Quantifying Lung Structure

Experimental Design and Biologic Variation in Various Models of Lung Injury[1-4]

LING-YI CHANG, ROBERT R. MERCER, KENT E. PINKERTON, and JAMES D. CRAPO

SUMMARY The lung is a complex organ composed of a large number of different cell types of varying size and shape. Quantification of lung structure requires an understanding of how the distribution of specific cells and their characteristics affect the accuracy of measurement made on them and how to optimize experimental design for a morphometric study. We have studied lung structural modifications in a variety of lung injuries over the last decade. Extensive quantitative data from EM morphometric studies of pulmonary tissue have been collected. These data provide a unique opportunity to study the accuracy and efficiency of methods used to quantitate lung structure. We present and discuss novel computation-intensive methods for the estimation of biologic variability, sampling error, and measurement error. A new concept, unnested analysis of variance for stratified sampling and the use of computer-based methods for statistical analysis (the bootstrap method) and optimizing experimental design (nonlinear minimization procedure) are described in this report. Examples of experimental designs with their corresponding levels of accuracy and cost are also provided. The number of samples needed for a given level of precision is affected by the volume density of the structure being measured. The most important determinant for the overall accuracy of a morphometric study is the number of animals studied. Biologic variations between samples within an animal and among animals can vary significantly as a function of the model of injury studied.

AM REV RESPIR DIS 1991; 143:625-634

Introduction

Morphometry is a valuable tool for the assessment of lung structural modifications in pulmonary pathology and toxicology. Quantitative measurements of lesions in terms of thickness of the alveolar wall, size of the alveolus, and numbers of inflammatory cells unambiguously define the nature and the severity of an injury. Morphometry done using electron microscopy (EM) can further define individual cell responses. For instance, increases in the number and thickness and a corresponding decrease in the surface area of type I epithelial cells usually indicates an increase in the transient population of cells differentiating from type II epithelial cells to type I epithelial cells (1-4). When early and critical pathologic events are identified by morphometry, it helps to elucidate the mechanism of disease development (4, 5). Furthermore, EM morphometry is highly sensitive for the detection of small but biologically significant injuries. In recent years it has been used extensively for the evaluation of the effects of environmental pollutants (1-14). It is necessary to achieve a high degree of accuracy to meet the stringent requirements of setting standards for clean air. Given the low ambient levels of most pollutants and the many factors that affect population responses to pollutants, this practice can be very costly because accuracy, from a statistical point of view, is best attained by studying more animals. Whereas a number of new stereologic methods have been described in the past few years that provide unbiased and accurate measurements of biologic structure, many of these techniques require specialized sampling procedures, such as dissectors (15, 16) or vertical sections (17). These procedures cannot be readily applied to certain pathologic samples, for instance, human biopsies. In addition, pathologists interested in evaluating complex disease or toxicant effects are frequently confronted with multiple tissues or cells that may have different characteristics and cannot be studied with the same degree of accuracy using a single sampling strategy. For the majority of pathologic and toxicologic investigations, random sampling remains the most convenient technique. It is essential to understand how to maximize design for this type of study.

The design of experiments to quantify subtle changes in lung structure is difficult because EM morphometry of the lung invariably involves multiple sites (airways and gas-exchange regions) and/or cell types that have different sizes, shapes, orientations, and characteristics of distribution. The purpose of this article is to provide a practical analysis of the common morphometric techniques used to study lung structure. In the past decade our laboratory has accumulated a substantial volume of quantitative data on lung structure from normal and injured animals using EM morphometry (1-5, 7, 18). These data provide a unique opportunity to examine sampling variability and measurement error when quantifying lung structure. We undertook to study the coefficient of variation in the measurement of the volume densities of two alveolar tissues: large, flat, and homogeneously distributed type I epithelium and the small, cuboidal, heterogeneously distributed type II cells. A novel method that estimates the sampling and measurement variability was used. These estimates of the experimental vari-

(*Received in original form July 23, 1990 and in revised form October 5, 1990*)

[1] From the Departments of Medicine and Pathology, Duke University Medical Center, Durham, North Carolina, and the Department of Anatomy, School of Veterinary Medicine, University of California, Davis, California.

[2] Supported in part by Grants PO1 HL 31992, 5 T23 HL 07538, and T32 ES 0703 from the National Institutes of Health, Environmental Protection Agency Cooperative Agreements CR810348 and CR807255, and contracts from Southern California Edison Company and R.J.R. Nabisco.

[3] Although the information in this document has been funded in part by the United States Environmental Protection Agency under Cooperative Agreements CR807255 and CR810348, it does not necessarily reflect the views of the Agency and no official endorsement should be inferred.

[4] Correspondence and requests for reprints should be addressed to Dr. Ling-Yi Chang, Box 3177, Duke University Medical Center, Durham, NC 27710.

ability were used to derive optimal experimental designs for complex EM morphometric studies. Furthermore, we found that the sampling and measurement variability of pulmonary tissues estimated from various models of lung injury reflect the distribution of the injury in the lung.

Methods

Measurements of Variances

Morphometric studies using the electron microscope usually involve volume density *point counting* performed on *micrographs* taken from *sections* of tissue blocks that are selected from the *animals* being studied (19). Variables exist between animals, sections, and pictures (sampling levels). There is also a measurement error associated with each sampling stage. To optimize experimental designs these variances and errors must be estimated. One method of estimation is to use equations (1) through (4) and incorporate data from preliminary experiments (15, 20):

$$\text{Animals } OS^2_a = S^2_a + OS^2_b \cdot 1/n_b \quad (1)$$
$$\text{Sections } OS^2_b = S^2_b + OS^2_p \cdot 1/n_p \quad (2)$$
$$\text{Pictures } OS^2_p = S^2_p + OS^2_{po} \cdot 1/n_{po} \quad (3)$$
$$\text{Points } OS^2_{po} = S^2_m \quad (4)$$

OS^2 is the observed variance of the sample, and S^2 is the true sample variance. S^2_m is the measurement error resulting from point counting. There is no sample variability between points. n_b, n_p, and n_{po} are the number of sections per animal, pictures per section, and points per picture, respectively. Note that observed variance consists of both the true sample variance and a measurement error arising from the inability to study all components of an entity (that is, the whole population, organ, or section). For this reason there is no clear separation of the sample variances or between sample variance and measurement error. Witness that OS^2_b in equation (1) is estimated from equation (2), which contains a measurement error $OS^2_p \times 1/n_p$. It is difficult to study directly how sample size affects observed variances due to this propagation of measurement errors.

We devised a method to measure sample variance and measurement error by unnesting the analysis of variance. A primary sample can be studied with a large number of subunits so that there is essentially no error propagation between sampling levels. The observed sample variance at each level can be studied in relation to its own sample unit. The equations for the unnested variances are

$$\text{Animals } OS^2_{sg} = S^2_a + OS^2_{ma} \cdot 1/n_a \quad (5)$$
$$\text{Sections } OS^2_{sa} = S^2_b + OS^2_{mb} \cdot 1/n_b \quad (6)$$
$$\text{Pictures } OS^2_{sb} = S^2_p + OS^2_{mp} \cdot 1/n_p \quad (7)$$
$$\text{Points } OS^2_{sp} = S^2_{mpo} \cdot 1/n_{po} \quad (8)$$

Take equation (7), for example: OS^2_{sb} is the observed variance for repeated measurements of pictures within a section. If every picture is counted with a large number of points, the propagation of measurement error due to a finite number of points becomes insignificant. The observed variability within a single section is derived only from the variation between pictures S^2_p and a measurement error S^2_{mp} due to the sampling of pictures. The importance of this relationship is that determination of OS^2_{sb} versus n_p allows the sample variance and measurement error to be estimated by regression analysis. The other two sampling levels can be studied in the same general manner.

