



UNIVERSITÀ DEGLI STUDI DI TORINO

This is an author version of the contribution published on:

Questa è la versione dell'autore dell'opera:

*Cell Tissue Res. 2015 Apr;360(1):5-12. doi: 10.1007/s00441-015-2143-6. Epub 2015
Mar 6.*

The definitive version is available at:

La versione definitiva è disponibile alla URL:

<http://link.springer.com/article/10.1007%2Fs00441-015-2143-6>

Update on stereology for light microscopy

Stefano Geuna^{1*} and Celia Herrera-Rincon²

¹ Neuroscience Institute of the Cavalieri Ottolenghi Foundation & Department of Clinical and Biological Sciences, University of Turin, Italy

² Neurocomputing and Neurorobotics Research Group, Universidad Complutense de Madrid, Spain

Abstract

The quantitative investigation of images taken from light microscopy observation is one of the pillars of biological and biomedical investigation. The main objective is the count of objects, usually cells. In addition, the measurement of several morphological parameters, such as the diameter of cells, the length of vessels, etc., can also be important for the quantitative assessment of the features of a tissue. Whereas counting and measuring histological elements may appear easy, especially today with the availability of dedicated software, in fact it is not, since what we can count and measure on light microscopy images are not the true histological elements but actually profiles of them. Obviously, the number and size of profiles of an object do not correspond to the object number and size and thus significant mistakes can be made in the interpretation of the quantitative data obtained from profiles. To cope with this problem, over the last decades a number of design-based stereological tools have been developed in order to obtain unbiased and reliable quantitative estimates of cell and tissue elements that originate from light microscopy images. This paper reviews the basic principles of the stereological tools from the first disector applications through some of the most recently devised methods.

Key Words

Morphometry, total number, cell size, fiber length, tissue section thickness, disector.

Introduction

The introduction of stereology in the biomedical field has been a major advance over the last 30 years (Geuna 2005). Although the application of stereological principles to biomedicine has been described already in the 1970s (Cruz-Orive 1976a,b), it has been only after the publication of the seminal paper entitled “The Unbiased Estimation of Number and Sizes of Arbitrary Particles Using the Disector”, by an author using the pseudonym D.C. Sterio (1984), that the “revolution of counting tops” began to spread in the scientific community (Geuna 2005).

Since then, stereology saw a progressive, though slow, spread in the scientific community (Gundersen et al. 1988a,b; West 1999; Benes and Lange 2001; von Bartheld 2002; Schmitz and Hof 2005; Kristiansen and Nyengaard 2012; Walloe et al. 2014). To estimate the spread of stereological methods, we carried an usage survey applying the same approach used by Coggeshall and Lekan in 1996 and von Bartheld in 2002, namely 100 research articles published in *Journal of Neuroscience*, *Journal of Comparative Neurology* and *Brain Research* were analyzed in order to determine the use of different counting methods (in our survey the sampled articles were the first 100 published in 2014 while in previous survey the reference year were 1994 and 2001 respectively). As can be seen (Tab. 1) usage of design-based methods is increased in comparison to the previous surveys, though biased profile-based counting (see next paragraph) is still the most frequently employed procedure.

It is beyond the aim of this paper to describe all the many stereological estimators that have been developed so far. Instead, the aim of this review is to briefly review the basic stereological principles and methods starting from the first disector application through some of the more recent advancements in this field.

The physical disector principle and the concept of design-based sampling

The disector is a three-dimensional counting probe that allows to create a small sample of histological particles (usually cells) which are representative of the entire particles' population thus

allowing reliable *statistical inference*, i.e. the process of extension from quantitative data obtained in the particular to conclusions that refer to the general (Cassel et al. 1993). If the sampling strategy is inadequate, the investigator will infer erroneous conclusions.

The disector principle is based on sampling particles on sets of pairs of parallel histological sections placed at a given distance from each other, thus creating a 3D sampling probe (Fig. 1a). The investigator selects only particles that appear in one of the two sections, the so called *reference section*, not in the other, the so called *look-up section* (Fig. 2). In other words, the investigator selects the “tops” of the particles, i.e. their first edge point that encounters the progressing plane of observation (Coggeshall 1992; Geuna 2005). Being a “point”, the top is a-dimensional and it has no shape and orientation that can influence the probability to be sampled or not. Each top has the same probability of being sampled thus meeting the “equal opportunity rule”, the basic requirement for random sampling (Geuna 2000).

