Morphological Quantitation of Emphysema: A Debate

The following series of letters to the editor concerns a recently published morphologic method proposed by Parameswaran and colleagues for sensitive, early detection of emphysema (J Appl Physiol 100: 186–193, 2006). The validity of the proposed method was critiqued by Ewald Weibel, and, in turn, this critique was rebutted by Parameswaran et al. Additional brief commentaries were contributed by scientists working in the field. Further comments on this important topic are welcome.

The following is the abstract of the article discussed in the subsequent letter:

Parameswaran, H., A. Majumdar, S. Ito, A. M. Alencar, and B. Suk. Quantitative characterization of airspace enlargement in emphysema. J Appl Physiol 100: 186–193, 2006.—The mean linear intercept (Lm) can be used to estimate the surface area for gas exchange in the lung. However, in recent years it is most commonly used as an index for characterizing the enlargement of airspaces in emphysema and the associated severity of structural destruction in the lung. Specifically, an increase in Lm is thought to result from an increase in airspace sizes. In this paper, we examined how accurately Lm measures the linear dimensions of airspaces from histological sections and a variety of computer-generated test images. To this end, we developed an automated method for measuring linear intercepts from digitized images of tissue sections and calculate Lm as their mean. We examined how the shape of airspaces and the variability of their sizes influence Lm as well as the distribution of linear intercepts. We found that for a relatively homogeneous enlargement of airspaces, Lm was a reliable index for detecting emphysema. However, in the presence of spatial heterogeneities with a large variability of airspace sizes, Lm did not significantly increase and sometimes even decreased compared to its value in normal tissue. We also developed an automated method for measuring the area and computed an equivalent diameter of each individual airspace that is independent of shape. Finally, we introduced new indexes based on the moments of diameter that we found to be more reliable than Lm to characterize airspace enlargement in the presence of heterogeneities.

Unbiased stereology is better

To the Editor: To understand the functional effects of structural damage in emphysema, it is important to quantitatively assess the changes in the geometry and size of airspaces. In their recent paper, Parameswaran et al. (7) purport to present a new advanced method for the morphometric assessment of pulmonary emphysema, but it uses flawed and biased methods that disregard some of the essential principles of stereology.

It is of primary importance that the images used for morphometric analysis faithfully represent the underlying structure. The automated image segmentation used for this study, as shown in Fig. 1, leads to a totally arbitrary skeleton supposed to represent air space walls. The image segmentation algorithm evidently does not allow for “free ends”: they are either erased or connected to other outlines in a totally arbitrary fashion. But free ends (on section) are an essential structural feature: they are the alveolar walls attached to alveolar ducts, thus forming openings of alveoli that allow ventilation to reach the alveolar surface.

Another central issue in this type of study is sampling. Nothing is said about guarding against bias in sampling, not even that in a two-dimensional (2D) section representing the three-dimensional (3D) lung large blebs, as they occur in emphysema, have a higher chance of being sampled than small blebs or normal alveoli, and how this could be avoided. There is no information about the reference space, the volume of the lung, an essential piece of data to avoid the so-called reference trap: the size of air spaces depends on the degree of lung inflation.

I was perplexed by Fig. 2 where mean linear intercept length (Lm) is plotted against the area-to-perimeter ratio (AIP). This is “self-correlation” because, on geometric grounds, Lm is directly proportional to AIP, as actually noted in Eq. 3! The sentence “Equations 3 and 4 make Lm a useful method for estimating perimeter and surface of a single object with known area and volume, respectively” makes no sense; it is a wrong statement. In principle, considering the analytical process of this study, the area of a closed profile is the sum of all intercept lengths multiplied with the distance between test lines, in this case presumably the pixel diameter. The perimeter is proportional to the number of end points of the intercepts, or the number of intersections with the test lines, but it is not the simple sum of these points times the pixel size; one must account for the orientation of the contour line, that is, where π (or 4 in the 3D case) comes from, from geometric probability (1, 8), it is not a shape factor. By the way, Lm is not shape dependent as the section heading suggests: the Crofton-Cauchy formula only demands isotropic random orientation of the test lines to the surface in 3D, or to the contour line in 2D (1, 8).

