The proportionator: unbiased stereological estimation using biased automatic image analysis and non-uniform probability proportional to size sampling

By

J.E. Gardi, J.R. Nyengaard, H.J.G. Gundersen

Stereology and Electron Microscopy Research Laboratory and MIND Center, University of Aarhus, Aarhus, Denmark

Corresponding Author:

Jonathan E. Gardi
Stereology and Electron Microscopy Research Laboratory and MIND Center
University of Aarhus
Ole Worms Allé 1185
DK - 8000 Aarhus C.
Denmark

E-mail: Jonathan.Gardi@ki.au.dk
Tel.: (+45) 8942 2945
Fax: (+45) 8942 2952

Keywords:
Automatic image analysis, PPS-sampling, section quantization, simple random, smooth fractionator, specific stains, stereology, systematic uniform random, tissue inhomogeneity.

Abstract

The proportionator is a novel and radically different approach to sampling with microscopes based on well-known statistical theory (probability proportional to size - PPS sampling). It uses automatic image analysis, with a large range of options, to assign to every field of view in the section a weight proportional to some characteristic of the structure under study. A typical and very simple example, examined here, is the amount of color characteristic for the structure, marked with a stain with known properties. The color may be specific or not. In the recorded list of weights in all fields, the desired number of fields are sampled automatically with probability proportional to the weight and presented to the expert observer. Using any known stereological probe and estimator, the correct count in these fields leads to a simple, unbiased estimate of the total amount of structure in the sections examined, which in turn leads to any of the known stereological estimates, including size distributions and spatial distributions. The unbiasedness is not a function of the assumed relation between the weight and the structure, which is in practice always a biased relation from a stereological (integral geometric) point of view. The efficiency of the proportionator depends, however, directly on this relation to be positive. The sampling and estimation procedure is simulated in sections with characteristics and various kinds of noises in possibly realistic ranges. In all cases examined, the proportionator is 2- to 15-fold more efficient than the common systematic, uniformly random sampling. The simulations also indicate that the lack of a simple predictor of the coefficient of error (CE) due to field-to-field variation is a more severe problem for uniform sampling strategies than anticipated. Because of its entirely different sampling strategy, based on known but non-uniform sampling probabilities, the proportionator for the first time allows the real CE at the section level to be automatically estimated (not just predicted), unbiased - for all estimators and at no extra cost to the user.

Introduction

In most available computer aided systems for stereological purposes, the fields of view to be examined by the user are selected using systematic, uniformly random sampling (SURS). The user first delineates the section or the relevant part of it, and the computer then samples a specified fraction of all fields of view, roughly equidistantly spaced. The sampling scheme is sometimes referred to as ‘meander sampling’ (actually referring to one of the ways of implementing it manually).

The essence is that fields of view are sampled with a predetermined, constant probability, irrespective of the content of the fields (there may be no tissue present in some fields, for example). Moreover, in inhomogeneous tissue, many fields contain none or very few positive events and a large number of fields must therefore be examined to obtain a reasonable precision. In the extreme case of rare events, the examination of a very large number of empty fields is almost the whole workload, and the resulting precision may still be very low.

It is a consequence of such a sampling scheme that any inhomogeneity in the tissue is noise, an unwanted characteristic of the tissue that, without exception, means more work (in order to obtain a given precision). For decades, numerous attempts have therefore been made to improve the

performance using automatic image analysis, generally without any success because of the low and varying contrast and almost insurmountable difficulties in defining the relevant image segmentation in biological materials.

The sampling paradigm presented here is different. It relies on the fact that all staining methods provide the structure of interest with some particular stain. The stain may occasionally be a specific antibody, but is often of low specificity like the standard hematoxylin-eosin, for example. With that stain, all cell nuclei are varying shades of blue and many other components are reddish. The specificity in counting cell nuclei of a well defined cell type, for example, is provided by the expert observer who recognizes the cell under study using texture, configuration, neighbor relations and many other aspects that vary from one cell type to another.

Nevertheless, any field of view without blue stain cannot contain the cell nucleus under study (the field may be outside the section or in a region that happens not to contain any cells). On the contrary, a field with much blue may contain many cells of interest. Crudely: no blue, no interest; much blue, much interest.

This is the idea of proportionator sampling. Initially, the computer is used for automatically collecting some relevant information about all fields in the section. Using some arbitrary, predefined algorithm the amount of information is ‘measured’ (the total amount of blue, for example). The computer then selects a predetermined number of the fields, each with a probability strictly proportional to this amount (hence the name of the sampling paradigm). In the selected fields the expert user then makes the specific, ‘correct’ count of cell nuclei using all the usual clues, as outlined above - and the disector [1], physical or optical. Finally, the correct count in fields sampled with a known probability provides an unbiased estimate of the total number of cell nuclei in the section.

Note that under this sampling scheme any inhomogeneity of the tissue (with respect to the selected characteristic) becomes a signal that is used for making the sampling more efficient, the opposite of the ordinary situation described above.

The present report is an explorative study of the proportionator and a preliminary report of its performance: the genuine test of a sampling strategy is evidently to study it under realistic circumstances in real tissue. The proportionator is based on an unequivocally unbiased principle, but its efficiency under all kinds of problems and real distributions of structures in biological material is impossible to predict. We have therefore made this simulation study, with all relevant details as close to reality as possible, in order to test the strategy’s robustness when pushing the envelope of the simulation in various directions.

The mathematical basis of proportionator sampling is first presented, then the simulation is briefly outlined and finally the simulation results are presented and discussed.

The proportionator: sampling fields with probability proportional to size

This is a well-known statistical sampling technique often used in survey sampling [2]. Among statisticians, it is generally referred to as PPS-sampling. ‘Size’ is quite abstract, it may be any feature or characteristic that may be quantitated, at least crudely, and which has some positive relation or association to the objects under study. We shall use the amount z_i of a specific color in each field of view (FOV) as an example; it is also used in the simulation.

z_i must be known for all fields in the whole section, for which reason it evidently must be obtainable automatically. The sum over all N fields is denoted Z. The N fields are listed in any arbitrary order and the accumulated sum Fz is computed for this ordering.

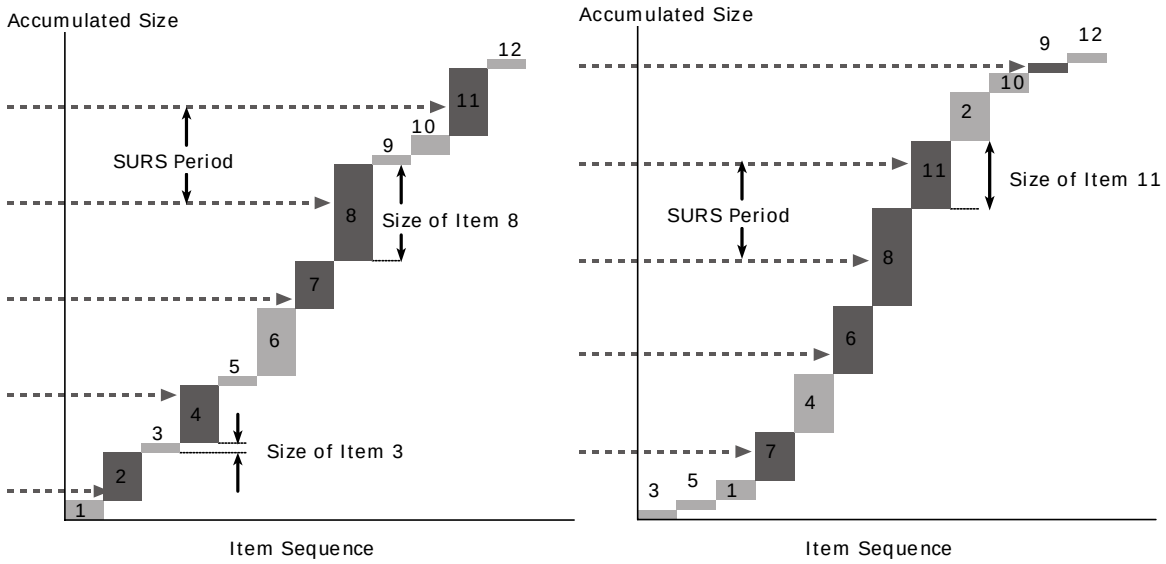

Figure 1. **Proportionator sampling**. The complete set of all fields is listed in any arbitrary order (1 through 12 in the example to the left), the ordinate shows the accumulated amount of color Fz. Sampling on the ordinate is systematic, the sampling period is Z/n, and the start is uniformly random (UR). Each sampling point on the ordinate uniquely identifies a field. The field’s probability for being sampled is strictly proportional to its recorded amount of color. To the right the fields are arranged in a sequence as for the smooth fractionator before SURS. This has no consequence for the probability of sampling, but stabilizes the total color content of the complete sample (of 5 fields in the example). This ordering is used for the simulation.

The sample size wanted is n, and the quantity Z/n serves as a sampling period under SURS, see Fig. 1. In the ordered set of all fields, sampling is uniform in the accumulated sum Fz, the ordinate of Fig. 1. Since Fz is a step function, any number between 0 and Z uniquely identifies a FOV (with known co-ordinates) among all fields. First a random start is generated between 0 and Z/n. The remaining fields to be sample are then identified by adding Z/n to the random sampling point of the previous selection. In short, this is ordinary SURS, like taking every seventh section, except that sampling is not among integer-indexed physical items but in the real-valued function Fz. As an unusual consequence, the sample size is a fixed constant (in ordinary SURS the sample size is a random variable).

One would usually study several sections from the same specimen. Optimally, they should be studied as one assembly, but it depends on the technical possibilities (more sections on one glass slide, several slides on the microscope stage, adequate software control of the microscope). If they can only be studied separately, one should aim at roughly a constant sampling period ***Z/n***, thereby sampling most cells from the sections with most of the indicative color. The adequate sampling period for obtaining a predetermined precision has to be determined in the pilot study, just like in SURS where one has to figure out the adequate sampling distances (step lengths) prior to making the serious observations.

