ACTA STEREOL 1999; 18/2: 219-224
ORIGINAL SCIENTIFIC PAPER

ON NUMERICAL ESTIMATES FROM HISTOLOGICAL MATERIAL: THE QUESTION OF CALIBRATION

Richard E. COGGESHALL

Department of Anatomy & Neurosciences, Marine Biomedical Institute, The University of Texas Medical Branch, Galveston, Texas 77555-1069

ABSTRACT

The present work concerns the evaluation of various types of methods for determining numerical data from histologic sections. Major points are:

1. Simple profile counts or estimates from assumption based methods require validation by calibration if the estimates are to be accepted.

2. Calibration is thus a useful thing, but it should be understood that one is not calibrating methods, one is determining whether particular estimates from any particular method are unbiased.

3. Whenever conditions change, assumption based estimates must be recalibrated.

4. Since calibration requires comparison against unbiased estimates, it is usually simpler to use the unbiased estimates.

Key words: assumption based methods, stereology

INTRODUCTION

One purpose of a special journal issue on stereology is to present new developments in the theory of stereology or exciting findings that result from presently available stereological technology. An equally important goal is evangelical, to explain extant theory and procedures so that efficient unbiased numerical estimates become the rule rather than the exception. This paper falls in the second category. The discussion will focus on two earlier papers (Coggeshall, 1984; Pover et al., 1991) that illustrate some of the common problems faced by investigators attempting to obtain unbiased numerical estimates for their material.

General Comments

The major problems in estimating numbers of objects from histological material are that many of the objects are split during the sectioning process so that there are more profiles than objects (a profile is what appears of an object in a histological section), and that the probability of 'hitting' an object and thus generating a profile is related to object 'height' which in turn is dependent on object size, shape and orientation as well as object numbers. Three types of methods are used to derive object counts from histological profiles; 1) simply counting profiles (profile counts), 2) using assumptions that convert profile counts to object counts (assumption based methods) and 3) reconstructions (Coggeshall, 1992; Coggeshall et al., 1996). It is clear that serial or disector reconstructions give unbiased estimates if proper sampling is used (Coggeshall et al., 1996). Because of ease of use, however, most investigators use profile or assumption based methods (Oorschot, 1994; Coggeshall et al., 1996). There is disagreement as to whether estimates from these methods are unbiased. This is an important issue because unbiased estimates are necessary for solid conclusions.

Profile Counts

Profile counts can be regarded as an assumption based method with the main assumption being that each profile uniquely identifies an object. Unfortunately this assumption is false whenever a microtome knife splits an object, which is almost always. Frequently, however, investigators do not use profile counts to estimate numbers of objects, rather they compare ratios of profile counts in control and experimental situations, the assumption being that the same biases are present in both situations and so cancel (Oorschot, 1994; Coggeshall et al., 1996). To this one can only say that a further assumption is now necessary, which is that objects do not change in any way in going from the control to the experimental situation (such changes change numbers of profiles per object). It is also impossible to tell whether changes have occurred or not from inspection of profiles. So the biases usually compound, and one cannot tell how severe the biases are from examining profiles. Thus, in my opinion, calibration (comparison of profile based estimates with unbiased estimates from reconstruction methods) must be done if profile estimates are to be accepted.

Assumption Based methods

Assumption based methods follow common patterns: 1) profiles are counted, 2) some parameters other than profile numbers (e.g., nuclear diameters) are measured and 3) assumptions are made that allow the estimation of numbers of objects. As an example of the most commonly used assumption based method (Abercrombie, 1946; Williams et al., 1988), nuclear profiles (n) are counted, mean nuclear diameters (D) and section thicknesses (T) are measured, nuclei are assumed to be spherical, and then numbers are estimated by the formula $N=nxT/T+D$ (Abercrombie, 1946). This and similar methods are efficient, but how would one know whether the assumptions are met? For example, nuclei are never perfectly spherical, and there is no easy way to determine if the nuclei are spherical enough.