The unnested equations described here represent a group of independent experiments designed to evaluate the four morphometric procedures specifically. They cannot replace the nested equations in analyzing experimental data whose primary purpose is to study a biologic phenomena.

Optimizing Experimental Design

Statistical principles and practical examples dealing with the allocation of resources have been documented (21–25). If the cost per item is known and the sample variances can be estimated, the optimum number of blocks per animal is determined by

$$n_b = \sqrt{\frac{S^2_b \cdot C_a}{S^2_a \cdot C_b}} \quad (9)$$

c_a and c_b are the costs of studying one animal and one section, respectively. n_p can be derived with a similar equation. If a desired level of variance of the group mean S^2_g is known or, alternatively, if a total cost C_T is imposed, the number of animals needed, n_a, can be solved by using either equation (*10*) or (*11*):

$$n_a = \frac{S^2_a}{S^2_g} + \frac{S^2_b}{n_b n_p S^2_g} + \frac{S^2_p}{n_b n_p S^2_g} + \frac{S^2_{mpo}}{n_b n_p n_{po} S^2_g} \quad (10)$$

$$n_a = (c_a + c_b n_b + C_p n_b n_p + c_{po} n_b n_p n_{po})\ 1/C_T \quad (11)$$

One shortcoming of this approach is the lack of flexibility. A sufficiently efficient experimental design may be one that costs substantially less than the imposed cost with little loss of accuracy. To compensate for this we made use of a procedure called nonlinear function minimization (26). All the combinations that satisfy a certain cost (± 1%) are used to calculate the observed variances. The combinations that yield group variances within 0.1% of the lowest value are the optimal experimental designs at that cost. In this manner a choice can be made to better suit the conditions of an individual laboratory after the optimal designs for a series of costs are determined and compared.

Data Collection

Data obtained from lung tissue of male Fisher 344 or Sprague-Dawley rats, 6 to 12 wk old, were used in this study (1–4, 18). Two groups of control animals and three groups of treated animals were studied. The control animals were exposed to room air only. Alveolar tissue (19, 25) was sampled from one control group, and a specifically defined region, the alveolar tissue closest to the terminal bronchioles (proximal alveolar region, PAR) (1–3), was sampled from the other control group. PAR are obtained from lung tissue by microdissection (2). Alveolar tissues from animals exposed to 85% oxygen for 7 days (4) or asbestos for 1 yr (18) and the proximal alveolar tissue from animals exposed to O_3 for 6 wk were also used (1). Tissue sampling was accomplished by a stratified random strategy (19, 25). Care was taken that tissue blocks that contained large nonparenchymal structures were not excluded (25).

The volume density of alveolar type I epithelium in a sample, $V_{V_{EPI}}$, was chosen to be the parameter on which our analyses were based. Type I epithelium (figure 1A) is relatively homogeneously distributed throughout the alveolar region of the lung and has a volume density of approximately 1% (1, 2, 4, 19). It thus provides a test system for a relatively "rare" tissue component for which the issues are not confused by a substantial amount of heterogeneity. Error analysis on the estimation of $V_{V_{EPI}}$ may be regarded as representative of many lung tissues because the volume density of all tissues in the lung are low as a result of the presence of proportionally large alveolar airspaces. In addition type I epithelium is sensitive to environmental insults, and changes in its volume occur in a variety of types of pulmonary injury (1–5, 18). It provides a suitable parameter for the determination of treatment effect on experimental design. Type II epithelium (figure 1B), on the other hand, is a tissue component that is more heterogeneous in distribution. The effect of this distribution on experimental design was examined using samples from the alveolar region (AR) of control rat lungs. Data collection at each sampling level was carried out as described subsequently. The animals, tissue blocks, and pictures studied had values of type I and type II epithelial fractions that are realistic for either normal or injured rat lungs.

Points. Micrographs taken at an original magnification of ×2,000 and printed on 11 × 14 inch photographic paper with a final magnification of ×8,500 were used. A total of 16 micrographs were analyzed for type I epithelium. Each micrograph was counted with overlays of various systematically spaced point densities (12, 56, 112, 224, 448, and 944) for the estimation of V_V of type I epithelium and tissue (19, 27). Every picture was counted six times at random orientations with each point density. The variance of the six repeated measurements was calculated for each picture. The mean variance from the 16 pictures was used to represent OS^2_{po}. A total of 40 micrographs were studied for the analysis of sampling error for type II epithelium.

This procedure was adopted because type II epithelium shows up on less than 20% of