In this view, the disector is a design-based sampling method, i.e. a procedure aimed at ascertaining that all particles in the sampling space have the same chance of being sampled. A system of sampling rules (the “design”) is adopted so that the morphological variability of the object (their size, shape and orientation as well as their isotropic distribution in the histological structure) do not influence the probability of each object being sampled. Design-based sampling can thus be adopted without making any preliminary assumptions about the morphology of the tissue/organ under analysis.

The introduction of design-based sampling constituted a clear breakpoint in quantitative morphology since the commonly methods used previously were based on model-based sampling. A “model” is represented by a theoretical construct built up based on a priori assumptions about the histological variables of a tissue/organ. The model allows to deal with the differential sampling probability of particles by “weighting” the rough numerical data. An example is Abercrombie’s method (Abercrombie 1946; Hedreen 1998) that allows to weigh the number of cells sampled based on the mean diameter (measured on the z axis) of that cell population. Since the variability in the

extension on the z-axis influences the probability of being sampled in more sections, data on cell counts are corrected based on the mean height of cells measured on the z-axis. Noteworthy, while most authors refer and use to the “model-based” Abercrombie method (Abercrombie 1946, pag. 240), this author also describes a second method (pag. 244) which, in fact, can be seen as the first unbiased “design-based” methods for particle counting in histological sections. This method is based on cutting alternate sections at two thicknesses which are as different as possible (the author made the example of 5 μ m and 12 μ m). Then, the particle profiles (e.g. cell nuclei) are counted in both sections. The difference in the profile counts at these two thicknesses is the true number of particles in 7 μ m (the difference between the thicker and the thinner section).

While Abercrombie’s methods were suitable for coping with size-related bias in most cases (Geuna 2000), what makes the disector a seminal tool which has revolutionized the approach to quantitative morphology is its clear advance over the still widely used simple profile sampling (while the Abercrombie’s methods were still based on profile counts). Simple profile-based sampling, i.e. the sampling of objects’ profiles on one section, is based on the assumption that the number of cross-sectional profiles is directly proportional to the number of objects and thus no correction factor is necessary for converting profile number to object number. Clearly, this assumption is wrong since the number of profiles is almost always larger than the number of objects and thus number estimates will be biased due to size-related differences in the probability of being sampled (larger objects have a higher probability of being detected in more than one section). In practical terms, an increase in the size of cells will be erroneously interpreted as an increase in their number (West 1993; Coggeshall and Lekan 1996).

A potential source of confusion is the use of different terms which focus on the different features/properties of the disector (Benes and Lange 2001; West and Slomanka 2001), namely: (i) “disector probe” refers to its property of creating a 3-D volume and the set of rules which allow to determine when an object is inside or outside the volume (in order to avoid counting the object more than once). (ii) “disector method” refers to the possibility to use the disector’s principles for the

unbiased estimation of the total number of objects in a tissue/organ; (iii) “disector sampling” refers to the use of disector probes to select a representative sample of objects ensuring that each object has the same probability of being sampled.

The optical disector

The disector procedure is based on the use of pairs of parallel histological sections (and has been lately renamed “physical disector”); though reliable, it proved to be very time-consuming and inefficient in an age when digital histological images were not available. Thus, an important advancement was made by the development of the optical disector, which is represented by a 3D sampling probe created by means of successive focal planes in a thick section (Fig. 1b); the particles are then sampled when they first come into focus within the sampling volume (Gundersen et al. 1988b) i.e. when their “top” meets the observation plane moving along the z-axis. This procedure makes the sampling of tops faster and thus the procedure more efficient. However, it should be pointed out that the optical image of the top is not a-dimensional and sometimes not easy to be unequivocally detected (Guillery 2002).