The authors then introduce the “equivalent diameter” method to estimate variability of airspace size by calculating the first and second moment of the diameter distribution. The problem is that this diameter represents the section image, or actually the artificial lattice of automatically segmented “cells” (see above), and not the real structure. As mentioned, the size distribution thus derived is biased because large blebs have a greater chance of being sampled by the section (1, 8). It is, furthermore, meaningless, in fact not permissible, to compare the moment estimates D1 and D2 with Lm (Fig. 6) to see which measure better separates emphysematous from normal lungs. Lm is by definition a mean, and it is trivial to say that a mean cannot distinguish between different degrees of variance! A mean, the first moment, is a statistical measure designed to extract the central tendency of distributed data against their variation.

The main problem is that this paper disregards the relevant literature on stereology (1, 2, 6, 8, 9). The linear intercept is a good measure, mainly because it allows the unbiased estimation of 3D parameters of structure, in contrast to the equivalent diameter that only refers to section profiles and is therefore heavily biased toward large blebs. But instead of the mean linear intercept, one could determine intercept length distributions (8) or use some of the new stereological methods that allow the direct and unbiased estimation of the variance of air space size (3D diameter!) from the estimation of number- and volume-weighted mean particle volumes using point-sampled linear intercepts (4, 9). These could be used to assess the 3D structural changes in emphysema with state-of-the-art methods of quantitative morphological analysis (3, 5, 10) that are no more laborious than the biased equivalent diameter method proposed here.

REFERENCES

Letters to the Editor


Ewald R. Weibel

Institute of Anatomy

University of Berne

Berne, Switzerland

e-mail: weibel@ana.unibe.ch

**REPLY**

*To the Editor:* Today, in the era of advanced computerized image analysis, emphysema is still evaluated manually or, at best, semiautomatically from two-dimensional histological slides of the lung parenchyma. In the January issue of the Journal, we propose a method capable of detecting air space enlargement in emphysema without the bias arising from manual analysis of two-dimensional images (10). We are honored that Prof. Weibel finds our paper noteworthy. However, we are surprised that he does not find our method useful. This appears to be a consequence of his misinterpretation of our objectives and methods.

The primary objective of this study was to develop a robust image-processing algorithm that automatically detects abnormal air space enlargement from two-dimensional sections of emphysematous lung tissue. Additionally, we also aimed at clarifying some issues related to using the mean linear intercept (Lm). If correctly analyzed, the Lm is an unbiased estimator of the ratio of lung volume to internal surface area (3). However, because of its simplicity, the Lm has been increasingly used by many laboratories to characterize size, specifically, air space enlargement in emphysema (2, 4, 5, 7, 11, 13). Our results provide evidence that Lm is not appropriate for this purpose because it depends on shape (Eq. 7 and Fig. 3) and it is a biased estimator of air space size (Fig. 3). Therefore, in this paper we also introduced a method for measuring the area of air spaces in histological sections. Being independent of shape, the area measurements can be used to construct indexes, which also capture the heterogeneity of the underlying structure.

Prof. Weibel’s main criticism is based on his claim that our method is biased. Such a criticism would be valid had we used our new index, as in “unbiased stereology,” to estimate three-dimensional parameters from two-dimensional sections. Stereological techniques invoke the use of test objects and counting their intersections with the objects of interest (9). Our algorithm was designed to detect abnormally large air spaces by measuring two-dimensional size from images without resorting to any of the tools used in stereology. Therefore, the argument that our method is biased in the stereological sense is not relevant.

Why is our index (D2 in Fig. 6) successful in differentiating emphysematous from normal tissue? Because heterogeneities invariably arise in emphysema, larger “blebs” or enlarged air spaces are likely to appear together with many smaller ones on histological slides. A trained observer will search for such regions because it is the very presence of these blebs that signifies emphysema. In this work, we put this intuitive approach in an algorithmic form. We find that we need to give more weight to the enlarged air spaces than the smaller ones. This preferential weighting is the single most important factor that is essential to automatically identify abnormal air space enlargement under heterogeneous conditions.

