Statistical derivation of a standardised procedure for counting expectorated alveolar macrophages as indicators of occupational air pollution

A M NILSEN1 AND S ENGEN2

From the Department of Botany, Section for Cell Biology,1 and Department of Mathematic and Statistics,2 University of Trondheim, N-7055 Dragvoll, Norway

ABSTRACT A biological test in which counts of expectorated alveolar macrophages are used to evaluate lung irritation attributable to occupational air pollution has been applied to workers exposed to dust and gas pollution. To determine the optimum allocation of resources when performing the test, the sampling and counting procedures have been evaluated statistically. Alveolar macrophages (AM) were recovered from workers at an iron works, an aluminium plant (one large and one small group), and from a small group of non-exposed smokers; the number of subjects was 213. Sampling was repeated with the small group at the aluminium plant (27 subjects) on three consecutive days and the small group of non-exposed smokers on five (3 subjects). AM were counted from smear slides using a light microscope. A standard random effects model was used as the basis for estimating the expectation of the log-transformed overall means (μ̂) and the variances (var(μ̂)) of the two categories at the aluminium plant and of the non-exposed smokers. The costs connected with finding μ̂ consist of the financial expenses and the working efforts and time used to perform the test. The precision and reproducibility of the test are closely related to the variability in the results, var(μ̂). To optimise the costs but still obtain reproducible results, AM should be counted in four drops of expectorate from each of three samples from at least 10 to 20 subjects.

The analysis of air filter samples has been used routinely to describe the degree of dust exposure in polluted industries, although many factors make it difficult to express human exposure on such a basis. Our aim has been to investigate alveolar macrophages (AM) from samples of expectorate in order to evaluate the degree of lung irritation in workers exposed to occupational air pollution12 (AM Nilsen et al, unpublished data).

The alveolar macrophage is the resident mononuclear phagocyte of the lung and functions as the primary defence against inhaled particulate matter.3-5 The number of AM in the respiratory tract varies with the quantity and the quality of the inhaled dust.6-7 Specific sampling and counting procedures have been examined in order to design a routine test system in which the number of AM is the main parameter determined. The rationale behind the biological test has been described in other papers covering investigations of different industries. During these studies we discovered the need for a statistically based standardised procedure for counting AM. On the basis of a standard random effects analysis of a variance model and a cost function, we have been able to determine the counting procedure that minimises the cost for a given precision in the estimates. The solution requires that the variances of the random components are known. Estimates of these variances are given. We also suggest a counting procedure for the less than optimal cases when the number of available subjects is 10 to 20. A standardised procedure for counting AM in such small groups according to the underlying statistical model is presented.

Materials and methods

Samples of expectorate were obtained from 49 workers at an iron works, two groups with 134 and 27 workers at an aluminium plant, and three non-

Received 8 December 1983
Accepted 2 July 1984
exposed smokers. Sampling was repeated with the small group at the aluminium plant on three consecutive days and the non-exposed smokers on five.

Smear slides were made from drops of expectorate according to the method earlier suggested by Mylius and Gullvåg. Two slides were made from each drop. AM from workers at the iron works were counted using a light microscope in parallel lines of field of vision in the central region across each slide. Counting of AM from workers at the aluminium plant and the non-exposed smokers was made in only one field of vision across each slide. The objective lens magnification was 40 times and the total magnification 500 times. The method is shown schematically in fig 1.

Results and discussion

The results of the counts are presented and discussed elsewhere. Briefly summarised they show an increasing number of AM in the expectorate in parallel with an increase in the sum of inhaled pollutants. Both smoking and air pollution at the workplace add to the AM response.

THE MODEL

Our work shows that the standard deviations of repeated samples (expectorate samples and drops) are roughly proportional to the mean. Hence a log transformation is suitable for stabilising the variances. Let $X_{ijk}$ denote the count for subject $i$, sample $j$, and drop $k$, and write $Y_{ijk} = \ln(X_{ijk} + 1)$. We then consider a standard random effects model (see for example Scheffé and Snedecor and Cochran).

$$Y_{ijk} = \mu + \alpha_i + \beta_{ij} + \gamma_{ijk},$$

where $\mu$ is a constant, $\alpha_i = E\beta_{ij} = EY_{ijk} = 0$, $\text{var}(\alpha_i) = \sigma^2_{\alpha}$, $\text{var}(\beta_{ij}) = \sigma^2_\beta$ and $\text{var}(\gamma_{ijk}) = \sigma^2_\gamma$. Our aim is to estimate the expectation $\mu = E(Y_{ijk})$ with maximum precision. In the model the $\alpha_i$ represent the variation in counts among subjects, $\beta_{ij}$ the variation among samples from a given subject, and $\gamma_{ijk}$ the variation in drops.