The selected FOVs are presented to the user. For each field, the user enters the correct stereological count, x_i, and the unbiased estimator of the total content X in the section is simply

$$X := \sum_i^n \frac{x_i}{\left[\frac{z_i}{Z/n}\right]} = \frac{Z}{n}\sum_i^n \frac{x_i}{z_i} \qquad (1)$$

This is an instance of the general, unbiased Horvitz-Thompson estimator of a population total [3]

$$\text{PopulationTotal} := \sum^n \frac{\text{ItemContent}}{\text{ItemSamplingProbability}} \qquad (2)$$

Note that the unbiasedness of the estimator has no relation to the fact that the amount of color in a field is clearly biased information with respect to the correct nuclear count in 3D according to the disector counting rule - or with respect to any other stereological estimator. At the time of sampling, the number, z_i = 0.123, for example, assigned to the field because it is the sum of particularly colored pixels, just guarantees that that particular field has a known and fixed sampling probability under proportionator sampling: $probability_i = \frac{z_i}{(Z/n)} = \frac{0.123}{(200/20)} = 0.0123$ for ***n*** = 20 and ***Z*** = 200. In other words, once it is recorded, z_i is just a fixed number that controls a correct sampling and estimation procedure and is not required to have anything to do with what is estimated. What is estimated depends on what the user counts, but if the count is according to an unbiased stereological principle, Eq. 2 (with the correct stereological constants) provides an unbiased estimate of the corresponding total quantity in the section. Continuing the example, x_i = 2 nuclei, and this field therefore contributes $\frac{x_i}{z_i} = \frac{2}{0.0123} = 162$ nuclei to the estimate of total number of nuclei in the section.

Note also, that the unbiasedness does not depend on the assumed relation between the amount of color and the structure under study. The estimator is unbiased even if the relation turns out to be non-existent or negative: more color actually indicating fewer specific nuclei, because irrelevant nuclei dominate and they are in other parts of the organ than one of interest, for example.

Contrarily, the efficiency is closely related to the relation between color and nuclear count. If the relation is negative, the proportionator is very likely less efficient than Simple Random (SR) sampling. If the relation is positive, it is likely more efficient. In short: unbiasedness is guaranteed by design, superior efficiency is not. Using examples from the simulation, the efficiency of the proportionator is further discussed below.

Like many other complex sampling schemes, the statistical efficiency, roughly proportional to $\mathbf{1/CE^2}$, cannot be computed from the sample alone. In this case the situation is slightly worse, because even prediction of the ***CE*** requires information not available under any realistic circumstances. We have therefore implemented direct (and unbiased) estimation of the ***CE*** using independent repetition, as described below.

The phrase ‘statistical efficiency’ is used above because to the user the total time spent to obtain a certain precision is what really matters. This is difficult to simulate realistically, but we notice that delineation of the section is unnecessary (unless the containing space is just a part of the section) and that is actually a rather large fraction of the total time spent analyzing sections with the present technology. As a surrogate variable for the time to do the complete analysis we have used the total number of fields that has to be analyzed in order to obtain a predetermined total count.

The proportionator is an example of non-uniform sampling (of FOV) with varying probability. Such techniques [4] use extra information (like the proportionator) or some (realistic) assumptions about the items to be sampled. They are just beginning to emerge among the stereological estimators and show encouraging results for providing more efficient estimators (as we shall try to demonstrate below for the proportionator). The fact that the varying probability needs to be known for each sampled item is not a problem in practice, and it is this varying probability that potentially make these methods more efficient.

The point sampled intercept estimator of volume weighted mean particle volume,

$$\bar{v}_V := \left(\frac{\pi}{3}\right)\overline{\ell_0^3} \quad (3)$$

where ℓ_0 is the intercept length through a particle [5] also relies on non-uniform sampling among particles, but the unknown and varying sampling probability (proportional to unknown individual particle volume) means that the mean value can only be computed for the volume distribution of particle size, not for the much more common number distribution of particle size.

The simulation

A detailed technical description of the simulation is provided in [6]. Briefly, sampling of fields and estimation in a section of irregular shape is simulated, see Fig. 2. The section is divided UR into roughly 400 fields of view, each of an aspect ratio 4:3 (like a computer monitor).

The section contains 2500 rectangular ‘cells’ (profiles) of pixels of one color with a saturation of 50%. When two or more cells overlap, the pixels of the overlap therefore contribute at most one and a half to the amount of color in the field. The amount of color in a field is the sum of the contributions from all its pixels, shown in Fig. 2C.

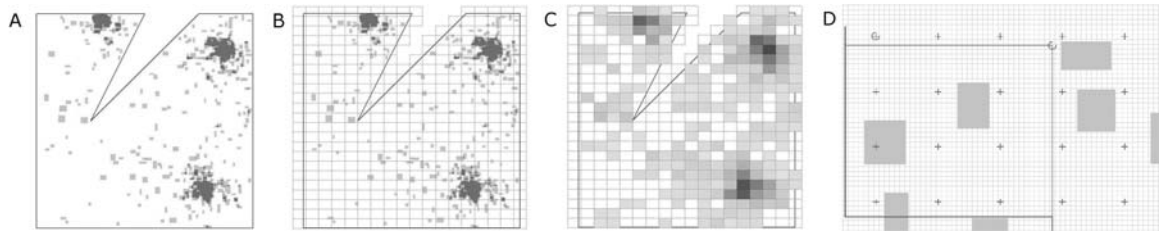

Figure 2. **Illustration of the simulation set-up**. A section with cells (A) is randomly tiled with fields of view (B). Each field of view is assigned a weight according to the amount of colored pixels in it; the amount of color is mapped as intensity (C). Panels A to C show the same clustered distribution of 1000 cells, mean area 290μm^2 size distribution ~0.3. Panel D shows a field of view with the counting frame and the set of grid points used. Panel D show cells with average area of 70μm^2, ~1/20 of the frame, the coefficient of variance (CV) of the cell size distribution is ~0.3.

Two stereological estimators are simulated: estimates of the total number of 2D cells (profiles) and of the total and average cell area in the section, respectively. Cells are counted in unbiased counting frames of an area of ~50% of the FOV, cf. Fig. 2D. On average, each frame captures 3 to 4 cells. For area estimation, a point grid with 20 points per field is used. The whole set-up is like a small 25-μm-thick section of layer IV in human neocortex examined at an objective magnification of 100X, for example.

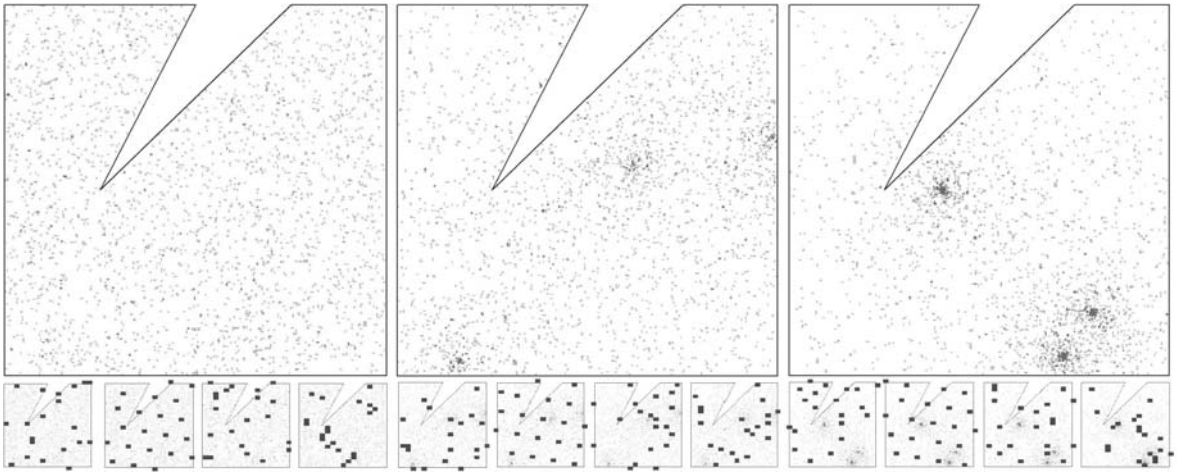

Figure 3. **The distributions and sampling paradigms simulated**. The three graphs show the cells in a uniform (homogeneous) distribution and two degrees of clustering (always with three centers in UR positions), respectively. Below each distribution, examples of samples of FOVs using the four sampling paradigms is shown: simple random (SR), systematic, uniformly random in 2D space (SURS), smooth fractionator (Smooth), and sampling with probability proportional to size (proportionator), where 'size' means total amount of color in each field. In the right-most example of a clustered distribution, all three UR strategies completely miss all three clusters (which only constitute less than 5 percent of the total area) whereas proportionator actually samples from the centers of all three clusters. Average cell size is 12μm^2, the CV of the cell size distribution is ~0.3.

The performances of four strategies for sampling fields of view are studied, cf. Fig. 3:

1. Simple independent random sampling, SR, with replacement. This is the basic UR sampling strategy, and a general baseline for efficiency considerations. It has the advantage that every aspect is known analytically, but it is also one of the most inefficient sampling strategies known, and therefore almost never used in stereology.
2. Systematic, uniformly random sampling, SURS, in the 2D geometric space of the section, used by most computer assisted sampling systems as outlined above. In 2D space the efficiency of

this sampling strategy is not predictable, but, in the absence of long-range gradients, it is expected to be only marginally more efficient than SR sampling.

3. Smooth fractionator sampling, where the amount of color per field is used as the associated variable for sorting and re-indexing all fields according to the smooth fractionator sampling scheme, as illustrated in Fig. 1, right, and in [7]. The sampling probability is the same constant as in the above two strategies, and the variance within a set of sampled fields is therefore the same as for the other two strategies, see Fig. 4. However, the strength of the smooth fractionator is its ability to reduce the variation between samples [6].
4. Proportionator: sampling with probability proportional to size. SURS is done the smoothed order of the accumulated weight, as described above and illustrated in Fig 1, right panel.

