OPTIMAL DESIGN AND ADAPTIVE DESIGN IN STEREOLOGY

by

Wei Zhang

B.S., University of Science and Technology of China, China, 2003

M.A., University of Pittsburgh, USA, 2007

Submitted to the Graduate Faculty of

the Department of Statistics in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2009

UNIVERSITY OF PITTSBURGH

STATISTICS DEPARTMENT

This dissertation was presented

by

Wei Zhang

It was defended on

October 22, 2009

and approved by

Allan R. Sampson, Ph.D., Professor

Leon J. Gleser, Ph.D., Professor

Robert A. Sweet, M.D., Professor

Yu Cheng, Ph.D., Assistant Professor

Dissertation Director: Allan R. Sampson, Ph.D., Professor

OPTIMAL DESIGN AND ADAPTIVE DESIGN IN STEREOLOGY

Wei Zhang, PhD

University of Pittsburgh, 2009

Stereology is the science that uses geometric probability to extract the internal quantitative properties of a three dimensional object based on lower dimensional information. It is a valuable research tool in biological science and relies heavily on statistical principles. In this dissertation, we focus on studies that examine the number of neurons in a brain region of interest using stereological techniques in order to compare subjects in different diagnostic groups, e. g., subjects with schizophrenia and control subjects. A large number of counting frames are usually used to obtain a prespecified precision for an individual in these kinds of studies. Typically, researchers determine the number of counting frames for each individual by controlling the coefficient of error for the individual. However, the researchers from the Conte Center for the Neuroscience of Mental Disorders (CCNMD) at University of Pittsburgh primarily focus on comparing biomarkers among different diagnosis groups rather than evaluating individuals. A design goal for such stereological studies is to keep study cost within budget and time constraints, while maintaining sufficient statistical power to address the research aims. Statistical power can be increased by either adding more subjects or more counting frames. And the cost of a study can be approximated by a linear combination of the number of subjects and number of counting frames. To address this need, we have developed new technologies that enable researchers to design a cost efficient study balancing the number of subjects with the number of counting frames for each subject.

We also develop adaptive designs to conduct stereological studies. Adaptive designs allow the opportunity to look at the data at an interim stage, and to modify the design based on

the information obtained from the first stage data. In our adaptive design, we estimate the stereological variance without breaking the blind of the Stage I data, and re-design the second stage based on the stereological variance estimator obtained from the first stage. Based on our procedure, we show researchers can cost-effectively modify the design while maintaining the desired study power.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

Writing a dissertation was an unforgettable experience for me. During the last six years, there were numerous people kindly offering me their help. My most sincere acknowledgment goes to my esteemed advisor Dr. Allan Sampson. His mentoring, research support, patience and encouragement throughout my graduate studies are the crucial parts for me finishing my dissertation. He was always there to listen and to give advice. He not only taught me the important lessons to be successful in academia, but also the lessons to be successful in life.

I was fortunate to have the opportunity to work under Dr. Sampson's supervision for the Conte Center for the Neuroscience of Mental Disorders (CCNMD) at University of Pittsburgh. I enjoyed every moments that I worked with them. I want to thank everybody who I worked with in CCNMD, especially Dr. Robert Sweet. He introduced the world of neuroscience to me, and provided his expert advise to my research. I really enjoyed his expertise and friendship. I also want to thank Dr. Karl-Anton Dorph-Petersen and Dr. Glenn Konopaske for kindly providing the data for my research.

I gladly express my gratitude to my committee members Dr. Leon Gleser and Dr. Yu Cheng for their time and guidance. Dr. Gleser always welcomed my visit, and inspired my thought by sharing his knowledge and wisdom. Dr. Yu Cheng was willing to share her experience of the dissertation writing. Without their help, I won't be able to accomplish my work.

Finally, it is impossible to have my research without my husband's love, support and understanding, as well as my family and friends' encouragement. This dissertation is dedicated to all of them.

1.0 INTRODUCTION

Psychiatrists and neuroscientists devote a considerable amount of effort to explore how biological structures in various brain regions of persons with mental diseases differ from those in normal individuals. Studies in animals may also be used to detect the effects of pharmacologic treatments on various types of brain cells in particular regions. Neuroscientists in the Department of Psychiatry at the University of Pittsburgh often use post-mortem tissue samples from the Brain Bank Core of the Conte Center for the Neuroscience of Mental Disorder (CCNMD) to assess biological alterations in subjects with schizophrenia (e.g., Dorph-Petersen et al. (2007)), or to understand the neuropharmacologic effects of treatment by performing animals studies (i.e., Konopaske et al. (2007), (2008)). The neurobiological measurements, such as the number of particular type of neurons or cells, the volume of brain regions and the density of specific cell types, are often the main focus of these studies. In this dissertation, we concentrate on studies that examine the number of neurons in particular brain regions. Neuron number ($\mathcal{N}$) is an important indicator of neurobiological alterations in schizophrenia. A pathologic manifestation of schizophrenia appears to be reflected in the reduction of the number of functional neurons in particular cortical regions. For example, subjects with schizophrenia show deficits in visual perception and one of CCNMD studies (Dorph-Petersen et al. 2007) found a substantial reduction in neuron number of the primary visual cortex in the postmortem tissue from subjects with schizophrenia.

It would be an intractable task to physically count all the neurons of interest for a subject in these types of studies. For example, there are about $2 \sim 5$ million neurons in the human primary visual cortex (Dorph-Petersen et al. (2007)). Therefore, it is necessary to sample tissue specimens to estimate the neuron number for an individual subject. Usually, the brain regions to be analyzed are cut into sections in a systematic, uniformly random way, and

then, counting frames within each section are also collected in a uniformly systematic way. Stereological techniques are applied to estimate the number of particular type of neurons or cells ($\mathcal{N}$) in the brain region of interest for a subject using the chosen sampling scheme. Stereology uses geometric probability to extract the internal quantitative properties of a three dimensional object based on lower dimensional information. In particular, stereological procedures provide the mathematical and statistical framework for developing techniques and methodologies for constructing estimates of various biological quantities (Jensen (1987)). Unbiasedness considerations for stereological estimators were studied and discussed by Miles and Davy (1976), Cruz-Orive (1980), Cruz-Orive and Weibel (1981), Jensen and Sunberg (1986) and Jensen (1987). But to evaluate the precision of a stereological estimator when the sampling scheme, as is typical, is based on uniform systematic samples is a complex problem. There is no simple and exact formula for the stereological variance.

According to the typically used double systematic sampling scheme, the stereological variance can be decomposed into two parts, between sections variance and within section variance. The first part of the stereological variance depends on a *covariogram* function which is defined by the true distribution of the neuron number in the region of interest. Gundersen and Jensen (1987), Gundersen et al. (1999) gave two estimators of between sections variance based on two assumed simple forms of the *covariogram*. Determination of the second part of the stereological variance depends on a two dimensional version of the *covarigram* which is very difficult to obtain due to the irregularly shaped sections and insufficient information. In order to obtain an approximation of the second part of the stereological variance, different procedures have been developed. In Cruz-Orive and Geiser (2004), the authors suggest a Poisson model to fit stereological data.

From 2005 to 2007, Dr. Konopaske performed a series of studies (i.e., Konopaske et al. (2007), (2008)) on chronic exposure of macaque monkeys with two antipsychotic treatments to assess whether or not treatment with antipsychotic medication contributes to the disturbances in the number of neurons, glial cells and subtypes of glial cells in individuals with schizophrenia. There were 18 male macaque monkeys (4.5-5.3 years of age) which were divided into 3 experimental groups ($n = 6$ per group). Two experimental groups were treated for 27 months with either haloperidol or olanzapine, and a third was given a sham treatment.

Haloperidol is an antipsychotic medication that has been used for more than 50 years, while olanzapine is a fairly new atypical antipsychotic. For each of several studies, Dr. Konopaske and his colleagues spent a number of months collecting the data. In these studies, the sample tissues are ready for use; however, they need to decide how much effort they should put in measuring each subject before beginning a study.

In this dissertation, we further develop the statistical theory behind the current variance estimator of Cruz-Orive and Geiser (2004) (Sections 2.2 and 2.3). Using the data that formed the basis for the analysis in Konopaske et al. (2007), we illustrate the methodology (Section 2.4). We then show that the validity of the Poisson assumption of Cruz-Orive and Geiser (2004) is questionable. The Poisson assumption may be appropriate for some cases but, as we show, the Konopaske data indicates evidence of overdispersion relative to the Poisson distribution (Section 2.4). This overdispersion of the data motivates us to find another model which fits the data better. In our research, we investigate a special Cox process, called the *Ammeter* process, to obtain a more accurate estimate of the stereological variance (Section 2.5). Basically, the *Ammeter* process is a Poisson process on $\mathbb{R}^1$ where the parameter λ is a random parameter instead of a constant. To apply the idea of the *Ammeter* process to the stereology setting, we extend the *Ammeter* process to a two-dimensional version to the tissue section data to take into account the sampling features of stereology. Based on the *Ammeter* process and the estimators of between section variance given by Gundersen and Jensen (1987) and Gundersen et al. (1999), we obtain a new estimator of the stereological variance (Section 2.5). Then we use the maximum likelihood estimates (MLE) of the *Ammeter* process model parameters to derive an improved stereological variance for the Konopaske data. In addition, as a possible alternative approach to estimate the stereological variance, we briefly consider a Bootstrap approach (Section 2.6). The stereological variance estimators obtained by the Cruz-Orive and Geiser method and our two new approaches are compared in Section 2.7.

In Chapter 3, we focus on designing a stereological study, where one of the important considerations is to select the appropriate number of subjects and number of sampling frames. In a stereological study, researchers generally decide on the number of sampling frames that are needed for each subject by what is required to precisely estimate the quantity of interest for each tissue specimen for a given subject. They do so by controlling the coefficient of

error (CE = standard deviation / mean) of the stereological estimate of ($\mathcal{N}$) based on each individual subject's tissue sample. In Konopaske et al. (2007), about 1600 counting frames were used for each animal and this took several months of work to prepare and observe all the sampling frames using a microscope. For a setting where it is clinically required to obtain the best estimate of a particular value for a given subject, it is reasonable to focus on this level of accuracy. However, for studies designed for CCNMD, the interest is more focused on comparing the biological measurements among different population groups. In the statistical context, the aim of an experimental design is to ensure that comparisons among the different populations are unbiased and, moreover, as precise and powerful as possible given the experimental cost and time constraints.

Our ultimate goal is to develop a procedure for planning a cost efficient stereological study. Statistical power can be increased by either adding more subjects or more sampling frames. The cost of a study can be approximated by a linear combination of the number of subjects and the number of sampling frames (Section 3.1.1). In our research, we consider designs with a fixed power, and obtain an algorithm to find the the number of subjects and the number of sampling frames that minimize the cost function (Section 3.1.3.1). Another approach we obtain considers a fixed cost budget, and we provide the combination of subjects and sampling frames that provide the maximum statistical power (Section 3.1.3.2). Because a fairly standard CCNMD design is to use matched pairs, we simplify our power considerations by examining in detail the paired t-test. In addition, one needs to consider the effect of unequal stereological variances among groups under the alternative hypothesis.

To design a stereological study, we require the information about the magnitude of the true stereological variance. However, as shown in Chapter 2, the stereological variance depends on the shape of the region of interest and also on the neuron density. It is usually difficult to prespecify the stereological variance before undertaking a study which focuses on a particular type of neuron. To avoid inefficient use of the resources, we apply adaptive design in the stereological studies to choose the optimal numbers of sampling frames to maintain power and keep certain costs as low as possible. In Chapter 4, we first review some existing literature on the topic of both blinded and unblinded adaptive designs (Proschan and Hunsberger (1995), Gould and Shih (1998), Kieser and Friede (2001), Shun (2001) and

Liu and Chi (2001)).

We use the approach of blinded adaptive procedures for stereological studies. Specifically, in our adaptive approach, we only estimate the stereological variance without breaking the blind of the Stage I data. We develop an approach that at the end of Stage I allows us to update the assumptions about the stereological variance that were used in the planning stage. In the setting, we consider the number of subjects is fixed, and we don't stop the study earlier. Based on the updated stereological variance, we change the number of counting frames to be used in Stage II while maintaining the power. Because stereological variance of the second stage differs from the first stage whenever the number of sampling frames changes in Stage II, the statistical procedure used to test group effect needs to be handled with care. In Section 4.3.5 we obtain an adjusted t-statistic to use for hypothesis testing at the end of the study. We use simulation to show that the type I error rate of our procedure is protected.

2.0 STEREOLOGICAL ESTIMATE AND VARIANCE

Stereological methods extract the internal quantitative information of a three dimensional object based on lower dimensional information. Stereology relies heavily on statistical principles, especially random sampling and sampling inference (Baddeley and Vedel Jensen (2005)). In performing an experiment involving tissue sampling using tissue stereological techniques, researchers usually decide on the number of sampling frames that are needed based on what is required to precisely estimate the quantity of interest. The coefficient of error (CE = standard deviation / mean) of the stereological estimate for the quantity of interest based on a tissue sample is usually controlled to be about 5%. This criterion apparently grew from the desire to estimate with precision the desired quantitative information for a specific individual or patient. This has been carried over in studies which compare parameters based upon samples from different populations. By this CE criterion, thousands of sampling frames may be required for each individual, which can require researchers to do a large amount of work requiring the use of programmable microscopes. For example, Dr. Konopaske performed a series of studies in the Lewis lab from 2005 to 2007. He spent about four months collecting data which consisted of about 25,000 observations (sampling frames) of cell types in the 18 monkeys that were in each study. One of the studies that was completed in 2005 is discussed in Section 2.4 in detail. The main goal of these studies in the CCNMD is always to detect the possible neurobiological difference or the treatment effect between groups of interest, which means that statistical inference about the population is more important than the inference about an individual. More specifically, the subject to subject variability of the primary outcome tends to be more important than the individual subject's measurement variability when we power stereological studies.

To do experimental design, we obviously need to connect the study power with the

stereological variability and the biological variability. To better understand the stereological variability, we consider and evaluate stereological techniques carefully in this chapter.

2.1 STEREOLOGICAL SAMPLING AND STEREOLOGICAL ESTIMATION FOR COUNTING

It is a well known fact that there are about a hundred billion neurons in a human brain. Therefore, it is impossible to physically count all the neurons of a particular type in a region of interest. It is necessary to sample tissues to estimate the number of neurons. Only comparatively very small pieces of tissue can be observed. One of the standard sampling schemes commonly used in stereology to count neurons or other cells within a region is uniformly systematic sampling, which proceeds in multiple stages. A brain region of interest with length L is exhaustively divided to M sections with thickness $\frac{L}{M}$ and m sections are selected for the study. The location of the initial section related to the one end of the brain region can for technical cutting reasons be viewed as having a uniform distribution on [0, $\frac{L}{M}$). The first sampled section is chosen from the first $\frac{M}{m}$ sections with equal probability. For simplicity, we assume M is an integer multiple of m. We further assume X_1 to be the position of the first selected section relative to the one end of the brain region, and $X_j = X_1 + (j-1)\frac{L}{m}$, $j = 2, \ldots, m$, to be the positions of the remaining $m-1$ sections selected by systematic sampling. Considering the inherent randomness of where the tissue of interest actually begins in the first section physically cut, and the independently chosen the first sampled section, the position of the first selected sample section can be viewed as uniformly randomly chosen between 0 and $\frac{L}{m}$, that is, $X_1 \sim U[0, \frac{L}{m})$, and $X_j = X_1 + (j-1)\frac{L}{m}$, $j = 2, \ldots, m$. The section sampling fraction (τ_s) which is sampled using systematic sampling, is $\frac{m}{M}$. In Figure 2.1, the ellipse in the left graph represents a three dimensional region of interest intersected by a series of systematic uniform random cutting planes, represented by a set of lines. The graph on the left represents a series of selected sections which are the grey stripes inside the ellipse. The darkened regions in the right hand graph (Schmitz and Hof (2007)) are the brain region of interest. In this example, a total of 50 sections ($M = 50$) were cut, and 10 sections ($m =$

10) were systematically selected. The first selected systematic sampling section is at $\frac{3L}{50}$ and then every fifth section ($\tau_s = 1/5$) subsequently.

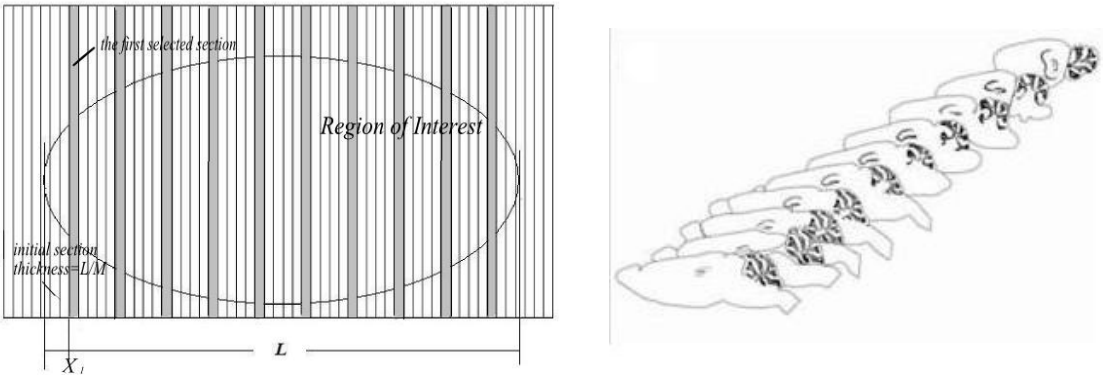


Figure 2.1: *Section Sampling Fraction* $\tau_s = \frac{m}{M}$

Second, a rectangular lattice grid, which has a fixed distance between cross lines called the u-step (a_u) and the v-step (a_v), is randomly superimposed on each section. The same u-step and v-step distances are used for all the sections within each individual. Within each rectangular grid a smaller rectangle is that the number of particles can be obtained for that smaller piece of tissue. This is done because the larger rectangular grid may include too many cells to measure. Oftentimes, one chooses this smaller rectangular frame, which is called the counting frame, to be the upper left corner of the rectangular grid, but in fact it can be any consistent area within the rectangular grid. The physical frame examined actually has depth and has a "cubic" structure, that is, it's three dimensional. The size of the rectangle is related to the area sampling fraction to be described shortly. The number of rectangular grids determines the number of counting frames that can be obtained from that tissue cross-section. Clearly, depending on the area of the cross-section, the numbers of counting frames may differ from section to section. Furthermore, for a given section, changing a_u and a_v will also change the number of counting frames. For each section j, we define a random vector (U_{1j}, V_{1j}) as follows. Find the leftmost rectangular grid entirely within the cross section. If two or more such rectangular grids exist, choose the uppermost. For this specific grid, let U_{1j} be the distance from the upper left corner to the tissue edge moving along the u-direction, and V_{1j} be the same in the v-direction. Then the counting

frame for that grid is the upper-left smaller square, and all other counting frames are chosen within any rectangular grid, as long as the upper left hand corner of the counting frame remains within the tissue cross section. The set of upper-most left hand corners of these counting frames is given by

$$\mathcal{P}_j = \{(U_{1j} + ka_u, V_{1j} + la_v) : (U_{1j} + ka_u, V_{1j} + la_v) \in R_j : k, l \in \mathbb{Z}\}$$

where R_j is the range of the j^{th} section, for $j = 1, \cdots, m$. The position (U_{1j}, V_{1j}) determines

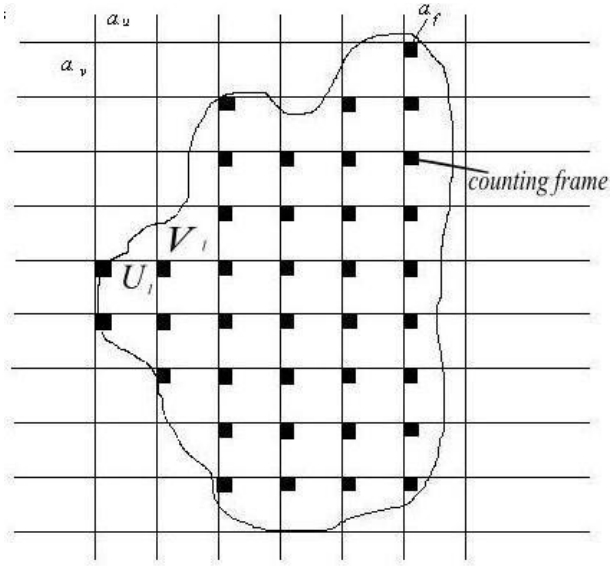


Figure 2.2: *Area Sampling Fraction* $\tau_a = a_f^2/(a_u \cdot a_v)$

the position of the entire grid $\mathcal{P}_j$ in the j^{th} section. Note that by construction, U_{1j}, V_{1j} can be considered to be uniformly randomly selected within $[0, a_u) \times [0, a_v)$. The initial quantities U_{1j} and V_{1j} for each tissue cross-section are chosen independently of each other and the position of the first selected section X_1. More specifically, $(U_{11}, V_{11}), \cdots, (U_{1,m}, V_{1,m})|X_1 = x$ are conditionally *i.i.d.* according to $(U, V) \sim Unif([0, a_u) \times [0, a_v))$ for any x. Hence $(U_{11}, V_{11}), \cdots, (U_{1,m}, V_{1,m})$ are independent of X_1. The little black square in Figure 2.2 illustrates a counting frame with length a_f is chosen at the upper left corner of each rectangle, and as noted is the specific tissue observed under the microscope. The area of the counting frame is fixed and the area associated with each u, v step (a_u multiplied by a_v) is also given. The area sampling fraction (τ_a) of this systematic sample can then be calculated as

$a_f^2/(a_u a_v)$, and it is the same among all sections within each individual. Figure 2.2 gives a two dimensional version of a cross-section with counting frames illustrated, and counting frames selected within each cross-section is a two dimensional systematic random sampling.

Third, to allow unbiased counting rules based on the thick sections with thickness $\frac{L}{M}$, a smaller thickness h is examined in detail where $h \leq \frac{L}{M}$. The height of the counting frame is assumed known relative to the thickness of the sections. The height sampling fraction (τ_h) can then be calculated as $h/\frac{L}{M}$. The numbers of particular type of neurons or cells within the selected counting frame between X_j and $X_j + h$ are recorded. Figure 2.3 illustrates a three dimensional version of a rectangle and a counting frame in Figure 2.2.

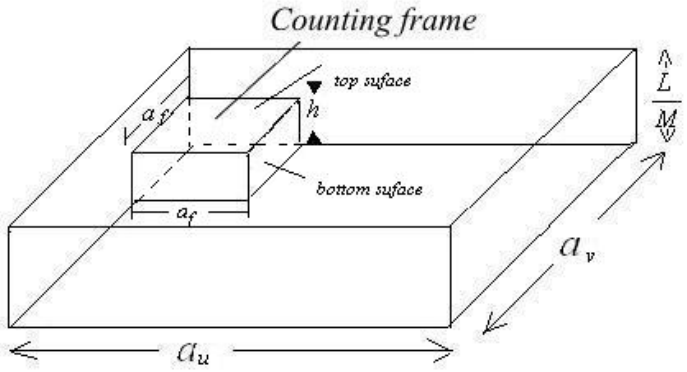


Figure 2.3: *Height Sampling Fraction* $\tau_h = h/\frac{L}{M}$

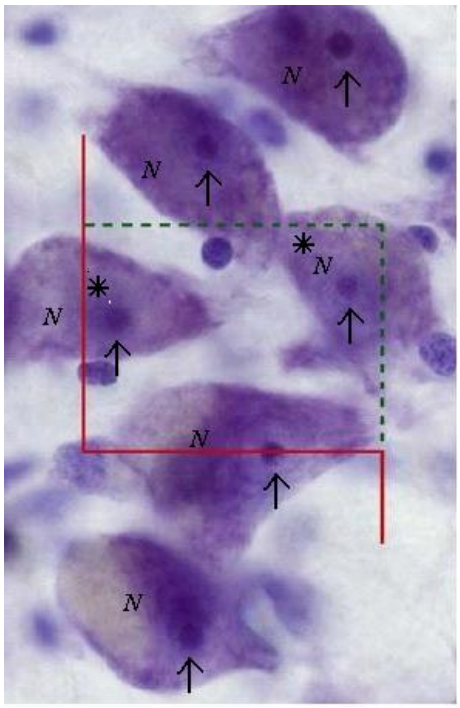
The micrograph in Figure 2.4 is from a study of the neuron number in the primary visual cortex conducted by Dorph-Petersen et al. (2007). An image of a typical counting frame is shown with a solid (red) exclusion line and a dotted (green) inclusion line. A neuron is counted if: (1) its nucleolus is in focus fully or partially inside the counting frame without touching the exclusion line, and (2) its nucleolus is fully or partially below the top surface without touching the bottom surface. As shown in Figure 2.4, there are six typical magnocellular neurons in total. The darker spot marked by "↑ " within the neuron is the nucleolus. In this micrograph, the two neurons in the middle with label "∗" are counted.