ESTIMATION

The overall mean $\mu$ is estimated by

$$\hat{\mu} = \frac{1}{lmn} \sum_{i=1}^{l} \sum_{j=1}^{m} \sum_{k=1}^{n} Y_{ijk} = \bar{Y},$$

the variance of which is

$$\text{var}(\hat{\mu}) = \frac{1}{lmn} \left( \frac{1}{\text{var}(\alpha_i)} + \frac{1}{\text{var}(\beta_{ij})} + \frac{1}{\text{var}(\gamma_{ijk})} \right)$$

Estimates of the variances are obtained by standard mean square methods. Confidence limits can be derived for $\sigma^2_\gamma$ and $\sigma^2_\beta$ by standard application of the $\chi^2$ and the F-distributions. Intervals for $\sigma^2_\alpha$ and $\sigma^2_\gamma$ cannot be derived by standard methods. We have checked by rough calculations that the conclusions of the paper are not affected by the uncertainty of the estimators.

The table shows $n$, $m$, $l$, $\bar{\sigma}^2_\alpha$, $\bar{\sigma}^2_\beta$ and $\bar{\sigma}^2_\gamma$ for the three categories investigated. Hence $\text{var}(\hat{\mu})$ may be estimated for any choice of $(l, m, n)$ by replacing the variances in equation (3) by their estimates.

MINIMISATION OF VARIANCES

Let us first discuss the number of lines of field of vision on each slide. In the counts from the iron works samples a high positive correlation was found between the results from the individual lines of field of vision on the same slide (fig 2). The mean correlation coefficient was 0.89, and even between every 25th line the coefficient was 0.88. There is, therefore, no need to count AM in more than one line of field of vision across each slide.

Now let us define cost as

- $c_1$: the cost of getting one subject to take part in the investigation,
- $c_2$: the cost of taking one sample from one subject, and
- $c_3$: the cost of making a drop and AM count by the above method.

Fig 1 Diagram to show method for counting AM from samples of expectorate.
Then the total cost in the procedure of finding \( \mu \) is

\[
c = c_1 l + c_2 l m + c_3 l m n
\]  

(4)

The cost is minimised for values

\[
m^* = \frac{\sigma_\beta}{\sigma_\alpha} \sqrt{\frac{c_1}{c_2}}
\]  

(5)

\[
n^* = \frac{\sigma_\gamma}{\sigma_\beta} \sqrt{\frac{c_2}{c_3}}
\]  

(6)

and with \( l^* \) chosen to give the required total cost (4) (see Snedecor & Cochran pp 531-4*). Then the solution to the optimisation problem is positive integers in the neighbourhood of \((l^*, m^*, n^*)\). Note that only the ratios \( c_1/c_2 \) and \( c_2/c_3 \) are required to find \( m^* \) and \( n^* \).

In our investigation of different types of industry we usually find, however, that \( c_1 \) is small compared with \( c_2 \) and \( c_3 \). From experience with this type of work we will estimate the ratios \( c_1/c_2 \) and \( c_2/c_3 \) to be roughly of the order 1:5:50. Hence the optimal procedure is \( n^* = m^* = 1 \) according to (5) and (6). Since, however, \( l^* \) in practice is limited, \( l^* \leq l_0 \), and \( l_0 \) is of the order of 10 to 20 subjects. Because the workers are divided into groups according to worksites, smoking habits, and use of personal protection equipment, \( n \) and \( m \) must be chosen to be larger than 1 to obtain the precision required. The value of \( n \) should be chosen to be equal to 1 by (6) and \( m \) large enough to make \( \text{var}(\mu) \) (equation 3) small enough. But because the workers are not available more than three to four consecutive days owing to workshifts, \( m \) cannot be chosen to be larger than 3 to 4. In these cases a sufficient precision is obtained by choosing \( n = 4 \) (fig 3). The term "sufficient precision" must be considered in relation to the maximum obtainable precision with \( l_0 \) subjects, which is \( \sigma_\alpha/l_0 \). We can see from figs 3 and 4 that a further increase in either \( m \) or \( n \) will be rather inefficient since \( \text{var}(\mu) \) approaches \( \sigma_\alpha/l_0 \) very slowly.