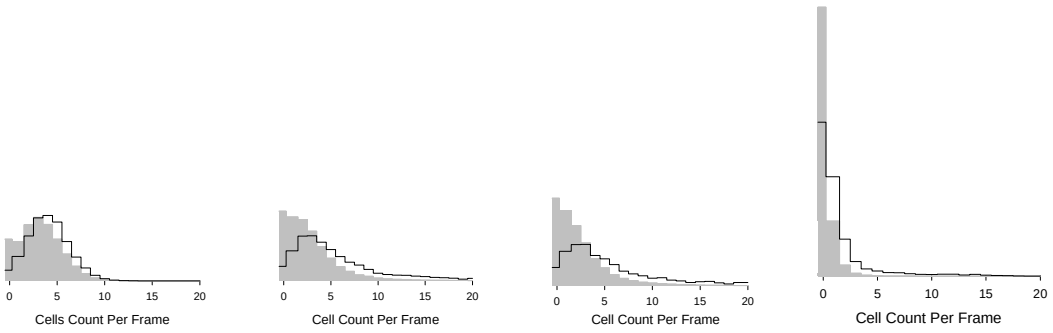

Figure 4. **The distributions of individual samples from different sampling strategies**. The first three panels show the distributions of the cell count per frame in the three simulated spatial cell distributions in Fig. 3, the fourth is that of the sparse distribution in Fig. 7 below. The grey histograms are the distributions of cell counts for the first three sampling paradigms, all using the same, constant sampling probability from the same total population of 2500 cells; the means of the first three gray histograms are thus identical. The full drawn histograms are those of proportionator sampling. The means and ***CV***s are shown in Table 1. Average cell size is ~12μm^2 (~1/100 of frame area), the CV of cell size distribution ~0.33. With such small cells there is virtually no overlap of cells upon each other.

As illustrated in Fig. 4, the four sampling strategies provide two different distributions of FOVs for each spatial distribution. The three UR, fixed probability strategies (SR, SURS, and Smooth) provide the same sampling distribution (which is the global distribution of FOVs), the gray histograms. The proportionator, with its non-uniform sampling probability proportional to size, provides size-weighted distributions of FOVs, the full-drawn histograms.

	Homogeneous		**Intermediary**		**Clustered**		**Sparse**	
	Mean	**CV**	**Mean**	**CV**	**Mean**	**CV**	**Mean**	**CV**
SR counts and estimates	3.25	0.67	3.23	1.55	3.01	1.88	0.32	2.93
Proportionator counts, relative to SR	127%	76%	263%	93%	344%	83%	467%	63%
Proportionator estimates, relative to SR		58%		32%		28%		75%

Table 1. **Characteristics of samples from different sampling strategies**. The upper row shows the means and ***CV***s of the four grey distributions in Fig. 4. The middle row shows the means and ***CV***s of the full-drawn proportionator distributions as a percentage of the corresponding SR ones. The lower row is the variation of the contributions to the proportionator estimate of the total, cf. Fig. 5 right, also in percent of the SR ***CV***.

As tabulated in Table 1, two distinct features characterize the proportionator sampling distributions. First, their mean is always higher than that of the UR distributions, which means that fewer fields need to be studied to obtain a given count. Secondly, the Prop sampling distributions have a smaller ***CV***, which means that the total sample (total count) is less varying. Both features translate to a greater efficiency.

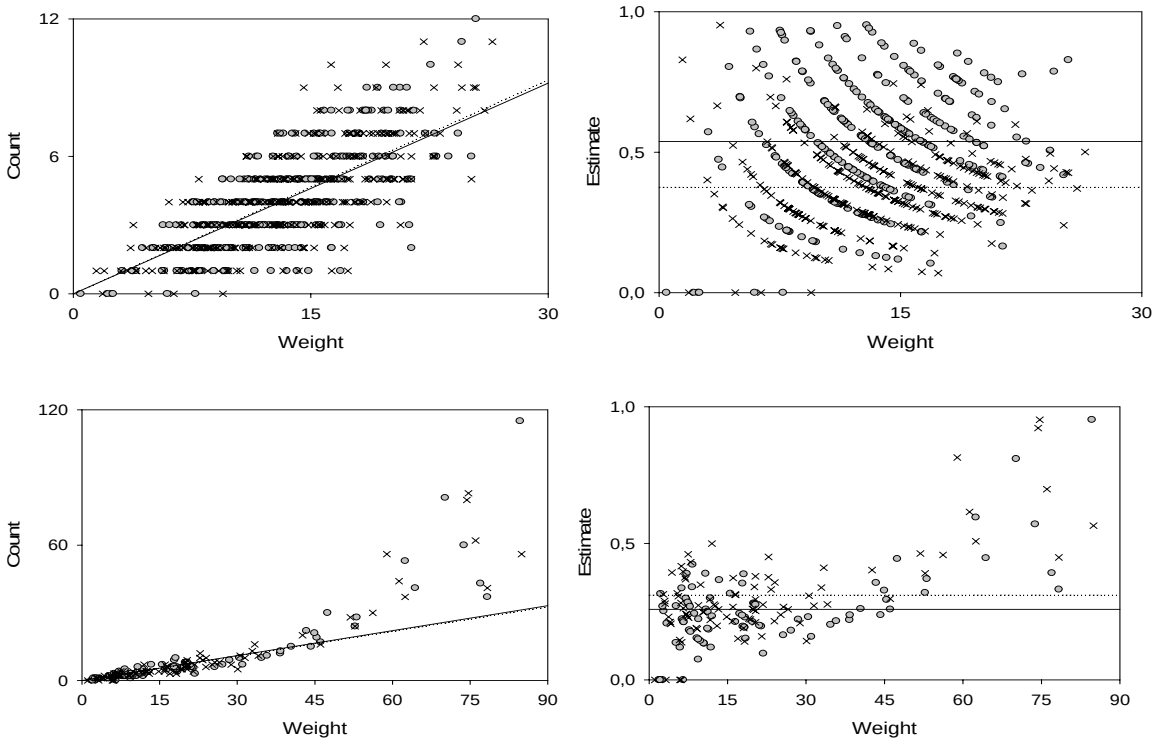

Figure 5. **The bivariate Prop sampling distribution and the contribution to the estimated total**. To the left is shown corresponding values of correct count and total amount of color (weight) in a few hundred fields of view from a homogeneous (top row) and a clustered (lower row) spatial distribution. For each data-point, the estimate of the contribution to the total is proportional to the slope of a line from the origin to the point. In this transform, the variability of the counts is irrelevant, only the scatter around the line trough origin (with a slope of the mean estimate to the right) matters. The estimates are shown to the right (the ordinate is fraction of maximal estimate). Symbols: o—o cell count, ×- - -× point count. For each data set the mean estimate is shown. The linear ordinates to the left are very different.

The proportionator is, however, a distinct 2-step procedure: sampling of fields proportional to size followed by estimation of their contributions proportional to the correct counts but inversely proportional to sizes, cf. Fig. 5. This has a profound effect on the distribution of contributions to the total estimate. The counts in the upper row of Fig. 5 come from a homogeneous spatial distribution with a moderate field-to-field variation of ***CV*** = 0.67, but the proportionator reduces the field-to-field variation to ***CV*** = 0.51. The estimates (to the right in Fig. 5), however, only have a ***CV*** of 0.30. The corresponding ***CV***s for the clustered distribution at the bottom of Fig. 5 are 1.88, 1.56, and 0.44. Both steps of the proportionator

may thus result in a substantial increase in precision. As already mentioned, the first step also increases the sample mean and thereby decreases the number of fields one has to study.

Note that the ***CV*** of the estimate contributions is completely independent of the spatial field-to-field variation. In the ideal case of a one-to-one correspondence between the count in a field and its color content, all data to the left in Fig. 5 would lie on a straight line from the origin, and the ***CV*** would be 0 - despite a any large field-to-field variation. The ***CV*** of the estimate contributions is solely due to the scatter of the values around the line through origin.

The upper right panel of Fig. 5, with some of the highest estimates associated with very low weights, also illustrates one of the ways the proportionator may potentially break down. If a relative low weight is associated with a relatively high count, the CV increases. Just a count of 1 in a field of a really low weight can reduce the efficiency remarkably. We have not, however, been able to provoke this phenomenon with realistic parameters of the simulation, but a single outlier of this type has occasionally been observed.

With the above set-up of the simulation, the sensitivity of the sampling strategies with respect to a number of characteristics or features of the section and the cells are studied:

- The spatial distribution of cells in the section: homogeneous (2D Poisson), moderately and quite clustered distributions are used, all illustrated in Fig. 3.
- The cell density: a relatively normal density of 3 to 4 cells per frame and a moderately sparse density of 1 cell per every second or third frame are studied. The simulation set-up is not well suited for studying real rare-events-sampling, but that is not necessary either, because the result is given: nothing competes with finding the rare events, automatically.
- The noise level is varied from no noise (corresponding to staining with a highly specific antibody) over a quite normal signal-to-noise ratio of 1:1 to a bad case of a signal-to-noise ratio of 1:3. Noise is particularly important in studying the relative performances of proportionator and smooth fractionator sampling because it has no influence at all on the other sampling strategies, which are independent of the information about color. Noise is simulated by adding cells that contribute fully to the color amount per field but are not counted, like a cell type not studied, but nevertheless present and stained.
- The influence of cell size is studied over a wide range from average cell size of 0.9% to 25% of the size of the frame. The cells have an approximately log-normal size distribution with a CV of about 0.3.

Two measures of the performance of the sampling strategies are generated.