The overall fraction of the brain region sampled from these three sampling fractions is $\tau_s \cdot \tau_a \cdot \tau_h$. The number of a particular type of neurons within a counting frame can be obtained, using a programmable microscope for sampling and visualization of the particular

neurons of interest. Let Q_j denote the total number of neurons actually counted by using the microscope in the j^{th} section of this selected random systematic sampling. The total number of neurons actually counted in all of the counting frames is $Q = \sum_{j=1}^{m} Q_j$. Then the total number of neurons in the region is estimated directly from

$$\hat{\mathcal{N}} = \cdot \frac{1}{\tau_s} \cdot \frac{1}{\tau_a} \cdot \frac{1}{\tau_h} Q. \tag{2.1}$$

$\hat{\mathcal{N}}$ is called the fractionator estimator, and is a primary outcome of the stereological studies. The estimator $\hat{\mathcal{N}}$ can be shown to be an unbiased estimator of the true total number of neurons $\mathcal{N}$ irrespective of tissue shrinkage or swelling that may occur when the tissue is processed (Dorph-Petersen et al. (2001)).



N - Neurons, ✳ - Counted, ↑ - Nucleolus.

Figure 2.4: *Image of a counting frame (Dorph-Petersen et al. (2007))*

2.2 THE DIFFICULTY OF VARIANCE ESTIMATION

From the theory of systematic sampling (Cochran (1977)), it is known that it is impossible to develop an unbiased estimate of the stereological variance based upon a single systematic sample. Since the variance of the fractionator estimator is important for the studies concerning group differences of the neurobiological measurements, we would like to better understand the stereological variance. Based on the stereological sampling described in the previous section, we have two systematic samples for each tissue block: the random systematic section sampling and the random systematic area sampling within a section. Therefore, the stereological variation is from these two sources, which are called the between sections variation and the within section variation. The first component is due to the random choice of the first section, i.e., choosing X_1, while the second component is due to the random choice of the position of the first counting frame on the u and v axis, i.e., choosing (U_{1j}, V_{1j}), $j = 1$, $\ldots$, m, under the random systematic sampling method.

We now consider the stereological variance problem in further detail. In general, the true number of a particular type of neurons can be represented as an integral of neuron's density function $s(x, u, v)$ in three-dimensions:

$$\mathcal{N} = \int_0^L \left\{ \iint_{(u,v)\in R(x)} s(x, u, v) \mathrm{d}u\mathrm{d}v \right\} \mathrm{d}x, \tag{2.2}$$

where $s(x, u, v)$ is a density function of a particular type of neuron over the section at x, and $R(x)$ is the range of cross section values of (U, V) at position x. To simplify the problem, let us consider the problem of a one-dimension integral first,

$$\mathcal{N} = \int_0^L f(x) \mathrm{d}x, \tag{2.3}$$

where $f(x)$ is a nonnegative measurement function over the interval $[0, L]$,

$$f(x) = \iint_{(u,v)\in R(x)} s(x, u, v) \mathrm{d}u\mathrm{d}v. \tag{2.4}$$

The function $f(x)$ represents the number of neurons at position x. Let $X_1 \sim U[0, \frac{L}{m})$ and $X_j = X_1 + (j-1)\frac{L}{m}$. Then, if we were to measure $f(x)$ without error at each of

$X_1, X_2, \cdots, X_m$, an approximation of the integral would be given by

$$\begin{aligned}\widetilde{\mathcal{N}} &= \frac{L}{m}\sum_{j=1}^{m} f(X_j) \\ &= \frac{L}{m}\sum_{j=1}^{m} f(X_1 + (j-1)\frac{L}{m}).\end{aligned} \tag{2.5}$$

Since

$$\begin{aligned}E[\widetilde{\mathcal{N}}] &= \int_0^{\frac{L}{m}} \frac{L}{m}\sum_{j=1}^{m} f(x_1 + (j-1)\frac{L}{m})\frac{dx_1}{\frac{L}{m}} \\ &= \sum_{j=1}^{m}\int_0^{\frac{L}{m}} f(x_1 + (j-1)\frac{L}{m})dx_1 \\ &= \sum_{j=1}^{m}\int_{(j-1)\frac{L}{m}}^{j\frac{L}{m}} f(x)dx \\ &= \int_0^{L} f(x)dx \\ &= \mathcal{N},\end{aligned} \tag{2.6}$$

$\widetilde{\mathcal{N}}$ is an unbiased estimator of $\mathcal{N}$. Let $\hat{f}(X_1 + (j-1)\frac{L}{m}, U_{1j}, V_{1j})$ (which we abbreviate as $\hat{f}_j$) be the estimator of the total number of neurons per unit of thickness at the section determined by X_j given both X_1 and (U_{1j}, V_{1j}) as the position of the first section and the first counting frame, respectively. Thus, using the same argument as in Section 2.1, we have

$$\hat{f}(X_1 + (j-1)\frac{L}{m}, U_{1j}, V_{1j}) = \frac{1}{\tau_a h}Q_j, \tag{2.7}$$

where Q_j is the total number of neurons actually counted in all counting frames within the section between $X_j = X_1 + (j-1)\frac{L}{m}$ and $X_j + h = X_1 + (j-1)\frac{L}{m} + h$. Then the stereological

estimate $\hat{\mathcal{N}}$ in (2.1) can be denoted as,

$$\begin{aligned}\hat{\mathcal{N}} &= \frac{1}{\tau_s}\cdot\frac{1}{\tau_a}\cdot\frac{1}{\tau_h}\sum_{j=1}^{m} Q_j \\ &= \frac{1}{\frac{m}{M}}\frac{1}{\tau_a}\frac{1}{\frac{h}{\frac{L}{M}}}\sum_{j=1}^{m} Q_j \\ &= \frac{L}{m}\sum_{j=1}^{m}\frac{1}{\tau_a h}Q_j \\ &= \frac{L}{m}\sum_{j=1}^{m}\hat{f}(X_1+(j-1)\frac{L}{m}, U_{1j}, V_{1j}). \end{aligned} \tag{2.8}$$

By a similar argument to that in (2.6), we can show that $\hat{\mathcal{N}}$ is an unbiased estimator of $\mathcal{N}$. By the standard variance decomposition, the variance of $\hat{\mathcal{N}}$ is,

$$\begin{aligned} var[\hat{\mathcal{N}}] &= var[E(\hat{\mathcal{N}})|X_1] + E[var(\hat{\mathcal{N}}|X_1)] \\ &= var[E[\frac{L}{m}\sum_{j=1}^{m}\hat{f}(X_1+(j-1)\frac{L}{m}, U_{1j}, V_{1j}]|X_1]+ \\ &\quad E[var[\frac{L}{m}\sum_{j=1}^{m}\hat{f}(X_1+(j-1)\frac{L}{m}, U_{1j}, V_{1j}]|X_1]. \end{aligned} \tag{2.9}$$

In (2.9),

$$\begin{aligned} E[\frac{L}{m}\sum_{j=1}^{m}\hat{f}(X_j, U_{1j}, V_{1j})|X_1] &= \frac{L}{m}\sum_{j=1}^{m}E[\hat{f}(X_1+(j-1)\frac{L}{m}, U_{1j}, V_{1j})|X_1] \\ &= \frac{L}{m}\sum_{j=1}^{m}\iint_{(u,v)\in R(X_1+(j-1)\frac{L}{m})} s(X_1+(j-1)\frac{L}{m}, u, v)\mathrm{d}u\mathrm{d}v \\ &= \frac{L}{m}\sum_{j=1}^{m} f(X_1+(j-1)\frac{L}{m}) \\ &= \widetilde{\mathcal{N}}, \end{aligned} \tag{2.10}$$

so that the first term in the right hand side of (2.9) is $var[\widetilde{\mathcal{N}}]$.

The second term on the right hand side of (2.9) is

$$E[var[\frac{L}{m}\sum_{j}^{m}\hat{f}(X_j,U_{1j},V_{1j})|X_1]] = E[(\frac{L}{m})^2\sum_{j=1}^{m}var[\hat{f}(X_1+(j-1)\frac{L}{m},U_{1j},V_{1j})|X_1]] \quad (2.10a)$$

$$= (\frac{L}{m})^2\sum_{j=1}^{m}E[var[\hat{f}(X_1+(j-1)\frac{L}{m},U_{1j},V_{1j})|X_1]] \quad (2.10b)$$

$$= (\frac{L}{m})^2\sum_{j=1}^{m}E[var[\hat{f}(X_1+(j-1)\frac{L}{m},U,V)|X_1]] \quad (2.10c)$$

$$= (\frac{L}{m})^2\sum_{j=1}^{m}\int_0^{\frac{L}{m}}var[\hat{f}(x_1+(j-1)\frac{L}{m},U,V)|x_1]\frac{\mathrm{d}x_1}{\frac{L}{m}} \quad (2.10d)$$

$$= \frac{L}{m}\sum_{j=1}^{m}\int_{(j-1)\frac{L}{m}}^{j\frac{L}{m}}var[\hat{f}(x,U,V)|x]\mathrm{d}x \quad (2.10e)$$

$$= \frac{L}{m}\int_0^L var[\hat{f}(x,U,V)|x]\mathrm{d}x, \quad (2.10f)$$

where $(2.10a)$ follows from $\hat{f}(X_1+(j-1)\frac{L}{m},U_{1j},V_{1j})$ being independent of $\hat{f}(X_1+(k-1)\frac{L}{m},U_{1k},V_{1k})$, $j\neq k$, for given X_1. Since $(U_{11},V_{11}),\cdots,(U_{1,m},V_{1,m})$ are *i.i.d.* according to $(U,V)\sim Unif([0,a_u)\times[0,a_v))$, $(2.10c)$ follows from the fact that the marginal distribution of $U,V|X$ does not depend on X. $(2.10d)$ follows from the uniform distribution of X_1; and $(2.10e)$ follows from a change of variables.

Hence, combining (2.10) and $(2.10f)$, we obtain

$$\sigma_{st} = var[\widetilde{\mathcal{N}}] + \frac{L}{m}\int_0^L var[\hat{f}(x,U,V)|x]\mathrm{d}x, \quad (2.11)$$

where we use σ_{st} as the notation of the stereological variance of the stereological estimator $\hat{\mathcal{N}}$, and where σ_{st1} and σ_{st2} are respectively the notation for the first and second component terms of the stereological variance in (2.11). Note that the between-section variance σ_{st1} is the same as the variance of $\widetilde{\mathcal{N}}$. The estimation of the between section variance can be examined by considering the variation of approximating the integral of function $f(x)$ over the interval $[0,L]$ using a systematic sample. To calculate this variance, usually the *covariogram* function $g_f(t)$ defined by (2.12) is introduced. The within section variance σ_{st2} is the integral over X of the conditional variance of $\hat{f}_j$ given $X_1=x$ and is determined by the shape of the density function $s(x,u,v)$ in the j^{th} section, $j=1,\ldots,m$. Note that this quantity depends

on the conditional distribution of $U_{1j}, V_{1j}|X_1$. As we later discuss, to compute this variance would require a notion of a two dimensional *covarigram* over a complex region, so other approaches are needed. To be clear, even if we know the true value of $f(X_j)$, which means the second term on the right hand side of (2.11) is zero, there is still variation in terms of $var[\widetilde{\mathcal{N}}]$.

To calculate the between section variance, $var[\widetilde{\mathcal{N}}]$, we need the *covariogram* function $g_f(t)$ defined as

$$g_f(t) = \begin{cases} \int_0^{L-t} f(x+t)f(x)\mathrm{d}x & 0 \le t \le L \\ \int_{-t}^{L} f(x+t)f(x)\mathrm{d}x & -L \le t \le 0 \\ 0 & otherwise. \end{cases} \tag{2.12}$$

the function $g_f(t)$ reflects in some sense the correlation of the measurement function f between two slices separated by a distance t.

Then $var[\widetilde{\mathcal{N}}]$ is given by

$$var[\widetilde{\mathcal{N}}] = \frac{L}{m} g_f(0) + 2\frac{L}{m}\sum_{l=1}^{m} g_f(l\frac{L}{m}) - 2\int_0^L g_f(y)\mathrm{d}y, \tag{2.13}$$

(see Correa (2001)).

A standard simplifying assumption is that $g_f(t)$ has the form $at^2 + bt + c$, $b < 0$ (see Gundersen and Jensen (1987)). Then it can be shown that

$$var[\widetilde{\mathcal{N}}] = -\frac{1}{6}(\frac{L}{m})^2 b. \tag{2.14}$$

Alternatively, if it is assumed that $g_f(t)$ has the form $b_3t^3 + b_2t^2 + b_0$ (see Gundersen et al. (1999)), then it can be shown that

$$var[\widetilde{\mathcal{N}}] = \frac{1}{60}(\frac{L}{m})^4 b_3. \tag{2.15}$$

We denote the variances in (2.14) and (2.15), respectively, by $\sigma^a_{st1}[\widetilde{\mathcal{N}}]$ and $\sigma^b_{st1}[\widetilde{\mathcal{N}}]$, both of which can be viewed approximations of σ_{st1}. The standard approach is to approximate

the *covariogram* function $g_f(t)$ at $0, \frac{L}{m}, 2\frac{L}{m}, \cdots, (m-2)\frac{L}{m}$ by $\widehat{g_f}(t)$, where

$$\widehat{g_f}(k\frac{L}{m}) = \frac{L}{m}\sum_{j=1}^{m-k} f_j f_{j+k}. \tag{2.16}$$

To estimate the coefficients a, b and c in the quadratic function and to obtain an estimate of $\sigma^a_{st1}[\widetilde{\mathcal{N}}]$, the standard approach uses, the estimates of the first three *covariogram* terms and solves

$$\begin{aligned}
&\widehat{g_f}(0) = \hat{c},\\
&\widehat{g_f}(\frac{L}{m}) = \hat{a}(\frac{L}{m})^2 + \hat{b}(\frac{L}{m}) + \hat{c},\\
&\widehat{g_f}(2\frac{L}{m}) = 4\hat{a}(\frac{L}{m})^2 + 2\hat{b}(\frac{L}{m}) + \hat{c}.
\end{aligned} \tag{2.17}$$

The solution for b can be obtained as

$$\hat{b} = -\frac{3\widehat{g_f}(0) - 4\widehat{g_f}(\frac{L}{m}) + \widehat{g_f}(2\frac{L}{m})}{2\frac{L}{m}}, \tag{2.18}$$

and hence, we estimate σ_{st1} by

$$\widehat{\sigma}^a_{st1}[\widetilde{\mathcal{N}}] = \frac{L^2}{12m^2}(3\sum_{\hat{j}=1}^{m} f_j^2 + \sum_{j=1}^{m-2} f_j f_{j+2} - 4\sum_{j=1}^{m-1} f_j f_{j+1}), \tag{2.19}$$

where $f_j = f(X_j, U_{1j}, V_{1j})$, $X_j = X_1 + (j-1)\frac{L}{m}$, and $j = 1, \ldots, m$.

Similarly, we can obtain the solution for b_3 in (2.15) as

$$\widehat{b_3} = \frac{3\widehat{g_f}(0) - 4\widehat{g_f}(\frac{L}{m}) + \widehat{g_f}(2\frac{L}{m})}{4(\frac{L}{m})^3}, \tag{2.20}$$

so that, we can also estimate σ_{st1} by

$$\widehat{\sigma}^b_{st1}[\widetilde{\mathcal{N}}] = \frac{L^2}{240m^2}(3\sum_{j=1}^{m} f_j^2 + \sum_{j=1}^{m-2} f_j f_{j+2} - 4\sum_{j=1}^{m-1} f_j f_{j+1}), \tag{2.21}$$

where $f_j = f(X_j, U_{1j}, V_{1j})$, $X_j = X_1 + (j-1)\frac{L}{m}$, and $j = 1, \ldots, m$.

We now consider in further detail the second term in the right hand side of the variance

formula (2.11). Recall that

$$f(x) = \iint_{(u,v)\in R(x)} s(x,u,v)\mathrm{d}u\mathrm{d}v, \tag{2.22}$$

so that

$$\begin{aligned}&\hat{f}(X_j, U_{1j}, V_{1j})\\ &\quad = a_x a_y \sum\sum_{\{k,l:(U_{1j}+(k-1)a_x, V_{1j}+(l-1)a_y)\in R_j\}} s(X_j, U_{1j}+(k-1)a_x, V_{1j}+(l-1)a_y),\\ &j = 1,\ldots,m.\end{aligned} \tag{2.23}$$

Hence by extending the approximations underlying (2.5) and (2.6) to a two dimensional systematic sample, we have that $\hat{f}(X_j, U_{1j}, V_{1j})$ is an unbiased estimator of $f(X_j)$. From the random systematic sample and the idea of the one dimensional *covariogram* we see that the between section variation is a function of the density function $s(x,u,v)$. To compute $var[\hat{f}(X_1+(j-1)\frac{L}{m}, U_{1j}, V_{1j})|X_1]$, we need a two-dimensional version of *covariogram* and a method to deal with the problem of irregularly shaped sections. It is very difficult to work with a two dimensional *covariogram*. Furthermore, to actually estimate the within section variance, a bivariate polynomial would be needed to approximate the bivariate *covarigram* in the spirit that Gundersen and Jensen (1987) and Gundersen et al. (1999) did for the univariate case.

To avoid using the two dimensional version of the *covariogram* function $g_s^*(x,y)$, Cruz-Orive and Geiser (2004) tried to make some simplifying assumptions on the neurons' distribution within the sections, which we discuss in the following section.

2.3 CRUZ-ORIVE AND GEISER'S STEREOLOGICAL VARIANCE ESTIMATOR

Cruz-Orive and Geiser (2004) stated that no simple formula has been developed for the estimator of the second component of the stereological variance. In fact, they note in their paper that the estimation of the within section variance based on a two dimensional systematic

sample was viewed by Cruz-Orive (1999) as an open problem. To avoid the problem of a two dimension *covariogram* function applied on a irregularly shaped section, Cruz-Orive and Geiser (2004) made the following approximation. Let n_j be the number of counting frames in the j^{th} section and Q_{jk} be the observed number of neurons in the k^{th} counting frame of the j^{th} section, $k = 1, \ldots, n_j$, $j = 1, \ldots, m$, so that

$$Q_j = \sum_{k=1}^{n_j} Q_{jk} \tag{2.24}$$

is the total number of neurons observed in the j^{th} section. Let

$$K = \sum_{j=1}^{m} n_j \tag{2.25}$$

denote the total number of the counting frames in all sections.

They assume that given the first position of X_1, the numbers of neurons in each counting frame in the j^{th} section are $i.i.d$ from a Poisson distribution with parameter λ_j, i.e., for each j, that is $Q_{jk}|X_1, k = 1, \cdots, n_j \sim^{i.i.d} Poisson(\lambda_j)$. As the counting frames all have equal size, then $Q_j|X_1, j = 1, \cdots, m, \sim^{indep} Poisson(n_j\lambda_j)$. Their assumption requires that neurons follow a homogeneous process over the region of interest.

Since the Poisson distribution is independent of the location, this means Q_{jk} is independent of U_{1j} and V_{1j}.

Thus, within the j^{th} section,

$$var[\hat{f}(X_1 + (j-1)\frac{L}{m}, U_{1j}, V_{1j})|X_1] = \frac{1}{\tau_a^2 h^2} var[Q_j] = \frac{1}{\tau_a^2 h^2} \sum_{k=1}^{n_j} var[Q_{jk}] = \frac{1}{\tau_a^2 h^2} n_j \lambda_j, \tag{2.26}$$

so that the ML estimator of the variance is

$$\widehat{var}[\hat{f}(X_1 + (j-1)\frac{L}{m}, U_{1j}, V_{1j})|X_1] = \frac{1}{\tau_a^2 h^2} Q_j. \tag{2.27}$$

Hence, Cruz-Orive and Geiser (2004) obtained the following estimator of σ_{st2}

$$\begin{aligned}\widehat{\sigma}_{st2}^{C-O} &= \frac{L^2}{m^2}\sum_{j=1}^{m}\widehat{var}[\hat{f}(X_1+(j-1)\frac{L}{m},U_{1j},V_{1j})|X_1] \\ &= \frac{L^2}{m^2}\frac{1}{\tau_a^2 h^2}\sum_{j=1}^{m}Q_j \\ &= \frac{1}{\tau_s^2\tau_a^2\tau_h^2}Q. \end{aligned} \tag{2.28}$$

Combined with the between section variance estimator (2.19) or (2.21), Cruz-Orive and Geiser then provided an estimate of the stereological variance σ_{st}.

2.4 MOTIVATING DATA

In 2005, Dr. Konopaske performed a study on chronic exposure of macaque monkeys with two antipsychotic treatments to assess whether or not treatment with antipsychotic medication contributes to the disturbances in the number of a particular type of cells, called glial cells, previously observed in individuals with schizophrenia (Konopaske et al. (2007)). In Konopaske's study, there were 18 male macaque monkeys which had been divided into 3 experimental groups ($n = 6$ per group): haloperidol, olanzapine and sham, where the monkeys were matched across experimental groups as triples. Actually a more simplified cutting scheme uses large slabs cut initially to make the small width sections cuts feasible and at the same time produce sections which follow the model. For each monkey, the parietal lobe was cut in a systematic, uniformly random manner producing $12 \sim 15$ slabs ($m = 12 \sim 15$) with a mean width of $\frac{L}{m} = 2.5mm(= 2500\mu m)$. The thickness of a small width section is $80\mu m$. There are about 31 small width sections for each slab, which yield approximately 375 to 469 small width sections for the entire region when the sampled sections are viewed to be exhaustively cut from the region. The section sampling fraction is 0.0320 (= 80/2500). The area of the counting frames is always kept as $219.7\mu m^2$. The length of u-step (a_u) and v-step (a_v), $a_u = a_v$, was kept constant within each monkey, but may differ across monkeys, varying from 700 to 900 μm. Dr. Konopaske has kindly provided us access to the individual

counting frame data from this study, where the number of glial cells was counted in each counting frame.

Uneven shrinkage in section thickness can introduce biases when using the classical optical fractionators (West and Gundersen (1991)). However, such potential biases are eliminated by using the optical fractionators based on a mean section thickness that is number-weighted ($\bar{t}_Q$) (Dorph-Petersen et al. (2001)). Mean section thickness is number-weighted as follows: $\bar{t}_Q = (\sum_j \sum_k (t_{jk} Q_{jk}))/(\sum_j \sum_k Q_{jk})$ where t_{jk} is the local section thickness of the j^{th} section and k^{th} counting frame having a count of Q_{jk}. The height of the counting frame is fixed, $h = 8\mu m$. Total cell numbers were estimated as:

$$\hat{\mathcal{N}} = \frac{1}{\tau_s} \cdot \frac{1}{\tau_a} \cdot \frac{1}{\tau_h} \cdot Q, \tag{2.29}$$

where the section sampling fraction (τ_s) is $\frac{m}{M}$, the area sampling fraction (τ_a) is $\frac{a_f^2}{a_x a_y}$, the height sampling fraction (τ_h) is $\frac{h}{\bar{t}_Q}$, and Q is the number of neurons counted in all sampled counting frames.