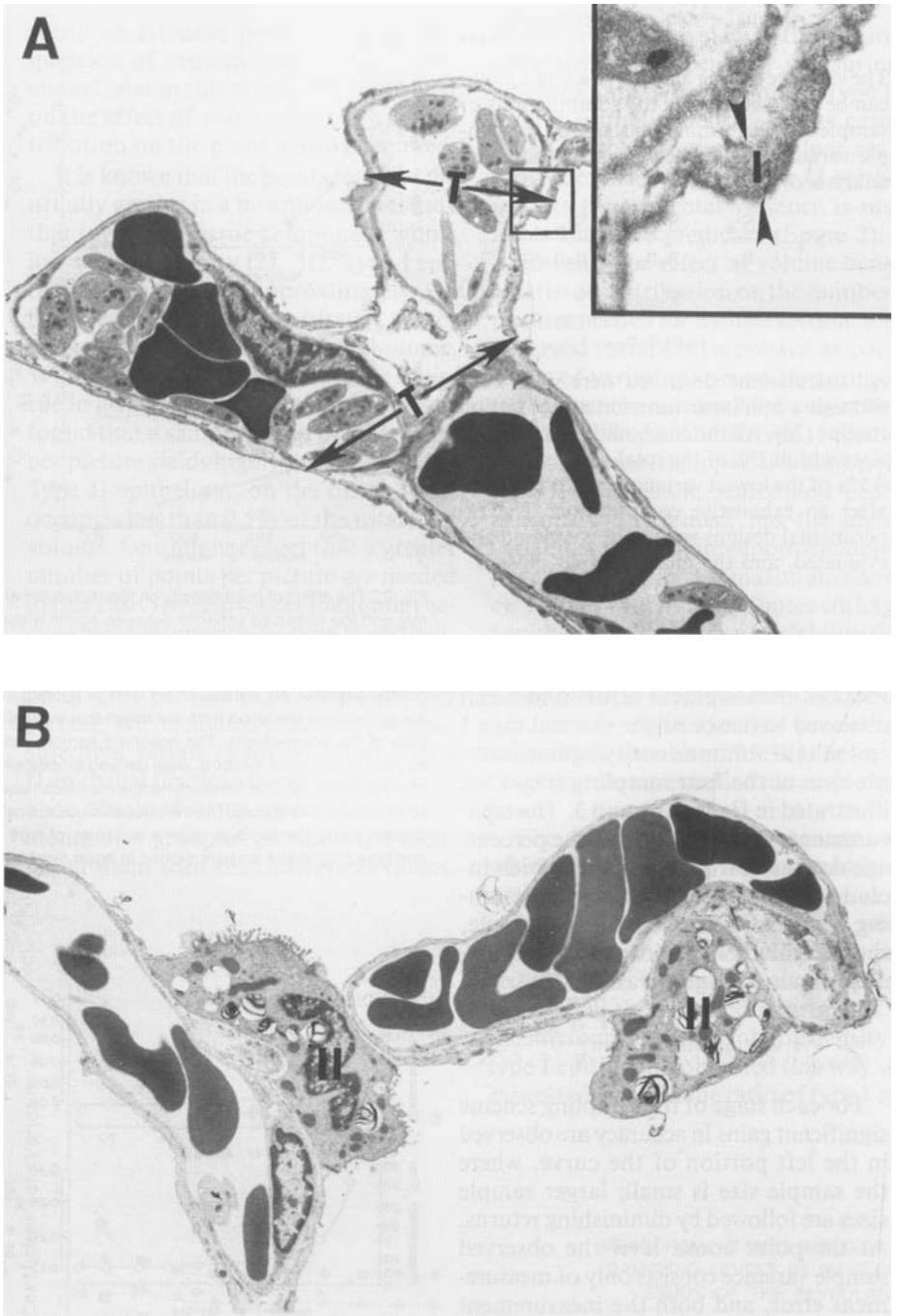

Fig. 1. Pulmonary tissues used in this study. (*A*) Total alveolar tissue and type I epithelium coexist in the lung and are homogeneously distributed. Inset depicts the thin, attenuated nature of type I epithelium. T = tissue; I = type I epithelium. (*B*) Type II epithelium (II) is represented by the lamellar body containing type II cells. These cells are cuboidal.

the micrographs. If the structure of interest does not exist in the containing space of the sample, the measurement variability is zero regardless of the sample size. The utmost accuracy on the estimation of type II epithelial cell density within a picture can be obtained by a single point on the 80% of the micrographs that do not contain type II epithelium. To use a micrograph that contains the mean volume density of type II epithelium for point density determination would be very misleading. We consequently chose to use a group of pictures that have a typical occurrence of type II epithelium to study the relationship between number of points and observed variance within a single picture. The group of pictures also has an average type II epithelial volume density close to that in normal rat lung. The same procedure was adapted to study type I epithelium volume density variation within a single picture.

Pictures. Six step sections from a tissue block were cut and placed on 200 mesh grids individually. Each grid was inserted into the microscope at a random orientation. The grid was surveyed beginning at the top of the grid while moving from left to right. Picture taking ensued when the first grid square completely covered by the section was encountered. Two pictures, one at the upper left and one at the lower right corner of each grid square, were photographed at ×2,000. A field that contained no tissue was not photographed but recorded as one picture. A set of 40 micrographs covering the first 20 grid squares encountered was obtained from each section, resulting in a total of 240 pictures. Micrographs were printed as described earlier and counted with a 448-point multipurpose overlay. The volume densities of type I epithelium, type II epithelium, and tissue were calculated from each picture. We also sampled the sections with an area-weighted procedure. A total of 240 pictures were taken using a method similar to that described by Cruz-Orive and Myking (28). The micrographs were counted with a 448-point counting overlay. Points falling on tissue and type I epithelium were tallied separately, and the volume ratio of type I epithelium to tissue was calculated from each picture.

Sections. Ideally, a large number of sections from one animal should be used. The effort for such an undertaking is prohibitively expensive. To circumvent this requirement we chose to make use of a large amount of section data from previous experiments. Volume density estimates of type I and type II epithelium from 24 normal rats studied under identical conditions in our laboratory were available. Four sections per animal were studied, giving a total of 96 sections. Each section was studied with 30 to 40 pictures counted with 224 points. To eliminate the between-animal variability, each single section estimate of volume density was multiplied by a correction factor that equals the ratio of the group mean to the mean of the animal from which the section was sampled. The 96 normalized values of type I epithelial density were then analyzed. Volume densities of type I epithelium from 48 sections from rats exposed to O_2, 56 sections from rats exposed to asbestos, and 64 PAR sections each from air- and O_3-exposed rats were also studied.

Animals. Animal means from the same rats were used. Each animal was studied with four sections. In addition to the 24 control rats 12 O_2-exposed animals, 14 asbestos-exposed animals, 16 O_3-exposed animals (PAR), and 16 PAR control rats were examined.

Data Manipulation

The bootstrap sampling technique was used to create large sets of simulated data from a single original set. The distribution of sample variability generated can be treated as if it were constructed from real samples. In this way the observed variability from a large number of simulated "observations" can be estimated. The resampling procedure was described by Efron (29). The method requires so many numerical calculations that it is feasible only with the aid of a computer. When

this procedure was applied to the estimation of observed sample variance at a certain sampling level all the single measurements were first copied into the computer, which was then instructed to draw combinations of multiple measurements (n = 2 to n = i), in a random fashion, a specified number of times, that is, 1,000, from the original measurements. For every sample size, therefore, 1,000 combinations were generated. An estimate of the mean was calculated for each combination of a given sample size, and an estimation of the variance of the mean was derived. This procedure was applied to the picture, section, and animal sampling levels.

Data Analysis

Points. Equation (*8*) shows that the observed variance between points in a single picture is proportional to the measurement variance at the point count level. We calculated observed variance either experimentally by regression analysis with data obtained from repeated measurements made on representative pictures or theoretically as described by Weibel (27, 30). Since the numerical value of volume density measured with one point is either 1 or zero, the observed sample variability resulting from point counting can be predicted. The formula used to calculate predicted (theoretical) variance for mean volume density at the point count level is

$$\sigma^2 = \frac{1}{(P_c - 1)} P_{pa} (1 - P_{pa}) \qquad (12)$$

P_c is the total test points applied to the containing space c and P_{pa} is the mean volume density of structure a in the containing space. In this equation structure a is assumed to be homogeneously distributed within the containing space. When the bootstrap procedure was carried out to calculate OS^2_{sp} using combinations of single point estimates, the resulting estimation of S^2_{mpo} from regression analysis was exactly the same as the predicted value. However, the experimental and theoretical values may be different, as is discussed subsequently.

Pictures, sections, or animals. Repeated sampling of the pool of pictures, sections, and animals in a series of sample sizes were performed with the bootstrap technique. The highest sample size attempted was 120 for pictures, 60 for sections, and 20 for animals. A mean estimation was calculated to determine observed sample variance. Regression analysis was then carried out for 1/n and OS^2 to derive the biologic and the measurement components of the observed sample variances at each of the three sampling levels.

Optimizing experiment design. The cost in time to sample one point, one picture, one tissue section, and one animal was determined by pooling estimates obtained from the experienced technical staff in our laboratory. Since the item costs are known and the total cost can be set, all the combinations of n_a, n_b, n_p, and n_{po} that satisfied a certain total cost can be found by the equation

$$C_T = c_a n_a + c_b n_a n_b + c_p n_a n_b n_p + c_{po} n_a n_b n_p n_{po} \qquad (13)$$

The variance among animals in a single group can be calculated using the combinations of sample sizes determined earlier and the sample variances calculated from the unnested analysis of variance:

$$S^2_g = \frac{S^2_a}{n_a} + \frac{S^2_b}{n_a n_p} + \frac{S^2_p}{n_a n_b n_p} + \frac{S^2_{mpo}}{n_a n_b n_p n_{po}} \qquad (14)$$

All calculations described here were done through a nonlinear function minimization routine (26). All the combinations of sample sizes within 1% of the total cost and within 0.5% of the lowest variance were printed out after an exhaustive computation. The experimental designs were then compared and evaluated, and the most suitable chosen.