With regard to the specific critiques, we note the following. First, we disagree with the criticism of the image segmentation method. We applied the watershed transform (1), which is a widely used method for image segmentation. Having thoroughly tested the method, we are sure that our automated image processing produces meaningful results. Indeed, if the skeleton in Fig. 1D was superimposed on the original image, then one would find that what Prof. Weibel calls a “totally arbitrary skeleton” actually runs inside the walls seen on the original image (Fig. 1A). Furthermore, Prof. Weibel appears to unnecessarily mix physiological concepts with purely methodological ones, namely, the “free ends” on histological images vs. the open alveoli for gas exchange. Our method indeed requires the air spaces to be closed in order to determine the size of the alveoli. However, this has no relation to gas exchange.

The assertion that Fig. 2 is superfluous because it is a “self-correlation” is unclear. The Lm and the area-to-perimeter ratio were measured by independent methods, which are outlined in the relevant section. The correlation seen in Fig. 2, as well as the value of the slope, validates the methods used for evaluating each. In addition, the theoretical derivations of Eqs. 3 and 4 rely on the convexity of measured objects, especially when the spatial dimension is greater than two (8). Because air spaces are not always convex, Fig. 2 also confirms that the lack of convexity of the air spaces is not a problem in two dimensions.

Prof. Weibel suggests the use of linear intercept distributions or other stereological methods. We have shown that the linear intercept distribution cannot be used as an indicator of size because it is influenced by shape and not by size alone (Fig. 3 and Appendix A). The stereological methods to estimate the three-dimensional diameter of air spaces would be robust for large data sets (6). These methods also rely on assumptions that may become unrealistic in the emphysematous lung. For example, the progression of alveolar destruction has been shown to be influenced by mechanical forces, which leads to spatial correlations and air space sizes whose distribution exhibits a

*J Appl Physiol* • VOL 100 • APRIL 2006 • www.jap.org
heavy tail (12). These features of the alveolar structure may violate the assumption of uniformity and randomness invoked in standard stereological methods. As demonstrated in Fig. 6, our method is sufficient for detecting emphysema from two-dimensional slices without such assumptions about the underlying structure.

Finally, as with any scientific method, our indexes will likely be improved or replaced by newer and better ones. Meanwhile, our method could lead to a better understanding of structure-function relations and perhaps help uncover the underlying mechanisms of the progressive nature of emphysema.

REFERENCES


Harikrishnan Parameswaran
Arnab Majumdar
Satoru Ito
Adriano M. Alencar
Béla Suki
Dept. of Biomedical Engineering
Boston University
Boston, Massachusetts
e-mail: bsuki@bu.edu

COMMENT

Morphologic quantification of heterogeneous parenchyma

To the Editor: I appreciate the opportunity to comment on this rather controversial exchange between Drs. Weibel and Suki relating to a very timely and important issue. On the basis of the comments in each letter, it is clear that there is need for some better understanding of both the problem and solution, perhaps to be discussed at a future scientific workshop. In his letter, Dr. Weibel makes a strong case for the use of proper stereological analysis of histologic sections of the fixed lung. Of course, as the one who wrote the book (both figuratively and literally) on how to do this (6), it would be foolish for anyone to take serious issue with him. However, the issue that is raised by Dr. Parameswaran and colleagues (5) is what is the best way to determine the progression of heterogeneous emphysema, and the answer to this question may not necessarily require conventional stereology.

The issue emphasized by Dr. Weibel of biased and unbiased sampling is clearly important in proper stereology, but in this case it is perhaps a red herring. Of course nobody wants their data analysis to be called biased, but despite the concern that statisticians rightly have with biased statistical samples, there may be practical utility with a biased measurement if it reliably can help detect pathologic changes. Indeed it might even be desirable. The issue is how one defines bias, and here one must be careful to distinguish between statistical bias and sampling bias. A sampling bias would select regions of a section for analysis that look emphysematous, and this is clearly unacceptable. However, although the mean linear intercept is conventionally considered to provide an unbiased estimate of alveolar surface area, the selection of tissue regions that exclude airways and blood vessels itself introduces a sampling bias, because alveoli immediately adjacent to these airways and blood vessels may be under substantially different stresses than other alveoli (1). Thus it may not be unreasonable to use a metric that deliberately introduces mathematical bias to distort the image to emphasize certain anatomic or pathologic features.