We now apply the Cruz-Orive and Geiser (2004) method to estimate the stereological variance, using monkey #256 as an example. For monkey #256, 12 sections and 1626 counting frames in total were sampled and these data are provided in Table 2.1. The length of the u-step and v-step are 750 μm and the mean counting frame thickness is 32.5 μm. The fractions τ_s, τ_a and τ_h were 0.0320 (= 80/2500), 0.0004 (= 219.7/(750*750)) and 0.4065 (= 8/19.7), respectively. A total of 931 neurons were counted in all the sampled counting frames, so that the estimated number of the glial cells in the parietal lobe by (2.29) was

$$\hat{\mathcal{N}} = \frac{1}{\tau_s} \cdot \frac{1}{\tau_a} \cdot \frac{1}{\tau_h} \cdot Q = \frac{1}{0.0320} \cdot \frac{1}{0.0004} \cdot \frac{1}{0.4065} \cdot 931 = 183.2 \cdot 10^6.$$

Applying (2.21) and (2.28) to calculate the between and within section variance, respectively, we obtain the stereological variance estimator for monkey #256 as

$$\widehat{\sigma}_{st} = \frac{1}{240} \frac{1}{\tau_s^2 \tau_a^2 \tau_h^2} [(3 \sum_{j=1}^{m} Q_j^2 - Q) - 4 \sum_{j=1}^{m-1} Q_j Q_{j+1} + \sum_{j=1}^{m-2} Q_j Q_{j+2}] + (\frac{1}{\tau_s \tau_a \tau_h})^2 Q = 36.8 \cdot 10^{12}.$$

Table 2.1 gives the section data of monkey #256.

In this study, there were more than 30,000 counting frames collected over the 18 ani-

Table 2.1: *Details for the Stereological Variance Formula: Monkey #256 (Konopaske et al. (2007))*

Section	n_j	Q_j	Q_j^2	Q_jQ_{j+1}	Q_jQ_{j+2}
1	42	25	625		
2	77	46	2116	1150	
3	92	56	3136	2576	1400
4	89	38	1444	2128	1748
5	95	44	1936	1672	2464
6	141	70	4900	3080	2660
7	181	85	7225	5950	3740
8	245	150	22500	12750	10500
9	218	145	21025	21750	12325
10	212	126	15876	18270	18900
11	177	107	11449	13482	15515
12	57	39	1521	4173	4914
Total	1626	931	93753	86981	74166

mals. Of interest to us was to examine whether the necessary Poisson assumption is valid to use Cruz-Orive and Geiser's approximation for the stereological variance models holds. To do this, we examined whether or not the Poisson assumption holds within each section for all animals. As we show, our results indicate that the Poisson assumption seems not to be appropriate. We considered typical examples of three monkeys, one from each group: monkey #256 from olanzapine, #261 from sham and #263 from haloperidol group. In Table 2.2 for each monkey, the variables in the first three columns are: animal ID (ID), section number (Section), and number of sampled counting frames for each section (n_j). For each corresponding section, the fourth to fifth columns are the sample mean $\bar{Q}_j$ over counting frames within that section, and the sample variance s_j^2 over counting frames of the number of the glial cells for each corresponding section. The sixth column shows the ratio of the variance to the mean ($s_j^2/\bar{Q}_j$), which under the Poisson distribution assumption should be approximately equal to 1.0. The variances within each section, however, are about 10-20 percent larger than the means. This suggests that the data have more variability within each section than the Poisson distribution predicts. Additionally, Pearson χ^2 goodness-of-fit tests are used to check the Poisson assumption and the corresponding p-values obtained by using Pearson χ^2 goodness-of- fit for each section are listed in the last column (p-value) of Table 2.2. The goodness-of-fit test shows that in about 50% of the sections the p-values are less than 0.10. Thus, our analyses suggest that we would be underestimating the true variability by using the Cruz-Orive and Geiser method. This motivates us to seek an approach to better fit the data.

Table 2.2: *Validation of Poisson Assumption*

ID	Section	n_j	$\bar{Q_j}$	s_j^2	$s_j^2/\bar{Q}_j$	p-value
	1	42	0.595	0.686	1.15	0.783
	2	77	0.597	0.848	1.42	0.123
	3	92	0.609	0.724	1.19	0.160
	4	89	0.427	0.361	0.85	**0.069**
	5	95	0.463	0.421	0.91	0.384
256	6	141	0.497	0.780	1.23	**0.031**
	7	181	0.470	0.650	1.39	**0.010**
	8	245	0.612	0.755	1.23	**0.012**
	9	218	0.665	0.887	1.33	0.483
	10	212	0.594	0.754	1.27	**0.022**
	11	177	0.605	0.899	1.49	**0.011**
	12	57	0.684	0.719	1.05	0.832
	1	31	0.871	0.983	1.13	0.728
	2	48	0.583	0.546	0.94	0.650
	3	62	0.548	0.481	0.88	0.806
	4	72	0.583	0.725	1.24	0.664
	5	81	0.457	0.726	1.59	**0.015**
	6	103	0.621	0.904	1.46	**0.023**
	7	124	0.589	0.832	1.41	**0.029**
261	8	176	0.534	0.799	1.50	**0.014**
	9	230	0.504	0.522	1.03	**0.075**
	10	169	0.521	0.549	1.05	0.968
	11	239	0.557	0.718	1.29	**0.001**
	12	152	0.599	0.811	1.36	**0.061**
	13	88	0.659	0.664	1.01	0.415

Note, the p-values listed in the last column are obtained by using Pearson χ^2 goodness-of- fit.

ID	Section	n_j	$\bar{Q}_j$	s_j^2	$s_j^2/\bar{Q}_j$	p-value
	1	12	0.500	0.455	0.91	0.818
	2	50	0.360	0.398	1.11	0.711
	3	71	0.592	0.617	1.04	0.736
	4	86	0.372	0.425	1.14	0.355
	5	110	0.473	0.490	1.04	0.928
263	6	144	0.583	0.720	1.23	**0.028**
	7	156	0.615	0.754	1.23	**0.005**
	8	146	0.699	1.053	1.51	**0.092**
	9	185	0.578	0.713	1.23	0.133
	10	299	0.448	0.517	1.15	0.199
	11	195	0.564	0.660	1.17	**0.071**
	12	116	0.509	0.896	1.76	**0.006**

Note, the p-values listed in the last column are obtained by using Pearson χ^2 goodness-of- fit.

2.5 AMMETER PROCESS

Due to Konopaske's data not conforming to the assumption of a Poisson point process, we would like to find a stochastic point process that would more appropriately fit the data. Since the counting frame data within sections are collected by a systematical sampling scheme, ideally we would like to find a non-homogeneous Poisson process which is independent of the location to avoid the use of two-dimensional *covariogram*; however, no such process exists. Instead we suggest the use of a special Cox process which appears to fit the data better than the Poisson.

A Cox process with a piecewise constant intensity, sometimes called an *Ammeter* process, is a Poisson process where λ is generated from a random variable instead of being constant (see Grandell (1997)). Consider a one dimensional range of interest, and assume that there are K fixed known mutually exclusive sub-range intervals $L_1, L_2, \ldots, L_K$. Let $\lambda_1, \lambda_2, \ldots, \lambda_K$ be $i.i.d$ from a common distribution U. Then the process given $\lambda_1, \lambda_2, \ldots, \lambda_K$ is a Poisson

process for each sub-range where the process with intensity parameter λ_k over the interval L_k, that is,

$$\lambda(t) = \lambda_k \qquad for \quad t \in L_k, k = 1, \ldots, K. \tag{2.30}$$

Denote $E[\lambda_k]$ by μ_A and $var[\lambda_k]$ by σ_A^2. An instance of the intensity process in the *Ammeter* process is illustrated in Figure 2.5.

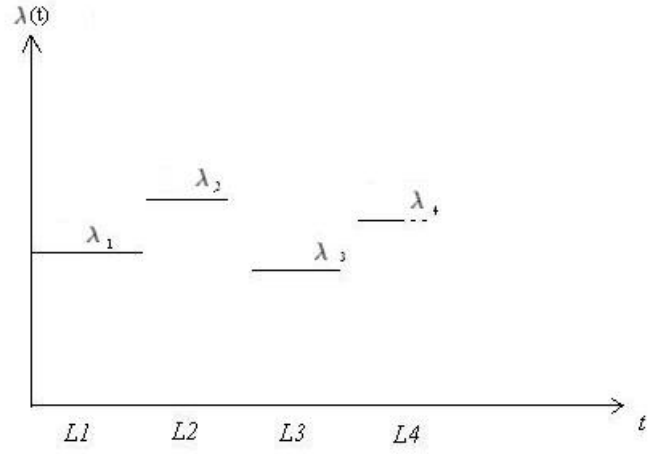


Figure 2.5: *Illustration of the Intensity in the Ammeter Process*

For application to the tissue section data, we need to extend the *Ammeter* process to a two dimensional version. For the range of interest, we assume that there are $I * K$ fixed known mutually exclusive square sub-ranges $L_{11}, L_{12}, \ldots, L_{IK}$ having common square length $\mathcal{L}$. Let $\lambda_{11}, \lambda_{12}, \ldots, \lambda_{IK}$ be $i.i.d$ from a common distribution U. Then the process given $\lambda_{11}, \lambda_{12}, \ldots, \lambda_{IK}$ is a two dimensional Poisson process for equal area size with intensity parameter λ_{ik} over the square L_{ik}, that is,

$$\lambda(t) = \lambda_{ik} \qquad for \quad t \in L_{ik}, \quad i = 1, \ldots, I \quad and \quad k = 1, \ldots, K. \tag{2.31}$$

Denote $E[\lambda_{ik}]$ by μ_A and $var[\lambda_{ik}]$ by σ_A^2.

If the grid length used under the microscope is approximately the same the common square length $\mathcal{L}$ in the *Ammeter* process, then each systematic sample falls within the square $[i\mathcal{L}, (i+1)\mathcal{L}] \times [k\mathcal{L}, (k+1)\mathcal{L}]$. Moreover, this assumption is still valid, if $\mathcal{L}$ is ρ times the grid length where ρ is an arbitrary positive integer. With these assumptions we can recalculate

the within section variances based on the *Ammeter* process assumption. Without these assumptions, there is the possibility that the counting frame would fall across two or more sub-ranges, vastly complicating the distribution theory. However, the ratio of the area of the counting frames to the area of the rectangle is very small, which is about 1/2000, so there is very little possibility that this happens. The estimate $\hat{\mathcal{N}}$ of the number of neurons would be given by the same formula as (2.1), and the estimate is an unbiased estimator of $\mathcal{N}$.

To estimate the stereological variance, we calculate the between section variation by (2.19) or (2.21), only change the calculation of the within section variation.

If we assume that P, the number of neurons counted in a counting frame, follows the *Ammeter* process and the distribution of the intensity U is a Gamma distribution, then P would follow a negative binomial (NB) distribution. The distribution of the resulting P can be expressed in terms of a mean μ and a dispersion parameter ϕ, so that the probability of observing q neurons is

$$Pr(P = q) = \frac{\Gamma(\phi + q)}{q!\Gamma(\phi)}(\frac{\mu}{\mu + \phi})^q(1 + \frac{\mu}{\phi})^{-\phi}. \tag{2.32}$$

The variance of this NB distribution is $\mu(1 + \frac{\mu}{\phi})$. When $\phi \to \infty$, the NB distribution converges to a Poisson distribution. The maximum likelihood (ML) estimator $\hat{\mu}$ of μ is the sample mean, and the ML estimator $\hat{\phi}$ of ϕ is determined by numerical maximization of the profile log-likelihood function $L(\hat{\mu}, \phi)$.

Under the *Ammeter* process assumption, if we assume that Q_{jik}, the observed values within the counting frame, in the square L_{ik} of the j^{th} section follow a NB distribution with parameters (μ_{Aj}, ϕ_{Aj}), then we have

$$\begin{aligned} var[\hat{f}(X_1 + (j-1)\frac{L}{m}, U_{1j}, V_{1j})|X_1] &= var(\sum_{i,k} \frac{1}{\tau_a h} Q_{jik}) \\ &= \frac{1}{\tau_a^2 h^2} n_j \mu_{Aj}(1 + \frac{\mu_{Aj}}{\phi_{Aj}}). \end{aligned} \tag{2.33}$$

The MLE of μ_{Aj} is $\sum_{i,k} Q_{jik}/n_j = Q_j/n_j$. Combining the MLE of $var[\hat{f}(X_1 + (j-1)\frac{L}{m}, U_{1j}, V_{1j})|X_1]$ in (2.33) with the between section variance estimator (2.19) or (2.21), the

stereological variance in (2.11) turns out to be

$$\begin{aligned}\widehat{\sigma}_{st} &= \widehat{\sigma}_{st1} + \frac{L^2}{m^2} var[\hat{f}(X_1 + (j-1)\frac{L}{m}, U_{1j}, V_{1j})|X_1] \\ &= \widehat{\sigma}_{st1} + \frac{L^2}{m^2}\frac{1}{\tau_a^2 h^2}\sum_{j=1}^{m} n_j \hat{\mu}_{Aj}(1 + \frac{\hat{\mu}_{Aj}}{\hat{\phi}_{Aj}}) \\ &= \widehat{\sigma}_{st1} + \frac{1}{\tau_s^2\tau_a^2\tau_h^2}\sum_{j=1}^{m} Q_j(1 + \frac{Q_j}{n_j\hat{\phi}_{Aj}}). \end{aligned} \tag{2.34}$$

The *Ammeter* based estimators of the variance are included in Table 2.3. In Section 2.7, we compare these variances with the stereological variances obtained under several different procedures.

2.6 BOOTSTRAP METHOD TO ESTIMATE THE STEREOLOGICAL VARIANCE

The bootstrap method can be applied to estimate the stereological variance. Without any distributional assumption, using section as a stratifying variable, independent bootstrap samples can be selected within the sections based on sampling the counting frame counts with replacement. It is clear that for homogenous processes and "partially" homogenous process like *Ammeter* for neuron development that the bootstrap variance will tend to overestimate the true variance derived based on the process.

For example, if we have 12 sections, with n_j counting frames in the j^{th} section, let $X = (x_{1,1}, x_{1,2}, \ldots, x_{1,n_1}, \ldots, x_{12,1}, \ldots, x_{12,n_{12}})$ denote the counting frame level data, so that x_{jk} is the number of neurons counted in the j^{th} section of the k^{th} counting frame. A bootstrap sample, $X^* = (x^*_{1,1}, x^*_{1,2}, \ldots, x^*_{1,n_1}, \ldots, x^*_{12,1}, \ldots, x^*_{12,n_{12}})$, is chosen where $x^*_{j,1}, \ldots, x^*_{j,n_j}$ are selected independently with replacement from $\{x_{j,1}, \ldots, x_{j,n_j}\}$, $j = 1, 2, \ldots, 12$. From the bootstrap sample, we can accordingly estimate the total number $\mathcal{N}^*$. We then simulate B independent bootstrap samples $X^{*1}, X^{*2} \ldots, X^{*B}$ and obtain stereological estimates $\hat{\mathcal{N}}^{*1}, \hat{\mathcal{N}}^{*2} \ldots, \hat{\mathcal{N}}^{*B}$, respectively. Then the within section variance of the stereological esti-

mate of $\hat{\mathcal{N}}$ ($\widehat{\sigma}^{BS}_{st2}$) can be estimated by

$$\widehat{\sigma}^{BS}_{st2} = var[\hat{\mathcal{N}}^*], \tag{2.35}$$

where $var[\hat{\mathcal{N}}^*]$ is the sample variance of B stereological estimates. Combined with the between section variance estimator (2.19) or (2.21), the estimated stereological variance in (2.11) turns out to be

$$\widehat{\sigma}^{BS}_{st} = \widehat{\sigma}_{st1} + var[\hat{\mathcal{N}}^*]. \tag{2.36}$$

2.7 COMPARING THE STEREOLOGICAL VARIANCE ESTIMATES

Using the Konopaske data, we compare the stereological variance estimates calculated by the three different methods. Table 2.3 lists the results for each animal. The column $\widehat{\sigma}_{st1}$ is the first component of the stereological variance estimator based on (2.21). The columns $\widehat{\sigma}^{C-O}_{st2}$, $\widehat{\sigma}^{AM}_{st2}$ and $\widehat{\sigma}^{BS}_{st2}$ give the second component of the stereological variance estimator based on the Cruz-Orive and Geiser assumption, *Ammeter* assumption and Bootstrap approach, respectively. The column $\widehat{\sigma}^{C-O}_{st}$ is the sum of $\widehat{\sigma}_{st1}$ and $\widehat{\sigma}^{C-O}_{st2}$, which is the stereological variance estimator by the Cruz-Orive and Geiser method. The next column $\widehat{\sigma}^{AM}_{st}$ is the sum of $\widehat{\sigma}_{st1}$ and $\widehat{\sigma}^{AM}_{st2}$, which is the stereological variance estimator by the *Ammeter* method. The last column $\widehat{\sigma}^{BS}_{st}$ is the sum of $\widehat{\sigma}_{st1}$ and $\widehat{\sigma}^{BS}_{st2}$, which is the stereological variance estimator by the Bootstrap method. Note that when comparing between sections variation with within sections variation, the between section variance estimator $\widehat{\sigma}_{st1}$ shows much less section-to-section variability than within section variability as shown by the estimators $\widehat{\sigma}^{C-O}_{st2}$, $\widehat{\sigma}^{AM}_{st2}$ or $\widehat{\sigma}^{BS}_{st2}$. The literature indicates that treating random systematic samples as simple random samples will overestimate the variability. Here, the Bootstrap method treats the counting frames as simple random samples within sections, and the stereological variance seems to be overestimated by the Bootstrap method perhaps because of the positive dependence between counting frames within sections. On the other hand, as our analysis summarized in Table 2.2 indicates, we believe that for the Konopaske data, the stereological variance is underestimated by the Cruz-Orive and Geiser method. Note that the stereological variance

estimators given by the *Ammeter* method are between these other two methods in most of the cases. In Table 2.3, we also provide the stereological estimate, the total counted number of neurons and the coefficient of error (under Cruz-Orive and Geiser assumption) for each animal (Columns $\hat{\mathcal{N}}$, Q and CE, respectively).

Table 2.3: *Comparison of Estimates of the Stereological Variance (Konopaske et al. (2007))*

ID	Triad	Group	$\hat{\mathcal{N}}$	Q	CE	$\widehat{\sigma}_{st1}$	$\widehat{\sigma}_{st2}^{C-O}$	$\widehat{\sigma}_{st2}^{AM}$	$\widehat{\sigma}_{st2}^{BS}$	$\widehat{\sigma}_{st}^{C-O}$	$\widehat{\sigma}_{st}^{AM}$	$\widehat{\sigma}_{st}^{BS}$
260	1	H	141.9	862	0.036	2.4	23.3	27.2	27.6	25.7	29.6	29.9
273	2	H	121.1	525	0.044	0.1	28.0	32.2	33.0	28.0	32.3	33.1
257	3	H	132.1	661	0.040	1.3	26.4	30.7	35.1	27.7	32.0	36.4
266	4	H	170.6	805	0.037	3.9	36.2	42.8	43.8	40.1	46.7	47.7
270	5	H	125.1	607	0.041	0.6	25.8	30.7	32.0	26.4	31.3	32.6
263	6	H	169.9	842	0.035	0.4	34.3	42.9	43.0	34.7	43.3	43.5
259	1	O	165.4	845	0.035	1.5	32.4	37.1	38.5	33.9	38.7	40.0
274	2	O	159.3	734	0.037	0.6	34.6	39.9	38.7	35.2	40.5	39.3
256	3	O	183.2	931	0.033	0.8	36.1	45.9	50.7	36.8	46.7	51.5
267	4	O	138.8	769	0.037	0.8	25.1	29.5	28.7	25.8	30.2	29.5
271	5	O	140.6	771	0.040	5.9	25.6	29.9	31.6	31.5	35.8	37.5
265	6	O	155.5	766	0.036	0.5	31.6	39.5	37.3	32.1	40.0	37.8
261	1	S	192.0	885	0.034	1.7	41.7	52.1	61.0	43.3	53.8	62.7
272	2	S	153.7	737	0.037	0.5	32.1	36.5	37.6	32.6	37.1	38.1
258	3	S	214.9	916	0.034	2.0	50.4	60.3	66.6	52.4	62.2	68.5
268	4	S	122.3	630	0.040	0.2	23.8	27.1	27.8	24.0	27.3	28.0
269	5	S	196.1	772	0.037	3.1	49.8	58.5	61.4	52.9	61.6	64.5
264	6	S	172.5	750	0.037	1.4	39.7	49.1	49.9	41.1	50.5	51.3

Note: H, O and S denote for the haloperidol, olanzapine and sham group, respectively. The units for $\hat{\mathcal{N}}$ *is* 10^6 *and the units for* $\widehat{\sigma}_{st1}$, $\widehat{\sigma}_{st2}^{C-O}$, $\widehat{\sigma}_{st2}^{AM}$,$\widehat{\sigma}_{st2}^{BS}$, $\widehat{\sigma}_{st}^{C-O}$, $\widehat{\sigma}_{st}^{AM}$, $\widehat{\sigma}_{st}^{BS}$ *are* 10^{12} *in the Table 2.3.*

Table 2.4 lists the summary statistics for the Konopaske data by groups. The biological variance is estimated by subtracting the stereological variance from the average of group variances. The stereological variance (by the Cruz-Orive and Geiser method) is estimated by averaging columns $\widehat{\sigma}_{st2}^{C-O}$ in Table 2.3 as 34.7. The estimated variance of the primary experimental outcome, $\hat{\mathcal{N}}$, is based on the pooled variance, 620.7, (average of the three within group variances). Hence, the average biological variance is 586.0 (= 620.7 - 34.7). We see that the stereological variance is small relative to the biological variance. In implementing experimental designs like this, researchers usually choose the number of counting frames for each subject by controlling the coefficient of error (CE) of the stereological estimate for each subject, in an attempt to assess each individual as precisely as possible. Usually, they require the CE to be less than 5%. To meet this criterion for the monkey study, Dr. Konopaske and his colleagues had to collect about 1600 counting frames for each animal. If the number of counting frames could have been reduced, the research studies could have saved several months' laboratory work. This motivates us to consider in the next two sections more efficient experiment designs which still maintain enough power to detect the treatment effects of interest. For the biological experiments which test between groups differences, usually the between-subject variation is larger than within subject variation mainly due to stereological procedures. Thus a reasonable allocation of sampling effort is to sample a large number of subjects and spend relatively little effort on measuring the data in each subject. Gundersen and Osterby (1981) were aware of this reality and described it as "Do More, Less Well".

Table 2.4: *Summary of Konopaske's Data.*

Group	Number	Mean	Variance
Haloperidol	6	143.46	480.05
Olanzapine	6	157.14	273.24
Sham	6	175.30	1108.89

3.0 OPTIMAL DESIGN

For obvious reasons of budgetary and time constraints, a goal of research in designing post-mortem tissue stereological studies is to keep study cost within budget and time constraints while maintaining sufficient statistical power to address the research aims. While the Konopaske study used tissue from a previous study, one could imagine designing a new monkey study for the sole purpose of a study like Konopaske. For such a study, we note that macaque monkeys are expensive study subjects, and that numerous labor hours could be involved in creating and observing the counting frames from each animal. One could reduce the cost of studies like this by decreasing either the number of subjects or the number of counting frames for each subject. Increasing the number of counting frames yields more precise estimates of the neuron number for each animal, and hence increases study powers. One trade-off would be to reduce the number of subjects, and then add counting frames for each subject in order to maintain the study power. Alternatively, one could increase the numbers of subjects and reduce the numbers of counting frames per subject to maintain power. There appears to be little written in the stereology literature about how to select the combination of sample size and numbers of counting frames for stereological study designs. The purpose of this chapter is to create a framework for finding the combination of sample size and number of counting frames that will not only maintain sufficient power to address research aim but also minimize the cost of the study.