Systematic random sampling

As previously mentioned, randomness (i.e. to assure that all particles in the sampling space have the same chance of being sampled) is the main goal of design-based sampling. Whereas the use of disector probes can guarantee randomness in each pair of histological section (for the physical disector) or each single histological section (for the optical disector), randomness should be also guaranteed with regard to the selection of the section pairs or single sections. This most efficient method is *systematic random sampling* (Gundersen et al. 1999) that is based on the systematic selection of every n^{th} section of the tissue/organ from one randomly selected starting section (where

n is the distance between serial sections of the whole tissue/organ that is preliminary decided upon in relation to the amount of sampling required).

The fractionator

The combined use of disector (physical or optical) probes and systematic random sampling allows to easily accomplish the unbiased estimation of the total number of objects in a given anatomical/histological region. This is obtained by calculating the mean density of objects in the randomly selected disector volumes and then by multiplying the density by the total volume of the region in which the objects are distributed.

Unfortunately, it is not always easy to clearly determine and measure the volume of the region under investigation, making thus more complex the counts of objects based on their relative density in disector volumes. Thus, the *fractionator* technique has been developed based on the combination of disector probe counting with a fractionator sampling design (Gundersen et al. 1988b). Its aim is to obtain objects' total number estimation without the need of measuring the total volume in which the objects are distributed. In fact, the samples of objects are collected so that they constitute a known fraction of the whole object population and then the number of objects is simply estimated by dividing the number counted in the sample by the fraction.

An interesting modification of the fractionator is the isotropic fractionator (Herculano-Houzel et al. 2015), which allows fast and inexpensive quantification of total numbers of cells in a whole organ, with the only main disadvantage that it provides no spatial information on the cellular location.

The proportionator

Whereas the “traditional” stereological tools, such as the disector/fractionator (both physical and optical) have proven to be solid and reliable approaches for quantitative morphology of all tissues and organs, they present some limitations in terms of efficiency. Therefore, new

methods have been more recently developed to improve the efficiency of stereological tools. One example is the *proportionator* (Gardi et al. 2008). The method takes advantage of today's availability of software and workstations for automatic image analysis and it is based on a two-step approach. First the software automatically collects some relevant information about all parts of a section and, using some predefined algorithms measures automatically the amount of information (for instance the amount of a specific staining). Then, the software selects a number of the microscopy fields, each with a probability proportional to the amount of information. In the second step, the researcher uses manually the sampling probe (e.g. the disector) in each selected field in order to estimate total cell number.

Counting versus measuring

The by far most frequent goal in quantitative morphology is the counting of objects, usually cells, for which several reliable and efficient stereological tools have been devised as described in the previous paragraph. However, besides counting, also measuring objects on histological slides can provide valuable information on the processes that are taking place in cells and tissues.

Several stereological tools are available also for the reliable and efficient measurement of cells and tissues. In the following paragraphs, we will describe one of the first methods described for cell and tissue measurement (the *nucleator*) together with a more recently devised method (the *spatial rotator*).

The nucleator

One of the first and still most used stereological size estimators is the nucleator (Gundersen 1988b). Once an object (e.g. a cell) is sampled using an unbiased probe (e.g. the disector), its size is estimated by placing one or more pairs of perpendicular lines on it and then identifying the intersection points between the lines and object boundaries (Fig. 3a,c).

In the original application the determination of the intersection points was made by the researchers manually. Today, this can be done automatically using dedicated software thus allowing to calculate a large number of intersections and increasing the precision of the estimates (Jensen 2000). This method has been named the *integrated nucleator* (Hansen et al. 2011) and it is based on the assumption that the identification boundaries made by the software are correct, a condition that should always be carefully verified since it might not be met in case of irregularities in the staining of cells and tissues.

It has been shown that the classical nucleator is sufficiently precise when the reference point used for sampling objects is centrally positioned and the objects have a spherical shape (Jensen 2000). By contrast, when objects are sampled using a reference point that is not uniformly located and/or the objects have an irregular shape, the integrated nucleator should be preferred.

More recently an intermediate option (the semi-automatic nucleator) between the classical and the integrated nucleator has been described (Hansen et al. 2011): In this method, first, boundary intersections are automatically identified by the software. Then, the researcher verifies whether or not the identification of the intersections is satisfactory, with the possibility to correct it in case the identification is judged as incorrect.