1. The real ***CE*** of each unbiased estimator is computed directly from repetitions of the simulation. The uncertainty (due to the finite number of repetitions) of this simulated real ***CE*** is in all cases a very small fraction; it is visible in the figures below as small bumps on the curvilinear relations, which are all ideally completely smooth.
2. The number of fields examined in order to obtain a specified total count.

Finally, we have implemented a direct and unbiased estimate of the ***CE*** of the proportionator at the level of fields of view, given the section. As previously mentioned, there is not enough information available for predicting the precision of the non-uniform sampling in the proportionator. This is similar to the smooth fractionator where prediction of the precision from samples turns out to be useless. For that purpose direct and unbiased estimation of the ***CE*** was introduced with the description of the smooth fractionator [7] and is described below.

The concept is simple and forceful: instead of sampling ***n*** fields and obtaining the estimate ***EST***, sample first ***n/2*** fields and then generate a new sample of ***n/2*** fields using the same sampling strategy, but independent of the first sample (with replacement). The second random start is chosen independently of the first one. The two samples produce two estimates, $\boldsymbol{est_1}$ and $\boldsymbol{est_2}$. The estimator used is now the average $est = \frac{est_1 + est_2}{2}$, accompanied by an unbiased estimate of its $CE = \frac{SD(est_1, est_2)}{est * \sqrt{2}}$. It is the complete independence of the two samples that guarantees that the divisor in this equation is $\boldsymbol{\sqrt{2}}$ and that the ***CE*** (quite unusually) is an unbiased estimate.

In order to make comparisons simple and conclusions unambiguous, the simulations are always made with identical set-ups on identical spatial distributions for all four sampling paradigms. Moreover, since the counting noise in number estimation [8], with a ***CE*** of $1/\sqrt{\sum Q^-}$ under UR sampling is a large and sometimes dominating source of variability, comparisons are always made for the same total count. Consequently, one should be aware of the fact that for none of the four sampling strategies has the set-up been optimized. As an example, SURS in much clustered cells might perform better than shown below in Fig. 6 if the frame size is decreased and more fields are studied. In other words, because of the purpose of the study, we have performed the simulations so that they are strictly comparable, but we have not tried to optimize any of the strategies because that is only relevant for another purpose (and probably mostly relevant in real examples, not in simulated ones).

A common practice while observing biological tissue is to count approx 100-200 cells (or counting points) per animal [9]. To make things easier, the user will probably tune the system during the common SURS to observe approximately 50 fields of view (changing the lens magnification and counting frame or area per point size) to achieve this count. This will result in approximately 1 to 4 cells per counting frame (keeping in mind that cell features must be recognized). Most simulations sample size (of fields of view) was tuned to have a total count of few cells to around fifth of the placed cells (i.e. from ~10 to approx ~500 when 2500 cells were placed, and ~2 to 50 when 250 sparse cells were placed) on each synthesized image.

Results

The influence of the spatial distribution of cells on the precision of number estimation using the four sampling strategies is illustrated in Fig. 6. Qualitatively, the result is independent of the spatial distribution: SR is uniformly the least precise and SURS is at most a few times more precise than SR.

Smooth is always more precise than SURS, and proportionator is always most precise of all, except in a perfectly homogeneous distribution where Smooth has a slightly smaller CE.

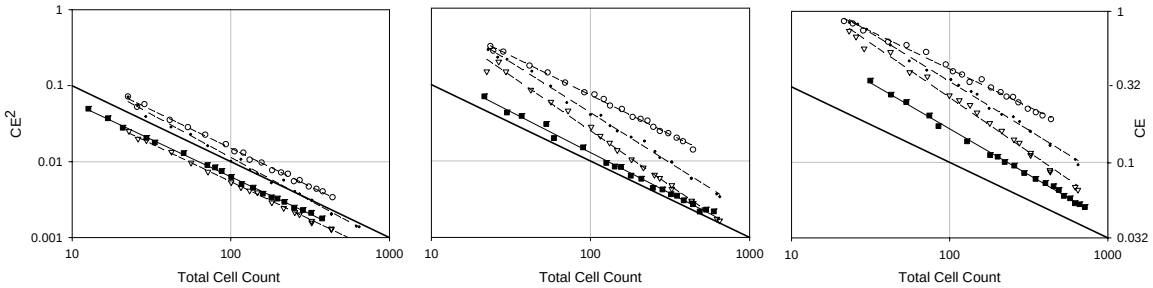

Figure 6. **The impact of the spatial distribution**. The ***CE***s of the four sampling strategies for number estimation in three different spatial distributions of the cells, illustrated in Fig. 3. The heavy line is the relative counting noise with a slope of -1, i.e. it decreases in direct proportion to the total count. The horizontal line indicates CE = 0.1, the vertical line indicates a total count of 100. Mean cell size is 70 μm^2, CV of cell size distribution ~0.36. Symbols: SR o- -o, SURS •- -•, Smooth ∇- -∇, Prop ■—■.

The influence of more and more clustering is also consistent: it uniformly reduces the precision. This is evidently to be expected, more clustering means a larger variability from field to field, cf. Fig. 4. The impact of more variability among fields is not, however, uniform over the strategies. For the simple UR ones, SR and SURS, the variability among fields adds to the degradation of the precision. For the two strategies that use additional information about the content, that is not the case. From a uniform distribution to a much clustered one, the precision of proportionator relative to SR increases from a factor of ~2 to ~10, i.e. it is ~5-fold less sensitive to the increasing spatial inhomogeneity.

Actually, the degrading effect of spatial inhomogeneity on proportionator is likely more exaggerated in the simulation. Inspection of Fig 5 (bottom panel) shows that the noise is due to a number of fields with very large counts but not quite correspondingly large color values, they are thus all above the line through origin. This is not an effect of the clustering; it is due to the high saturation of 50% whereby multiple overlaps contribute too little to the color value.

It is quite illustrative to the concept of using the information from an associated variable that both Smooth and proportionator are more efficient than simple UR in a UR spatial distribution. The simple fact is that even in a homogeneous spatial distribution some fields happen to have more cells and it pays off to focus on those: the increase in precision is a respectable factor of 2 to 3. The reason that SURS is marginally more precise in a UR distribution than SR is the edge effect: the only deviation from complete uniformity is the presence of frames on the edges with a lower-than-average count.

In a homogeneous spatial distribution the simple UR strategies are of necessity close to the counting noise, but the ***CV*** is inflated a bit due to the effect of section edges under uniform sampling. This edge effect is more pronounced in the simulation than in most real observations because the section is rather small. It is, therefore, even more remarkable, that both Smooth and proportionator can deliver samples of a ***CE*** below the counting noise for a homogeneous distribution, the edge effect included, as shown in Fig. 6, left. As explained above, the proportionator is not sensitive to the ordinary counting noise, what matters are only the deviations from the slope of Fig. 5 (upper left).

Fig. 6 also very clearly indicates the deficiency of the usual ***CE***-prediction based on just the counting noise, $1/\sqrt{\sum Q^{-}}$, but ignoring the field-to-field variation (for which no simple prediction is available). In all cases, except the exceptionally precise Smooth and proportionator in homogeneous sections, the real ***CE*** is several-fold larger than the counting noise - simply because of field-to-field variation. The prediction equations currently used, would predict the CE^2 of SURS to be the same in all three graphs of Fig. 6 - but it is more than 10-fold higher in a clustered than in a homogeneous distribution.

As described below, this deficiency is remedied for the proportionator using the direct estimator of the ***CE***.

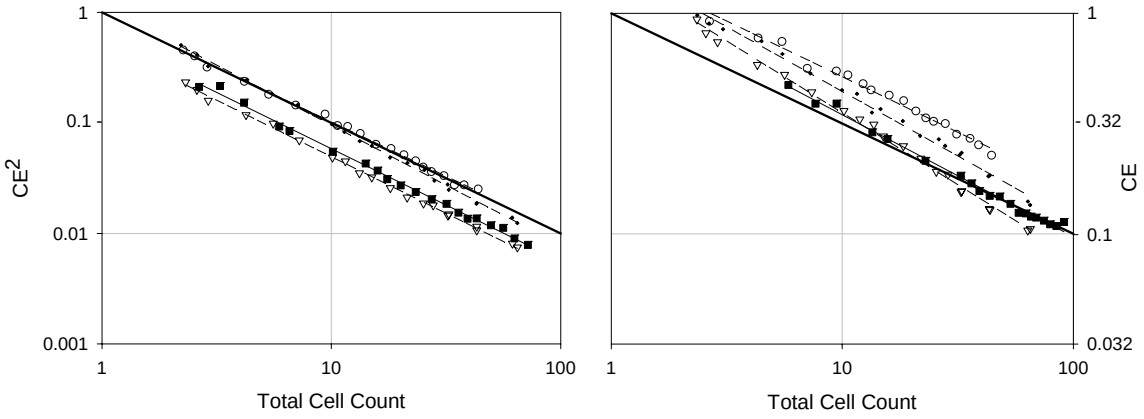

Figure 7. **The precision of sparse sampling**. The ***CE***s of the four sampling strategies for number estimation in a sparse sample from a total population of only 250 cells. Only a homogeneous (left panel) and a clustered spatial distribution are shown. Symbols as in Fig. 6.

The precision of sparse sampling from a small population of cells with a low density (Fig. 7) follows the same pattern as the ordinary sampling situation shown in Fig. 6, except the counting noise is relatively larger because of the small count. The large relative counting noise mostly obscures the impact of the spatial distribution and the field-to-field variation.

All the above statements are in relation to the statistical precision of the estimator, the CE^2. As already indicated, to the user the efficiency is also related to the number of fields, ***F***, to be studied in order to obtain a certain, predetermined sample size (total count). The necessary numbers of fields to obtain 'common' sample sizes are listed in Table 1 for the various sampling strategies.