There is established methodology for general repeated measure designs which deal with optimal trade-offs between sampling units and replicates for each sample. (For instance, see Liu and Liang (1992), Mentre et al. (1997), Vickers (2003)). Most of this work focuses on the covariance structure of the repeated measures. However, this literature is not directly applicable to the stereological problem due to several specific difficulties that become ap-

parent as we describe our research. In stereological studies dealing with counting, the first component of the stereological variance is always unchanged, so that one thing we can do to maintain power is to change the number of counting frames in order to adjust the within section variance. Also, we need to take into account the possibility that the stereological variances are different between groups.

3.1 MATHEMATICAL APPROACH

3.1.1 Cost Function and Power Function

We begin by using the idea of some standard designs employed in the CCNMD, where we want to compare two groups by using a simple linear model for testing, which for power considerations can be well approximated by a paired t test. In order to control for covariates, in human studies for example, each experimental subject and corresponding control subject are matched by covariates, such as age at death, gender, post-mortem interval and brain pH value. Pairing also helps to reduce variability due to tissue processing designs, a topic we don't discuss further here. A consideration of experimental design takes into account the stereological variance. Using the results described in Chapter 2 concerning the stereological variance, we discuss in Section 3.1.2 the amount of increase in the stereological variance caused by decreasing the number of counting frames. Then taking into account the subject-to-subject biological variability, we introduce in this section a linear cost function and also a power function, both of which depend on the number of subjects and the number of counting frames.

We consider a paired analysis to compare the difference of the number of neurons between two groups, control and treatment (or experimental). Let N be the number of pairs used in the study and K be the number of counting frames collected for each subject, assuming each subject has the same number of counting frames. We define a cost function as follows:

$$Cost(N, K) = C_0 + 2C_1N + 2C_2NK. \tag{3.1}$$

In the cost function, C_0 is the setup cost for a study. C_1 is the cost for each subject, i.e., the cost of tissue preparation in a post-mortem tissue study; in the animal studies, also including animal cost, treatment cost and nursing cost. We assume the cost is the same for each subject in the treatment group and control group. C_2 is the cost for each counting frame which includes constructing the counting frame, delineating the appropriate neurons and counting them. Since C_0, C_1 and C_2 are positive constants, it is sufficient to know the relative cost C_1/C_2 in order to minimize the study cost. Invariably the subject cost C_1 is relatively large compared to the counting frame cost C_2 .

We now provide a model for analyzing a paired study using paired differences. Let X_{ik} be the true number of neurons of subject k in group i, $i = c$ (control group) or $i = t$ (treatment group), $k = 1, \cdots, N$, then the typical model in such a study assumes

$$X_{ik} \sim^{i.i.d} N(\mu_i + p_k, \sigma_B), \tag{3.2}$$

where μ_i is the mean of group i, p_k is the effect of pair k (with $\sum_{k=1}^{N} p_k = 0$) and σ_B is the biological variance. We assume that the biological variance is the same for treatment and control. Obviously, we cannot observe X_{ik} for any subject, but must use stereological methods to "estimate"this quantity.

Let $\hat{\mathcal{N}}_{ik}$ be the stereological estimate of X_{ik}. For simplicity, we assume the shape of the brain region of interest is the same among subjects. In this chapter, we continue to use σ_{st} as the notation for the stereological variance, σ_{st1} and σ_{st2} as the notations of the two parts of the stereological variance in (2.11), respectively. The within section variance may be based on either Cruz-Orive and Geiser or *Ammeter* assumptions. Given X_{ik}, we assume that $\hat{\mathcal{N}}_{ik}$ is approximately normally distributed, and

$$\begin{aligned} &E[\hat{\mathcal{N}}_{ik}|X_{ik}] = X_{ik}, \\ &var[\hat{\mathcal{N}}_{ik}|X_{ik}] = \sigma^i_{st}, \quad i = c \text{ or } t, \end{aligned} \tag{3.3}$$

where σ^c_{st} and σ^t_{st} are the stereological variances of control and treatment groups, respectively. From the stereological variance formula in Chapter 2, we know that if there is a group difference the stereological variances are most likely different between groups.

The unconditional distribution of $\hat{\mathcal{N}}_{ik}$ is as follows

$$E[\hat{\mathcal{N}}_{ik}] = E[E[\hat{\mathcal{N}}_{ik}|X_{ik}]] = E[X_{ik}] = \mu_i + p_k,$$
$$var[\hat{\mathcal{N}}_{ik}] = var[E[\hat{\mathcal{N}}_{ik}|X_{ik}]] + E[var[\hat{\mathcal{N}}_{ik}|X_{ik}]] = var[X_{ik}] + E[\sigma^i_{st}] = \sigma_B + \sigma^i_{st}. \tag{3.4}$$

Hence,

$$\hat{\mathcal{N}}_{ik} \sim^{i.i.d} N(\mu_i + p_k, \sigma_B + \sigma^i_{st}),\ i = c \text{ or } t,\ k = 1, \cdots, N, \tag{3.5}$$

so that the pairwise difference $D_k = \hat{\mathcal{N}}_{ck} - \hat{\mathcal{N}}_{tk} \sim^{i.i.d} N(\mu_c - \mu_t, 2\sigma_B + \sigma^c_{st} + \sigma^t_{st}), k = 1, \cdots, N$. The sample variance of D_k given by $s^2_D = \frac{1}{N-1}\sum_{k=1}^{N}(D_k - \bar{D})^2$ has the following distribution: $(N-1)s^2_D \sim (2\sigma_B + \sigma^c_{st} + \sigma^t_{st})\chi^2_{N-1}$.

The hypotheses are

$$H_0 : \mu_c = \mu_t,$$
$$H_a : \mu_c \neq \mu_t. \tag{3.6}$$

It is well known that to test H_0 based on(3.5), we use the usual paired t-test statistic

$$T = \frac{\bar{\hat{\mathcal{N}}}_{c\cdot} - \bar{\hat{\mathcal{N}}}_{t\cdot}}{\sqrt{\frac{s^2_D}{N}}}. \tag{3.7}$$

Hence, the power function for the 0.05 level test of (3.6) is given by

$$\begin{aligned} Power &= Pr(\text{reject } H_0 |\ H_a) \\ &= Pr\left(|\ T\ | > t_{0.975,N-1} |\ \xi = \frac{\delta}{\sqrt{\frac{2}{N}(\sigma_B + \frac{\sigma^c_{st} + \sigma^t_{st}}{2})}} \right), \end{aligned} \tag{3.8}$$

where the test statistic T follows a noncentral t distribution with $N-1$ degrees of freedom, and noncentrality parameter ξ, and where $t_{0.975,N-1}$ is the 0.025 critical value of the t distribution with $N-1$ degrees of freedom, and δ is the pre-specified alternative.

Under the null hypothesis $\mu_c = \mu_t$, so that the stereological variances are the same for control and treatment group, that is, $\sigma^c_{st} = \sigma^t_{st}$. But under the alternative hypothesis $\mu_c \neq \mu_t$, the stereological variances are different for control and treatment group according to

the stereological variance formula (2.11). P-values are obtained based on the null hypothesis being true; thus, the paired t test gives the correct p-values. Statistical power is based on the alternative hypothesis, so the test statistic needs to account for the unequal variances. However, for the paired study, no matter whether the variances are equal or not between groups, the paired t statistic is appropriate to use, since the difference statistic depends on the sum of stereological variance of two groups.

3.1.2 Stereological Variance and Number of Counting Frames

In this section, we present the mathematical relationship between number of counting frames and stereological variance. Consider two stereological designs on the same region of interest with length L, where one has K and the other has K^* counting frames, respectively. We suppose the number of the sections respectively m and m^* to be the same between the two designs. The sizes of the square counting frames are the same among the sections and between designs, so that $a_f = a_f^*$. The height is also the same within and between designs, so that $h = h^*$. The difference between these two stereological designs is that they use different lengths of grids, where $a_u \neq a_u^*$ and $a_v \neq a_v^*$. The grid lengths are the same among sections in the same design. This difference in grid lengths allows us to vary the number of sampling frames.

Thus, the section sampling fraction and the height sampling fraction are the same for these two designs:

$$\begin{aligned} \tau_s &= \tau_s^*, \\ \tau_h &= \tau_h^*. \end{aligned} \tag{3.9}$$

Let the ratio of two numbers of counting frames be p, i.e., $p = \frac{K}{K^*}$, and $n_j = pn_j^*$ for $j = 1, \cdots, m$. Since the area of the j^{th} section, $Area_j = n_j a_u a_v$, is fixed, then

$$\begin{aligned} a_u a_v &= Area_j / n_j \\ &= Area_j / (p \cdot n_j^*) \\ &= \frac{1}{p} \cdot a_u^* a_v^*. \end{aligned} \tag{3.10}$$

Hence,

$$\tau_a = a_f^2/(a_u a_v) = a_f^2/(\frac{1}{p} \cdot a_u^* a_v^*) = p \cdot \tau_a^*. \tag{3.11}$$

Recall that in the stereological variance formula in (2.11) ($\sigma_{st} = var[\widetilde{\mathcal{N}}] + (\frac{L}{m})^2 \sum_{j=1}^{m} var[\hat{f}(X_1 + (j-1)\frac{L}{m}, U_{1j}, V_{1j})|X_1]$), the first term on the right hand side is the between section variation ($\sigma_{st1} = var[\widetilde{\mathcal{N}}]$). Hence,

$$\begin{aligned} \widetilde{\mathcal{N}} &= \frac{L}{m} \sum_{j=1}^{m} f(X_j) \\ &= \frac{L}{m^*} \sum_{j=1}^{m^*} f(X_j) \\ &= \widetilde{\mathcal{N}}^*, \end{aligned} \tag{3.12}$$

where $f(X_j)$ is the true number of neurons at X_j for $j = 1, \cdots, m$, so that the between section variances are the same for two designs when the two designs have the same number of sections, that is, $\sigma_{st1} = var[\widetilde{\mathcal{N}}] = var[\widetilde{\mathcal{N}}^*] = \sigma_{st1}^*$.

The between section variance only depends on region length L, section number m and the $f(X_j)$'s. It is important to understand that the values of L, m and $f(X_j)$'s are unchanged under these two stereological designs, and thus the between section variance component will be the same no matter how many counting frames are sampled.

The second component of the stereological variance in (2.11) is $\sigma_{st2} = (\frac{L}{m})^2 \sum_{j=1}^{m} var[\hat{f}(X_1 + j\frac{L}{m}, U_{1j}, V_{1j})|X_1]$.

Under the Cruz-Orive and Geiser assumption, we have

$$\begin{aligned} \sigma_{st2} &= (\frac{L}{m})^2 \sum_{j=1}^{m} \frac{1}{\tau_a^2 h^2} n_j \lambda_j \\ &= (\frac{L}{m^*})^2 \sum_{j=1}^{m^*} \frac{1}{(p\tau_a^*)^2 h^{*2}} p n_j^* \lambda_j \\ &= \frac{1}{p} (\frac{L}{m^*})^2 \sum_{j=1}^{m^*} \frac{1}{\tau_a^{*2} h^{*2}} n_j^* \lambda_j \\ &= \frac{1}{p} \sigma_{st2}^*. \end{aligned} \tag{3.13}$$

Similarly, assuming we were to use our *Ammeter* model, we would have

$$\begin{aligned}
\sigma_{st2} &= (\frac{L}{m})^2 \sum_{j=1}^{m} \frac{1}{\tau_a^2 h^2} n_j \mu_{Aj} (1 + \frac{\mu_{Aj}}{\phi_{Aj}}) \\
&= (\frac{L}{m^*})^2 \sum_{j=1}^{m^*} \frac{1}{(p\tau_a^*)^2 h^{*2}} p n_j^* \mu_{Aj} (1 + \frac{\mu_{Aj}}{\phi_{Aj}}) \\
&= \frac{1}{p} (\frac{L}{m^*})^2 \sum_{j=1}^{m^*} \frac{1}{\tau_a^{*^2} h^{*2}} n_j^* \mu_{Aj} (1 + \frac{\mu_{Aj}}{\phi_{Aj}}) \\
&= \frac{1}{p} \sigma_{st2}^*. \qquad (3.14)
\end{aligned}$$

Hence, for both the approaches the second component of the stereological variance (σ_{st2}) is inversely proportional to the number of counting frames.

In order to obtain our cost considerations, we make a number of simplifying assumptions. We assume that the shape and the length of brain region of interest are the same for all the subjects. In the stereological study, we also assume the same number of sections will be selected for all subjects. In practice, the size and height of the counting frames will be kept the same through the study. Then the between section variation (σ_{st1}) is unchanged according to the number of counting frames, while the remaining part of the stereological variance (σ_{st2}) and the number of counting frames have an inversely proportional relationship. The two parts of the stereological variance are assumed to be the same within each group, but not necessary to be the same between groups.

3.1.3 Theoretical Results

In the last section, we obtained the relationship between the number of counting frames and the stereological variance. The first term on the right hand side of (2.11) is unchanged under the systematic sampling design when the number of sections is the same. The second term on the right hand side of (2.11) and the number of counting frames are inversely proportionally related. In this section, we use the notation σ^i_{st1,K_0} to be the first term, and σ^i_{st2,K_0} to be the second term where the number of counting frames is K_0 for each subject in both groups the control and experimental.

When the number of pairs and the number of counting frames are N and K respectively,

the noncentrality parameter in the power function (3.7) is given by

$$\begin{aligned}\xi &= \frac{\delta}{\sqrt{\frac{1}{N}\left(2\sigma_B + (\sigma^c_{st,K} + \sigma^t_{st,K})\right)}} \\ &= \frac{\delta}{\sqrt{\frac{1}{N}\left(2\sigma_B + (\sigma^c_{st1,K_0} + \frac{K_0\sigma^c_{st2,K_0}}{K}) + (\sigma^t_{st1,K_0} + \frac{K_0\sigma^t_{st2,K_0}}{K})\right)}} \\ &= \frac{\delta}{\sqrt{\frac{1}{N}\left(2\sigma_B + (\sigma^c_{st1,K_0} + \sigma^t_{st1,K_0}) + \frac{K_0(\sigma^c_{st2,K_0}+\sigma^t_{st2,K_0})}{K}\right)}}. \end{aligned} \tag{3.15}$$

Hence the power function in (3.8), viewed as a function of N and K, is

$$\begin{aligned} Power(N,K) &= P(\text{reject } H_0 |\ H_a) \\ &= P\left(|T| > t_{.975,N-1} |\ \xi = \frac{\delta}{\sqrt{\frac{2}{N}\left(\sigma_B + \frac{\sigma^c_{st1,K_0}+\sigma^t_{st1,K_0}}{2} + \frac{K_0(\sigma^c_{st2,K_0}+\sigma^t_{st2,K_0})}{2K}\right)}} \right), \end{aligned} \tag{3.16}$$

where δ is the pre-specified alternative, and σ_B, σ^c_{st1,K_0}, σ^c_{st2,K_0}, σ^t_{st1,K_0}, and σ^t_{st2,K_0} are known parameters.

Thus the power function and cost function both depend on the number of pairs N and the number of counting frames K used in the study. Combining these two functions together, we have

$$\begin{cases} Power(N,K) = P\left(|T| > t_{.975,N-1} |\ \xi = \frac{\delta}{\sqrt{\frac{2}{N}\left(\sigma_B + \frac{\sigma^c_{st1,K_0}+\sigma^t_{st1,K_0}}{2} + \frac{K_0(\sigma^c_{st2,K_0}+\sigma^t_{st2,K_0})}{2K}\right)}} \right) \\ Cost(N,K) = C_0 + 2C_1N + 2C_2NK. \end{cases}$$

In the next section we consider two ways to optimize cost, two types of optimizations:

I. for a certain power, to minimize the cost of study;

II. for a fixed budget, to maximize the power of study.

3.1.3.1 The Type I Optimization For the type I optimization, let Ω_β be the set of all possible combinations of numbers of pairs and numbers of counting frames satisfying power

function $\geq 1-\beta$, at the alternative δ, that is

$$\Omega_\beta = \{(N,K) : Power(N,K) \geq 1-\beta\}. \tag{3.17}$$

Denote the type I optimal combination as $(N,K)_{opt,I}$ which is the combination minimizing the cost function $Cost(N,K)$ in Ω_β.

Given δ and β, let $\sigma_{\delta,\beta}(N)$ denote the variance of the pairwise difference in order for a study of N pairs subjects to reach at least $1-\beta$ power. $\sigma_{\delta,\beta}$(N) can be obtained from noncentral t power calculates for a given number of pairs N. When the number of counting frames for each subject is K, $\sigma_{\delta,\beta}(N)$ in the power function (3.16) can be given in terms of σ_B, K_0, σ^c_{st1,K_0}, σ^t_{st1,K_0}, σ^c_{st2,K_0}, and σ^t_{st2,K_0}, that is

$$\sigma_{\delta,\beta}(N) = 2\sigma_B + (\sigma^c_{st1,K_0} + \sigma^t_{st1,K_0}) + \frac{K_0(\sigma^c_{st2,K_0} + \sigma^t_{st2,K_0})}{K}. \tag{3.18}$$

Note that $\sigma_{\delta,\beta}(N)$ is a function of N, so that the solution of K is obtained as a function of N

$$K(N) = \left[\frac{K_0(\sigma^c_{st2,K_0} + \sigma^t_{st2,K_0})}{\sigma_{\delta,\beta}(N) - 2\sigma_B - (\sigma^c_{st1,K_0} + \sigma^t_{st1,K_0})}\right], \tag{3.19}$$

where $[a]$ is the smallest integer larger than a and $K > 0$. For a given N, K is adjusted to change the stereological variance, and hence the variance of the primary outcome ($\hat{\mathcal{N}}$) in order to achieve the desired power. Thus, among all the possible combinations of (N,K) that satisfy the statistical power, the optimal combination of (N,K) is given by

$$\begin{aligned}
(N,K)_{opt,I} &= argmin_{\Omega_\beta} Cost(N,K(N)) \\
&= argmin_{\Omega_\beta} \{C_0 + C_1 N + C_2 N K(N)\} \\
&= argmin_{\Omega_\beta} \{C_2(C_1/C_2 N + N K(N))\} \\
&= argmin_{\Omega_\beta} \{C_1/C_2 N + N K(N)\},
\end{aligned} \tag{3.20}$$

where C_0, C_1 and C_2 are positive constants in the cost function. Note that to find the optimal combination, it is sufficient to know C_1/C_2 which is the relative cost of a subject to the cost of a counting frame.

To illustrate the Type I optimal results one can obtain from the optimization approach,

Table 3.1: *Type I Optimal Combination of Number of Pairs and Number of Counting Frames* $(N, K)_{opt,I}$.

C_1/C_2	$1-\beta = 0.70$	$1-\beta = 0.75$	$1-\beta = 0.80$	$1-\beta = 0.85$	$1-\beta = 0.90$
10	(26, 52)	(28, 54)	(34, 49)	(38, 50)	(42, 53)
20	(22, 64)	(24, 66)	(26, 69)	(29, 70)	(34, 69)
50	(15, 110)	(17, 106)	(19, 106)	(21, 109)	(23, 118)
100	(13, 138)	(14, 144)	(15, 153)	(17, 151)	(19, 158)
500	(9, 286)	(9, 395)	(10, 350)	(11, 361)	(13, 329)
1000	(8, 391)	(8, 514)	(9, 472)	(10, 471)	(11, 516)
10000	(6,1506)	(7, 906)	(7, 1591)	(8,1209)	(9, 1203)

we consider an example where the results are given in Table 3.1. In this example, the parameters used for the optimal calculation are $\delta = 20$, $\sigma_B = 100$, $K_0 = 1000$, $\sigma^c_{st1,K_0} = 1$, $\sigma^t_{st1,K_0} = 2$, $\sigma^c_{st2,K_0} = 30$ and $\sigma^t_{st2,K_0} = 40$ in Table 3.1. The range of power is $(0.70, 0.75, 0.80, 0.85, 0.90)$ and a relatively large range of relative costs C_1/C_2 $(= 10, 20, 50, 100, 500, 1000, 10000)$ are considered. The optimal results in Table 3.1 show that at a fixed level of power as the relative cost C_1/C_2 increases, then the optimal number of pairs decreases and more counting frames are added. When the relative cost is small, the number of pairs is what mostly changes in order to achieve different levels of powers. When the relative cost is large, the number of counting frames is mostly adjusted for different powers. Clearly, for any relative cost and required statistical power, we can provide the optimal design for a stereological study.

Since the number of counting frames K is a function of the number of pairs N, the core part of the cost function $(C_1/C_2 N + NK(N))$ can also be considered as a function of N. The core part of the cost function versus the number of pairs for three different β's are illustrated in Figure 3.1, where the relative cost C_1/C_2 is 100. The diamond, circle and triangle denote power levels of 0.9, 0.8 and 0.7 respectively. The three dots mark the optimal combination for each power level. Figure 3.1 shows when the number of pairs is small, the cost of study is

large, because a huge amount of counting frames is required for each subject. On the other hand, too many pairs also increase the cost of study. The optimal cost and optimal number of pairs both increase as the desired power goes large.

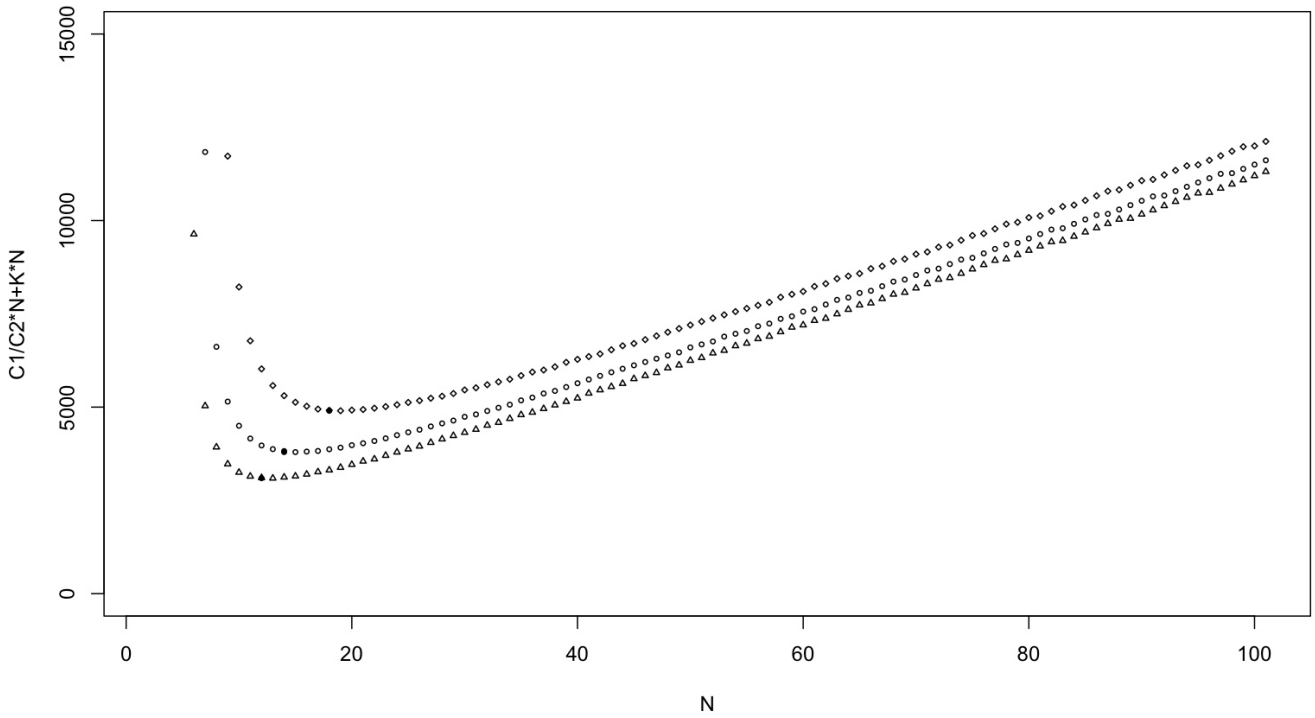


Figure 3.1: *The Core Part of Cost Function vs. the Number of Pairs for Power*$(= 0.9, 0.8, 0.7)$ *and Relative Cost* C_1/C_2 $(= 100)$.