Results

Accuracy of Measurements

Plots of measurement errors and mean observed variance of the normal type I epithelium volume density against sample sizes at the four sampling stages are illustrated in figures 2A and 3. The measurement error is defined as the percentage deviation from the mean, which includes 95% of the observations assuming a normal distribution. For example, the measurement error for the sampling of animals in a group is

$$\% \text{ deviation} = \frac{2\sqrt{S^2_{ma}}}{\bar{x}}$$

For each stage of the sampling scheme significant gains in accuracy are observed in the left portion of the curve, where the sample size is small; larger sample sizes are followed by diminishing returns. At the point count level the observed sample variance consists only of measurement error, and both the measurement error and the observed sample variance eventually become zero. For the remaining three sampling levels measurement errors become zero at extremely large sample sizes, but observed sample variance remains at the level of sample variances. The reduction in observed sample variance as more and more samples are taken is due to the reduction in measurement error only. It appears that to estimate the volume density of type I epithelium in normal rats with good (but not utmost) accuracy, a sampling grid of 300 to 500 points per picture, 10 to 20 pictures per section, 8 to 16 sections per animal, and 10 to 16 animals per group need to be employed. This is not neces-

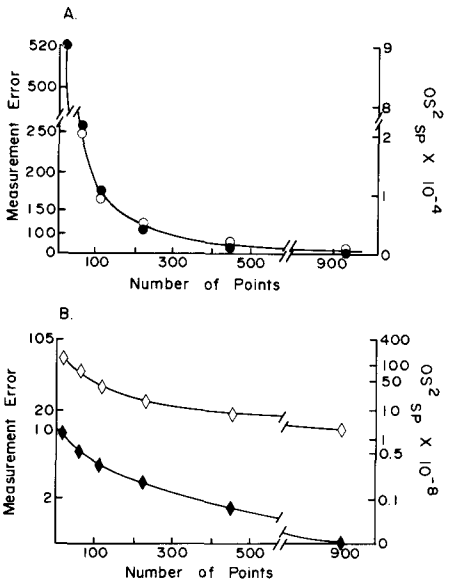

Fig. 2. The effect of point density on measurement errors and the observed variance between points within a single picture for the measurement of (*A*) the volume density of type I epithelium and (*B*) the volume density of type II epithelium. Measurement error is defined as the percentage deviation from the mean that includes 95% of the observations. The observed sample variances (OS^2_{sp}), solid symbols, were derived empirically using equation (*8*). The measurement variance can also be predicted theoretically. The theoretical values are calculated using the average volume densities of type I and type II epithelia and are plotted in open symbols.

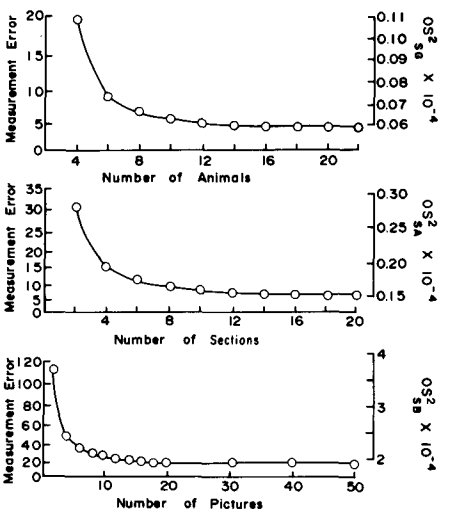

Fig. 3. The effect of sample size on measurement errors and the observed sample variance of type I epithelial density at three sampling stages. Measurement error is defined as the percentage deviation from the mean that includes 95% of the observations. (*A*) Changes in measurement error and the observed variance between animals in a single group in relation to the number of animals studied. (*B*) Changes in measurement error and the observed variance between sections from one animal in relation to the number of sections studied. (*C*) Changes in measurement error and the observed variance between pictures from one section in relation to the number of pictures studied.

sarily an efficient path to follow. The question of experimental design is discussed later in this report. We focus first on the effect of volume density and distribution on the point density required.

It is known that the points required are usually greater in a morphometric study that involves a tissue component with a low volume density (27, 31). Type I epithelium constitutes approximately 1% but the alveolar septum (tissue) makes up around 15% of the total lung volume. When the mean observed variance of tissue in a single picture was studied it was found that a sampling grid of 200 points per picture yields highly accurate results. Type II epithelium, on the other hand, occupies less than 0.5% of the total lung volume. One might expect that a greater number of points per picture are needed in this case. In reality, near maximum accuracy for estimation of type II epithelial density can be achieved with a sampling grid of less than 40 points per picture. We calculated the observed measurement variances for type I and type II epithelial fractions using values of variance obtained from repeated measurements of a group of pictures and compared them with the theoretical values.

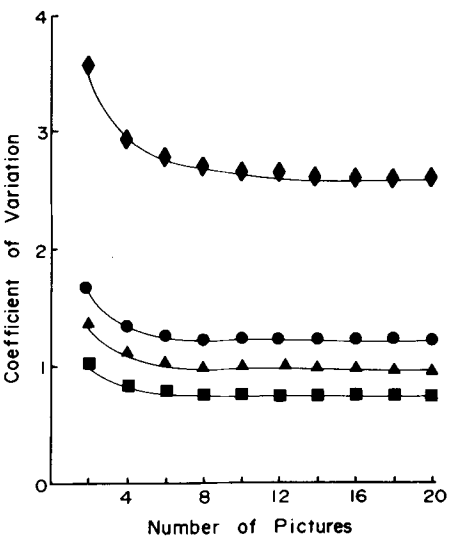

Fig. 4. The effect of volume density on the variability of type I epithelial density between pictures within a single section. Variability is expressed by the coefficient of variation as a fraction of the single section mean. Tissue of three volume densities: EPII, type II epithelium (0.0038); EPI, type I epithelium (0.0115); and the total alveolar tissue (0.135) were studied. An area-weighted procedure was also used to sample type I epithelium (AW-EPI). The resultant variability of the volume ratio of type I epithelium and tissue between pictures in a single section (area-weighted sampling) can be compared with that from randomly sampled type I epithelium. Diamonds = EPII; circles = EPI; triangles = tissue; squares = AW-EPI.

The latter are equivalent to the variances obtained from repeatedly measuring a single picture of representative type I or type II epithelial density. In the case of type I epithelium the two values are almost identical. For the type II epithelium the experimental variance is much lower than the predicted (figure 2).

To study the effect of volume density and tissue distribution on the number of pictures needed for a single section, mean observed variances (expressed as coefficients of variation) per section are plotted against the number of pictures counted for type I epithelium, type II epithelium, and tissue (figure 4). It is shown that type II epithelium, scarce and heterogeneously distributed, has the highest variability and requires approximately 12 pictures per section for maximum accuracy. Tissue, which codistributes with type I epithelium, has a lower variability than type I epithelium but requires just as many pictures per section for maximum accuracy. If a given level of precision is subscribed (that is, 1.25CV), however, fewer pictures are needed to sample total tissue.