In this regard, I admit to having always been somewhat skeptical of the accuracy and reliability of all published methods (including ones I have used) of determining alveolar volumes from histologic sections. In the best of these methods using serial sections with the selector or disector approaches (2–4), the number of alveoli sampled and measured varies from 30 to 200. Given that in the mouse, there may be some 20,000,000 alveoli, this is placing a lot of faith in conventional stereological analysis that look emphysematous, and this is clearly unacceptable. However, although the mean linear intercept is conventionally considered to provide an unbiased estimate of alveolar surface area, the selection of tissue regions that exclude airways and blood vessels itself introduces a sampling bias, because alveoli immediately adjacent to these airways and blood vessels may be under substantially different stresses than other alveoli (1). Thus it may not be unreasonable to use a metric that deliberately introduces mathematical bias to distort the image to emphasize certain anatomic or pathologic features.

In this regard, I admit to having always been somewhat skeptical of the accuracy and reliability of all published methods (including ones I have used) of determining alveolar volumes from histologic sections. In the best of these methods using serial sections with the selector or disector approaches (2–4), the number of alveoli sampled and measured varies from 30 to 200. Given that in the mouse, there may be some 20,000,000 alveoli, this is placing a lot of faith in the assumption that the parenchymal structure is homogeneous and isotropic. This assumption is surely not true in emphysema and, as already noted, may not even be true in a normal mouse lung. In such conditions, taking even random fractionator samples (3) of lung tissue and sections will yield biased results. One needs to know the nature of the anisotropy and heterogeneity to be able to apply conventional stereological analysis.

The problem stems from there being so little known about the details of three-dimensional lung structural changes in
Letters to the Editor

1422

emphysema, from the very early pathologic changes to the later stages with gross destruction. Do alveoli enlarge? Do the walls of alveoli retract? Do the alveolar mouths enlarge? Do the ducts enlarge? Do whole alveolar walls just collapse below a critical stabilizing stress? The processes in this pathology are not yet well enough understood to answer these questions at different time points, and until this ignorance is remedied, questions will remain as to how might one best detect the changes in emphysema. Thus an accurate partitioning of air space volume between alveoli and ducts or the calculation of alveolar number, becomes very difficult if not impossible with modern stereology. Parameswaran et al. (5) have accepted this limitation, but they have gone ahead to calculate an index that they claim to be more sensitive in its ability to detect early changes in emphysema. At present, it is not possible to tell how successful this new method of histologic image analysis will be, but if it can be validated by other investigators and used to reliably quantify the pathology in the face of such heterogeneity, then it may not matter whether or not it follows conventional stereological methods.

REFERENCES


Wayne Mitzner
Johns Hopkins University
Baltimore, Maryland
e-mail: wmitzner@jhsph.edu

COMMENT

To the Editor: A paper by Parameswaran et al. (1) that attempts to quantify heterogeneous airspace enlargement on histological sections by measuring the cross-sectional area of every visible airspace and deriving an equivalent diameter of a circle has triggered a critical letter by one proponent of stereology whose major criticism is that Parameswaran’s methods of sampling and measurement were biased. The authors responded that their methods did not use the basic tools of stereology; hence the criticism of stereological bias is irrelevant (1).

Without commenting directly on the above exchange, let us identify the fundamental approach of Parameswaran’s paper, namely the quantification of the physical property (size) of irregular 3-D objects (airspaces) from observations made on 2-D sections. This approach coincides with the precise definition of stereology! Because 2-D information incompletely describes 3-D objects, such information represents only probabilities of reality, leaving plenty of room for ambiguity and forming the basis of numerous “optical illusions” (Fig. 1). Since the 1600s, scientists from diverse disciplines have grappled with this ambiguity; from their collective insight, a set of statistical principles evolved specifically to minimize the risks of drawing biased 3-D conclusions from 2-D data. “Bias” refers to systematic errors that cause experimental estimates to deviate from the true value, such as those arising from the way the sections are sampled and measured. An unbiased method yields estimates that with repetition eventually converge on the true value with repetition. A biased method may initially yield similar results as an unbiased one but will not converge upon the true value. Sound stereological methods should characterize the 3-D structure without bias, using a minimum number of 2-D sections and yielding results that can clarify structure-function relationships.