3.1.3.2 The Type II Optimization For the type II optimization, let Ω_C be the set including all possible combinations of group size and numbers of counting frames satisfying the cost function $\leq C$. Ω_C can be denoted as

$$\Omega_C = \{(N, K) : Cost(N, K) \leq C\}. \tag{3.21}$$

The Type II optimal combination is the $(N, K)_{opt,II}$ which maximizes the power function $Power(N,\ K)$ in Ω_C.

For a fixed budget C,

$$C = Cost(N, K) = C_0 + 2C_1N + 2C_2NK = C_0 + 2N(C_1 + C_2K), \tag{3.22}$$

so that the group size N can be expressed as a function of the number of counting frames K, where

$$N(K) = \frac{C - C_0}{2(C_1 + C_2K)}. \tag{3.23}$$

The stereological variance σ_{st} also depends on the number of counting frames, given the relationship in Section 3.1.2.

Hence, the noncentral parameter ξ can also be given as a function of K,

$$\xi(K) = \frac{\delta}{\sqrt{\frac{2}{N(K)}(\sigma_B + \frac{\sigma^c_{st}(K)+\sigma^t_{st}(K)}{2})}}. \tag{3.24}$$

Now, the power function can be described as

$$\begin{aligned} Power(N, K) &= P(|T| > t_{0.975,\nu}|\xi) \\ &= \int_{-\infty}^{t_{0.025,\nu}} f(t)dt + \int_{t_{0.975,\nu}}^{\infty} f(t)dt, \end{aligned} \tag{3.25}$$

where $f(t) = \frac{\nu^{\nu/2}}{\sqrt{\pi}\Gamma(\nu/2)} \frac{e^{-\xi^2/2}}{(\nu+t^2)^{\nu+1}} \sum_{j=0}^{\infty} \frac{\nu+j+1}{2} \frac{\xi^j}{j!} \left(\frac{2t^2}{\nu+t^2}\right)^{j/2}$, and $\nu = N - 1 = \frac{C-C_0}{2(C_1+C_2K)} - 1$.

Since the power function now only depends on K, to maximize the power function, it is equivalent to solving

$$\begin{cases} \frac{dPower(K)}{dK} = 0 \\ \frac{d^2Power(K)}{dK^2} < 0. \end{cases}$$

When we consider the type II optimization, this algorithm requires more information about the three cost coefficients C_0, C_1 and C_2 than was needed for Type I optimization.

3.2 EXAMPLE

We consider an example which shows how to extend some of the ideas in Section 3.1 and how to potentially apply if we were to repeat a study like Dr. Konopaske. In the Konopaske study, 18 monkeys were matched by their body weights as triads and then within triads assigned at random into the three treatment groups: haloperidol, olanzapine, and sham control. Using the CE criterion, about 1600-1800 counting frames were collected from each animal.

In Konopaske et al. (2007) a two-way ANOVA model (with additive effects of group and triad) was used for the data set consisting of each monkey's stereologically estimated total glial cell number in its parietal lobe, in order to assess the effect of chronic antipsychotic exposure. The contrast of the combined antipsychotic-exposed groups vs the sham group was used to evaluate the effect of chronic antipsychotic exposure. One-sided testing of the contrast was done due to the directionality of expected reduction glial number in antipsychotic-exposed monkeys.

To first explore how the number of counting frames affects the final test statistic and resulting inference, we systematically deleted different proportions of counting frames within each animal and analyzed the reduced data set. First, we decreased the number of counting frames K in the data set by 10%. To roughly keep the systematical sampling characteristic, we systematically deleted the counting frames. Ten subsets for each animal were generated. For subset i, the i^{th} counting frame of each section was deleted and then every 10th thereafter is also deleted, $i = 1, \cdots, 10$. All the i^{th} subsets of each animal are combined together as a new sample of 18 animals. Therefore, ten such samples are generated. Then the two-way ANOVA model described in Konopaske et al. (2007) is applied to each sample. The estimate and statistical inference about the contrast between the combined antipsychotic-exposed and controls from the sample data sets do not change much comparing to the original data set. Table 3.2 lists the estimate of the average effect of the two chronic antipsychotic exposure groups minus that of the sham group, the test statistic and the p-values.

We then proceeded to decrease the number of counting frames by 20% and 30% sequentially. The statistical inference changed little (see Appendix A).

Table 3.2: *Comparing the Original Data Set and 10 Samples with* 10% *Reduction.*

Dataset	Estimate	StdErr	t-stat	p-value
Original	-25.0	12.94	-1.93	0.0413
Sample1	-24.5	12.85	-1.91	0.0427
Sample2	-25.8	13.60	-1.90	0.0436
Sample3	-23.2	12.63	-1.84	0.0479
Sample4	-26.4	12.90	-2.05	0.0339
Sample5	-24.5	13.06	-1.87	0.0453
Sample6	-23.7	12.66	-1.87	0.0454
Sample7	-26.2	12.95	-2.02	0.0355
Sample8	-25.1	13.46	-1.86	0.0460
Sample9	-25.1	12.49	-2.01	0.0364
Sample10	-25.4	13.11	-1.93	0.0410

Note: The degree of the t statistics is 10 and the p-value is based on a one-sided testing.

A second approach to exploring this issue used the bootstrap method to resample the data set. To obtain a bootstrap sample with $100p\%$ counting frames of the original data of each animal using section as a strata variable, $p \cdot n_j$ counting frames are independently selected within the j^{th} sections with replacement, for $j = 1, \cdots, m$. For each animal, we obtained a bootstrap sampling with $p \cdot K$ counting frames. We combined the bootstrap sampling from each animal to get a new sample of 18 monkeys, and performed the two-way ANOVA analyses. The results of this method are consistent with these obtained from the preceding systematic deletion (see Appendix B).

Both systematic reduction and bootstrap methods demonstrate that reducing the number of counting frames will only cause a slight loss of power. Moreover, in an actual study the lost power could by compensated for by increasing slightly the number of animals. Decreasing the number of counting frames by 30% in the Konopaske study means reducing 500 counting frames for each animal, which in turn translates into a reduction of more than a month of lab work in this study. We will later show in this chapter using the Konopaske data how we would provide an optimal design which also maintains the original data set's power.

3.2.1 Paired Design

In Konopaske study, we can estimate the stereological variance from the data set. If someone wants to repeat the study, we could provide the optimal design based on the information obtained from the Konopaske's study. For simplicity, suppose that we only have two groups matched as pairs in the new study, for example, haloperidol and sham groups. Due to the directionality of the treatment effect, a one-sided test is considered.

The hypotheses are

$$\begin{aligned} H_0 &: \mu_H = \mu_S, \\ H_a &: \mu_H < \mu_S. \end{aligned} \tag{3.26}$$

The test statistic in (3.7) is

$$T = \frac{\bar{\hat{\mathcal{N}}}_{H\cdot} - \bar{\hat{\mathcal{N}}}_{S\cdot}}{\sqrt{\frac{1}{N}(\frac{1}{N-1}\sum_{k=1}^{N}((\hat{\mathcal{N}}_{H,k} - \hat{\mathcal{N}}_{S,k}) - (\bar{\hat{\mathcal{N}}}_{H\cdot} - \bar{\hat{\mathcal{N}}}_{S\cdot}))^2)}}, \tag{3.27}$$

and the power function for the one-sided test is

$$\begin{aligned} Power &= P(\text{reject } H_0 | H_a) \\ &= P(T < t_{.05,N-1} |\ \xi = \frac{\delta}{\sqrt{\frac{1}{N}(2\sigma_B + \sigma_{st}^H + \sigma_{st}^S)}}), \end{aligned} \tag{3.28}$$

where $\hat{\mathcal{N}}_{i,k}$ is the stereological estimate of the total glial number in the parietal lobe for the monkey of pair k in group i, $\bar{\hat{\mathcal{N}}}_{i\cdot}$ is the mean of the stereological estimate of group i, σ_{st}^i is the stereological variance of group i, $i = H$ (haloperidol) or $i = S$ (sham), and σ_B is the biological variance, and δ is the numerical difference in total number that is to be detected.

In the data set, the average numbers of total counting frames used in haloperidol and sham groups are 1676 and 1646 respectively. We have estimated the two components of stereological variance for each group by averaging the two components of stereological variances of subjects within group, respectively. The estimates of σ_{st1}^H and σ_{st1}^S are 1.5 and 1.5, respectively, and the estimates of σ_{st2}^H and σ_{st2}^S are 29.0 and 39.6 (under the Cruz-Orive and Geiser assumption), respectively, where

$$\begin{aligned} \widehat{\sigma}_{st1}^H &= \frac{2.4 + 0.1 + 1.3 + 3.9 + 0.6 + 0.4}{6} = 1.5 \\ \widehat{\sigma}_{st2}^H &= \frac{23.3 + 28.0 + 26.4 + 36.2 + 25.8 + 34.3}{6} = 29.0 \\ \widehat{\sigma}_{st1}^S &= \frac{1.7 + 0.5 + 2.0 + 0.2 + 3.1 + 1.4}{6} = 1.5 \\ \widehat{\sigma}_{st2}^S &= \frac{41.7 + 32.1 + 50.4 + 23.8 + 49.8 + 39.7}{6} = 39.6, \end{aligned}$$

and the numbers can be obtained in Table 2.3.

The estimated noncentrality parameter for N pairs and K counting frames is

$$\begin{aligned}\hat{\xi} &= \frac{\delta}{\sqrt{\frac{1}{N}(2\sigma_B + (\hat{\sigma}^H_{st1} + \frac{1676}{K}\hat{\sigma}^H_{st2}) + (\hat{\sigma}^S_{st1} + \frac{1646}{K}\hat{\sigma}^S_{st2}))}} \\ &= \frac{\delta}{\sqrt{\frac{1}{N}(2\sigma_B + (1.7 + 29.2\frac{1676}{K}) + (1.4 + 38.8\frac{1646}{K}))}} \\ &= \frac{\delta}{\sqrt{\frac{1}{N}(2\sigma_B + 3.1 + \frac{112804}{K})}}. \end{aligned} \tag{3.29}$$

Based on the information obtained from the data set, we are able to decide the number of counting frames and sample size for a new study. The following table lists the numerical results for different pre-specified alternatives and relative costs of a study with 80% power.

Table 3.3: *Optimal Combination* $(N, K)_{opt,I}$ *for Konopaske Study - Paired Case.*

δ	σ_B	C_1/C_2	N	K
-20	20^2	100	29	116
		1000	19	342
		10000	16	821
-30	20^2	100	14	112
		1000	9	398
		10000	8	828
-20	15^2	100	20	153
		1000	12	500
		10000	10	1176
-30	15^2	100	11	123
		1000	7	338
		10000	6	610

In Konopaske data, the observed difference for haloperidol and sham group is -31.8 and the observed biological variance is 760.4. The optimal combination for the study is (12, 574) when the relative cost is 10000.

3.2.2 Matched Triads

In this section, we consider the issue involving in repeating the Konopaske's study exactly. Due to the assumption of unequal stereological variance among groups under the alternative hypothesis, the triad case is more complex than the paired case. We use Sattherwaite's approximation to adjust the test statistic.

The hypothesis of interest here is

$$\begin{aligned} H_0 &: \frac{\mu_H + \mu_O}{2} = \mu_S \\ H_a &: \frac{\mu_H + \mu_O}{2} < \mu_S. \end{aligned} \tag{3.30}$$

Let σ denote $var[\frac{\bar{\hat{N}}_{H\cdot}+\bar{\hat{N}}_{O\cdot}}{2} - \bar{\hat{N}}_{S\cdot}]$, so that

$$\begin{aligned} \sigma &= \frac{1}{4}var[\bar{\hat{N}}_{H\cdot}] + \frac{1}{4}var[\bar{\hat{N}}_{O\cdot}] + var[\bar{\hat{N}}_{S\cdot}] \\ &= \frac{1}{4}(\sigma_B + \sigma_{st}^H) + \frac{1}{4}(\sigma_B + \sigma_{st}^O) + (\sigma_B + \sigma_{st}^S) \\ &= \frac{3}{2N}\sigma_B + \frac{1}{4N}\sigma_{st}^H + \frac{1}{4N}\sigma_{st}^O + \frac{1}{N}\sigma_{st}^S. \end{aligned} \tag{3.31}$$

Let $MS_i = \frac{1}{N-1}\sum_{k=1}^{N}(\hat{N}_{ik} - \bar{\hat{N}}_{i\cdot})^2$, then $(N-1)MS_i \sim var[\bar{\hat{N}}_{i\cdot}]\chi^2_{N-1}$, $i = H, O$ and S. Using the Satterthwaite approximation, we have

$$\widehat{var}[\frac{\bar{\hat{N}}_{H\cdot} + \bar{\hat{N}}_{O\cdot}}{2} - \bar{\hat{N}}_{S\cdot}] = \frac{1}{4N}MS_H + \frac{1}{4N}MS_O + \frac{1}{N}MS_S, \tag{3.32}$$

so that $\widehat{var}[\frac{\bar{\hat{N}}_{H\cdot}+\bar{\hat{N}}_{O\cdot}}{2} - \bar{\hat{N}}_{S\cdot}]/\sigma$ can be approximated as a χ^2 distribution with degrees of freedom given by

$$\begin{aligned} df &= \frac{(\frac{1}{4N}MS_H + \frac{1}{4N}MS_O + \frac{1}{N}MS_S)^2}{\frac{(\frac{1}{4N}MS_H)^2+(\frac{1}{4N}MS_O)^2+(\frac{1}{N}MS_S)^2}{N-1}} \\ &= \frac{(\frac{1}{4}MS_H + \frac{1}{4}MS_O + MS_S)^2}{\frac{(\frac{1}{4}MS_H)^2+(\frac{1}{4}MS_O)^2+MS_S)^2}{N-1}}. \end{aligned} \tag{3.33}$$

A suitable test statistic of hypothesis (3.30) would be

$$T = \frac{\frac{\bar{\hat{\mathcal{N}}}_{H\cdot}+\bar{\hat{\mathcal{N}}}_{O\cdot}}{2} - \bar{\hat{\mathcal{N}}}_{S\cdot}}{\sqrt{\widehat{var}[\frac{\bar{\hat{\mathcal{N}}}_{H\cdot}+\bar{\hat{\mathcal{N}}}_{O\cdot}}{2} - \bar{\hat{\mathcal{N}}}_{S\cdot}]}}. \tag{3.34}$$

The resulting power function is

$$\begin{aligned} Power(N,K) &= P(\text{reject } H_0|H_a) \\ &= P(T < t_{.05,df}|\ \hat{\xi}), \end{aligned} \tag{3.35}$$

where given the alternative δ, T has a noncentral t distribution with a df given in (3.33) and the estimated noncentral parameter being

$$\hat{\xi} = \frac{\delta}{\sqrt{\frac{1}{N}(3\sigma_B + \frac{1}{4}(\hat{\sigma}^H_{st1} + \frac{K_0}{K}\hat{\sigma}^H_{st2}) + \frac{1}{4}(\hat{\sigma}^O_{st1} + \frac{K_0}{K}\hat{\sigma}^O_{st2}) + (\hat{\sigma}^S_{st1} + \frac{K_0}{K}\hat{\sigma}^S_{st2})}}. \tag{3.36}$$

The optimal algorithm is similar as the paired case, where we need the two components of stereological variances of the three groups. Optimal results for the triads case can be obtained as Table 3.3.

4.0 ADAPTIVE DESIGN

4.1 INTRODUCTION

Chapter 3 shows that to plan an optimal design we need the information about the sizes of the biological and stereological variances, which are unknown prior to study. Traditionally, researchers estimate such values from previous experience for the experimental design. We know from Chapter 2 that the stereological variance depends on the shape of the region of interest and on the number of neurons in it. Hence, it is difficult to prespecify the stereological variance before a study that is interested in a particular type of neuron. When the design parameters are incorrectly prespecified in the optimality calculation, the number of subjects and the number of counting frames may be incorrect. To avoid inefficient use of resources in stereological studies, we propose to introduce ideas from adaptive designs that have been widely studied and implemented in clinical trials during the past decade. An adaptive design allows us the opportunity to look at the data at an interim stage, and to modify the design based on the information obtained from the first stage data. The flexibility of adaptive design allows adjusting the sample size and the number of counting frames when the variability is mis-specified. Adaptive procedures can have cost advantages over standard fixed procedures. In this chapter we apply two stage adaptive procedures in stereology. While in the literature both blinded and unblinded procedures have been considered, we focus on blinded adaptive design procedures.

Several different unblinded adaptive procedures are described by Proschan and Hunsberger (1995), Shun (2001), Liu and Chi (2001). These designs unblind the treatment assignments at the end of Stage I, so that the group means and common variance can be estimated and used in the design of Stage II.

Blinded adaptive procedures (e.g., Stein (1945), Gould and Shih (1991), Gould and Shih (1998), Kieser and Friede (2001)) do not break the blind at the end of Stage I. Typically for these designs, only the common within-group variance is attempted to be estimated, and usually the sample size of Stage II is re-estimated based on the estimate of the variance.

In our adaptive procedure, we only estimate the stereological variance without breaking the blind of the Stage I data. Our goal is to design Stage II with the information obtained from Stage I. At the end of Stage I, we estimate the stereological variance, and adjust the number of counting frames in Stage II. We develop the appropriate test procedure and show that the type I error rate is controlled.

4.2 ADAPTIVE PROCEDURE

In the current chapter, we develop a blinded adaptive procedure in which the diagnostic codes for the subjects remain blinded until the end of the stereological study. Unlike traditional blinded procedure, we only estimate the stereological variance based on the blinded Stage I data instead of the variance of the primary outcome (number of neurons), which involves both biological variance and stereological variance components. Particularly, we consider the situation in which the design of Stage II, such as the number of counting frames for each subject and/or the sample size, depends on Stage I observations. In our procedure, we always assume that the biological variance is known and kept the same between stages. However, adjusting the number of counting frames in Stage II creates the difficulty that the variances at the subject level will be different between two stages . The usual paired t test is not appropriate for our procedure, and we propose to use an adjusted t-statistic.

The motivation for developing a two-stage adaptive procedure for the stereological study is due to the fact that it is difficult to prespecify the design parameters concerning the stereological variances prior to the beginning of the study which focuses on a particular type of neurons. The stereological variances can vary from study to study, since the stereological variation depends on the number of counting frames, and adjusting the number of counting frames changes the stereological variation.

A stereological study based on our adaptive design gives us the opportunity to look at the blinded data without revealing the treatment and pair assignments at the end of Stage I. The design of Stage II depends on the updated information concerning the stereological variance. We find that our two-stage adaptive procedure presents an advantage in terms of efficient use of resource with claimed power. One of the challenges for our statistical method is to maintain appropriate control of Type I error rate while permitting adaptation.

4.3 ADJUSTING THE NUMBER OF COUNTING FRAMES (K_2) OF STAGE II

4.3.1 Introduction

In this section, we consider only adjusting the number of counting frames for Stage II, while keeping the sample size of a study fixed. In fact, the total number of subjects that are available to be used in a study is often fixed by the budget or the available resources for many studies. For example, Dr. Konopaske's study had a fixed number of monkeys that were available. However, the number of counting frames can be adjusted to assure improved power for detecting a treatment effect under a proposed alternative. Under the uniformly

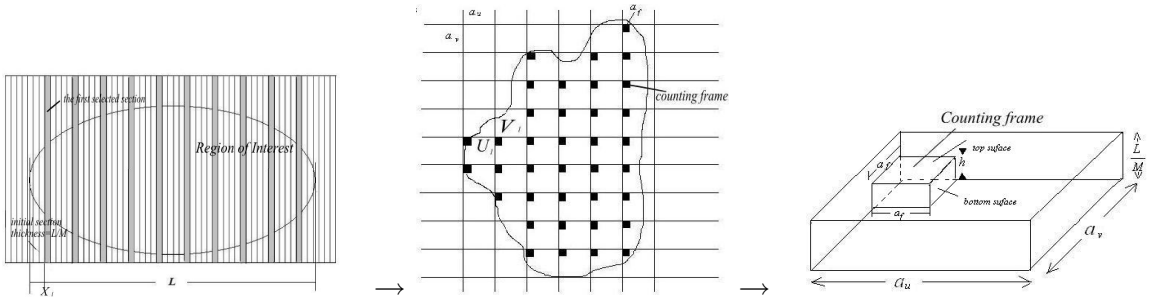


Figure 4.1: *Uniformly Systematic Sampling Scheme*

systematic sampling scheme of a stereological study (Figure 4.1), the section sampling step and the height sampling step remain unchanged for the two stages of the adaptive design, and only the area sampling step is adjusted to adjust the number of counting frames. The

way to adjust the number of counting frames is by altering the distances between the crossing lines which are called u-step (a_u) and v-step (a_v) of the grid. The other characteristics of the stereological design, such as the number of sections, the area of a counting frame and the thickness of a counting frame in a stereological study, are kept the same though the entire study. An advantage of keeping a fixed number of sections across stages is that the researchers can prepare the section slices in advance and don't need to wait until Stage I is finished. Figure 4.2 given by (Schmitz and Hof (2007)) illustrates a series of selected sections of brain tissues.

Figure 4.2: *Section samplings (Schmitz and Hof (2007))*

Let us suppose that the study is interested in a brain region with length L. The region is cut into M sections and m sections will be systematically chosen for examination. The thickness and the area of the counting frame are h and a_f^2, respectively. The areas of each section selected are denoted by $(Area_1, ..., Area_m)$, respectively. The section sampling fraction (τ_s) and height sampling fraction (τ_h) are given by

$$\tau_s = \frac{m}{M}$$
$$\tau_h = \frac{h}{L/M}.$$

In an actual stereological study, the researchers choose the u-step (a_u) and v-step (a_v) in order to obtain the required number of counting frames. They can decrease the number of

counting frames by enlarging the distances between the cross lines and making the rectangle larger, so that fewer number of counting frame are observed, and to increase the number of counting frames they decrease the relevant distances. The area sampling fraction is given by

$$\tau_a = \frac{a_f^2}{a_u a_v},$$

and the corresponding numbers of counting frames in each section are

$$n_j = \frac{Area_j}{a_u a_v},\ j = 1, \cdots, m. \tag{4.1}$$

Hence

$$\frac{n_j}{\tau_a} = \frac{Area_j}{a_f^2},\ j = 1, \cdots, m. \tag{4.2}$$

$Area_1, \cdots, Area_m$, and a_f^2 are fixed and known for a study; thus, the number of counting frames $K = \sum_{j=1}^{m} n_j$ and the sampling fraction τ_a have a proportional relationship.

