

Applications of Modern Stereology to Quantitative Microscopy

P. R. Mouton*

*Stereology Resource Center, Baltimore, MD, 21224

Modern stereology is carefully designed, hence the alias *design-based* stereology, to address a ubiquitous yet deceptively simple problem in quantitative microscopy: How to quantify the number, size, surface area, and length of biological objects present in 3-D tissue, based on measurements made on 2-D sections, *without the introduction of methodological bias*. Within the past two decades, research scientists have increasingly expressed preferences for these principles for studies that involve quantification of biological features on microscopic images. In 1996, Dr. Clifford Saper, Professor of Neurology at the Harvard Medical School and Editor-in-Chief of the *Journal of Comparative Neurology*, published an unequivocal and highly unusual editorial to the journal readership that stated, in effect, the journal's intention for future submissions with quantitative studies of biological structure to either: 1) use unbiased stereology approaches; or, 2) provide an explanation as to why these methods are not applicable [1]. Several other peer-review journals, including the *Journal of Microscopy*, had at that time similarly rigid though less explicit editorial guidelines in place, and several other journals have adapted similar approaches in the past decade. In the past decade, study section reviewers for federal and private funding agencies, as well as regulatory agencies that monitor drug development and safety, have reached similar conclusions with the recognition that data based on methodologically biased approaches fail to meet the highest standards available for quantification of biological morphology. Since journal publications, extramural funding, and regulatory approval provide important indices of progress and productivity in academic research, government agencies, and the pharmaceutical industry, modern stereology has become a *sine qua non* for morphological studies across a wide range of disciplines, including neuroscience, pathology, and toxicology. Among the applications of these approaches to human diseases and experimental research are accurate quantification of total numbers of discrete cells in a population, a measure of degeneration, apoptosis, or proliferation; total length of neurotic fibers and blood vessels to assess angiogenesis and sprouting; total surface area of absorptive membranes; and total volume of cells and tissues as measures of atrophy and hypertrophy.

Like a well-designed Gallop poll, unbiased stereology uses a combination of unbiased sampling and geometric probes to quantify first-order stereological parameters [Number (N), Length (L), Surface Area (S), Volume (V)] of macro- and microscopic structures [2]. The "unbiasedness" refers to the observation that in the absence of methodological bias, parameter estimates for biological features in a small sample of subjects (e.g., groups of $n = 5-10$ mice) provide adequate information for accurate estimates of the expected (true) value of parameters. This goal has been realized by the identification and elimination of all known sources of methodological bias from morphological studies, starting with a basis for stochastic geometry and probability theory, rather Euclidean geometry, to extract quantitative information from microscopic images.

Understanding the distinction between accuracy and precision helps to differentiate biased and unbiased approaches in the field of quantitative microscopy. When a quantitative method avoids all known sources of bias, increased sampling of tissues within a group of randomly selected subjects, e.g., the control group, generates data for individual subjects that clusters around the true central tendency, such as the mean total number of cells in a defined anatomical structure. With biased methods, i.e., approaches based on non-verifiable assumptions and faulty models, the data points for

each subject may also cluster around a central value; however, those data will cluster a point some unknown distance from the true value. Because this bias (systematic error) cannot be quantified, the bias cannot be eliminated by *post-hoc* “correction” factors. Instead, these Euclidean-based equations introduce further bias through false assumptions and faulty models (“assume cells are spheres”); random variations in biological structures; and, treatment effects and tissue processing artifacts. Through avoidance of these and other well-known sources of bias, design-based stereology generates accurate estimates of biological structure. With regard to precision, both biased and unbiased approaches can generate data with low variability, i.e., high levels of precision, which simply depends on the level of effort expended to sample the tissue. The critical distinction is that both biased or unbiased methods can generate *precise* results; however, only unbiased methods can generate *accurate* results.

Tissue preparation for design-based stereology follows traditional protocols, with a few exceptions to ensure unbiased sampling. For estimation of N of biological features, tissue is cut serially along any preferred axis (e.g., coronal) and then subsampled into ~ 8-12 sections. The next step is to define a clear and anatomically well-defined reference space, the region of tissue that contains the biological features of interest. For a typical reference space (e.g., mouse neocortex, rat liver), tissue may be cut into 200 sections at a microtome setting of 40 μ m. Sampling every 20th section with a random start in the first 20 sections will yield ten sets of sections with n=10 sections in each set; each of these 10 sets of sections contains a systematic-random sample through the reference space. One set of these sampled sections is stained to ensure that all biological features of interest can be reliably identified, thus avoiding recognition bias. Rather than counting all features in the tissue, the stereology uses a virtual 3-D probe, a disector, to make accurate estimates of the *probability of an intersection* between the probe and the feature of interest. Like all stereology probes, the disector does not allow variations in size, shape, and orientation of the features to adversely affect (bias) this intersection count; only the true number of features determines whether the disector intersects the feature. By counting the occurrence of either one of two random events, like flipping a coin to estimate the probability of “heads,” repeating the procedure about 100 and 200 times leads to a reliable estimate of the probability of a disector-feature intersection. From this known probability, a relatively simple formulas lead to an accurate (unbiased) estimate of the number of features through the reference space. This approach is entirely independent of the spatial distribution of the biological features, since more or less heterogeneous distributions of features means more or less sampling, i.e., 200 vs. 100 repetitions, respectively, to achieve a stable estimate of the intersection probability. However, regardless of the distribution of features in the tissue, like flipping a coin to estimate the probability of heads, repeating this sampling procedure more than 200 times in the same reference space is simply a waste of time and effort.

In the past decade the combination of video-microscopy with computer hardware and software systems has dramatically increased the throughput time for analysis of tissue sections using unbiased stereological approaches. Since this technology, combined with accurate, precise, and efficient stereological approaches, increases the number of hypotheses that can be tested per amount of work-effort, time, and material resources, computerized applications of unbiased stereology may help in the development of novel approaches for the therapeutic management of human diseases.

[1] Saper, CB. *J Comp Neurol* (1996) 364(1): 5.

[2] Mouton, P.R. *Principles and Practices of Unbiased Stereology: An Introduction For Bioscientists*” The Johns Hopkins University Press, Baltimore, 2002.