The next question addressed was whether accuracy can be improved in the measurement of $V_{V_{EPI}}$ if an area-weighted procedure is followed when selecting photographic fields. This procedure in effect increases the volume densities of all the tissue compartments on the micrographs. The volume density of type I epithelium obtained this way was expressed as a volume ratio of type I epithelium to tissue. After performing bootstrap analysis on the volume ratios and plotting the results (figure 4) it was found that the mean observed variance in a single site was indeed reduced with area-weighted sampling.

The sample variance and the measurement error of normal type I and type II epithelium determined by regression analysis of mean observed variance and the reciprocal of sample size are listed in table 1. Variability is expressed as a coefficient of variation. The actual values of the variances can be calculated using the group mean. The volume density of type I epithelium has a high measurement error at the point sampling level but lower variabilities at the remaining three levels. The volume density of type II epithelium has a low measurement error at the point count level but higher variabilities at the remaining sampling levels, particularly picture and section. The heterogeneous distribution of the type II epithelium is reflected in the high biologic variability reported for this cell.

The effects of three different treatments on animal and site variabilities are illustrated in figures 5 and 6. The coefficients of variation of single animal and single group means are plotted against the number of sections and the number of animals, respectively. A total of 8 animals and 36 sites were analyzed by the bootstrap method for groups of animals exposed to 85% oxygen or asbestos fibers. A total of 16 animals and 64 sites were used for the group exposed to 0.25

TABLE 1

MEASUREMENT AND BIOLOGIC VARIABILITY FOR EACH OF THE FOUR SAMPLING LEVELS IN AN ELECTRON MICROSCOPIC MORPHOMETRIC STUDY OF THE ALVEOLAR REGION OF NORMAL RAT LUNGS

	Coefficient of Variation*	
Epithelium	Measurement Error	Sample Variance
EPI†		
Point	9.08	—‡
Picture	1.70	1.11
Section	0.49	0.13
Animal	0.30	0.19
EPII§		
Point	1.61	—‡
Picture	3.65	2.45
Section	1.05	0.66
Animal	0.42	0.27

* Coefficient of variation (CV) is equal to the square root of variance divided by the mean. It is expressed as a fraction of the mean. The 95% confidence interval is equal to 2CV; in other words, 95% of the measurements lie within ± 2CV of the mean.

† Type I epithelium (EPI) has an overall volume density of 0.0115 in the group of animals studied and is reasonably homogeneous in distribution.

‡ There is no biologic variability for points.

§ Type II epithelium (EPII) has an overall volume density of 0.0038 in the group of animals studied and has a heterogeneous distribution.

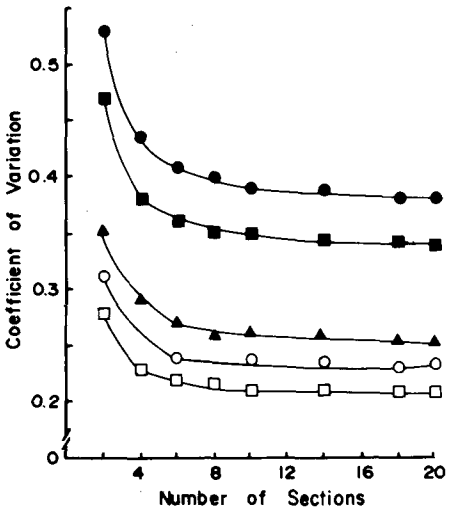

Fig. 5. The effect of treatment on the variability (CV) of type I epithelial density between sections in a single animal is evaluated by comparing the CV of SO^2_{SA} of treated animals to that of controls from two regions in the rat lung, the alveolar region (AR) and the proximal alveolar region (PAR). The effect of three treatments were studied: exposure to asbestos, exposure to 85% O_2, and exposure to O_3. The first two involve alveolar regions, and O_3 involves the proximal alveolar region. Closed squares = control-AR; closed circles = asbestos-AR; closed triangles = 85% O_2-AR; open squares = control-PAR; open circles = O_3-PAR.

ppm ozone. Animal variability is highest with the 85% O_2-exposed rats, and O_3-treated animals had a lower variability.

Another determinant for site and animal variabilities is the sampling technique. Comparison of sample variances of random alveolar tissue and proximal alveolar tissue from normal rat lungs shows that specimens from the PAR, a specific anatomic region in the lung, have lower sample variabilities. Proximal alveolar tissue was used to study the effect of low levels of oxidant pollutants, which produce lesions primarily at the bronchiolar-alveolar duct junction (1, 2). The morphometric characteristics of the tissues of this highly specific lung region appear to be less variable than is randomly selected alveolar tissue.

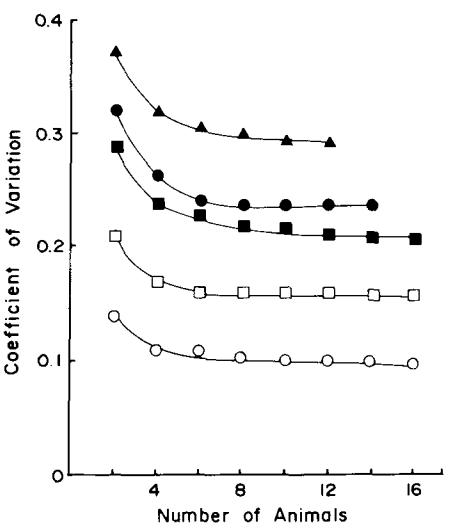

Fig. 6. The effect of treatment on the variability (CV) of type I epithelial density between animals in a single group. The same treatments and anatomic units studied for section variability in figure 5 are plotted here. Closed squares = control-AR; closed circles = asbestos-AR; closed triangles = 85% O_2-AR; open squares = control-PAR; open circles = O_3-PAR.

TABLE 2

ITEMIZED TASKS AND ESTIMATION OF COST* FOR A MORPHOMETRIC STUDY OF THE ALVEOLAR REGION

Levels	Tasks	Unit of Effort†
Animal	Fixation‡	90
	Measure lung volume	
	Embedding§	
Section	Sectioning	100
	Staining	
Picture	Photographing	5.0
	Developing	
	Printing	
Point	Counting	0.01

* Costs common to a whole group of animals regardless of the number of animals are regarded as fixed costs. Since this type of cost differs with treatment it is not considered part of this discussion.
† Cost is expressed as unit of effort proportional to the time used to conduct the tasks listed.
‡ Fixation includes surgical procedures as well as infusing fixative into the lung.
§ Effort estimated for embedding does not include block curing time. This is considered a fixed cost.

Optimizing Experimental Design

It has been pointed out time and again that although the variance of individual measurements at the lower sampling level is large, its contribution to the overall observed error is small. From this comes the concept "do more less well!" (32). On the other hand, the cost of counting extra points on a micrograph is substantially less than the cost of producing extra micrographs. The most efficient design balances cost and accuracy. Efficiency is approached by minimizing observed sample variance in a single group at a fixed cost. We used the time spent at each sampling stage as a measure of cost. In table 2 the various tasks involved and the total cost in units of effort accorded on each sampling stage are listed. These estimations reflect the relative degree of difficulty of these tasks when maximum productivity is achieved in our laboratory. The appropriate variance estimates (S^2_a, S^2_b, S^2_p, and S^2_{mpo}) and the cost estimates (C_{po}, C_p, C_b, and C_a) can be inserted into equations (*13*) and (*14*), respectively, to carry out the procedure of nonlinear minimization with various total costs and derive the optimal combination of n_a, n_b, n_p, and n_{po} for each cost.