Following the simple Cruz-Orive and Geiser approach, we assume that the number of neurons counted in each counting frame follows a Poisson distribution with the same λ among sections (see Section 2.3). Based on the Poisson assumption, the measurement function given by (2.4) on each section is $f_j = \frac{1}{\tau_a h} n_j \lambda = \frac{\lambda}{a_f^2 h} Area_j$, $j = 1, ..., m$. Then the two components of the stereological variance are given by

$$\begin{aligned}
\sigma_{st1} &= \frac{L^2}{240m^2}(3\sum_{j=1}^{m} f_j^2 + \sum_{j=1}^{m-2} f_j f_{j+2} - 4\sum_{j=1}^{m-1} f_j f_{j+1}) \\
&= \frac{L^2\lambda^2}{240m^2h^2a_f^4}(3\sum_{j=1}^{m} Area_j^2 + \sum_{j=1}^{m-2} Area_j Area_{j+2} - 4\sum_{j=1}^{m-1} Area_j Area_{j+1}),
\end{aligned} \tag{4.3}$$

and

$$\begin{aligned}
\sigma_{st2} &= \frac{1}{\tau_s^2\tau_a^2\tau_h^2}\lambda\sum_{j=1}^{m} n_j \\
&= \left[\frac{L^2\lambda}{m^2h^2a_f^4}\sum_{j=1}^{m} Area_j\right] a_u a_v.
\end{aligned} \tag{4.4}$$

As previously noted, for a stereological study, L, m, h, a_f and $Area_j$, $j = 1, \cdots, m$ are known and fixed. Thus, the first component (4.3) of the stereological variance is fixed for a

study, and the second component (4.4) of the stereological variance depends on the distance between the cross lines of the grid (a_u and a_v). The number of counting frames is determined by a_u and a_v.

In designing a stereological study one choses a_u and a_v based on pre-specified design parameters: type I error rate α, type II error rate β, total pairs of subject N, biological variance σ_B, alternative δ, and λ_0, where λ_0 is the initial guess of the true parameter λ. These will determined the number of counting frames, K_1, in the Stage I. The details for this calculation are discussed in Section 4.3.2. Let K_2 denote the number of counting frames re-calculated for Stage II. The procedure of choosing K_2 will be discussed in Section 4.3.3.

4.3.2 Planning Stage

In a matched pair study, the numbers of control and treatment subjects are balanced within each stage. In fact, the tissue specimens are often processed in batches and that pairs always appear in the same batch (to protect against batch to batch variation). We assume that the region of interest is basically the same among subjects for both treatment and control groups. We also assume that the stereological design parameters, such that the section sampling fraction, the area sampling fraction and the height sampling fraction, are the same for both stages. Hence, the same number of counting frames for each subject are collected within each stage.

In a study designed for CCNMD, control and treated subjects are pair-matched by covariates, such as age at death, gender, post-mortem interval and brain pH value. Let $\hat{\mathcal{N}}_{i,\gamma}$ be the stereological estimate of the neuron number of pair γ from i = control or treatment groups, respectively. Then

$$\hat{\mathcal{N}}_{i,\gamma} \sim^{i.i.d} N(\mu_c + p_\gamma, \sigma_B + \sigma^i_{st}), \tag{4.5}$$

where μ_i is the mean of group i ($i = c$ or $i = t$), p_γ is the effect of pair γ satisfying $\sum_{\gamma=1}^{N} p_\gamma = 0$ and σ_B is the biological variance. We assume that the biological variance is the same for treatment and control groups, but the stereological variances may be different due

to the differing neuron numbers for treatment and control. The hypotheses that one test are

$$H_0 : \mu_c = \mu_t$$
$$H_a : \mu_c \neq \mu_t. \tag{4.6}$$

To design a stereological study with a fixed number of pairs, the number of counting frames needs to be decided. It is important to note that the difference statistic only depends on the average of the stereological variances of treatment and controls for the paired studies (see Section 3.1.1). An initial guess of the average intensity of the Poisson distribution of treatment and control is λ_0. A simple paired t test is used to determine a_u and a_v. The power function for a two-sided α level test of (4.6) can be expressed by

$$\begin{aligned} &Power = P(\text{reject } H_0 |\ H_a) \\ &= P\left(|\ T\ | > t_{1-\alpha/2,N-1} |\ ncp = \frac{\delta}{\sqrt{\frac{1}{N}((\sigma_B + \sigma^c_{st1} + \sigma^c_{st2}) + (\sigma_B + \sigma^t_{st1} + \sigma^t_{st2}))}} \right) \\ &= P\left(|\ T\ | > t_{1-\alpha/2,N-1} |\ ncp = \frac{\delta}{\sqrt{\frac{2}{N}(\sigma_B + \bar{\sigma}_{st1} + \bar{\sigma}_{st2})}} \right), \end{aligned} \tag{4.7}$$

where α denotes the nominal type I error rate, $1-\beta$ denotes the planned power, $t_{1-\alpha/2,N-1}$ denotes the $1-\alpha/2$ critical value of the t distribution with $N-1$ degrees of freedom, δ is the pre-specified alternative, σ_B is the biological variance, ncp the corresponding non-centrality parameter of the noncentral t-distribution for T, and $\bar{\sigma}_{st1}$ $(= (\sigma^c_{st1} + \sigma^t_{st1})/2)$ and $\bar{\sigma}_{st2}$ $(= (\sigma^c_{st2} + \sigma^t_{st2})/2)$are the average of the two components of the stereological variances (i.e., (4.3) and (4.4)) for treatment and control groups, respectively. Substituting (4.3) and (4.4) for the two components of the stereological variance in the power function (4.7), we have that the resulting equations directly, relates power to $a_u a_v$ with all parameter values fixed by design, so that by a suitable choice of $a_u a_v$, we can obtain the desired power.

Note, there can be situations when σ_B is too large, so that power is unachievable. For a fixed N pairs subjects, to detect a difference of δ, $\xi_{\{\delta,\beta,N\}}$ which is the effect size to achieve $1-\beta$ power is also fixed. If the biological variance is relative large compared to the alternative δ, such that $\delta/\sqrt{2\sigma_B} \leq \xi_{\{\delta,\beta,N\}}$, the study power can't be obtained by adjusting the number

of counting frames, even with a very large number of counting frames. Also, when the average of the first component of the stereological variance $\bar{\sigma}_{st1}$ is too large, the study power can be unachievable.

Now, we give a simple hypothetical example to demonstrate how to select the u-step (a_u) and v-step (a_v) of grid for a stereological study. Table 4.1 gives an example of a set of

Table 4.1: *An Example*

N	σ_B	δ	$L(mm)$	$a_f^2(\mu m^2)$	$h(\mu m)$	m	λ_0
20	400	20	25	512	30	10	0.636

Note: The units for σ_B and δ are 10^{12}and10^6, respectively.

parameters for a stereological study. The length of the region of interest is 25 mm and 10 sections will be used in the study. The areas of the 10 sections are 204.8 mm^2, 230.4 mm^2, 256.0 mm^2, 281.6 mm^2, 307.2 mm^2, 307.2 mm^2, 281.6 mm^2, 256.0 mm^2, 230.4 mm^2 and 204.8 mm^2, respectively. The area and the thickness of a counting frame are 512 μm^2 and 30 μm, respectively.

With the initial guess of the average intensity of the Poisson distribution being 0.636, the average of the first components of the stereological variance is

$$\begin{aligned}\bar{\sigma}_{st1} &= \frac{L^2\lambda_0^2}{240m^2h^2a_f^4}(3\sum_{j=1}^{m} Area_j^2 + \sum_{j=1}^{m-2} Area_j Area_{j+2} - 4\sum_{j=1}^{m-1} Area_j Area_{j+1}) \\ &= \frac{(25*10^3)^2*0.636^2}{240*10^2*30^2*512^2}(3*668467.2*10^{12} + 564920.3*10^{12} - 4*623902.7*10^{12}) \\ &= 3.3*10^{12}.\end{aligned}$$

And the average of the second component of the stereological variance is

$$\begin{aligned}\bar{\sigma}_{st2} &= [\frac{L^2\lambda_0}{m^2h^2a_f^4}\sum_{j=1}^{m} Area_j]a_u a_v \\ &= \frac{(25*10^3)^2*0.636}{10^2*30^2*512^2} * 2560*10^6 a_u a_v \\ &= 43131510 a_u a_v.\end{aligned}$$

In this example, the total number of pairs N and the biological variance σ_B are assumed

to be 20 and $400 * 10^{12}$, respectively. To detect a difference of δ for a study with N pairs subjects, $\xi_{\{\delta,\beta,N\}}$ is the effect size to achieve $1-\beta$ power. $\xi_{\{\delta,\beta,N\}}$ can be obtained from noncentral t power calculates. and is 0.66 in our case. Accordingly the effect size is related to $a_u a_v$ by

$$\begin{aligned}\xi_{\{20,0.2,20\}} &= \frac{\delta}{\sqrt{2(\sigma_B + \bar{\sigma}_{st1} + \bar{\sigma}_{st2}))}} \\ &= \frac{20}{\sqrt{2(400 + 3.3 + 43131510 a_u a_v * 10^{-12})}}.\end{aligned}$$

Usually, we use the same u-step (a_u) and v-step (a_v), so that

$$\begin{aligned}a_u = a_v &= \sqrt{\left(\left(\frac{20}{.66}\right)^2 /2 - 400 - 3.3\right) /(43131510 * 10^{-12})} \\ &= 1.131\ mm.\end{aligned}$$

When the u-step and v-step is 1.131 mm, then the number of counting frames to be collected is

$$\begin{aligned}K_1 &= \frac{\sum_{j=1}^{m} Area_j}{a_u a_b} \\ &= \frac{2560}{1.131 * 1.131} \\ &= 2000.\end{aligned}$$

Thus, we use the u-step and v-step being 1.131 mm and have 2000 counting frames in this study. This example will also be used in the simulation study in Section 4.3.6.

Note that in the planning stage, we use the average intensity of treatment and control in the calculation for simplicity. In fact, for the second component of the stereological variance σ_{st2},

$$\begin{aligned}\bar{\sigma}_{st2} &= \frac{\sigma^c_{st2} + \sigma^t_{st2}}{2} \\ &= (\frac{L^2\lambda_c}{m^2h^2a_f^4}\sum_{j=1}^{m} Area_j a_u a_v + \frac{L^2\lambda_t}{m^2h^2a_f^4}\sum_{j=1}^{m} Area_j a_u a_v)/2 \\ &= \frac{L^2\lambda_0}{m^2h^2a_f^4}\sum_{j=1}^{m} Area_j a_u a_v,\end{aligned}$$

where λ_c and λ_t is the intensity for control and treatment group, respectively. And the average intensity $\lambda_0 = (\lambda_c + \lambda_t)/2$. But the average of the first component of the stereological variance,

$$\begin{aligned}\bar{\sigma}_{st1} &= \frac{\sigma^c_{st1} + \sigma^t_{st1}}{2} \\ &= \frac{L^2}{240m^2h^2a_f^4}(3\sum_{j=1}^{m} Area_j^2 + \sum_{j=1}^{m-2} Area_j Area_{j+2} - 4\sum_{j=1}^{m-1} Area_j Area_{j+1})(\lambda_c^2 + \lambda_t^2)/2,\end{aligned}$$

which requires the specification of both λ_c and λ_t. However, σ_{st1} is usually small compare with the σ_{st2}. In the previous example, σ_{st1} and σ_{st2} are $3.3*10^{12}$ and $55.8*10^{12}$, respectively.

4.3.3 Details of the Blinded Adaptation

The two components of the stereological variance in (4.7) are never known exactly in practice, and so in planning a study can only be guessed from previous experiments. Clearly, the number of counting frames will be either too large or too small when an initial choice of stereological variance is mis-specified. It is wasteful of time and money using more counting frames than necessary, and using an inadequate number of counting frames will increase the likelihood of an inconclusive study, which also wastes time and money.

Countering the uncertainty of the initial guess of stereological variance motives us to apply adaptive procedures to stereological studies. After obtaining first N_1 pairs subjects, the data are kept blinded and the two components of the stereological variance are estimated using interim observations. Unlike some blinded sample size re-estimation procedures used in clinical trials, such as Gould and Shih (1992), the biological variance is not re-estimated in our procedure. At the end of Stage I, we don't have any information about the treatment and pair assignments. Only the counted number of neurons in each of the K_1 counting frames is available for each subject. However, this information is sufficient for us to estimate the stereological variance, because the variance of pairwise difference depends on the average of treatment and controls stereological variance components , and the average stereological variance components can be estimated from the paired Stage I data without breaking the blind. Our goal in using blinded data is to avoid paying the typical penalties caused by breaking the blind. As in clinical trials, procedures for estimating variability based on

blinded data before the completion of the study tend to be more scientifically acceptable. Also, in the clinical trials literature, a number of authors have found that the inflation of type I error rate is at most slight when modifying designs using blinded data. We employ blinded estimation of the stereological variance to adjust the number of counting frames in Stage II, if necessary, to provide the required power against the null hypothesis when the alternative hypothesis is true.

4.3.4 Two Stage Data Structures

Before we present the strategy for picking the number of counting frames for Stage II, we give the structure of the two stage data. Note that the stereological variances in the two stages are different wherever the number of counting frames changes in Stage II. At this point to establish the appropriate notation, we are assuming that K_2 doesn't depend on the first stage data. Also assumed is that the distribution of the stereological estimates follows a normal distribution. N_1 and N_2 denote the numbers of pairs in two stages, respectively, and $N = N_1 + N_2$. N, N_1 and N_2 are all fixed in our procedure. Let $\hat{\mathcal{N}}^{\zeta}_{i,\gamma}$ be the stereological estimates of neuron number for control or treatment subjects ($i = c$ or t), in the ζ^{th} Stage, $\zeta = 1, 2$, for the γ^{th} pair, $\gamma = 1, \cdots, N_\zeta$, where

$$\hat{\mathcal{N}}^{\zeta}_{i,\gamma} \sim N(\mu_i + p_{\zeta,\gamma}, \sigma_B + \sigma^i_{st,K_\zeta}), \tag{4.8}$$

Then the pairwise differences between treatment and control subject within a pair in each of the two stages are

$$D_{\zeta\gamma} = \hat{\mathcal{N}}^{\zeta}_{c,\gamma} - \hat{\mathcal{N}}^{\zeta}_{t,\gamma} \sim N(\mu_c - \mu_t, 2\sigma_B + \sigma^c_{st,K_\zeta} + \sigma^t_{st,K_\zeta}), \quad \gamma = 1, \cdots, N_\zeta, \ \zeta = 1, 2, \tag{4.9}$$

and their average

$$\bar{D}_\zeta = \frac{1}{N_\zeta} \sum_{\gamma=1}^{N_\zeta} D_{\zeta\gamma} \sim N\left(\mu_c - \mu_t, (2\sigma_B + \sigma^c_{st,K_\zeta} + \sigma^t_{st,K_\zeta})/N_\zeta\right), \quad \zeta = 1, 2. \tag{4.10}$$

where K_ζ denotes the number of counting frames used in Stage ζ, $\zeta = 1, 2$.

The difference between treatment and control at the end of study is

$$\begin{aligned}\bar{D} = \bar{\hat{\mathcal{N}}}_c - \bar{\hat{\mathcal{N}}}_t &= \frac{1}{N}[(N_1\bar{\hat{\mathcal{N}}}_c^1 + N_2\bar{\hat{\mathcal{N}}}_c^2) - (N_1\bar{\hat{\mathcal{N}}}_t^1 + N_2\bar{\hat{\mathcal{N}}}_t^2)] \\ &= \frac{N_1}{N}(\bar{\hat{\mathcal{N}}}_c^1 - \bar{\hat{\mathcal{N}}}_t^1) + \frac{N_2}{N}(\bar{\hat{\mathcal{N}}}_c^2 - \bar{\hat{\mathcal{N}}}_t^2) \\ &= \frac{N_1}{N}\bar{D}_1 + \frac{N_2}{N}\bar{D}_2, \end{aligned} \tag{4.11}$$

so that

$$var[\bar{D}] = \frac{N_1}{N^2}(2\sigma_B + \sigma^c_{st,K_1} + \sigma^t_{st,K_1}) + \frac{N_2}{N^2}(2\sigma_B + \sigma^c_{st,K_2} + \sigma^t_{st,K_2}). \tag{4.12}$$

Hence,

$$\bar{D} \sim N(\mu_c - \mu_t, \frac{N_1}{N^2}(2\sigma_B + \sigma^c_{st,K_1} + \sigma^t_{st,K_1}) + \frac{N_2}{N^2}(2\sigma_B + \sigma^c_{st,K_2} + \sigma^t_{st,K_2})). \tag{4.13}$$

If $K_1 = K_2$, then $\bar{D} \sim N(\mu_c - \mu_t, \frac{1}{N}(2\sigma_B + \sigma^c_{st,K_1} + \sigma^t_{st,K_1}))$, which has the same distribution as the simple paired case.

In the next section we show how to choose K_2, but now we are proposing the test statistic which will be used after the completion of the study. Since the variances in the two stages are different, the simple paired t-test which is the basis for the power calculation in (4.7) is not appropriate. The distribution given in (4.13) suggests to consider a linear combination of sample variances in the two stages to estimate the variance of $\bar{D}$.

The sample variances of difference in each of the two stages are

$$s^2_{D_\zeta} = \frac{1}{N_\zeta - 1}\sum_{\gamma=1}^{N_\zeta}(D_{\zeta\gamma} - \bar{D}_\zeta)^2, \tag{4.14}$$

where $\frac{N_\zeta - 1}{2\sigma_B + \sigma^c_{st,K_\zeta} + \sigma^t_{st,K_\zeta}} s^2_{D_\zeta} \sim \chi^2_{N_\zeta - 1}$, for $\zeta = 1, 2$.

Then $E[s^2_{D_\zeta}] = 2\sigma_B + \sigma^c_{st,K_\zeta} + \sigma^t_{st,K_\zeta}$ and $Var[s^2_{D_\zeta}] = \frac{2(2\sigma_B + \sigma^c_{st,K_\zeta} + \sigma^t_{st,K_\zeta})^2}{N_\zeta - 1}$ follow from the moments of the χ^2 distribution.

Now let

$$s^2 = \frac{N_1}{N^2}s^2_{D_1} + \frac{N_2}{N^2}s^2_{D_2}, \tag{4.15}$$

then

$$\begin{aligned} E[s^2] &= \frac{N_1}{N^2}E[s^2_{D_1}] + \frac{N_2}{N^2}E[s^2_{D_2}] \\ &= \frac{N_1}{N^2}(2\sigma_B + \sigma^c_{st,K_1} + \sigma^t_{st,K_1}) + \frac{N_2}{N^2}(2\sigma_B + \sigma^c_{st,K_2} + \sigma^t_{st,K_2}). \end{aligned} \tag{4.16}$$

Thus s^2 is an unbiased estimate of the variance of $\bar{D}$.

Then we have

$$\begin{aligned} Var[s^2] &= \left(\frac{N_1}{N^2}\right)^2 Var[s^2_{D_1}] + \left(\frac{N_2}{N^2}\right)^2 Var[s^2_{D_2}] \\ &= \left(\frac{N_1}{N^2}\right)^2 \frac{2(2\sigma_B + \sigma^c_{st,K_1} + \sigma^t_{st,K_1})^2}{N_1 - 1} + \left(\frac{N_2}{N^2}\right)^2 \frac{2(2\sigma_B + \sigma^c_{st,K_2} + \sigma^t_{st,K_2})^2}{N_2 - 1}. \end{aligned} \tag{4.17}$$

So s^2 can be used as the denominator of an approximate t-statistic. We want to find the degree of freedom df such that $\frac{df \cdot s^2}{E[s^2]}$ has approximately χ^2_{df} distribution. The Satterthwaite is one such approximation, and attempts to choose df such that the variance of $\frac{df \cdot s^2}{E[s^2]}$ matches that of the χ^2_{df} distribution where

$$\begin{aligned} Var\left[\frac{df \cdot s^2}{E[s^2]}\right] &= \frac{df^2 \cdot Var[s^2]}{E^2[s^2]} \\ &= 2 \cdot df. \end{aligned} \tag{4.18}$$

Note that by definition, the mean of $\frac{df \cdot s^2}{E[s^2]}$ already matches that of the χ^2_{df} distribution.

Then

$$\begin{aligned} df &= \frac{2 \cdot E^2[s^2]}{Var[s^2]} \\ &= \frac{\left[\frac{N_1}{N^2}(2\sigma_B + \sigma^c_{st,K_1} + \sigma^t_{st,K_1}) + \frac{N_2}{N^2}(2\sigma_B + \sigma^c_{st,K_2} + \sigma^t_{st,K_2})\right]^2}{\left(\frac{N_1}{N^2}\right)^2 \frac{(2\sigma_B + \sigma^c_{st,K_1} + \sigma^t_{st,K_1})^2}{N_1 - 1} + \left(\frac{N_2}{N^2}\right)^2 \frac{(2\sigma_B + \sigma^c_{st,K_2} + \sigma^t_{st,K_2})^2}{N_2 - 1}}. \end{aligned} \tag{4.19}$$

Thus, the test statistic

$$T = \frac{\bar{D}}{\sqrt{s^2}} \tag{4.20}$$

has approximately a t_{df} distribution with df defined by (4.19).

When $K_1 = K_2$, the degrees of freedom df is

$$\begin{aligned} df &= \frac{\left[\frac{N_1}{N^2}(2\sigma_B + \sigma^c_{st,K_1} + \sigma^t_{st,K_1}) + \frac{N_2}{N^2}(2\sigma_B + \sigma^c_{st,K_1} + \sigma^t_{st,K_1})\right]^2}{\left(\frac{N_1}{N^2}\right)^2 \frac{(2\sigma_B+\sigma^c_{st,K_1}+\sigma^t_{st,K_1})^2}{N_1-1} + \left(\frac{N_2}{N^2}\right)^2 \frac{(2\sigma_B+\sigma^c_{st,K_1}+\sigma^t_{st,K_1})^2}{N_2-1}} \\ &= \frac{\left(\frac{N_1}{N^2} + \frac{N_2}{N^2}\right)^2}{\left(\frac{N_1}{N^2}\right)^2 \frac{1}{N_1-1} + \left(\frac{N_2}{N^2}\right)^2 \frac{1}{N_2-1}} \\ &= \frac{N^2}{\frac{N_1^2}{N_1-1} + \frac{N_2^2}{N_2-1}}. \end{aligned}$$

Hence when $N_1 = N$, $df = N - 1$ which is the same as simple paired study. When $N_1 = 0.5N$, $df = N - 2$ which is 1 degree of freedom less than the paired one, which is the cost for carrying out a two stage procedure.

For the moment ignoring the adaptation, we are able to assume $s^2_{D_1}$ and $s^2_{D_2}$ are independent random variables. When performing an unblinded adaptive design, the concern is doing the analysis at the end of the study ignoring the adaptation might inflate the type I error rate. Nonetheless, even after K_2 is chosen adaptively based on blinded Stage I data we propose in Section 4.3.5, to use the t-test in (4.20) to test whether or not the treatment and control population neuron counts differ. However, in related settings with blinded sample size adjustments, other researchers have found no inflation of the type I error in an approach like ours. For example, Gould and Shih (1991) described an approach to recalculating sample size that allowed estimating the variance without unblinding the data at the end of the first stage, and the approach protected type I error rate. The type I error rate for our procedure is discussed in Section 4.3.6.1.

4.3.5 Choice of K_2

In this section we show how to choose the number of counting frames K_2 for Stage II. We note that the power function for the t-statistic in (4.20) assuming K_2 is not a function of Stage I is

$$1 - \beta = Power = P(|T| > t_{0.975,df}|ncp). \tag{4.21}$$

Under the alternative δ, the non-centrality parameter is

$$ncp = \frac{\delta}{\sqrt{\frac{N_1}{N^2}(2\sigma_B + \sigma^c_{st,K_1} + \sigma^t_{st,K_1}) + \frac{N_2}{N^2}(2\sigma_B + \sigma^c_{st,K_2} + \sigma^t_{st,K_2})}}, \tag{4.22}$$

and the degrees of freedom is defined by (4.19).