The spatial rotator

Among the development of size estimators that have been more recently proposed (which are all based on the identification of boundary intersection points) the spatial rotator is particularly powerful, since it does not require randomization in the sectioning process and/or viewing direction.. In addition, the spatial rotator uses also information available in the 3D space. The method is based on using test lines in several planes at different optical depths in thick sections. In contrast to an original method devised by Tandrup et al. (1997) and named *optical rotator*, the spatial rotator is based on the use of only one test line in each focal plane (Fig. 3b,d), a

feature that makes its use much faster and efficient (Rasmusson et al. 2013). Once intersections are identified, the Cavalieri principle is then used to estimate the volume of the object.

Debated issues

Although the experience of about 30 years tells us that stereological methods are valuable tools for the quantitative morphology of all tissues and organs, nonetheless it should be always kept in mind that there is no absolutely correct procedure for solving problems that involve human inductive reasoning (Smith 1994); thus, stereological tools should not be considered a priori better than other methods and other methods should not be rejected a priori, and stereological data too should be always dealt with caution (Geuna 2005).

One of the limitations that affects stereological methods is related to the problem of tissue shrinkage and z-axis distortion (Guillery 2002; Gardella et al. 2003). The measurement of disector thickness along the z-axis can be influenced by many variables related to both the slice (in particular tissue shrinkage and irregularity in the section surface) and the optics (e.g. thickness of coverslips and type of lens). Yet, the assumption that a top of an object is a point and thus a-dimensional is true in theory but it is usually not true in practice when considering the object's optical image that is what the investigator has to deal with (Guillery 2002). It has been proposed that potential bias originating from non-uniform z-axis distortion/shrinkage could be coped with by adopting laser confocal microscopy since it allows the creation of defined optical slices with better localization of particles inside the slices (Johnson 2001; Kubinova and Janacek 2001; Mura et al. 2004; Kubinova and Janacek 2015). However, whereas the use of laser confocal microscopy is in theory superior to traditional light microscopy with regard to z-axis distortion/shrinkage, it should be pointed out that confocal microscopy is limited by the need for fluorochrome staining of objects that can generate bias because of the variability in the intensity of fluorescence that often occurs due to both variable tissue binding of many antibodies (especially when indirect immunohistochemistry is adopted) and

photobleaching at the time of quantitative analysis (a problem that is much reduced for traditional histochemical staining that are much more reproducible).

Another limitation of stereological methods, as for any other method in microscopy, is the observer's eye, a potential source of bias that always has to be taken into account, especially when comparing data from different laboratories. In this view however, it should be noted that though computers can certainly make quantitative morphology easier and faster (Williams and Rakic 1988 Dolapchieva et al. 2000), a recent evaluation of the performance of automated cell detection algorithms revealed that manual approach is still the most adequate method for stereologic cell counting (Schmitz et al. 2014).

The still unsolved problems that affect stereological morpho-quantitative estimates have raised a debate about the use of the term “unbiased” to label stereological estimates in contrast to other morphoquantitative methods (Guillery and Herrup 1997). Actually, although stereological methods may “in theory” lead to unbiased estimates, the existence of the above-mentioned limitations in its practical application may generate a bias in the stereological results (Farel 2002; Hatton and von Bartheld 1999; Hyman and Gomez-Isla 1994; Popken and Farel 1996, 1997; von Bartheld 2002). Therefore, it has been realistically proposed that the term “unbiased” might be used to label an ideal aim that should be sought by the systematic analysis of the potential sources of bias and by the selection of the most appropriate procedure to cope with them taking into consideration the unavoidable methodological limitations and interpreting the results within those limitations (Geuna 2000; Saper 1999). Within these limits, stereological tools are able to approach more closely the unbiasedness goal than any other profile-based method and should thus be preferentially adopted (Pover and Coggeshall 1991).

Another debated issue is related to the sample size that should be adopted for stereological studies. In fact, some authors recommend a relatively small sample of 100–200 particles (Gundersen et al. 1988a,b). However, it appears that this sample size might be too small for a heterogeneous population, and other authors have recommended that a greater sample size must be adopted (several

hundreds) in order to cope with heterogeneity in the distribution of objects in histological sections (Benes and Lange 2001; Schmitz and Hof 2000). It appears thus that it is not feasible to pre-determine a “golden” sampling size in stereological research. However, it is important to carefully determine the sample size depending on the type of cells and tissues under analysis and especially on their heterogeneity. In this view, another point that should be emphasized is that employment of stereological methods by no means prevents the need for a good experimental design, i.e. for asking appropriate biological questions in the experiments (Hyman et al. 1998).

