *Recommendations for Control of Occupational Safety and Health,* (DHHS (NIOSH) Publication No. 85-116) Ohio.

[12] O'Brien T. R, Flanders W. D, Decoufle. (1991), "Use of the Mantel-Haenszel χ^2 overestimates precision in studies with sparse data," *Journal of Occupational Medicine,* **33**, 1081-83.

[13] Rappaport S. M. (1984), "The rules of the game: an analysis of OSHA's enforcement strategy," *American Journal of Industrial Medicine,* **6**, 291-303.

[14] Rappaport S. M, Selvein S. (1987), "A method for evaluating the mean exposures from a lognormal distribution," *Journal of the American Industrial Hygiene Association,* 48(4), 374-9.

[15] Rappaport S. M, Spear R. C. (1988), "Physiologic damping of exposure variability during brief periods," *Annals of Occupational Hygiene,* **32**, 21-33.

[16] Roach S. A. (1966), "A more rational basis for air sampling programs," *Journal of the American Industrial Hygiene Association,* **27**, 1-12.

[17] Robins J. (1987), "A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods," *Journal of Chronic Disease,* **40**(suppl 2), 139s-161s.

[18] Straub W. E, Rostrand R. (1976), *Health hazard evaluation determination* report 74-97-286. US DHEW CDC NIOSH.

산업보건연구에서의 통계학적 분석

백도명, 최 정 근, 손 미 아

요 약

산업보건연구의 통계학적 분석은 작업환경측정에 대한 평가방법과 산업보건에 특유한 혼란변수의 통제 등 건강상태와의 연관성을 분석하는 평가방법에 있어 아른 보건통계분야와 구별되는 특성을 지니고 있다. 본 논문에서는 주물공장 근로자들의 호흡기 건강상태와 작업환경에서 폭로되는 유해물질에 대한 조사를 통하여 산업보건연구에 사용되는 통계학적 분석에 대한 기술을 하였다.

조사된 환경측정결과의 일부는 허용폭로기준을 초과하고 있었는 바, 폭로기준의 준수여부를 판정하기 위하여 작업환경측정결과와 그들의 대수변환치들로 부터 얻은 산술평균과 대수평균들이 폭로기준과 다른지에 대한 t-검정을 실시하였다.환경측정을 비롯한 위험요인들과 그들로 인한 건강상태와의 관계 분석을 위해, 범주적 건강측정 변수인 경우에는 χ-square 검정과 다변량 logistic분석을 시행하였고, 연속적 변수인 경우에는 다변량 회귀분석을 시행하였다.

작업환경내의 오염물질의 농도는 그 측정장소와 측정시점에 따라 매우 가변적이다. 이러한 작업환경의 측정결과를 평가하는데 있어 사용될 수 있는 서로 다른 여러 지수들의 장단점과 가변적 측정결과들로 인한 오차를 보정할 수 있는 통계학적 분석 방법에 대한 논의를 하였다. 본 조사의 폐기능검사 결과는 직업에서 폭로되는 정도와 아무 연관성을 보이고 있지 않은 바, 이는 "건강한 근로자 효과"에 기인한 것으로 추정되고 있다. 이러한 "건강한 근로자 효과"를 비롯한 측정시점에 따라 변화하는 혼란변수를 보정하기 위한 통계적 분석 방법이 논의 되었다.

산업보건 연구에서 기본적으로 사용될 수 있는 전국적 생정통계와 같이 광범위하며 쉽게 비교되는 외부 대조군 내지는 질병의 예측기대치에 대한 통계가 한국의 경우에는 아직 없다. 이러한 경우, 그 분석의 범위가 매우 제한되어 있으나, 문제되는 질병 내지는 임상검사결과가 다른 일반적 질환과 비교하여 차지하는 비율을 서로 다른 집단간에 비교함으로써 서로 다른 작업환경에 폭로되는 집단들에 대한 통계학적 분석을 시도할 수 있다. 현재 한국에서는 일반검진과 특수검진이 정규적으로 실시되고 있으며, 이러한 검진결과에 대한 체계적 통계학적 분석이 앞으로 필요할 것으로 기대된다.