Our adaptive approach uses the stereological variance estimators based on Stage I and combines the two stages data using the t-test of (4.20), ignoring the fact that we have done an adaptation. The stereological design for Stage II is adjusted according to the Stage I estimates. In particularly, the estimated stereological variance at the end of Stage I will be used in determining the number of counting frames of Stage II. Our approach is to adjust the number of counting frames for Stage II by updating the stereological variance components in the power function (4.7). The power function with this estimation is now given by

$$1 - \beta = P(|T| > t_{1-\alpha/2,\widehat{df}(K_2)}|\widehat{ncp}(K_2)), \tag{4.23}$$

where

$$\begin{aligned}\widehat{ncp}(K_2) =& \delta/\{\frac{N_1}{N^2}[2\sigma_B + (\widehat{\sigma}^c_{st1,K_1} + \widehat{\sigma}^t_{st1,K_1}) + (\widehat{\sigma}^c_{st2,K_1} + \widehat{\sigma}^t_{st2,K_1})] + \\ & \frac{N_2}{N^2}[2\sigma_B + (\widehat{\sigma}^c_{st1,K_1} + \widehat{\sigma}^t_{st1,K_1}) + \frac{K_1}{K_2}(\widehat{\sigma}^c_{st2,K_1} + \widehat{\sigma}^t_{st2,K_1})]\}^{-\frac{1}{2}},\end{aligned} \tag{4.24}$$

and estimated degrees of freedom

$$\begin{aligned}\widehat{df}(K_2) =& \{\frac{N_1}{N^2}[2\sigma_B + (\widehat{\sigma}^c_{st1,K_1} + \widehat{\sigma}^t_{st1,K_1}) + (\widehat{\sigma}^c_{st2,K_1} + \widehat{\sigma}^t_{st2,K_1})] \\ & + \frac{N_2}{N^2}[2\sigma_B + (\widehat{\sigma}^c_{st1,K_1} + \widehat{\sigma}^t_{st1,K_1}) + \frac{K_1}{K_2}(\widehat{\sigma}^c_{st2,K_1} + \widehat{\sigma}^t_{st2,K_1})]\}^2 \\ & /\{\left(\frac{N_1}{N^2}\right)^2 \frac{[2\sigma_B + (\widehat{\sigma}^c_{st1,K_1} + \widehat{\sigma}^t_{st1,K_1}) + (\widehat{\sigma}^c_{st2,K_1} + \widehat{\sigma}^t_{st2,K_1})]^2}{N_1 - 1} \\ & + \left(\frac{N_2}{N^2}\right)^2 \frac{[2\sigma_B + (\widehat{\sigma}^c_{st1,K_1} + \widehat{\sigma}^t_{st1,K_1}) + \frac{K_1}{K_2}(\widehat{\sigma}^c_{st2,K_1} + \widehat{\sigma}^t_{st2,K_1})]^2}{N_2 - 1}\}.\end{aligned} \tag{4.25}$$

Note that the number of counting frames K_1 in Stage I cannot be changed, but that we can change the number of counting frames K_2 in Stage II, based on our estimate of the average stereological variances from Stage I. Observe that both the non-centrality parameter ($\widehat{ncp}(K_2)$) and degrees of freedom ($\widehat{df}(K_2)$) are functions of K_2. For the pre-specified alter-

native δ and fixed sample size, the power function only depends on the number of counting frames of Stage II (K_2). Our adaptive design for the second stage of the study selects K_2 to satisfy the power equation (4.23). In essence, we are estimating one of the design parameters which we may have mis-specified in the initial study design, and then recomputing the power taking into account that the number of counting frames in Stage I is already fixed. Simulation results which are provided in Section 4.3.6 indicate that our approach preserves the type I error rate well.

In order to obtain K_2 counting frames for each subject in Stage II, the u-step and v-step are adjusted to $\sqrt{\frac{K_1}{K_2}}a_u$ and $\sqrt{\frac{K_1}{K_2}}a_v$, respectively. Consider the example in Section 4.3.2, if K_2 required in Stage II is 1000, then the u-step and v-step are adjusted to $\sqrt{2000/1000} * 1.131 = 1.599\ mm$.

It is important to reiterate that the distribution (4.13) of the difference statistic only depends on the average of the stereological variances of treatment and controls. Thus, the two stereological variance components can be estimated by the average of the treatment and control components from the paired Stage I data which can be done without breaking the blind. Applying the method described in Chapter 2, the two components of the stereological variance can be obtained separately for each animal using (2.20) and (2.27). Then the two components of the stereological variance estimate is given by the average, for $\upsilon = 1, 2$,

$$\widehat{\sigma}_{st,\upsilon,K_1} = \frac{\bar{\widehat{\sigma}}^{c}_{st,\upsilon,K_1} + \bar{\widehat{\sigma}}^{t}_{st,\upsilon,K_1}}{2} = \frac{\sum_{\gamma=1}^{N_1} \widehat{\sigma}^{c\gamma}_{st,\upsilon,K_1} + \sum_{\gamma=1}^{N_1} \widehat{\sigma}^{t\gamma}_{st,\upsilon,K_1}}{2N_1}, \tag{4.26}$$

where $\widehat{\sigma}^{i\gamma}_{st,\upsilon,K_1}$ is the υ^{th} component ($\upsilon = 1$ corresponds to between section variance and $\upsilon = 2$ to within section variance) of the stereological variance estimate of the γ^{th} pair in the group i ($i = c$ or $i = t$), and $\bar{\widehat{\sigma}}^{i}_{st,\upsilon,K_1}$ is the average of the υ^{th} component of the stereological variance estimate of the group i ($i = c$ or $i = t$). The stereological variance estimates in (4.26) are the updated stereological variances at the end of Stage I and used for designing of Stage II. The two components of the stereological variance are used in the power function (4.23).

4.3.6 Simulation

To demonstrate that the null hypothesis distribution of the proposed test statistic (4.20) is appropriately approximated by the corresponding t distribution with df defined in (4.19), we conduct a simulation study. Our simulation continues the example in Section 4.3.2. The parameters used in the simulation are listed in Table 4.2. Following Section 4.3.2, for a fixed sample size stereological study, we pick $K_1 = 2000$ based on our initial guess ($\lambda_0 = 0.636$) of the average intensity of Poisson distribution under the Cruz-Orive assumption. The numbers of counting frames in each section are calculated based on (4.1), which are $(160, 180, 200, 220, 240, 240, 220, 200, 180, 160)$, respectively, in Stage I.

4.3.6.1 Type I error rate In the simulation study, we generate the counting frame data for each subject. Under the null hypothesis, the population numbers of neurons in the control and experimental groups are the same and two groups have the same population intensities. Note that the true stereological variability depends on the true population intensity, λ, of neurons in each counting frame which as noted under the null is the same for both treatment and control groups. In this simulation study, we explore the effects of differing λ (= 0.4, 0.5, 0.6, 0.7,0.8) on the type I error rate. For each subject, the numbers of neurons counted in each counting frame is independently generated according to Poisson(λ).

First, we generate the counting frame data for Stage I with N_1 pairs. In the simulation, we explore the effects of various choices of N_1 from $5 = 25\%N$ to $15 = 75\%N$.

Stage I Simulation In Stage I, for subject l_1 in group i , the counted number $Q_{jk}^{1,l_1,i}$ in k^{th} counting frame j^{th} section follows a distribution of Poisson(λ), where $i = c, t$, $l_1 = 1, \cdots, N_1$, $k = 1, \cdots, n_j$ and $j = 1, \cdots, 10$, and $K_1 = \sum_{j=1}^{m} n_j$.

- We generate $Q_{jk}^{1,l_1,i}$ independently from Poisson(λ).
- The counting frame observation for the $(l_1, i)^{th}$ subject is $(Q_{11}^{1,l_1,i}, \cdots, Q_{1,n_1}^{1,l_1,i}, \cdots, Q_{10,1}^{1,l_1,i}, \cdots, Q_{10,n_{10}}^{1,l_1,i})$,
- The stereological estimate $\hat{\mathcal{N}}^{1,l_1,i}$ for the $(l_1, i)^{th}$ subject is estimated directly from (2.1),

Table 4.2: *Simulation Parameter List*

$L\ (mm)$	25
m	10
$(Area_1, \cdots, Area_m)$ (mm^2)	(204.8, 230.4, 256.0, 281.6, 307.2, 307.2, 281.6, 256.0, 230.4, 204.8)
$a_f^2\ (\mu m^2)$	512
$h\ (\mu m)$	30
$a_u = a_v\ (mm)$	1.131
α	0.05
β	0.20
$\delta(10^6)$	20
$\sigma_B(10^{12})$	400
λ_0	0.636
N	20
K_1	2000

$$\hat{\mathcal{N}}^{1,l_1,i} = \frac{1}{\tau_s\tau_a\tau_h}\sum_{j=1}^{10}\sum_{k=1}^{n_j} Q_{jk}^{1,l_1,i}$$
$$= \frac{La_xa_y}{mha_f^2}\sum_{j=1}^{10}\sum_{k=1}^{n_j} Q_{jk}^{1,l_1,i}.$$

Note σ_B is fixed by the example.

- To generate the primary outcomes, the biological variability (4.8) should be taken into account. To do so we generate a random number $\epsilon^{1,l_1,i}$ independently from $N(0,\sigma_B)$, and add that value to the stereological estimate $\hat{\mathcal{N}}^{1,l_1,i}$ to obtain the primary outcome $\hat{\mathcal{N}}^{*1,l_1,i}$, that is

$$\hat{\mathcal{N}}^{*1,l_1,c} = \hat{\mathcal{N}}^{1,l_1,c} + \epsilon^{1,l_1,c},$$
$$\hat{\mathcal{N}}^{*1,l_1,t} = \hat{\mathcal{N}}^{1,l_1,t} + \epsilon^{1,l_1,t}.$$

  Note, in reality, the data we obtain in the Stage I is $\hat{\mathcal{N}}^{*1,l_1,i}$ which includes the biological variation, and $\hat{\mathcal{N}}^{1,l_1,i}$ is never known. One wary to actually generate the data is using the random intensity λ for each subject. The variation of the intensity would be the source of the biological variance. However, since such a simulation is under Cruz-Orive and Geiser assumption where the counting frame data follows a Poisson distribution, the validity of normality distribution assumption in (4.8) would be in doubt. In our adaptation procedure, since our focus is to re-estimate the stereological variance, we use the approximation of adding the biological variance to the stereological estimators of the neuron number to avoid the conflict with distribution assumption of primary outcome.

- The two components of the stereological variance of the simulated stereological sample are calculated based on formula (2.20) and (2.27) in Chapter 2,

$$\begin{aligned}
\hat{\sigma}^{1,l_1,i}_{st1,K_1} &= \frac{1}{240\tau_s^2\tau_a^2\tau_h^2}\left(3\sum_{j=1}^{10}(Q_j^{1,l_1,i})^2+\sum_{j=1}^{8}Q_j^{1,l_1,i}Q_{j+2}^{1,l_1,i}-4\sum_{j=1}^{9}Q_j^{1,l_1,i}Q_{j+1}^{1,l_1,i}\right)\\
&= \frac{L^2a_u^2a_v^2}{240m^2h^2a_f^4}\left(3\sum_{j=1}^{10}(Q_j^{1,l_1,i})^2+\sum_{j=1}^{8}Q_j^{1,l_1,i}Q_{j+2}^{1,l_1,i}-4\sum_{j=1}^{9}Q_j^{1,l_1,i}Q_{j+1}^{1,l_1,i}\right),\\
\hat{\sigma}^{1,l_1,i}_{st2,K_1} &= \frac{1}{\tau_s^2\tau_a^2\tau_h^2}\sum_{j=1}^{10}Q_j^{1,l_1,i}\\
&= \frac{L^2a_u^2a_v^2}{m^2h^2a_f^4}\sum_{j=1}^{10}Q_j^{1,l_1,i},
\end{aligned}$$

where $Q_j^{1,l_1,i}(=\sum_{k=1}^{n_j}Q_{jk}^{1,l_1,i})$ is the sum of the counted neuron in j^{th} section for subject l_1 in group i.

At the end of Stage I, use the average of the $2\cdot N_1$ stereological variance estimates as the estimated stereological variance for each component,

$$\begin{aligned}
\widehat{\sigma}_{st1,K_1} &= \frac{\sum_{i=c,t}\sum_{l_1=1}^{N_1}\hat{\sigma}^{1,l_1,i}_{st1,K_1}}{2N_1},\\
\widehat{\sigma}_{st2,K_1} &= \frac{\sum_{i=c,t}\sum_{l_1=1}^{N_1}\hat{\sigma}^{1,l_1,i}_{st2,K_1}}{2N_1}.
\end{aligned}$$

Then we can calculate the number of counting frames of the second stage (K_2) based on (4.23) with the two components of the stereological variance $\widehat{\sigma}_{st1,K_1}$ and $\widehat{\sigma}_{st2,K_1}$. Thus, K_2 solves the power equation

$$1-\beta = P(|T| > t_{1-\alpha/2,\widehat{df}}|\widehat{ncp}),$$

where

$$\widehat{ncp}(K_2) = \frac{\delta}{\sqrt{\frac{N_1}{N^2}[2(\sigma_B+\widehat{\sigma}_{st1,K_1}+\widehat{\sigma}_{st2,K_1})]+\frac{N_2}{N^2}[2(\sigma_B+\widehat{\sigma}_{st1,K_1}+\frac{K_1}{K_2}\widehat{\sigma}_{st2,K_1})]}},$$

and estimated degrees of freedom

$$\widehat{df}(K_2) = \frac{\left[\frac{N_1}{N^2}(\sigma_B+\widehat{\sigma}_{st1,K_1}+\widehat{\sigma}_{st2,K_1})+\frac{N_2}{N^2}(\sigma_B+\widehat{\sigma}_{st1,K_1}+\frac{K_1}{K_2}\widehat{\sigma}_{st2,K_1})\right]^2}{\left(\frac{N_1}{N^2}\right)^2\frac{(\sigma_B+\widehat{\sigma}_{st1,K_1}+\widehat{\sigma}_{st2,K_1})^2}{N_1-1}+\left(\frac{N_2}{N^2}\right)^2\frac{(\sigma_B+\widehat{\sigma}_{st1,K_1}+\frac{K_1}{K_2}\widehat{\sigma}_{st2,K_1})^2}{N_2-1}}.$$

Then, we generate the counting frame data for Stage II with $N_2 = N - N_1$ pairs.

Stage II Simulation In Stage II, the counted number $Q_{jk}^{2,l_2,i}$ is also from Poisson distribution with same λ, where $l_2 = 1, \cdots, N_2$, and the steps are essentially the same as Stage I, except the number of counting frames is adjusted.

- We generate $Q_{jk}^{2,l_2,i}$ independently from Poisson(λ).
- The counting frame observation for the $(l_2, i)^{th}$ subject is $(Q_{11}^{2,l_2,i}, \cdots, Q_{1,\frac{K_2}{K_1}n_1}^{2,l_2,i}, \cdots, Q_{10,1}^{2,l_2,i}, \cdots, Q_{10,\frac{K_2}{K_1}n_{10}}^{2,l_2,i})$.
- The stereological estimate $\hat{\mathcal{N}}^{2,l_2,i}$ for the $(l_2, i)^{th}$ subject is estimated directly from (2.1),

$$\begin{aligned}\hat{\mathcal{N}}^{2,l_2,i} &= \frac{1}{\tau_s \tau_a \tau_h} \sum_{j=1}^{10} \sum_{k=1}^{\frac{K_2}{K_1}n_j} Q_{jk}^{2,l_2,i} \\ &= \frac{La_x a_y}{mha_f^2} \sum_{j=1}^{10} \sum_{k=1}^{\frac{K_2}{K_1}n_j} Q_{jk}^{2,l_2,i}.\end{aligned}$$

- To generate the primary outcomes, the biological variability is again accounted for, and we generate random numbers $\epsilon^{2,l_2,i}$ independently from $N(0, \sigma_B)$ and appropriately add them to obtain the primary outcomes.

Combining Stage At the end of Stage II, we calculate the p-value for the stereological study based on the combined data sets of $(\hat{\mathcal{N}}^{*1,l_1,c}, \hat{\mathcal{N}}^{*1,l_1,t})$ and $(\hat{\mathcal{N}}^{*2,l_2,c}, \hat{\mathcal{N}}^{*2,l_2,t})$, where the test statistic is defined by (4.20).

This process is repeated 1000 times, so that the type I error rate of our procedure is estimated by the proportion of p-value less than $\alpha = 0.05$. The average number of counting frames K_2 is also obtained. Table 4.3 gives the estimated type I error rates for different true λ's and varying values of N_1, the Stage I sample size. No substantive impact of λ and N_1 on the type I error rate is found in this simulation study.

4.3.6.2 Power

The power under blinded adaptive procedure can be obtained analogously to the type I error rate by simulation. Under the alternative hypothesis, the population numbers of neuron in the control and experimental groups are different , so that the two

Table 4.3: *Simulation – Type I Error Rate* (*1000 times,* $\lambda_0 = 0.636$)

N_1	$\lambda = 0.4$		$\lambda = 0.5$		$\lambda = 0.6$		$\lambda = 0.7$	
	α	K_2	α	K_2	α	K_2	α	K_2
5	0.052	1083	0.048	1445	0.047	1866	0.051	2369
6	0.046	1050	0.049	1417	0.045	1857	0.046	2398
7	0.051	1012	0.051	1386	0.048	1845	0.049	2435
8	0.053	972	0.053	1352	0.051	1834	0.056	2479
9	0.048	929	0.046	1312	0.052	1821	0.049	2538
10	0.052	881	0.053	1270	0.047	1808	0.051	2605
11	0.047	831	0.048	1219	0.045	1787	0.053	2693
12	0.051	774	0.042	1164	0.053	1763	0.047	2818
13	0.047	711	0.053	1097	0.049	1732	0.050	2990
14	0.045	642	0.051	1020	0.046	1696	0.044	3259
15	0.046	565	0.048	930	0.051	1646	0.048	3732

groups have different true population intensities. In simulating the power, we assume that the true population intensities of neurons in each counting frame for treatment and control group are λ_t and λ_c, respectively. Since

$$\begin{aligned}\delta &= \mu_c - \mu_t \\ &= \frac{1}{\tau_s\tau_a\tau_h}\sum_{j=1}^{m} n_j\lambda_c - \frac{1}{\tau_s\tau_a\tau_h}\sum_{j=1}^{m} n_j\lambda_t \\ &= \frac{L}{mha_f^2}\left[\sum_{j=1}^{m} Area_j\right](\lambda_c - \lambda_t), \end{aligned} \tag{4.27}$$

the difference of the two intensities has a proportional relationship with the alternative δ.

According to (4.27), the difference of λ_c and λ_t is $\frac{\delta}{\frac{L}{mha_f^2}\sum_{j=1}^{m} Area_j}$, which is independent of the number of counting frames. In our simulation, we assume the alternative δ is $20 * 10^6$, then

$$\begin{aligned}\lambda_c - \lambda_t &= \frac{\delta}{\frac{L}{mha_f^2}\sum_{j=1}^{m} Area_j} \\ &= \frac{20 * 10^6}{\frac{25*10^3}{10*30*512} * 2560 * 10^6} \\ &= 0.05.\end{aligned}$$

On the other hand, the overall mean density for the paired study is

$$\lambda = \frac{\lambda_c + \lambda_t}{2}.$$

Then we select the following pairs of parameters $(\lambda_c = 0.425, \lambda_t = 0.375)$, $(\lambda_c = 0.525, \lambda_t = 0.475)$, $(\lambda_c = 0.625, \lambda_t = 0.575)$ and $(\lambda_c = 0.725, \lambda_t = 0.625)$. The power is estimated by the proportion of times that we reject the null hypothesis in 1000 tests. The expected number of counting frames is very similar as the number we obtained in type I error rate simulations.

We also generate the counting frame data for each subject of treatment and control groups from Poisson distribution with λ_t and λ_c, respectively. However, all other parameters remain the same including σ_B and N_1. Table 4.4-4.7 give the desired power of the study under the alternative. Consider the case when the true λ is 0.4 so that there is about a

60% over-estimate of the intensity parameter used in the planning stage. If there are 5 pairs of subjects in Stage I, we find that on average 1083 counting frames are required in Stage II, which is about 55% of the number used in Stage I. If we followed the original design without an adaptation, the number of counting frames is 80000, while the average number of counting frames is 52490 in our procedure. On the other hand, when the initial guess of λ is too low, the number of counting frames is required to be increased to maintain power. When the true λ is 0.8, there is about a 20% under-estimate of the intensity parameter used in the design. If there are 5 pairs of subjects in Stage I, we find that on the average 2975 counting frames are required in Stage II in order to obtain the desired power, which is about 50% increase of number used in Stage I. This increase in the numbers of counting frames occurs because the original design is under powered.

Table 4.4: *Simulation - Power 1* ($\lambda_c = 0.425, \lambda_t = 0.375$)

N_1	5	6	7	8	9	10	11	12	13	14	15
Power	0.833	0.825	0.829	0.818	0.845	0.836	0.827	0.826	0.797	0.825	0.815
K_2	1084	1048	1012	972	929	882	830	774	711	642	566

Table 4.5: *Simulation - Power 2* ($\lambda_c = 0.525, \lambda_t = 0.475$)

N_1	5	6	7	8	9	10	11	12	13	14	15
Power	0.830	0.849	0.815	0.823	0.800	0.808	0.816	0.823	0.814	0.839	0.814
K_2	1445	1417	1387	1351	1313	1269	1220	1163	1097	1021	930

Table 4.6: *Simulation - Power 3* ($\lambda_c = 0.625, \lambda_t = 0.575$)

N_1	5	6	7	8	9	10	11	12	13	14	15
Power	0.844	0.817	0.822	0.827	0.831	0.844	0.832	0.816	0.811	0.826	0.817
K_2	1866	1858	1847	1835	1821	1806	1787	1762	1734	1696	1645

Table 4.7: *Simulation - Power 4* ($\lambda_c = 0.725, \lambda_t = 0.675$)

N_1	5	6	7	8	9	10	11	12	13	14	15
Power	0.819	0.838	0.823	0.831	0.828	0.849	0.817	0.841	0.828	0.823	0.828
K_2	2367	2397	2433	2480	2536	2603	2694	2819	2994	3264	3737

5.0 CONCLUSIONS AND FUTURE WORK

In Chapter 2, we develop a new procedure based on the *Ammeter* process for estimating the variance of the fractionator estimator obtained using uniformly systematic sampling. The assumption for our approach is that the density of the neurons is a random variable instead of being a constant. In fact, the distribution of neurons in the brain region is always non-homogenous. The Konopaske data shows the evidence of the overdispersed Poisson process. We also compare the stereological variance estimator using the *Ammeter* method with the estimators obtained by Cruz-Orive and Geiser's method and the Bootstrap method.

In Chapter 3, we consider experimental design for a stereological study. We develop procedures for planning a cost efficient study considering two types of optimization. For the standard paired studies designed for CCNMD, the research aim is to compare the number of neurons between two groups instead of evaluating individuals' neuron numbers. We provide an algorithm to find the the number of subjects and the number of counting frames that minimize the cost function while maintaining sufficient power to address the research aim. We also consider an algorithm to maximize statistical power for a fixed budget.