In spite of the critical points that have been raised, however, the theoretical “intrinsic strength” of stereological principles and methods have not been questioned and most of the debate has focused on the validity of the practical application of stereology to microscopic images. So far, in depth validation studies that have been carried out (Pover and Coggeshall 1991; Hatton and von Bartheld 1999; Kaplan et al. 2010) revealed that stereological methods are sufficiently reliable and should be regarded as the best possible options in the present state of the art. Nonetheless, implementation of stereological tools should be sought in order to overcome the shortcomings that have been identified. In this view, the technological progress in light microscopy informatics has definitely contributed to make design-based methods more user-friendly and reduce problems in their practical application.

Finally, the simple suggestion of performing a careful calibration/pilot study when a stereological approach method is used for the first time by a research group and/or it is applied to cells and tissues that have not yet been investigated in that lab (Farel 2002; Geuna 2000; von Bartheld 2002), it may be very useful for detecting and avoiding bias related to the practical application.

Conclusions

In spite of the enormous developments in investigation techniques, stereology still remains one of the pillars of quantitative biomedical research. Most studies on the normal and/or pathological phenomena occurring in animals base their main findings on the quantification of changes at the cell and tissue level. Unfortunately, in spite of the body of evidence accumulated on the pitfalls of counting and measuring on histological slides and on the need to adopt adequate methodological procedures to prevent bias, still many researchers use the inadequate morphometric approach based on the assumption that the number and size of histological profiles of tissue elements are equivalent to the number and size of those elements themselves (Table 1).

Stereology provides the methodological procedures needed to prevent this type of bias and, today, with the availability of several dedicated software packages and workstations, practical application of such procedures is much more easy and accessible to any researcher. Whereas most current stereological methods adopt a design-based approach, nonetheless, adoption of a model-based approach can be justified when a design-based approach is not applicable, such as in cases of precious human material and/or specimens already collected and processed (e.g., collections of slides) (Hyman et al. 1998).

As regards the parameters that are estimated using stereological tools, another important point that deserves particular attention is the frequent estimation of the density of objects instead of their total number. In fact, it should be clearly pointed out that even when a researcher is interested in only comparing relative numbers (i.e., % differences of cell types in a tissue) in different experimental conditions rather than comparing the absolute numbers, bias is not eliminated (Guillery and Herrup 1997; von Bartheld 2002). Therefore, the use of design-based stereological methods is not only necessary when absolute numbers are sought, but also for density estimation. In this view, when using an adequate design-based stereological probe (e.g. the disector), the estimation of total number can be directly obtained for the estimation of density (and vice versa); it is preferable to always report data on both morphological parameters. In fact, the

adoption of the density parameter alone for comparing cell and tissue populations makes difficult the interpretation of the data because density not only depends on the total number of objects but also on their size and distribution.

In conclusion, after more than thirty years of employment of design-based stereological methods in the scientific community, we feel confident in supporting the view that these methods should be the first choice for most research applications that quantify morphological parameters of cells and tissues in biology and biomedicine. It is important that researchers are aware of the high risk due to a methodological bias that can deeply influence their results leading them to infer erroneous scientific conclusions. It should also be emphasized that the adoption of a rigorous method for the statistical analysis of the morphometric data does not prevent nor correct the errors due to a methodological bias in sampling since biasedness cannot be detected from the data themselves. Once bias creeps into the estimates, the researcher will be completely blind about that and is prone to interpret numerical differences that are due to the bias as if they had occurred as true changes due to the experimental conditions. We believe that awareness about these concepts is very important for the correct production and interpretation of morpho-quantitative data and this paper is aimed at providing a contribution to the further spread of a mindful stereological approach.

Acknowledgments

This work has been supported by the European Community's Seventh Framework Programme (FP7-HEALTH-2011) under grant agreement n. 278612 (BIOHYBRID).