The implicit assumption in Parameswaran’s method, that cross-sectional area of airspaces reflects their true size, presents another type of optical illusion. An image taken from their paper (Fig. 1C) evidently shows heterogeneous airspace sizes, or does it? Could these irregular airspaces extend above and below the section plane such that their 3-D volumes are similar regardless of cross-sectional area? Is it possible for an airspace to be transected more than once by a given plane? Nonuniform airspace distortion could stem from external compression, pathological tissue properties, or technical problems in lung inflation, fixation, or processing. Potential errors resulting from the false assumption of 2D-3D equivalence could be overcome neither by counting more pixels or airspaces on the section nor by applying “correction algorithms” because essential information in the third dimension remains missing. In contrast, sound stereological methods purposely avoid such pitfalls by incorporating steps into the study design that ensure 1) each unit volume (not area) of the structure has an equal probability of being sampled, 2) the physical features (e.g., volume or surface area) are measured without geometric assumption, and 3) the measurement is related to a known reference volume (e.g., volume-to-volume or surface-to-volume ratio). The strictly planar measurement of Parameswaran et al. does not fulfill any of these criteria or yield results that clarify the pathophysiology of emphysema (e.g., loss of gas exchange surface or elastic recoil). Therefore, my bias would be to avoid bias whenever possible by employing unbiased stereological methods. After all, unbiased 3-D sampling is an innate cortical function that permits the “intuitive” recognition of false 2-D perception. If you doubt this fact, try counting the legs of the elephant in Fig. 2.

REFERENCES

To the Editor: Controversial discussion of a paper in the Journal of Applied Physiology (11) leads to the question “which methods should be used to assess morphologic changes in emphysema?” From this author’s perspective, two aspects need to be addressed: first, which parameters are necessary to demonstrate the presence of emphysema; second, which methodological strategies ensure that the data recorded are reliable.

Emphysema is defined as “abnormal permanent enlargement of the airspaces distal to the terminal bronchioles, accompanied by destruction of their walls” (1). Unfortunately, most studies rely on quantification of airspace enlargement alone assessing mean linear intercept length or mean chord length as indexes of airspace size. These parameters are highly sensitive to inflation during fixation and tissue shrinkage during embedding (14, 16). As airspace size increases with age (3, 12), “abnormal enlargement” can only be demonstrated compared with age-matched control lungs and “permanent enlargement” only when additional groups are implemented to demonstrate persistence of airspace enlargement. However, to conclude that emphysema is present, it is not sufficient to reveal abnormal permanent airspace enlargement alone but to demonstrate destruction of alveolar walls (2), e.g., as a decrease in total alveolar wall
volume, total alveolar surface area, total capillary length, and/or total number of alveoli (5, 7, 10, 17).

As airspaces are three-dimensional (3-D) structures, characterized by volume, surface, and number, each technique attempting to measure these structures must consider their 3-D nature. This can be done along two perspectives: “airspace” can be considered as one compartment characterized by global parameters, e.g., total volume, total surface area, and mean linear intercept length. Airspace can also be considered as a population of subunits characterized by average values or distributions of subunit parameters, e.g., mean volume or mean surface area of an alveolus.

Although imaging techniques have developed considerably the analysis of lung microstructure in the living organ with intact 3-D relationships, histological analysis is still the gold standard (18). Thus the lung must be transformed into 2-D sections. For efficiency, measurements are performed on a collection of sections rather than serial sectioning the whole lung. To ensure that the sections (as well as the fields of view) selected for analysis are representative of the whole organ, i.e., that each part has the same probability of being selected, systematic uniform random sampling strategies have been developed (4, 8, 9). The statement of Parameswaran et al. (11) that “the assumption of uniformity and randomness (is) invoked in standard stereological methods” is a clear misinterpretation of stereological concepts. The sampling scheme is designed to uniformly cover all areas of the object in a systematic way to obtain random samples, which ensures that heterogeneously distributed features are represented according to their contribution to whole lung structure. Such approaches were successfully applied in emphysema studies (5, 7, 10, 17).

The probability of a 3-D object being represented in a section (we may disregard effects of orientation in the context of alveoli) is proportional to its volume (13). This is no problem if the reference volume (e.g., fixed lung volume) is known and the end point is a global parameter. However, it becomes a major problem if the recorded parameter is an average characterizing a population of 3-D subunits. Because the likelihood of an individual alveolus being sampled with a single 2-D section inevitably depends on its individual volume, collecting alveolar profiles with single sections leads to samples and measurements that are systematically biased toward larger alveoli. This fact was ignored in the paper under debate (11). Stating that only 2-D parameter-like areas (the 2-D representation of a 3-D volume!) were measured using non-stereological tools and, therefore, the bias introduced by single section sampling can be ignored, is unacceptable. It is not stereology that dictates 3-D interdependencies, it is the 3-D nature of the airspaces that does. Stereology has developed strategies to perform measurements on 2-D sections without the necessity to ignore such fundamental relationships (8, 9, 15).