Given the cell density, the number of counts per field, ***Q/F***, is the same constant for all three UR sampling strategies (because the means of the grey distributions in Fig. 4 are identical). For proportionator, ***Q/F*** increases with increasing clustering. The variation in ***Q/F*** is thus the other aspect of efficiency: the more counts per field, the fewer fields have to be examined to obtain a predetermined total count. In a clustered distribution, the proportionator provides counts that are 4 to 5-fold higher than the UR strategies for both an ordinary and a low density.

Spatial Distribution	Sampling Strategy	Total Count, Q	Total Observed Fields, F	Q/F
Ordinary Sample Size	SR, SURS, Smooth	100	31	3.2
Homogeneous	Proportionator	100	24	4.2
Intermediate	Proportionator	100	12	8.3
Clustered	Proportionator	100	7	14.3
Sparse Sampling	SR, SURS, Smooth	25	78	0.3
Homogeneous	Proportionator	25	37	0.7
Clustered	Proportionator	25	16	1.6

Table 2. **The number of fields necessary to study in order to obtain a given sample size**. The three UR sampling strategies (SR, SURS, Smooth) all require the same number of fields, independent of the spatial distribution of the cells (the cell density is a constant), but for the proportionator the necessary number of fields depends on the spatial distributions (the full-drawn histograms in Fig. 4 have varying means). ***Q/F*** is the average count per frame, the mean of the distributions shown in Fig. 4.

It is, however, quite conceivable that a sampling strategy with a low ***F*** in addition has a high $\boldsymbol{CE^2}$, i.e. it is fast but imprecise. With the risk of oversimplifying, we have therefore combined the factors of efficiency into one expression

$$Efficiency = \frac{c}{CE^2 \times F} \qquad (4)$$

We do not know realistic values for the factor ***c*** of this proportionally. But assuming that ***c*** is a global constant, we may with some confidence express the efficiency of a sampling strategy relative to that of SR, the conventional base line for such comparisons:

$$Efficiency\ relative\ to\ SR\ = \frac{CE_{SR}^2 \times F_{SR}}{CE_Y^2 \times F_Y} \qquad (5)$$

where ***Y*** is any of the other sampling strategies.

The three graphs in Fig. 8 represent three realities: three different cell distributions, which cannot be changed. Faced with anyone of them, the user can choose between the three strategies (being uniformly the worst, SR is not a realistic option) to obtain a certain predetermined precision, the horizontal lines in Figs. 6 and 7, for example. All three sampling strategies have relative efficiency that depends somewhat on the total count, but, as it happens, none of them cross any other one within the 'realistic' range simulated. Independent of the spatial distribution, the proportionator is most efficient.

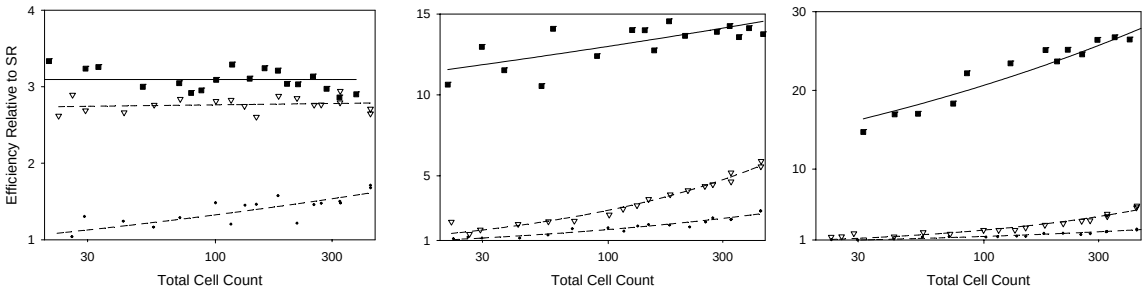

Figure 8. **The combined expression of efficiency relative to SR** takes into account both the statistical precision and the number of fields to obtain a predetermined total count. The three panels represent homogeneous, intermediate and clustered spatial distributions, respectively. Note that the ordinates are linear but have very varying scales. The curves are (exponential) regression lines. Symbols as in Fig. 6.

The center panel indicates that in the very realistic and common situation of a moderately non-uniform distribution (the center panel of Fig. 3), the total workload is 5 to 10-fold smaller using the proportionator than using the ordinary SURS. Even in the (rare) case of a perfectly homogeneous distribution the usual SURS represents about twice as much work as the proportionator.

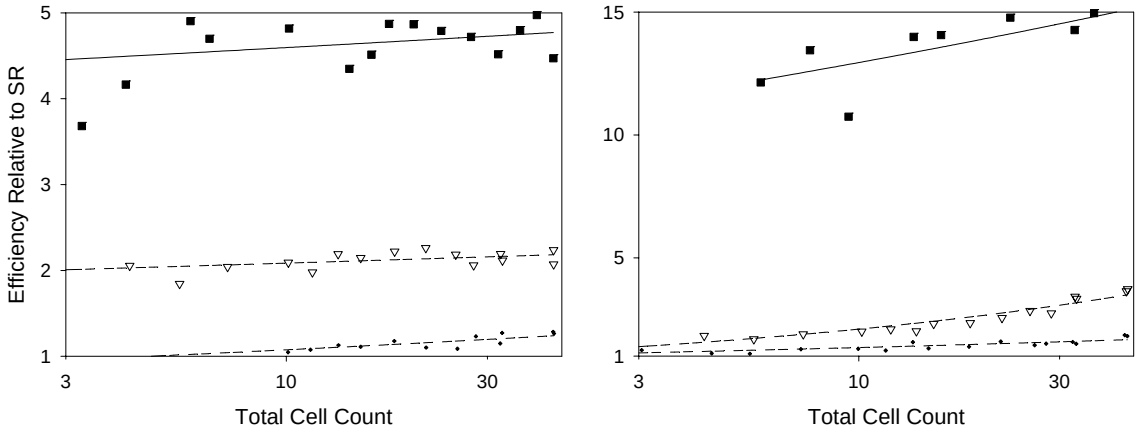

Figure 9. **Sparse samples are dominated by the counting noise of the low count.** The differences in efficiency relative to SR between homogeneous (left) and clustered (right) spatial distributions are consequently less than for samples of ordinary size, cf. Fig. 7. Note the varying scales of the ordinates. Symbols as in Fig. 6.

For very small samples from distributions with a low density, the counting noise is a pronounced fraction of all variation. Still, the proportionator finds the few cells with an efficiency which is 2 to 10-fold better than the UR sampling strategies, cf. Fig. 9.

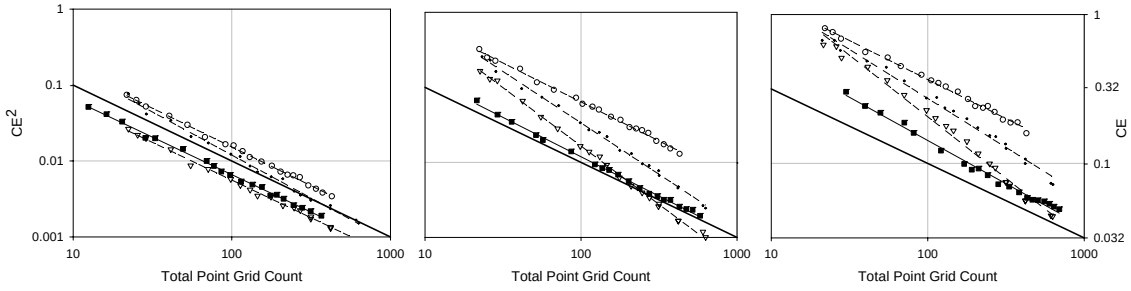

Figure 10**. The impact of the spatial distribution on area estimation**. The ***CE***s of the four sampling strategies for area estimation in three spatial distributions of the cells, from homogeneous over intermediate to clustered, respectively. The areal density of the object phase is 0.035. The point density (1 point per 135 μm^2) and the cells size (average area 70 μm^2) means that the point hits are almost independent. Symbols as in Fig. 6.

Qualitatively and quantitatively, the precision of the point counting area estimator behaves almost exactly as number estimation, compare Figs. 10 and 6.

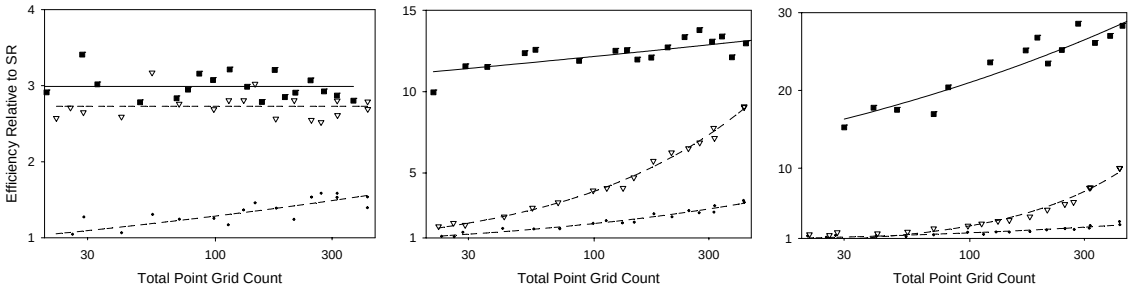

Figure 11**. The efficiency relative to SR of the other sampling strategies for area estimation**. In the homogeneous spatial distribution (left) Smooth, using the information about the content of fields, is almost as efficient as the proportionator, both are much more efficient than SURS. In the clustered spatial distribution (right) proportionator outperforms every other strategy 10 to 15-fold for an ordinary count of 200 to 300 points. Symbols as in Fig. 6.

The number of fields is also distributed essentially as in Table 1, and the efficiency relative to SR, Fig. 11, is therefore again almost as for counting 'cells'. The only noticeable difference is that Smooth area estimation is more efficient than Smooth cell counting, particularly for large total counts.