A more serious problem is that it is impossible to measure the necessary parameters from profiles in single sections. For example, 'D' in the method above is the height of the object being counted orthogonal to the plane of section (Abercrombie, 1946). If the objects are spheres, as is

assumed, the height is the mean diameter of the spheres. But investigators cannot measure these diameters, all they can do is measure diameters of nuclear profiles, and the two are not the same. As an example of one problem, measuring diameters implies that the measurements are from the centers of the spheres, which is impossible to validate if one is looking at single profiles. Proponents of assumption-based methods might protest by saying that the measured diameters are from nuclear profiles that are large and round and contain nucleolar profiles. But nucleoli are usually not in the center of the nucleus, and one does not have cut through the center of a nucleus to get a large round profile. To generalize, the parameters that are necessary for obtaining unbiased estimates from the assumption based methods are 3-dimensional (e.g. diameters or volumes of spheres), and these cannot be accurately assayed by 2-dimensional measurements (e.g. diameters or areas of profiles). Thus, assumption based assays are always biased by unknown amounts. These biases often lead to inaccurate estimates (Coggeshall et al., 1990; Pakkenberg et al., 1991). For these reasons, assumption based estimates should be calibrated.

The empirical method.

In an attempt to deal with the above problems, an 'empirical' method was proposed (Coggeshall, 1984). The reasoning is simple; total profile numbers (n) times the reciprocal of numbers of profiles per object in a sample of objects (Nsamp/nsamp) estimate numbers of objects (N=nxNsamp/nsamp). Unbiased profile numbers are easily estimated by counting in selected sections and multiplying by section separation. Mean numbers of profiles per object are easily determined by taking a small number of objects and counting the profiles seen in all sections of these objects (cf numbers of nuclear profiles in all profiles of selected neurons). Calibration showed that dorsal root ganglion cell numbers were accurately estimated (Coggeshall, 1984). Advantages are that one does not measure anything, one simply counts profiles, and the counts are done primarily on single sections, reconstructions being done only for a small population of objects to determine numbers of profiles per object. What then is 'wrong' with this procedure, and what lessons can we draw from this that would aid those interested in getting unbiased counts?

What is 'wrong' with this procedure is that the selection of objects in the sample is usually biased. The reason is that object size, shape and orientation influence the numbers of profiles when objects are cut during histological sectioning. For example, large objects are cut into more profiles than small ones. Consequently if one selects neurons (objects) by choosing clearly visible neuronal profiles containing nuclei, as is usually done, disproportionate numbers of large neurons will be chosen. Why then did calibration give accurate estimates for our material? The answer is that guided by experience we compensated by picking a sample of DRG profiles that represented the whole population. In doing this, however, we used one bias to remove another. So if we looked at another population of cells or if an experimental procedure changed cell sizes, this method would result in inaccurate estimates. If one desired to use this procedure, therefore, one would have to choose the sample population in which to count profiles by an unbiased 3-dimensional method such as a disector analysis or calibrate the estimates as for profile and other assumption-based estimates.

Several lessons can be learned from this. First, calibration is necessary to demonstrate that numerical estimates resulting from any of these methods are not biased. Secondly, one cannot

calibrate a method once and then use it indefinitely. Every time conditions change, for example after an experimental manipulation, recalibration must be done. To argue by analogy, one has no more justification in using an assumption based method indefinitely after a calibration than one does in using a spectrophotometer with no adjustments after the first blank is tested. Finally, since an unbiased estimate is necessary to calibrate the above estimates, it would usually be more efficient to simply take the unbiased estimate rather than using it to verify another type of estimate.

Calibrating disector analyses and a consideration of edge recognition (peripheral caps).

To consider these issues further, we compared disector and serial reconstruction analyses and reported that disector analyses resulted in inaccurate estimates when reference and look-up sections were adjacent (Pover et al., 1991). This paper is sometimes cited as evidence that disector analyses can produce biased estimates. Thus this issue needs to be reconsidered.

Disector analyses are of two types; physical and optical (a similar procedure to the latter being called 3-dimensional counting) (Gundersen et al., 1988a; Gundersen et al., 1988b; West, 1993; Williams et al., 1988; Williams et al., 1989). The physical disector compares one section with another, the optical disector counts between two optical planes within one section. Virtues of these procedures are that they remove biases that arise from object size, shape, and orientation that invalidate most profile and assumption based estimates, and they are suitable for sampling. An important point is that disector analyses and serial reconstructions are essentially the same in that if one does a total disector analysis, one is doing a serial reconstruction (Coggeshall et al., 1996). Thus when one is comparing disector estimates with numbers derived from serial reconstructions, one is comparing slightly different ways of doing the same thing. Yet in the study under consideration (Pover et al., 1991), the estimates from the two methods were different. So what lessons can we derive that would be of use to those interested in obtaining unbiased numerical estimates.