Different laboratory procedures may result in a different effort outlay for each task. Of the four sampling stages, section and point efforts can be regarded as relatively uniform from one laboratory to another. Time spent on printing pictures can vary a great deal depending on the equipment (dry print processor or trays and drum dryer) available. It is also possible to reduce the effort in producing pictures by recording ultrastructure on 35-mm films and directly projecting them onto an overlay screen, thus bypassing the need to print micrographs. Cost at the animal stage can also fluctuate relative to the possible treatment and surgical procedures involved in any experimental protocol. Several scenarios on the variation of effort and its effect on optimal sampling designs at a fixed cost are presented in table 3. The results, considering only type I epithelium, show that fluctuations in picture costs do not change the optimal number of sections to be studied in an animal. A higher picture cost makes increasing points counted per picture and decreasing the number of pictures per section the desired method to gain efficiency. A lower picture cost has the opposite effect. High costs for processing each animal left increasing the number of sections as the most effective way to reduce variation.

Changes in the optimal experimental design when the total cost is varied are reflected mainly at the animal level (table 4). The best way to spend increased resources appears to be by increasing the number of animals studied. The differences in the optimal number of points or the number of pictures is minimal in the case of either type I or type II epithelium. The distribution of effort, however, is quite different for the two types of pul-

TABLE 3

OPTIMIZED EXPERIMENTAL DESIGN* FOR VARIOUS ANIMAL OR PICTURE COSTS AT A FIXED TOTAL COST OF 8,000 UNITS OF EFFORT

Item	Cost†	Animals	Sections	Pictures	Points‡	CV§
Animal	480	6	4	21	201	0.12
	240	10	3	16	211	0.10
	90	18	2	16	141	0.09
Picture	10	15	2	11	241	0.10
	5	18	2	16	141	0.09
	2.5	19	2	21	151	0.08

* This analysis considered only the type I epithelium in the alveolar region of normal rat lungs.
† All other costs are those listed in table 2.
‡ The numbers of points per picture reported in tables 3 through 5 represent the size of the counting grid.
§ The coefficient of variance was calculated from the group variance and the group mean.

TABLE 4

OPTIMIZED EXPERIMENTAL DESIGN FOR MORPHOMETRIC STUDY OF TYPE I AND TYPE II EPITHELIUM IN THE ALVEOLAR REGION OF NORMAL RAT LUNGS AND THE EFFECT OF TOTAL COST ON DESIGN*

	Type I Epithelium			Type II Epithelium		
Total Cost	Design†		CV‡	Design†		CV‡
4,000	A	9	0.13	A	6	0.23
	S	2		S	3	
	P	16		P	21	
	Po	141		Po	41	
6,000	A	13	0.10	A	10	0.18
	S	2		S	3	
	P	16		P	16	
	Po	211		Po	51	
8,000	A	18	0.09	A	17	0.16
	S	2		S	2	
	P	16		P	21	
	Po	141		Po	41	
10,000	A	22	0.08	A	24	0.145
	S	2		S	2	
	P	16		P	16	
	Po	181		Po	41	

* Item costs for this analysis are those reported in table 2.
† Reported as a combination of animals (A) per group, sections (S) per animal, pictures (P) per section, and points (Po) per picture.
‡ The coefficient of variance was calculated from the group variance and the group mean.

TABLE 5

INTEGRATED EXPERIMENTAL DESIGN FOR TYPE I (EPI) AND TYPE II (EPII) EPITHELIUM IN TWO REGIONS OF THE NORMAL RAT LUNGS AND TYPE I EPITHELIUM IN POLLUTANT- OR OXYGEN-EXPOSED ANIMALS

Design	Region	Treatment and Tissue	CV*	Cost
10 Animals	Alveolar region	Control, EPI	0.098	7,180
3 Sections		Control, EPII	0.177	7,180
20 Pictures		Asbestos, EPI	0.101	7,180
200 Points		85% O_2, EPI	0.092	7,180
	Proximal† alveolar region	Control, EPI	0.083	8,875
		Control, EPII	0.198	8,875
		O_3, EPI	0.071	8,875

* The coefficient of variance was calculated from the group variance and the group mean.
† Microdissection used to select proximal alveolar region increases the site cost to 150 units of effort per section.

monary epithelium. Experiments with type II epithelium as the sole end point require less point-counting precision but more animals and/or sections than experiments focused on type I epithelium, owing mainly to the heterogeneous distribution of type II epithelium.

Most morphometric studies of lung structure involve more than a single tissue type. It is inconvenient as well as inefficient to install different protocols, such as a different counting overlay, for each end point. Attempts are made to devise an experimental design for studies of multiple end points. A large number of analyses similar to those in table 4 were carried out with data from normal as well as treated rat lungs. An integrated experimental design was selected that allows the coefficients of variation on type I epithelium to be equal or less than 0.1 and those of type II epithelium under 0.2 while keeping the total cost under 10,000 total units. The result is outlined in table 5. It is not possible to list all the information and reasonings that were used in the formulation of table 5. A large number of experimental designs with total costs between 7,000 and 10,000 were reviewed. Those that satisfied the CV requirement for all groups and tissues were selected. It was decided that gain in accuracy by using more than 10 animals was not cost efficient. Three sites per animal were selected because the low variability between proximal alveolar regions rendered its choice preferable for PAR groups, although only two sites were required by the AR groups. Likewise, 200 points per picture was chosen to preferentially optimize accuracy for type I epithelium since this is one of the more important injury indicators among pulmonary tissues. The optimum number of pictures did not vary greatly with various categories: 20 were decided upon to optimize accuracy for type II epithelium and alveolar macrophages. It should be cautioned that the results presented in table 5 are not the only mathematical solution.

Discussion

Using the volume densities of pulmonary epithelia as examples, we have presented a new approach to the analysis of sampling variances, demonstrated how the estimates were made, utilized them to optimize the allocation of resources, and described the effect of treatment on sample variability.

It is important to use unbiased sampling methods to obtain accurate stereologic measurements. When the whole lung is not used for sampling bias may result from such factors as heterogeneity of tissue distributions or regional differences in the size of alveoli and capillaries, particularly in large lungs where gravity may be a significant issue (22). Another potential source of bias in studies of lung

parenchymal tissue may occur if tissue blocks containing large nonparenchymal structures (25) are excluded from the study. Ideally, systematic or stratified random sampling methods (21, 25) should be used when conducting a morphometric study and attention to potential sources of bias carefully examined. When sampling bias cannot be eliminated it is worthwhile to note that Müller and colleagues reported that the internal structure of alveolar septa is very constant (33). We have found that simple random samples and stratified random samples yield very similar overall results when analyzing alveolar structures.

In the first part of this study our goals were to find out how observed sample variance changes in relation to sample size and to define an adequate range of sample sizes to obtain accurate measurements. As expected, increasing sample size diminishes variance and improves accuracy of measurement. Ultimate accuracy is obtained when the observed variance equals the biologic variance or when the sample size approaches infinity, but near maximum accuracy can be obtained with reasonably small sample sizes.