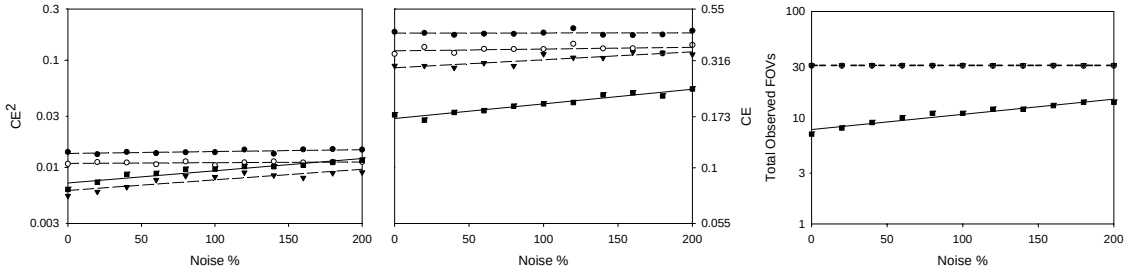

Figure 12**. The degrading effect of noise on total number estimation**, in homogeneous (left) and clustered (center) spatial distributions. To the right is shown the total number of fields studied for obtaining the total cell count of 100 (used in all noise levels) for the clustered distribution. Noise is invisible to SR and SURS and their precision is constant under varying noise. Average cell size is ~12μm^2. Symbols as in Fig. 6.

The effect of noise was studied by adding cells with the same color as the 'real' ones, but which could not be counted. Noise thus has a direct effect on the estimators using the information about image content, but has no influence at all on the SR and SURS estimators. Noise increases the ***CE*** of the proportionator, but also increases the number of FOV necessary for a given total count, cf. Fig. 12, right. The combined effect is that the efficiency of proportionator relative to SR decreases from 2.9 (no noise) over 1.7 (100% noise) to 1.5 at a noise level of 200% for a homogeneous spatial distribution. For a clustered distribution the corresponding relative efficiencies are factors of 26, 13, and 8, respectively.

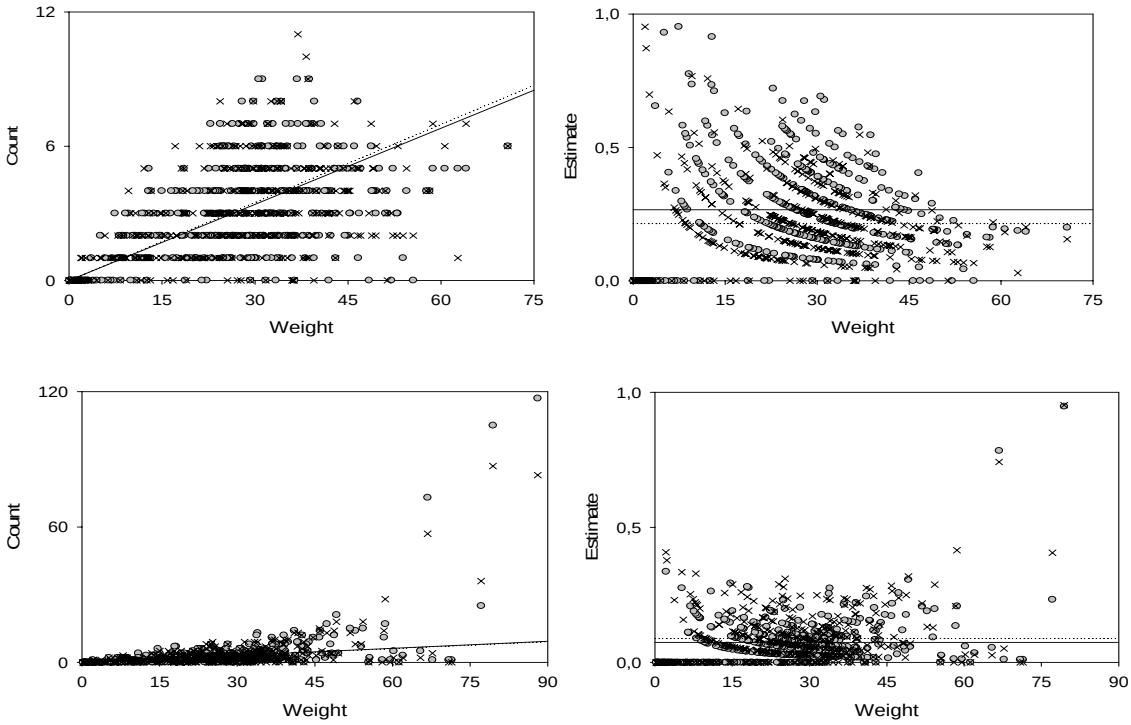

Figure 13**. The count-color relation is broadened by noise**, and the number of sampled fields with no count is increased. The top row is a homogeneous spatial distribution and lower row is clustered spatial distribution, both are with 200 % noise. Average cell size is ~70μm^2, CV of cell size distribution ~0.33. Symbols and lines as in Fig. 5.

Noise makes the relationship between color and count broader, mostly to the right towards larger color content, as illustrated in Fig. 13 (compare to Fig. 5). It also increases the probability of sampling fields with a count of 0. Both lead to an increase in the ***CV*** of estimates and reduce the average count when sampling proportional to color. The reduced average count in turn enforces a larger sample of fields in order to obtain the predetermined total count.

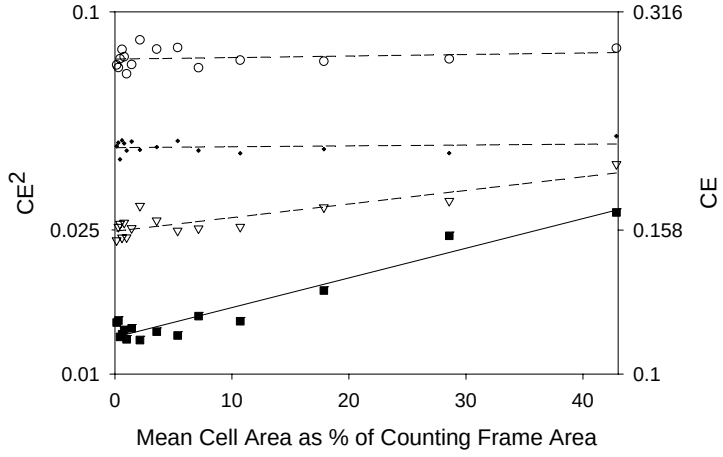

Figure 14**. A large cell area is another noise in number estimation** for the sampling strategies that use information from the field of view. Total cells count is fixed at 100 in all cases; the spatial distribution is intermediate. Symbols as in Fig. 6.

When the areas of the field of view and the frame are fixed, a larger mean cell area means that cells not counted in the frame provide a color signal not related to the local count, i.e. they are a source of noise.

This is particularly the case for cells on the left and lower margin of the field, which can never be counted just because they become larger.

As illustrated in Fig. 6 and several others, predicting the efficiency of any of the sampling paradigms from just the counting noise simply does not work. None of the common predictors of ***CE(N)*** takes the field-to-field variation into account (it is not known how to do that in simple ways, mainly because it is systematic sampling in a 2D universe).

Faced with an analogous problem with the smooth fractionator we introduced the direct estimation of the ***CE*** using two completely independent samples of reduced size [7], resulting in two independent estimates, which in turn is a firm basis for estimating the ***CE*** of their mean.

This strategy will work in all cases, but it has a price: the mean of 2 independent estimates ***est*** from 2 samples of size ***n/2*** may be less precise than an estimate ***EST*** from one sample of size ***n***. The decisive factor is the relation between $\boldsymbol{CE^2}$ and the sample size ***n***, $CE^2 \propto 1/n^a$ where ***a*** may be 1.0 or larger. For SR, ***a = 1.0*** (the slope is always -1.0), but for SURS and Smooth ***a > 1.0*** is often the case (due to negative covariances).

The proportionator primarily derives its efficiency not from a negative covariance but from selecting a different sampling distribution, cf. Fig. 4, and by exploiting the positive relation between the count and the weights, illustrated in Fig. 5. The resulting slope of the relation between sample variance and sample size is in all examples -1.00 (Fig. 6, left) or quite close to that (-1.07 and -1.21 in the center and right panels, respectively).

As illustrated in Fig. 15, the two estimators are therefore almost equally precise. The price for using the mean estimator ***est*** from two independent samples of size ***n/2***, with a direct and unbiased estimate of its ***CE***, is therefore essentially nil.

That is not the real situation, however. The real choice in the proportionator is between, on the one hand, an estimator with unknown precision, deviating from the noise-prediction by unknown factors in the range from ½ to 5 (or 10), for example, and on the other hand, an estimator with known precision and an efficiency which is close to 100%. For the other estimators, SURS and Smooth, the situation is worse; here the choice is between either an even more imprecise or an even less efficient estimators (because the slopes are well below -1.0, cf. Figs. 6 and 10).