For the specific issue as to what was wrong with the study under consideration (Pover et al., 1991), reexamination of the same sections on which our original estimates were done reveals that different criteria were used for recognizing edges of cells (tops). This was because when we followed profiles of cells in the serial reconstructions, it was clear which profiles belonged to cells and which were thickenings in the cytoplasm, edges of blood vessels, etc. By contrast when doing adjacent section disectors, we were not following cells through serial sections, so edges were more problematic and fewer edges (peripheral caps, tops, and counts) were identified. When one or more sections were between the look-up and reference sections, however, the cells were followed through at least 3 sections, so the same criteria were used for edges as in the serial reconstructions. Thus the criteria for identifying profiles were not constant, and so the estimates differed.

More general conclusions are first that the study under consideration (Pover et al., 1991) cannot be used to indicate that disector analyses are biased when look-up and reference sections are adjacent. Secondly, methods that convert profile counts into object numbers can give unbiased estimates if their assumptions are fully met, but it should be understood that these assumptions cannot be validated from examining profiles. In addition the necessary parameters

for the assumption-based methods cannot be measured from single profiles as is usually done. Finally, no matter what method is used, attention still has to be paid to such issues as sampling strategies, correct profile identification, etc.

Final conclusions

1. Calibration is necessary to validate estimates from profile counts or assumption-based methods, but it should be understood that the calibrators are not evaluating methods, they are evaluating particular estimates. As a corollary, the particular studies considered here cannot be cited as validating or invalidating methods, they are validating or invalidating particular estimates derived from these methods.

2. Since profile or assumption based estimates require calibrations against unbiased estimates, it is usually more efficient to use the unbiased estimates.

3. If estimates from different apparently unbiased procedures are compared against each other and they differ, one should search for inconsistencies in such things as recognition criteria and sampling strategies between the different analyses.

ACKNOWLEDGEMENTS

This work is supported by NIH grants NS10161 and NS11255.

REFERENCES

Abercrombie M. Estimation of nuclear population from microtome sections. Anat Rec 1946; 94: 239-247.

Coggeshall RE, Chung K, Greenwood D, and Hulsebosch CE. An empirical method for converting nucleolar counts to neuronal numbers. J Neurosci Meth 1984; 12: 125-132.

Coggeshall RE. A consideration of neural counting methods. TINS 1992; 15: 9-14.

Coggeshall RE, LaForte R, and Klein CM. Calibration of methods for determining neuronal numbers. J Neurosci Meth 1990; 35: 187-194.

Coggeshall RE and Lekan HA. Commentary on methods for determining numbers of cells and synapses. A case for more uniform standards of review. J Comp Neurol 1996; 364: 6-15.

Gundersen HJ, Bagger P, Bendtsen TF, Evans SM, Korbo L, Marcussen N, Moller A, Nielsen K, Nyengaard JR, Pakkenberg B, Sorensen FB, Vesterby A, and West MJ. The new stereological tools: Disector, fractionator, nucleator and point sampled intercepts and their use in pathological research and diagnosis. APMIS 1988a; 96: 857-881.

Gundersen HJ, Bendtsen TF, Korbo L, Marcussen N, Moeller A, Nielsen K, Nyengaard JR, Pakkenberg B, Sorensen FB, Vesterby A, and West MJ. Some new, simple and efficient stereological methods and their use in pathological research and diagnosis. APMIS 1988b; 96: 379-394.

Oorschot DE. Are you using neuronal densities, synaptic densities or neurochemical densities as your definitive data? There is a better way to go. Prog Neurobiol 1994; 44: 233-247.

Pakkenberg B, Moeller A, Gundersen HJ, Dam AM, and Pakkenberg H. The absolute number of nerve cells in substantia nigra in normal subjects and in patients with Parkinson's disease

estimated with an unbiased stereological method. J Neurol Neurosurg Psychiatry 1991; 54: 30-34.

Pover CM and Coggeshall RE. Verification of the disector method for counting neurons, with comments on the empirical method. Anat Rec 1991; 231: 573-578.

West MJ. New stereological methods for counting neurons. Neurobiol Aging 1993; 14: 275-285.

Williams RW and Rakic P. Three dimensional counting: An accurate and direct method to estimate numbers of cells in sectioned material. J Comp Neurol 1988; 278: 344-352.

Williams RW and Rakic P. Three-dimensional counting: an accurate and direct method to estimate numbers of cells in sectioned material. J Comp Neurol 1989; 281: 335.