A number of factors may influence precision and subsequently alter the sample size required for precise measurement. The first component of observed sampling variance, the measurement error, is related to the method used and is manifested at the point count level. The size, shape, and distribution of the structure within the area of the micrograph, as well as the counting procedure used, may affect the magnitude of the measurement error. Our data in general agree with previous conclusions that for a given level of precision a smaller sample size can be used for a compartment of higher volume density (31).

The effect of differences in homogeneity of distribution on point counts is more complex and must be considered together with the point lattice system used (27, 30). With clumped profiles (heterogeneously distributed in relation to the area of the picture) one may obtain higher accuracy by the use of systematically spaced points instead of random points. This is because the probability of a nearby point hitting the same structure is larger for a systematically spaced point. For the same reason the variance of measurements on clumped profiles is less than on measurements of homogeneously distributed but dispersed profiles with the same area density when a systematic point lattice is used. This accounts partially for the observation that type II epithelium does not need a denser point overlay. The major factor for the low measurement error of type II epithelium is that there is no variance for the measurement of type II epithelial density on pictures that do not contain type II cells.

At the picture sampling level the effects of volume density and distribution are more clearly demonstrated. In comparison to type I epithelium, tissue- and area-weighted sampling of type I epithelium exhibited lower observed sample variance between pictures. Because these tissues codistribute, high volume density alone is the factor that contributes to the lowering of variance. In the case of type II epithelium, heterogeneity of distribution is the main cause of high variability.

At the section sampling levels variance is demonstrated primarily by the distribution of the structure within the containing space. Heterogeneity of distribution increases variance. This is demonstrated by the high within-section variance of type II epithelium. Such procedures as the selection of proximal alveolar regions that restrict the type of sample taken reduced the relative variability in distribution. Injury may also influence measurements of variability. However, biologic variance among animals is the ultimate limiting factor to accuracy, as depicted in equation (14).

Our study of the effect of treatment on sample variance and measurement error revealed that differences in the magnitude of observed between-site (section) variance is reflected by the nature of the injury. The high section variability exhibited by asbestos-exposed rats may reflect the irregular distribution of inhaled particulates in the lung (18). The dose of O_3, a very reactive gas in low concentration, delivered to each PAR may also be different depending on the length of the preceding airway (3, 34). This results in a between-site variation higher than the control level. Oxygen at a concentration of 85%, on the other hand, distributes homogeneously within the lung (3). Curiously, the variability between sites of oxygen-treated animals is lower than it is in control sites. It may be that the normal type I cell population is more diversified in developmental, metabolic, or differentiation stages. The presence of a homogeneous high-level insult induced uniform cell responses, resulting in a more homogeneous population. The high between-animal variability observed with O_2-exposed rats may reflect differences in the effectiveness of individual antioxidant protective mechanisms. It is not clear why O_3-treated rats demonstrated a lower animal-to-animal variability. Ozone is a very reactive gas, and regional differences in O_3 dose in a lung can result from variations in path length, rate of ventilation, and ventilatory unit size (34). It is likely that the average O_3 dose given to an animal under controlled conditions (as indicated by the effect measured by averaged random selected sites) is more consistent than the dose to individual sites within a single animal.

In the second part of the study the numerical values of the various sample variances were estimated, and these were used to optimize experimental design. Computer-assisted, computation-intensive methods were used to create large data sets for minimization of error propagation (bootstrap) and to determine optimal experimental designs (linear minimization routine). We found that the lowest possible observed variance within a single group at any given cost can be estimated. In addition, a curve that plots cost against accuracy can be obtained. When this is done for a number of tissue types and treatment groups we found that the same general curve resulted every time and only the values of variance on the ordinate varied. This curve

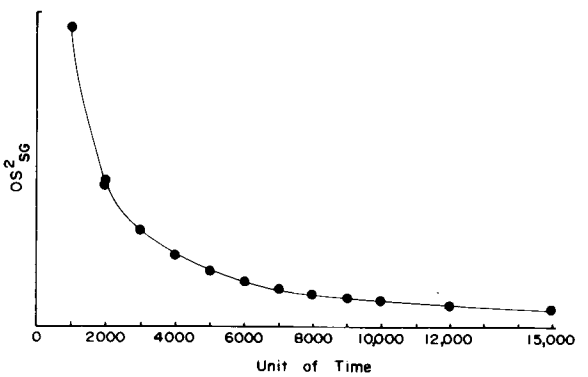

Fig. 7. The effect of cost (proportionally related to sample sizes) on the between-animal variability in a single group. Various treatment and tissues have the same cost effect.

is illustrated in figure 7. It can be seen that near maximum precision is obtained with an effort outlay of around 10,000 units of effort. In fact, precise and accurate results are attainable with a total cost of 6,000 to 8,000 units. This curve therefore establishes the cost limit as well as the cost range for efficiency. Alternatively, if an acceptable level of within-group variation is decided upon the corresponding minimum effort can be determined.

We have demonstrated that type II epithelium, a heterogeneously distributed structure, has a sample variance at the animal level that is close to 30% of the group mean. Alveolar macrophages are more scarce and also more heterogeneous. They are expected to express an even higher biologic variability. This means that to set a goal for a relative error of less than 10% on the estimation of the group mean for this type of cell is not realistic. The knowledge of a range of costs that yield near maximum accuracy for all tissue types prevents the investigator from wasting precious resources. Since the efficiency range for all tissues is the same on the cost axis, this type of cost-effective analysis must be performed only once.

At the beginning of the Results section we stated that for maximum accuracy a sampling grid of 300 to 500 points per picture, 10 to 20 pictures per section, 8 to 16 sections per animal, and 10 to 16 animals per group must be studied. This work load amounts to a cost range of 15,300 to 78,240 units and gives a coefficient of variation of about 0.06 for type I epithelium in control animals. Comparing this cost and error to a cost of 7,180 units and a CV of 0.098 listed for the experimental design reported in table 5 for the same tissue, we found that the extra work results in a gain of accuracy of only about 38% while increasing cost by 2 to 10 times the effort!

Another comment pertaining to the discussion of accuracy and efficiency is the merit of area-weighted sampling (25, 28). We have established that area-weighted sampling reduced the picture-to-picture variation within a single site. Nevertheless, the total effort spent was similarly increased in degree in our hands. Moreover, since the data obtained were in the form of a ratio to total tissue, it would be necessary to insert an additional sampling stage to determine the volume density of alveolar tissue in parenchymal lung. This new sampling level would add additional cost and contribute to more sampling error. We decided that the increase in cost in doing area-weighted sampling was not adequately compensated for by the improved accuracy. The topic was consequently not pursued further.

The issue of optimal design has been examined before (21, 23-25, 31, 32, 35-38), with emphasis on stereologic theory and techniques. The basic concepts utilized in this study are essentially the same as those in previous reports. Our aim, however, is to provide practical guidelines and actual estimations of sample variances for pulmonary pathologists in the process of designing a more morphometric study. Two novel approaches, the unnesting of the analysis of variances and the employment of computation-intensive statistical methods, were introduced in this study. These approaches and access to a large number of quantitative data on lung structure enabled us to directly study the effect of sample size on measurement accuracy at each sampling stage. The data analysis illustrated with this pulmonary model provides some insight into biologic variation and measurement error manifested by lung tissue, particularly in relation to models of lung injury. The optimal experimental designs described can be easily adapted for most EM morphometric studies of the lung with the same four-tiered sampling scheme. For a different sampling scheme the sample variances and measurement errors reported here for type I and type II epithelial volume density may be used for the formulation of suitable experimental designs.