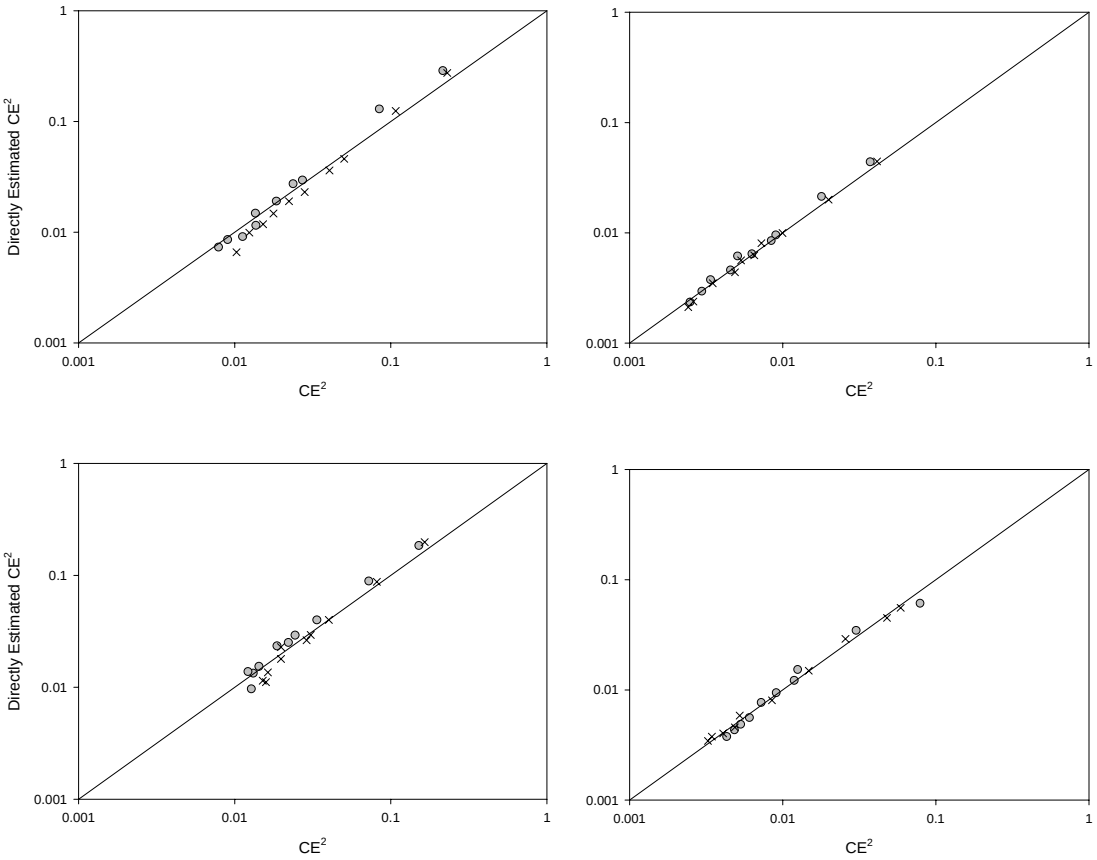

Figure 15**. The direct estimate of *CE²* for two independent samples of size n/2 compared to the real *CE²* for a sample of size n**. The values are from simulations in homogeneous (top row) and clustered (bottom row) spatial distributions; the left column is from sparse distributions, the right column from distributions of ordinary density. The identity line is shown. Symbols as in Fig. 5.

In ordinary practice, it is more useful to think of a direct estimation of the variance $Var_i(n) = \left[CE_i(n) \cdot n_i \right]^2$ of the estimate of the total number of particles n_i in the i'th section. For the estimate of total number in all m sections, Σn, one may then compute the overall CE:

$$CE_m\left(\sum n\right) := \frac{\sqrt{\sum_m Var(n)}}{\sum_m n} \qquad (6)$$

Discussion

We shall first discuss the overall scope and purpose of proportionator sampling and estimation, then discuss in detail its problems and possible solutions in the more narrow context of stereological estimators of geometric quantities.

The proportionator uses automatic image analysis to obtain information about the distribution of the structure or feature of interest from all fields of view in a section or a set of sections. This information is first used to generate a sample with more of the structure and with less variation with respect to the amount of structure. Secondly, the expert's 'correct' count (or any other correct calibration of the signal or information from the section) is converted to an unbiased estimate of the section total content. The variation of these estimates is smaller again, and does not depend on the variation among the sampled fields with respect to the content of the structure. Instead, the precision of the proportionator depends on how tight is the relation between correct count and 'size' (analogous to a regression analysis).

As a basis for estimation of section total quantities, the proportionator is guaranteed to be unbiased. Considering the general but often vague relation between the amount of color, of the observer's choice, and the amount of structure, it may generally be expected to be more efficient than other sampling paradigms in common use. Potentially, it can be much more efficient, particularly when the stain is very specific or the structure is sparse (or both). However, like the other two generally efficient sampling paradigms, SURS and Smooth, it is not guaranteed to be more efficient than simple random sampling. We shall therefore discuss the problems it may encounter in detail below.

The stereological scope of proportionator sampling and estimation is all the known estimators: volume, surface, length, and number and connectivity. We have used the two extreme examples, volume and number, because they represent the general range of relationship between content and count: there is a 1-to-1 correspondence between the amount of structure in the (thin) section and the probability of a point hit, whereas the amount of structure and the count of structural elements have a much looser relationship, because the amount of structure in the section depends equally on the number of elements and their mean size. The reason that this particular difference did not show up in the simulations is likely that we have been too particular in assigning the color a saturation of 50% in each pixel.

The proportionator is radically different from all other stereological sampling and estimation paradigms in that it provides the total quantity in the section or the set of sections. All the ordinary global stereological estimators, except the fractionator, provide densities, which are then converted to organ totals by multiplying with a separate estimate of the volume of the reference space.

Proportionator sampling thus fits exceptionally well into the fractionator estimator of total number: the simple sum of estimated section totals times the constant section sampling fraction is the organ total. In order to emulate the common ratio estimators, one would need separate estimates of the total areas of the reference space in the sections (technically, that would likely be obtained automatically during the initial scan of the entire section). The density would then be estimated as the ratio of sums of section totals, as usual.

A much more interesting alternative to ratio estimators is to use sections of a constant and known sampling intensity (like for the Cavalieri-estimator of total volume). The 2D density of any probe, which is always known, now has a known 3D density, and the sum of total counts directly estimates the organ total. As an example, structure surface is estimated using cycloids of a known and constant density ℓ / a

on VUR sections of a constant distance T. The sum of proportionator estimates of total intersection count, ΣI, in each section leads directly to the estimate of organ total surface [10]

$$S := 2\left(\frac{aT}{\ell}\right)\sum_{organ}\sum_{sect} I \qquad (7)$$

Analogous direct estimators of total quantities exist for all ratio estimators.

Combined with the local estimators of particle size, for example, the nucleators [11] and rotators [12], the proportionator would naturally provide the very informative absolute size distribution: the total number of cells in each size class. The commonly used relative size distribution (the percentage of cells in each size class) is a slippery basis for firm biological conclusions: a reduction in frequency in some classes is often interpreted to mean a loss of these cells, but could as well be caused by an increase in the number of cells in other classes. Such ambiguities are eliminated in the absolute size distribution.

Proportionator sampling and estimation is of course not restricted to stereological estimators of geometric or dimensional quantities. All kinds of histochemical and immuno-chemical quantities may naturally be estimated from 2D sections. Using either a known constant or a calibrated relationship between color and chemical content the result would be total chemical quantities per section. These are converted to total organ quantities as simply as described above for the fractionator.

A pivotal feature of proportionator sampling and estimation using microscopes is the definition of ‘size’. In the simulation we have used the total amount of color because it fits well into the common purpose of staining sections: to identify some structure of particular interest. ‘Size’ can, however, be many other features that can be quantitated and may have a positive relation to the feature under study. To mention a few, it can be

- The covariance between any template pattern or texture and the field of view,
- The steepness of a gradient with respect to any measurable quantity,
- The mass density in ultrasound microscopy (ultrasound scans),
- The extinction of the electron beam in a particular part of the spectrum related to the presence of plain contrast provided by heavy metals (in transmission electron microscopy) or the positive signal in distinct parts of the spectrum of secondary electrons indicative of the presence of particular atomic elements (electron energy loss spectroscopy).

In all cases the requirements are simply that ‘size’ can be estimated automatically using detectors or image analysis or similar techniques and that it is positively related to the quantity under study - which clearly leaves room for a vast range of modalities that can be used as input for sampling and estimating almost any well-defined quantity in 2D specimens or samples.

There is one use of microscopes that deserves special attention: the detection of rare events. A prime example is the detection and quantization of micrometastases in so-called sentinel lymph nodes, which are the primary ‘filters’ in the drain from a malignant tumor, but there are many other examples in both

pathology and toxicology. There is a range of common cancers where the presence of such metastases is of large prognostic and therapeutic significance. Their size relative to that of the lymph nodes is, however, such that a very large number of thin sections have to be examined in order to obtain the most important information, “none found”, with a sufficient degree of precision (the information may be decisive for using or not using some very unpleasant and high-risk treatment for a lethal condition). Even using specific immuno-stains providing a high contrast at a relatively low magnification the examination is very expensive in terms of skilled human labor and in clinical routine can only be performed on a rather limited number of sections per patient. One might evidently increase the real value of the procedure enormously by examining a very large number of sections automatically, and then let the expert judge a rather low number of signal-weighted fields sampled with the proportionator. Given the consequences of the information and the imperfect quality of immunological stains, no one is likely to let the last step be automatic. Although probably of secondary importance, it is a nice feature of such a procedure that a positive count would result in an unbiased estimate of the total number of micrometastases, that may turn out to be of independent prognostic significance.

Just as the simulations have indicated the magnitude of potential advantages using the proportionator, they have clearly indicated where there may be problems in the practical application and implementation of the procedure. It mostly has to do with the size-count relation, shown in Figs. 5 and 13.

The estimation procedure may break down if a count of 1 is made in a field of very low sampling probability, then the estimate from that field, the count divided by the probability, can be arbitrarily large (it is still an unbiased estimate). Evidently, fields of a very low sampling probability are rarely actually sampled, but if there are many of them, it may occur. We have not managed to generate a spatial distribution where this problem became significant. It is clearly possible to provoke the break down of the procedure if a significant fraction of cells are arbitrarily small; they would all have a count of 1 irrespective of their small size. We have not simulated this situation because it is biologically unrealistic: although cells may certainly vary much in size, their size has a definite lower limit. Another possibility is a clustered distribution where cell size and clustering is correlated such that at low densities the cell size is small (and vice versa). We consider, however, this process too special to be dealt with in this initial and explorative study. We use the expression ‘break-down’ for the effect of these problems because they are probably rare, but when present, they alone can make the procedure useless.