This author strongly believes that the scientific community in respiratory research needs a clear statement toward stereology-based methods such as that made by the Journal of the American Society of Nephrology requesting “that appropriate stereologic methods be used to quantify structures in tissue sections in all manuscripts submitted to the Journal” (6).

REFERENCES

Heinz Fehrenbach
Clinical Research Group “Chronic Airway Diseases”
Clinic of Internal Medicine (Respiratory Medicine)
Philippines-University of Marburg
Marburg, Germany
e-mail: fehrenba@med.uni-marburg.de

COMMENT
To the Editor: The current paper of Parameswaran et al. (1) introduces new ideas for morphometric characterization of pulmonary parenchyma that may be of use especially in quantitatively distinguishing emphysematous lungs from normal lungs. Weibel’s letter raises several challenges; Parameswaran et al. have offered rebuttals to these. A reconciliation of these issues may be found by focusing on the specific goals of this
work and the intrinsic ambiguities in both morphometry and stereology.

Goals. Weibel argues that, in contrast to the mean linear intercept $L_m$, the distribution and moments of area equivalent diameters suffer from sampling bias and, by definition, a restriction to the 2-D section. This is true, but it needs to be seen in the context of Parameswaran et al.’s explicit goal, which is to consider a new method to morphometrically quantify the severity of emphysema and not to make quantitative stereologic claims about the underlying 3-D structure. Given this, the utility of this approach is to be found and assessed empirically, especially as shown in their Fig. 6. Weibel’s comment about the unfairness of a comparison of their $D_1$ and $D_2$ indexes with $L_m$ is perfectly valid, because simple means carry no information about spatial heterogeneity. Open questions remaining include a comparison with other measures of geometric dispersion and determining the probabilistic structure of the equivalent isoarea diameter distribution for realistic model geometries of normal and emphysematous lungs.

Concepts. I believe a major, and underappreciated, difficulty here is conceptual ambiguity in the intuitive ideas of “airspace,” “size or dimensions,” and “shape.” Apart from examples such as the well-defined surface density or volumetric tissue density (per sample volume), other concepts are intrinsically (and hence mathematically) ambiguous. For example, the fact that, in the absence of frank airway closure, the real airspace is completely continuous implies that any kind of fractionation into distinct “spaces” requires a concept of bounding surfaces separating entities such as alveoli, ducts, and airways. Absent such clearly definable bounding surfaces, one is forced to include the morphometric methodology itself in their definitions. From this point of view, the 2-D “airspace” used by Parameswaran et al. is nothing more nor less than what they define it to be algorithmically.

But even given operationally defined airspaces bounded by closed curves postprocessing, the ideas of size or dimensions and shape are similarly ambiguous. Consider the apparently irreconcilable statements that $L_m$ does not depend on shape (Weibel) and that $L_m$ does depend on shape (Parameswaran et al.). In fact, both statements are true! The difference is in what features of size or dimensions are held fixed when shapes are changed. For example, if the structural property of volumetric surface density or areal boundary density is fixed, then it is true that $L_m$ does not depend on shape. By contrast, if the structural volume or section area is held fixed while shape is changed, then, because the surface area or boundary lengths must change to preserve the volume or area, $L_m$ must also change, showing that it does depend on shape. Clarity here is especially needed because of the importance of fixing or controlling lung volumes when assessing the degree of airspace “enlargement.”

In summary, both 2-D morphometry on sections and its stereological extension to 3-D will benefit from careful definitions of distinguishable airspaces (hopefully based on functional consequences to either mechanics or gas exchange), the important definable measures of size or shape, and how best to implement image analysis to characterize these.

ACKNOWLEDGMENTS

I thank F. G. Hoppin, Jr., for his, as always, insightful ideas and help in preparing this comment.

REFERENCES


James P. Butler
Physiology Program
Harvard School of Public Health
Boston, Massachusetts
e-mail: jbutler@hsph.harvard.edu