References

1. Barry BE, Miller FJ, Crapo JD. Effects of inhalation of 0.12 and 0.25 ppm ozone on the proximal alveolar region of juvenile and young adult rats. Lab Invest 1985; 53:692-704.
2. Chang L, Graham JA, Miller FJ, Ospital JJ, Crapo JD. Effects of subchronic inhalation of low concentrations of nitrogen dioxide. I. The proximal alveolar region of juvenile and adult rats. Toxicol Appl Pharmacol 1986; 83:46-61.
3. Crapo JD, Barry BE, Chang L, Mercer RR. Alterations in lung structure caused by inhalation of oxidants. J Toxicol Environ Health 1984; 301:121-41.
4. Crapo JD, Barry BE, Foscue H, Shelburne J. Structural and biochemical changes in rat lungs occurring during exposures to lethal and adaptive doses of oxygen. Am Rev Respir Dis 1980; 122: 123-43.
5. Chang L, Overby LH, Brody AR, Crapo JD. Progressive lung cell reactions and extracellular matrix production after a brief exposure to asbestos. Am J Pathol 1988; 131:156-70.
6. Barr BC, Hyde DM, Plopper CG, Dungworth PL. Distal airway remodeling in rats chronically exposed to ozone. Am Rev Respir Dis 1988; 137: 124-38.
7. Barry BE, Miller FJ, Crapo JD. Alveolar epithelial injury caused by inhalation of 0.25 ppm of ozone. In: Lee SD, Mustafa MG, Mehlman MA, eds. International symposium on the biomedical effects of ozone and related photochemical oxidants. Princeton, NJ: Princeton Scientific Publishers, 1983; 299-309.
8. Bartlett D, Faulkner CS, Cook K. Effect of chronic ozone exposure on lung elasticity in young rats. J Appl Physiol 1974; 37:92-6.
9. Boorman GA, Schwartz LW, Dungworth DL. Pulmonary effects of prolonged ozone insult in rats. Morphometric evaluation of the central acinus. Lab Invest 1980; 43:108-15.
10. Evans NJ, Dekker NP, Cabral-Anderson LJ, Sham SG. Morphological basis of tolerance to ozone. Exp Mol Pathol 1985; 42:366-76.
11. Fujinaka LE, Hyde DM, Plopper CG, Tyler WS, Dungworth DL, Lollthi LO. Respiratory bronchiolitis following long-term ozone exposure in bonnet monkeys: a morphometric study. Exp Lung Res 1985; 8:167-90.
12. Hyde DM, Orthoefer J, Dungworth DL, Tyler W, Carter R, Lum H. Morphometric and morphologic evaluations of pulmonary lesions in beagle dogs chronically exposed to high ambient levels of air pollutants. Lab Invest 1978; 38:455-69.
13. Lum H, Schwartz LW, Dungworth DL, Tyler WS. A comparative study of cell renewal after exposure to ozone or oxygen: response of terminal bronchiolar epithelium in the rat. Am Rev Respir Dis 1978; 118:335-45.
14. Wilson DW, Plopper CG, Dungworth DL. The response of the macaque tracheobronchial epithelium to acute ozone injury. A quantitative ultrastructural and autoradiographic study. Am J Pathol 1984; 116:193-206.
15. Sterio DC. Estimating number, mean sizes and variations in size of particles in 3-D specimens using dissectors. J Microsc 1984; 134:127-36.
16. Gundersen HJG. Quantitative analysis of 3-dimensional structure in neuroanatomy. In: Agnatiand LF, Fuxe K, eds. Quantitative neuroanatomy in transmitter research. London: MacMillan, 1985; 3-9.
17. Baddeley AJ, Gundersen HJG, Cruz-Orive LM. Estimation of surface area from vertical sections. J Microsc 1986; 142:259-76.
18. Pinkerton KE, Pratt PC, Brody AR, Crapo JD. Fiber localization and subsequent lung reaction to chronic inhalation of chrysotile asbestos in Fisher 344 rats. Am J Pathol 1984; 117:484-98.
19. Weibel ER. Morphometry of the human lung. Berlin: Springer-Verlag; New York: Academic Press, 1963.
20. Snedecor W, Cochran WG. Statistical methods. Ames, IA: Iowa State University Press, 1967.
21. Cruz-Orize LM, Weibel ER. Sampling designs for stereology. J Microsc 1981; 122:235-57.
22. Behr P, Wiebel ER. Morphometric estimation of regional differences in the dog lung. J Appl Physiol 1974; 37:648-53.
23. Gundersen HJG, Jensen EB. The efficiency of systematic sampling in stereology and its prediction. J Microsc 1987; 147:229-63.
24. Nicholson WL. Application of statistical methods in quantitative microscopy. J Microsc 1978; 113:223-9.
25. Weibel ER, Gehr P, Cruz-Orive LM, Müller AE, Mwangi DK, Haussener V. Design of the mammalian respiratory system. IV. Morphometric estimation of pulmonary diffusing capacity; critical evaluation of a new sampling procedure. Respir Physiol 44:39-59.
26. Nelder JA, Mead R. A simple method for func-

tion minimization. Comput J 1965; 7:308–13.
27. Weibel ER. Stereological methods, Vol. 1. Practical methods for biological morphometry. London: Academic Press, 1979.
28. Cruz-Orive LM, Myking AO. Stereological estimation of volume ratio by systematic sections. J Microsc 1981; 122:143–57.
29. Efron B. Computers and the theory of statistics: thinking the unthinkable. SIAM Rev 1979; 21:460–80.
30. Henning A. Fehlerbetrachtungen zur volumenbestimmung aus det integration ebener schnitte. In: Weibel ER, Elias H, eds. Quantitative methods in morphology. Berlin: Springer-Verlag, 1967; 99–129.
31. Bolander RP, Paumbartner D, Losa G, Muellener D, Weibel ER. Integrated stereological and biochemical studies on hepatic membranes. I. Membrane recoveries in subcellular fractions. J Cell Biol 1978; 77: 565–83.
32. Gundersen HJG, Osterby R. Optimizing sampling efficiency of stereological studies in biology: or 'Do more less well!' J Microsc 1981; 121:65–73.
33. Müller AE, Cruz-Orive LM, Gehr P, Weibel ER. Comparison of two subsampling methods for electron microscopic morphometry. J Microsc 1981; 12:35–49.
34. Mercer RR, Crapo JD. Anatomical modeling of microdosimetry of inhaled particles and gases in the lung. In: Crapo JD, Smolko ED, Miller FJ, Graham JA, Hayes AW, eds. Extrapolation of dosimetric relationships for inhaled particles and gases. San Diego: Academic Press, 1989; 69–78.
35. Shay J. Economy of effort in electron microscope morphometry. Am J Pathol 1975; 81:503–11.
36. Gupta M, Mayhew TM, Bedi KS, Sharma AK, White FH. Inter-animal variation and its influence on the overall precision of morphometric estimates based on nested sampling designs. J Microsc 1983; 131:147–54.
37. Cruz-Orive LM, Hunziker EB. Stereology for anisotropic cells: application to growth cartilage. J Microsc 1988; 143:47–80.
38. Mathieu O, Cruz-Orive LM, Hoppeler H, Weibel ER. Measuring error and sampling variation in stereology: comparison of the efficiency of various methods for planar image analysis. J Microsc 1981; 121:75–88.