**COMPOSITE SAMPLING: AN ALTERNATIVE METHOD?**

Composite sampling\(^1\) may be used as an alternative to our method if a description of the mean response to an environment alone is needed. The advantage of such a method is that the information about the mean response is increased and the total costs of the investigation are reduced by decreasing the total number of drugs of expectorated while increasing the number of subjects. Mixing the samples from subjects with uniform exposure and picking out drops

<table>
<thead>
<tr>
<th>Categories</th>
<th>( n )</th>
<th>( m )</th>
<th>( l )</th>
<th>( \sigma_\alpha )</th>
<th>( \sigma_\beta )</th>
<th>( \sigma_\gamma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aluminium plant</td>
<td>4</td>
<td>1</td>
<td>134</td>
<td>1-48</td>
<td>0</td>
<td>1-79</td>
</tr>
<tr>
<td>Large group</td>
<td>4</td>
<td>3</td>
<td>27</td>
<td>0-39</td>
<td>1-24</td>
<td>1-35</td>
</tr>
<tr>
<td>Small group</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>0-85</td>
<td>1-02</td>
<td>1-35</td>
</tr>
<tr>
<td>Three smokers</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>0-85</td>
<td>1-02</td>
<td>1-35</td>
</tr>
</tbody>
</table>

*Fig 2  Correlation coefficient between number of AM in two by two lines of field of vision on smear slides.*

*Regression line shows there is no need to counting AM in more than one line across each slide. Neighbouring lines have distance 1. Samples from workers at iron works.*

*Fig 3  \( \text{var}(\mu) \) (3) for the categories shown in table; \( m \) and \( l \) are as in table, and \( n \) varies (indicating that sufficient precision is obtained with four drops).*

- Large group at the aluminium plant.
- Small group at the aluminium plant.
- Three smokers.

*Fig 4  \( \text{var}(\mu) \) (3) for categories shown in table with \( m > 1 \) indicates that an acceptable level of precision is obtained with three repeated samples. \( l \) and \( n \) are as in table whereas \( m \) varies.*
from the bulk sample, however, means that all the
detailed information about the individuals is lost.
Therefore, if the variance is not known in advance, a
total mixing of samples may be disadvantageous, as,
for example, when the results of workers from dif-
f erent working environments need to be compared
with each other or with a control group. Moreover,
the possibility is lost of finding individuals whose
AM values indicate extra sensitivity. The purpose of
the study will determine which method is best
suited.

We thank A/S Norsk Jernverk and Lista
Aluminiumverk for their kind permission to let us
perform this study, and the plant physicians Dr O
Storsæther and Dr H Johannessen for their excellent
help during the collection of samples. We gratefully
acknowledge the practical help of R Vik and M
Denstad. We thank the Ministry of Local Govern-
ment and Labour, the Norwegian Research Council
for Science and the Humanities, the Norwegian
Council for Scientific and Industrial Research for
their financial support, and the Secretariat of
Environmental Studies of the Aluminium Industry,
Elkem Spigerverket A/S, and the Electrochemical
Employers' Association for a grant for attending to
the statistical work for six months.

References

1. Mylius EA, Gullvåg B. Alveolar macrophage count (AM-test) as
indicator of lung reaction to air industrial pollution. Acta Cytol
(in press).
2. Nilsen AM, Madslien O, Mylius EA, Gullvåg BM. Alveolar mac-
rophages from expectorate samples—a stress signal from
occupational pollution. Bull Environ Contam Toxicol
3. Bowden DH. The alveolar macrophage and its role in toxicology.
Defense mechanisms of the respiratory membrane. Am Rev
(suppl):59–72.
6. Brain JD. Free cells in the lungs. Some aspects of their role,
quantitation, and regulation. Arch Intern Med
7. Brain JD. The effects of increased particles on the number of
alveolar macrophages. In: Walton WH, ed. Inhaled particles
III. Proceedings of a symposium. London: Gresham Press,
8. Gullvåg B, Mylius E, Nilsen A. Counting alveolar macrophages
(AM) from expectorate samples of exposed workers as a test
for lung irritation from occupational exposure. Bull Environ
Contam Toxicol (in press).
Sons, 1959.
10. Snedecor GW, Cochran WG. Statistical methods. 6th ed. Ames:
11. Rohde CA. Batch, bulk and composite sampling. In: Cormack
RM, Patil GP, Robson DS, eds. Statistical ecology. Fairland,