What will certainly occur in practice is a degrading of the ideal size-count relation by noise in the relationship. Although the proportionator enjoys freedom from the usual field-to-field variation, it has some new problems of its own, since there are several sources of noise.

First of all, the count is a discrete variable, typically of low value, whereas field ‘size’ will often be an essentially continuous variable, so their relation cannot be a continuous curve. This source of noise is, however, insignificant compared to all others.

More to the point, the stereological probes generate random and noisy hit-and-miss transforms. Even with a rather large areal fraction of colored structure in the field, the randomness of the points means that hits can range from none to all of them. Similarly, a profile may be in the field but outside the frame or the

forbidden line may intersect it. The magnitude of this noise is illustrated in Fig. 14. Its impact on the size-count relation is indicated in Fig. 5, where it is the main source of noise (saturation is the only other significant one in that example, and only in the lower panels). The counting of 3D particles in disectors adds noise since a particle may be present but not counted because its position on the z-axis forbids that (it is present in the look-up section or only present in the guard zones of the optical disector). The nature of this noise is identical to the previous one, but it may be pronounced if large particles are studied in thin disectors.

The use of thick sections for optical disector counting further degrades the size-count relation because of saturation of the signal. Histochemistry is not ideal photometry, for a number of reasons, and two or more overlapping structures in a point of a thick section do not provide a signal that is the simple sum of the signals of individual structures. As previously mentioned, the choice of a saturation of 50% in each pixel for the simulation, may, however, have been too selective.

All the above sources of noise are unavoidable, but with freedom to design the study and the sampling constants of all probes, they may be kept at bay. As an example, the noise introduced by a large mean size of cells should evidently be avoided by using a lower magnification and reducing the sizes of both the field of view (the area over which the color signal is integrated) and of the frame - no cell need to be magnified to about half the frame for proper identification. This is not simulated, because it only affects Smooth and proportionator, and we have not included any optimization in the simulation study, as already mentioned.

What may really matter in practice is the noise due to the technical preparation of the sections and the stain and technical aspects of the imaging in the microscope. Overall, the dependence on automatic image analysis makes the proportionator very sensitive to all kinds of technical imperfections and problems. Folds in the section and large precipitations of stain, both generating very strong but spurious color signals, can clearly make real havoc in the color-count relation, as an example. They are also generally incompatible with unbiased estimation and simply have to be avoided (they are, after all, due to bad laboratory practice). Uneven staining of sections is another not uncommon technical problem that will generate noise.

The microscope itself will generate noise in several ways. Bad adjustment of the light source, the condenser, and the optics evidently generates noise. Particularly at low magnification, which is likely used for the automatic, initial scanning and assigning of color values to pixels, the illumination in the field of view is uneven (more bright centrally), a source of noise which may actually be removed automatically during the image analysis using simple algorithms.

Unspecific staining is obviously a source of noise. Very few chemical stains are specific for a particular cell type, and even very specific antibodies sometimes produce some staining of the ‘background’ and frequently stain the edges of sections somewhat. This type of problem, may, however, be reduced to a low level by paying attention to it. With present sampling strategies unspecific staining is mostly an esthetical problem, a problem that not so many try to minimize or eliminate.

The long list of (mostly technical) problems above may not have much impact when efforts are put into minimizing them. But there also problems of unspecificity that are difficult to solve. Using an unspecific chemical stain like the very popular hematoxylin-eosin stain, all cell nuclei are blue. If the particular cell type under study is a small fraction of all cells present the proportionator will only be efficient if the special cell has the same spatial distribution as the majority. As previously mentioned, if the relatively rare cell under study is spatially separated from the more frequent other cells, the color-count relation may be inverse, and the proportionator is distinctly inefficient. Only a specific stain can solve such problems.

The direct CE estimation for proportionator sampling and estimation in a section solves a long-standing problem in optimizing sampling for stereological purposes. The commonly used SURS sampling strategy is systematic in 2D and the sampled fields are thus not independent. The commonly used predictions of the ***CE*** for number estimation therefore does not include a term for the contribution from the field-to-field variation, but only a prediction of the contribution from the 1D systematic sampling of sections and the ubiquitous counting noise.

As already pointed out, in all situations in Fig.6 the actual $\boldsymbol{CE^2}$ in one section is anywhere in the range from half the counting noise alone (for Smooth) to about 15-fold larger than the prediction based on just counting noise. Admittedly, none of the sampling strategies are optimized, but even with a five times smaller frame and five times more fields (about 150) for SURS in the clustered spatial distribution, the real $\boldsymbol{CE^2}$ would still be several fold larger than the prediction. In short, for the only common sampling strategy, the lack of a realistic idea about the ***CE*** seems to be a severe problem (which to a large degree we have hitherto overlooked). The direct estimation of the ***CE*** of SURS might in principle be implemented, but that has at least two problems:

- For the estimator based on two smaller samples to be as efficient as that of one large sample of size ***n***, the smaller samples must be larger than ***n/2***, i.e. the user must perform extra work in order to know just the precision.
- How much larger than ***n/2*** each of the smaller samples has to be is not known in general, and it would therefore be necessary to make a separate study of that for each application, a rather heavy investment.

On this background, it is adequate that proportionator sampling allows realistic estimates of the real ***CE*** (which includes the counting noise) without extra cost. It is first of all a prerequisite for not being overly optimistic and thereby unintentionally inflates the variation between animals (the observed variance between animals = the biological variance + the real $\boldsymbol{CE^2}$).

Moreover, the fact that proportionator sampling in roughly homogeneous tissue may have a $\boldsymbol{CE^2}$ lower than the counting noise, known from unbiased estimates of the ***CE***, not from assumed correct predictions, can of course be exploited profitably: aiming at a ***CE*** of 0.1, one may then count not 100 cells in 24 fields (Table 2) but ~70 cells in about 16 fields with a known ***CE*** of 0.1 due to field-to-field variation and counting noise, a rather welcome bonus.

Finally, the directly estimated, real ***CE*** is available for all stereological estimators (and all other estimators within the scope of proportionator sampling and estimation), including those for which predictions simply do not exist. As an example, the length of test lines for intersection count and the curvature of all structure boundaries exclude a simple prediction of the ***CE*** of surface estimation (it would depend on the curvature density).

We are not unaware of the fact that all estimates of efficiency in this study are based on simulation - another predictor! We have, however, as carefully as we could, not made the parameters of the simulation optimistic in favor of proportionator sampling (and in the same spirit we have not simulated sampling of really rare events, where proportionator sampling is essentially guaranteed to be arbitrarily more efficient than any blind sampling technique). Finally, preliminary studies of its performance on authentic tissue under true circumstances in microscopes with all kinds of noise, while measuring the time spent doing it (including outlining of reference spaces), indicate efficiencies relative to SURS in the range 8- to 25-fold (in a specific stain with severe staining of the background and a specific stain with a clean background, respectively). An ordinary chemical stain of cell nuclei in a dominating cell type with a rather inhomogeneous spatial distribution showed a relative efficiency compared to SURS of 25-fold [13].

It is therefore our belief that microscopic examinations of sections for a large range of scientific purposes, in particular stereological ones - but not only those - may be made much more efficient and user-friendly with proportionator sampling and estimation. All the necessary technology is already off-the-shelf, and thousands of specific antibody stains and similar probes are also readily available. In the near future, the fast development of hardware/software and the growing list of specific markers may make today's commonly used sampling strategy for scientific microscopy rather obsolete.

Acknowledgments

The MIND Center is supported by the Lundbeck Foundation. The project is supported by Novo Nordisk Foundation, Foundation of 17-12-1981, Eva & Henry Frænkels Mindefond, and Mindefonden for Alice Brenaa.
This work was supported by the Programme Commission on Nanoscience, Biotechnology and IT under the Danish Council for Strategic Research.

Reference List

1. D.C. Sterio, The unbiased estimation of number and sizes of arbitrary particles using the disector, J. Microsc. 134 (1984) 127-136.

2. M.M. Hansen and W.N.Hurwitz, On the theory of sampling from finite populations, Ann. Math. Statist. 14 (1943) 333-362.

3. D.G. Horvitz and D.J.Thompson, A Generalization of Sampling Without Replacement from A Finite Universe, J. Am. Statist. Assoc. 47 (1952) 663-685.

4. K.A. Dorph-Petersen, H.J.G.Gundersen, and E.B.Jensen, Non-uniform systematic sampling in stereology, J. Microsc. 200 (2000) 148-157.

5. H.J.G. Gundersen and E.B.Jensen, Stereological estimation of the volume-weighted mean volume of arbitrary particles observed on random sections, J. Microsc. 138 (1985) 127-142.

6. J.E. Gardi, J.R.Nyengaard, and H.J.G.Gundersen, Using biased image analysis for improving unbiased stereological number estimation - a pilot simulation study of the smooth fractionator, J. Microsc. 222 (2006) 242-250.

7. H.J.G. Gundersen, The smooth fractionator, J. Microsc. 207 (2002) 191-210.

8. H.J.G. Gundersen, Stereology of arbitrary particles. A review of unbiased number and size estimators and the presentation of some new ones, in memory of William R. Thompson, J. Microsc. 143 (1986) 3-45.

9. H.J.G. Gundersen, E.B.Jensen, K.Kieu, and J.Nielsen, The efficiency of systematic sampling in stereology - reconsidered, J. Microsc. 193 (1999) 199-211.

10. A.J. Baddeley, H.J.G.Gundersen, and L.M.Cruz-Orive, Estimation of surface area from vertical sections, J. Microsc. 142 (1986) 259-276.

11. H.J.G. Gundersen, The nucleator, J. Microsc. 151 (1988) 3-21.

12. E.B.V. Jensen and H.J.G.Gundersen, The Rotator, J. Microsc. 170 (1993) 35-44.

13. J.E. Gardi, J.R.Nyengaard, and H.J.G.Gundersen, Automatic sampling for unbiased and efficient stereological estimation using the proportionator in biological studies, In Press - J. Microsc. (2007).