In Chapter 4, we introduce the idea of adaptive design to stereological studies. For the optimal designs considered in Chapter 3, we require the information about the magnitudes of the true stereological variance. Stereological variances vary from study to study. It is usually difficult to pre-specify before a study which is interested in a particular type of neurons. We apply the approach of blinded adaptive design procedures to stereological studies specially to the case of a fixed number of subjects pairs, but where we can change the number of sampling frames. We develop an approach which at the end of Stage I allows us to update the assumption about the stereological variance used in the planning stage. Based on the update, we change the number of counting frames to be used in Stage II. We propose an

adjusted t-statistic to use for hypothesis testing at the end of the study. Simulation is used to show that the type I error rate of our procedure is protected, and also appropriate power is maintained.

5.1 FUTURE WORK

5.1.1 Adjusting Both the Counting Frames and Group Size of Stage II (K_2 and N_2)

In the future, we will consider the situation that the number of pairs can also be adjusted. Both K_2 and N_2 will be allowed to depend on Stage I data. The difference $\bar{D} = \frac{1}{N}[(N_1\bar{X}_1 + N_2\bar{X}_2) - (N_1\bar{Y}_1 + N_2\bar{Y}_2)]$ then will have data dependent weights. The adjusted t-statistic we proposed in (4.21) may not be appropriate to use.

Then, we would consider an alternative test statistic, which is a combination of Z-statistics,

$$Z = \sqrt{\phi}Z_1 + \sqrt{1-\phi}Z_2, \tag{5.1}$$

where Z_1 and Z_2 are the Z-statistics for Stage I and Stage II respectively, and ϕ is a pre-specified constant. This test statistic has a normal distribution and type I error rate is guaranteed. However, the strategy to pick ϕ need to be discussed. Under the alternative, $Z \sim N(\xi, 1)$, where $\xi = \sqrt{\phi}\frac{\delta}{\sqrt{\frac{2\sigma_B + \sigma^c_{st,K_1} + \sigma^t_{st,K_1}}{N_1}}} + \sqrt{1-\phi}\frac{\delta}{\sqrt{\frac{2\sigma_B + \sigma^c_{st,K_2} + \sigma^t_{st,K_2}}{N_2}}}$.

Consider the total cost for the study:

$$\begin{aligned} Cost &= C_0 + 2C_1N + 2C_2N_1K_1 + 2C_2N_2K_2 \\ &= (C_0 + 2C_1N_1 + 2C_2N_1K_1) + 2C_1N_2 + 2C_2N_2K_2 \\ &= C_0^* + 2C_1N_2 + 2C_2N_2K_2. \end{aligned} \tag{5.2}$$

where $C_0^* = C_0 + 2C_1N_1 + 2C_2N_1K_1$. The cost of the first stage is fixed with fixed (N_1, K_1), so that to obtain an optimal design we need to optimize the second stage.

We plan to develop the strategy to pick K_2 and N_2 as part of our future research goals.

5.1.2 *Ammeter* assumption validity

We estimated the stereological variance under the *Ammeter* assumption in Chapter 2. However, the assumption's validity has to be checked. We note that the assumption of an *Ammeter* process can be compared to a Poisson model using a likelihood ratio test. To do this, we use $\alpha = 1/\phi$ in the NB distribution. Because $\alpha \to 0$ implies that the NB distribution goes to a Poisson distribution, the validity test can be represented as testing $\alpha = 0$.

5.1.3 Matched Triads Design

Dr. Konopaske uses a triads study to detect the antipsychotic exposure effect on the number of glial cells. Due to the unequal stereological variances among groups under the alternative hypothesis, to apply the adaptive procedure in designing a triads study is more complex. Furthermore, the two components of the stereological variance can't be estimated by blinded data. The unblinded approach may be considered in triads case. We will examine this situation more carefully in the future.

APPENDIX A

SYSTEMATIC DELETION

A.1 SYSTEMATIC DELETION OF DATA SET WITH 20% REDUCTION

Table A1: *Comparing the Original Data Set and 45 Samples with* 20% *Reduction*

Dataset	Estimate	StdErr	t-stat	p-value
Original	-25.0	12.94	-1.93	0.0413
Sample1	-27.9	13.55	-2.06	0.033
Sample2	-27.4	13.03	-2.10	0.031
Sample3	-26.6	13.28	-2.01	0.036
Sample4	-24.5	13.54	-1.81	0.050
Sample5	-25.6	13.60	-1.89	0.044
Sample6	-27.0	13.93	-1.93	0.041
Sample7	-27.6	14.21	-1.94	0.040
Sample8	-27.2	13.75	-1.98	0.038
Sample9	-24.3	12.65	-1.92	0.042
Sample10	-27.0	12.47	-2.16	0.028
Sample11	-26.2	12.72	-2.06	0.033
Sample12	-24.1	12.99	-1.85	0.047
Sample13	-25.2	13.07	-1.93	0.041
Sample14	-26.5	13.34	-1.99	0.037
Sample15	-27.2	13.63	-1.99	0.037
Sample16	-26.8	13.19	-2.03	0.035
Sample17	-23.9	12.07	-1.98	0.038
Sample18	-25.7	12.13	-2.12	0.030
Sample19	-23.6	12.51	-1.88	0.044
Sample20	-24.7	12.58	-1.97	0.039
Sample21	-26.0	12.79	-2.04	0.034
Sample22	-26.7	13.20	-2.02	0.035

A1: (*continued*) *Comparing the Original Data Set and 45 Samples with* 20% *Reduction.*

Dataset	Estimate	StdErr	t-stat	p-value
Sample23	-26.3	12.62	-2.09	0.032
Sample24	-23.4	11.60	-2.02	0.036
Sample25	-22.8	12.69	-1.80	0.051
Sample26	-24.0	12.83	-1.87	0.046
Sample27	-25.3	13.05	-1.94	0.041
Sample28	-25.9	13.34	-1.94	0.040
Sample29	-25.6	12.94	-1.98	0.038
Sample30	-22.6	11.82	-1.91	0.042
Sample31	-21.8	13.17	-1.66	0.064
Sample32	-23.1	13.38	-1.73	0.057
Sample33	-23.8	13.75	-1.73	0.057
Sample34	-23.4	13.19	-1.77	0.053
Sample35	-20.5	12.19	-1.68	0.062
Sample36	-24.3	13.42	-1.81	0.050
Sample37	-24.9	13.83	-1.80	0.051
Sample38	-24.6	13.30	-1.85	0.047
Sample39	-21.6	12.28	-1.76	0.054
Sample40	-26.2	14.02	-1.87	0.045
Sample41	-25.9	13.55	-1.91	0.043
Sample42	-22.9	12.40	-1.85	0.047
Sample43	-26.5	13.78	-1.92	0.042
Sample44	-23.6	12.80	-1.84	0.048
Sample45	-23.2	12.32	-1.89	0.044

Note: The degree of the t statistics is 10 and the p-value is based on a one-sided testing.

A.2 SYSTEMATIC DELETION OF DATA SET WITH 30% REDUCTION

Table A2: *Comparing the Original Data Set and 120 Samples with* 30% *Reduction*

Dataset	Estimate	StdErr	t-stat	p-value
Original	-25.0	12.94	-1.93	0.0413
Sample1	-29.2	13.11	-2.23	0.025
Sample2	-28.3	13.42	-2.11	0.031
Sample3	-25.8	13.63	-1.90	0.044
Sample4	-27.2	13.68	-1.99	0.038
Sample5	-28.7	14.14	-2.03	0.035
Sample6	-29.4	14.39	-2.04	0.034
Sample7	-29.0	13.94	-2.08	0.032
Sample8	-25.6	12.62	-2.03	0.035
Sample9	-27.7	12.72	-2.18	0.027
Sample10	-25.3	13.07	-1.93	0.041
Sample11	-26.6	13.12	-2.03	0.035
Sample12	-28.1	13.49	-2.08	0.032
Sample13	-28.8	13.89	-2.07	0.032
Sample14	-28.4	13.27	-2.14	0.029
Sample15	-25.1	12.06	-2.08	0.032
Sample16	-24.4	13.29	-1.84	0.048
Sample17	-25.7	13.42	-1.92	0.042
Sample18	-27.2	13.81	-1.97	0.038
Sample19	-27.9	14.06	-1.99	0.037
Sample20	-27.5	13.66	-2.02	0.036
Sample21	-24.2	12.34	-1.96	0.039
Sample22	-23.3	13.73	-1.70	0.060
Sample23	-24.8	14.10	-1.76	0.055
Sample24	-25.5	14.46	-1.76	0.054
Sample25	-25.1	13.86	-1.81	0.050
Sample26	-21.7	12.67	-1.71	0.059
Sample27	-26.1	14.10	-1.85	0.047
Sample28	-26.8	14.52	-1.85	0.047

Table A2: (*continued*) *Comparing the Original Data Set and 120 Samples with* 30% *Reduction.*

Dataset	Estimate	StdErr	t-stat	p-value
Sample29	-26.4	13.95	-1.89	0.044
Sample30	-23.1	12.74	-1.81	0.050
Sample31	-28.3	14.86	-1.90	0.043
Sample32	-27.9	14.37	-1.94	0.040
Sample33	-24.6	13.01	-1.89	0.044
Sample34	-28.6	14.56	-1.97	0.039
Sample35	-25.3	13.40	-1.89	0.044
Sample36	-24.9	12.90	-1.93	0.041
Sample37	-27.3	12.10	-2.25	0.024
Sample38	-24.8	12.46	-1.99	0.037
Sample39	-26.1	12.53	-2.09	0.032
Sample40	-27.6	12.82	-2.15	0.028
Sample41	-28.3	13.25	-2.14	0.029
Sample42	-27.9	12.64	-2.21	0.026
Sample43	-24.6	11.41	-2.15	0.028
Sample44	-23.9	12.68	-1.89	0.044
Sample45	-25.3	12.83	-1.97	0.039
Sample46	-26.8	13.15	-2.03	0.035
Sample47	-27.5	13.41	-2.05	0.034
Sample48	-27.1	13.04	-2.08	0.032
Sample49	-23.7	11.70	-2.03	0.035
Sample50	-22.8	13.15	-1.73	0.057
Sample51	-24.3	13.45	-1.81	0.051
Sample52	-25.0	13.82	-1.81	0.050
Sample53	-24.6	13.24	-1.86	0.046
Sample54	-21.3	12.05	-1.76	0.054
Sample55	-25.6	13.47	-1.90	0.043
Sample56	-26.3	13.90	-1.89	0.044
Sample57	-25.9	13.35	-1.94	0.040
Sample58	-22.6	12.13	-1.86	0.046
Sample59	-27.8	14.18	-1.96	0.039
Sample60	-27.4	13.71	-2.00	0.037
Sample61	-24.1	12.32	-1.96	0.040
Sample62	-28.1	13.90	-2.03	0.035
Sample63	-24.8	12.73	-1.95	0.040
Sample64	-24.4	12.25	-1.99	0.037
Sample65	-23.4	12.04	-1.94	0.040
Sample66	-24.7	12.18	-2.03	0.035
Sample67	-26.2	12.40	-2.11	0.030
Sample68	-26.9	12.83	-2.10	0.031
Sample69	-26.5	12.28	-2.16	0.028
Sample70	-23.2	11.04	-2.10	0.031

Table A2: (*continued*) *Comparing the Original Data Set and 120 Samples with* 30% *Reduction.*

Dataset	Estimate	StdErr	t-stat	p-value
Sample71	-22.2	12.65	-1.76	0.055
Sample72	-23.7	12.85	-1.85	0.047
Sample73	-24.4	13.38	-1.83	0.049
Sample74	-24.0	12.62	-1.90	0.043
Sample75	-20.7	11.55	-1.79	0.052
Sample76	-25.1	12.87	-1.95	0.040
Sample77	-25.8	13.46	-1.92	0.042
Sample78	-25.4	12.73	-1.99	0.037
Sample79	-22.0	11.64	-1.89	0.044
Sample80	-27.3	13.64	-2.00	0.037
Sample81	-26.9	13.00	-2.07	0.033
Sample82	-23.5	11.71	-2.01	0.036
Sample83	-27.6	13.34	-2.07	0.033
Sample84	-24.2	12.29	-1.97	0.038
Sample85	-23.8	11.63	-2.05	0.034
Sample86	-21.4	12.86	-1.66	0.064
Sample87	-22.9	13.09	-1.75	0.056
Sample88	-23.6	13.47	-1.75	0.055
Sample89	-23.2	12.93	-1.79	0.052
Sample90	-19.8	11.74	-1.69	0.061
Sample91	-24.2	13.19	-1.83	0.048
Sample92	-24.9	13.62	-1.83	0.049
Sample93	-24.5	13.12	-1.87	0.046
Sample94	-21.2	11.91	-1.78	0.053
Sample95	-26.4	13.82	-1.91	0.043
Sample96	-26.0	13.41	-1.94	0.041
Sample97	-22.7	12.01	-1.89	0.044
Sample98	-26.7	13.59	-1.97	0.039
Sample99	-23.4	12.42	-1.88	0.045
Sample100	-23.0	12.00	-1.91	0.042
Sample101	-21.7	13.58	-1.60	0.070
Sample102	-22.4	14.11	-1.59	0.071
Sample103	-22.0	13.41	-1.64	0.066
Sample104	-18.7	12.35	-1.51	0.081
Sample105	-23.9	14.31	-1.67	0.063
Sample106	-23.5	13.69	-1.72	0.058
Sample107	-20.2	12.45	-1.62	0.068
Sample108	-24.2	13.98	-1.73	0.057
Sample109	-20.9	12.97	-1.61	0.069
Sample110	-20.5	12.33	-1.66	0.064
Sample111	-25.3	14.33	-1.76	0.054
Sample112	-24.9	13.75	-1.81	0.050

Table A2: (*continued*) *Comparing the Original Data Set and 120 Samples with* 30% *Reduction.*

Dataset	Estimate	StdErr	t-stat	p-value
Sample113	-21.5	12.48	-1.72	0.058
Sample114	-25.6	14.10	-1.81	0.050
Sample115	-22.2	13.06	-1.70	0.060
Sample116	-21.8	12.46	-1.75	0.055
Sample117	-27.1	14.35	-1.89	0.044
Sample118	-23.7	13.13	-1.81	0.050
Sample119	-23.3	12.61	-1.85	0.047
Sample120	-24.0	12.94	-1.86	0.046

Note: The degree of the t statistics is 10 and the p-value is based on a one-sided testing.

APPENDIX B

BOOTSTRAP DELETION

B.1 BOOTSTRAP DELETION OF DATA SET WITH 10% REDUCTION

Table B1: *Comparing the Original Data Set and 50 Bootstrap Samples with* 10% *Reduction*

Dataset	Estimate	StdErr	t-stat	p-value
Original	-25.0	12.94	-1.93	0.0413
Bootstrap1	-25.0	15.12	-1.66	0.064
Bootstrap2	-28.7	12.28	-2.34	0.021
Bootstrap3	-27.5	14.25	-1.93	0.041
Bootstrap4	-29.1	13.31	-2.19	0.027
Bootstrap5	-29.0	12.74	-2.28	0.023
Bootstrap6	-23.4	10.82	-2.16	0.028
Bootstrap7	-23.6	12.88	-1.83	0.049
Bootstrap8	-29.8	15.14	-1.97	0.039
Bootstrap9	-23.4	12.04	-1.95	0.040
Bootstrap10	-26.6	13.77	-1.93	0.041
Bootstrap11	-24.0	12.70	-1.89	0.044
Bootstrap12	-23.5	10.85	-2.17	0.028
Bootstrap13	-29.7	11.07	-2.69	0.011
Bootstrap14	-25.2	11.44	-2.20	0.026
Bootstrap15	-24.5	13.38	-1.83	0.049
Bootstrap16	-29.7	14.15	-2.10	0.031
Bootstrap17	-29.6	13.90	-2.13	0.029
Bootstrap18	-23.3	13.38	-1.74	0.056
Bootstrap19	-22.2	12.04	-1.84	0.047
Bootstrap20	-26.7	14.40	-1.86	0.046
Bootstrap21	-29.7	13.84	-2.14	0.029
Bootstrap22	-28.1	12.99	-2.16	0.028

B1: (*continued*) *Comparing the Original Data Set and 50 Bootstrap Samples with* 10% *Reduction.*

Dataset	Estimate	StdErr	t-stat	p-value
Bootstrap23	-25.0	12.06	-2.07	0.033
Bootstrap24	-28.6	13.39	-2.14	0.029
Bootstrap25	-27.8	14.22	-1.95	0.040
Bootstrap26	-27.7	11.48	-2.41	0.018
Bootstrap27	-29.4	14.16	-2.07	0.032
Bootstrap28	-26.7	13.36	-2.00	0.037
Bootstrap29	-26.1	14.72	-1.78	0.053
Bootstrap30	-26.4	13.37	-1.97	0.038
Bootstrap31	-25.9	14.03	-1.85	0.047
Bootstrap32	-26.0	13.58	-1.91	0.042
Bootstrap33	-29.5	13.20	-2.23	0.025
Bootstrap34	-24.3	13.73	-1.77	0.054
Bootstrap35	-25.8	12.81	-2.02	0.036
Bootstrap36	-22.9	11.39	-2.01	0.036
Bootstrap37	-30.7	13.53	-2.27	0.023
Bootstrap38	-30.1	13.92	-2.16	0.028
Bootstrap39	-28.2	13.25	-2.13	0.030
Bootstrap40	-23.5	13.86	-1.70	0.060
Bootstrap41	-31.2	11.81	-2.64	0.012
Bootstrap42	-25.8	12.01	-2.14	0.029
Bootstrap43	-26.0	13.02	-1.99	0.037
Bootstrap44	-23.8	13.01	-1.83	0.049
Bootstrap45	-25.8	13.43	-1.92	0.042
Bootstrap46	-32.3	12.68	-2.55	0.014
Bootstrap47	-31.7	12.90	-2.45	0.017
Bootstrap48	-25.0	12.75	-1.96	0.039
Bootstrap49	-22.7	11.21	-2.02	0.035
Bootstrap50	-26.4	12.79	-2.07	0.033

Note: The degree of the t statistics is 10 and the p-value is based on a one-sided testing.

B.2 BOOTSTRAP DELETION OF DATA SET WITH 20% REDUCTION

Table B2: *Comparing the Original Data Set and 50 Bootstrap Samples with* 20% *Reduction*

Dataset	Estimate	StdErr	t-stat	p-value
Original	-25.0	12.94	-1.93	0.0413
Bootstrap1	-24.3	13.44	-1.81	0.050
Bootstrap2	-26.3	14.37	-1.83	0.048
Bootstrap3	-30.0	15.05	-1.99	0.037
Bootstrap4	-28.6	13.23	-2.16	0.028
Bootstrap5	-26.7	12.51	-2.14	0.029
Bootstrap6	-25.7	13.32	-1.93	0.041
Bootstrap7	-27.4	13.68	-2.01	0.036
Bootstrap8	-26.6	13.73	-1.93	0.041
Bootstrap9	-23.8	12.27	-1.94	0.040
Bootstrap10	-24.9	14.19	-1.75	0.055
Bootstrap11	-26.6	13.24	-2.01	0.036
Bootstrap12	-26.3	14.30	-1.84	0.048
Bootstrap13	-28.8	12.76	-2.25	0.024
Bootstrap14	-24.9	11.19	-2.22	0.025
Bootstrap15	-22.7	15.21	-1.50	0.083
Bootstrap16	-23.8	12.20	-1.95	0.040
Bootstrap17	-31.4	13.84	-2.27	0.023
Bootstrap18	-24.6	12.23	-2.01	0.036
Bootstrap19	-29.1	13.18	-2.21	0.026
Bootstrap20	-26.9	12.80	-2.10	0.031
Bootstrap21	-29.4	12.67	-2.32	0.021
Bootstrap22	-25.5	13.22	-1.93	0.041
Bootstrap23	-32.0	12.29	-2.60	0.013
Bootstrap24	-26.8	15.37	-1.75	0.056
Bootstrap25	-28.7	14.77	-1.95	0.040
Bootstrap26	-29.8	11.78	-2.53	0.015
Bootstrap27	-22.2	12.12	-1.83	0.049
Bootstrap28	-29.7	14.48	-2.05	0.034

Table B2: (*continued*) *Comparing the Original Data Set and 50 Bootstrap Samples with* 20% *Reduction.*

Dataset	Estimate	StdErr	t-stat	p-value
Bootstrap29	-25.4	13.85	-1.83	0.048
Bootstrap30	-24.2	12.03	-2.01	0.036
Bootstrap31	-23.6	11.64	-2.03	0.035
Bootstrap32	-25.0	13.07	-1.91	0.042
Bootstrap33	-28.1	12.76	-2.20	0.026
Bootstrap34	-21.5	11.68	-1.84	0.047
Bootstrap35	-21.5	13.73	-1.57	0.074
Bootstrap36	-28.4	11.28	-2.52	0.015
Bootstrap37	-29.2	13.78	-2.12	0.030
Bootstrap38	-22.7	13.06	-1.74	0.056
Bootstrap39	-25.9	14.19	-1.83	0.049
Bootstrap40	-26.6	13.38	-1.99	0.037
Bootstrap41	-22.9	12.35	-1.86	0.046
Bootstrap42	-31.0	15.08	-2.05	0.034
Bootstrap43	-27.7	12.82	-2.16	0.028
Bootstrap44	-27.2	13.88	-1.96	0.039
Bootstrap45	-23.6	15.27	-1.55	0.076
Bootstrap46	-32.1	12.99	-2.47	0.016
Bootstrap47	-30.4	12.89	-2.36	0.020
Bootstrap48	-24.7	15.29	-1.61	0.069
Bootstrap49	-27.7	12.20	-2.27	0.023
Bootstrap50	-23.0	11.72	-1.97	0.039

Note: The degree of the t statistics is 10 and the p-value is based on a one-sided testing.

B.3 BOOTSTRAP DELETION OF DATA SET WITH 30% REDUCTION

Table B3: *Comparing the Original Data Set and 50 Bootstrap Samples with* 30% *Reduction*

Dataset	Estimate	StdErr	t-stat	p-value
Original	-25.0	12.94	-1.93	0.0413
Bootstrap1	-23.9	12.06	-1.98	0.038
Bootstrap2	-27.1	13.51	-2.01	0.036
Bootstrap3	-26.8	14.73	-1.82	0.049
Bootstrap4	-21.3	13.01	-1.64	0.066
Bootstrap5	-27.0	13.07	-2.07	0.033
Bootstrap6	-23.3	10.74	-2.17	0.027
Bootstrap7	-24.9	14.06	-1.77	0.053
Bootstrap8	-26.2	13.40	-1.95	0.040
Bootstrap9	-26.3	13.42	-1.96	0.039
Bootstrap10	-31.9	14.37	-2.22	0.025
Bootstrap11	-27.0	11.82	-2.28	0.023
Bootstrap12	-30.9	14.07	-2.20	0.026
Bootstrap13	-23.7	12.83	-1.84	0.048
Bootstrap14	-25.3	10.89	-2.33	0.021
Bootstrap15	-27.7	12.62	-2.19	0.027
Bootstrap16	-23.3	12.16	-1.91	0.042
Bootstrap17	-26.8	13.98	-1.91	0.042
Bootstrap18	-30.8	14.88	-2.07	0.033
Bootstrap19	-23.5	12.36	-1.90	0.043
Bootstrap20	-29.4	13.92	-2.11	0.031
Bootstrap21	-35.3	12.76	-2.77	0.010
Bootstrap22	-26.8	14.03	-1.91	0.043
Bootstrap23	-24.7	13.54	-1.82	0.049
Bootstrap24	-26.6	12.69	-2.10	0.031
Bootstrap25	-31.2	15.66	-1.99	0.037
Bootstrap26	-24.7	12.80	-1.93	0.041
Bootstrap27	-24.6	11.97	-2.05	0.034
Bootstrap28	-35.2	16.57	-2.12	0.030

Table B3: (*continued*) *Comparing the Original Data Set and 50 Bootstrap Samples with* 30% *Reduction.*

Dataset	Estimate	StdErr	t-stat	p-value
Bootstrap29	-26.5	12.34	-2.14	0.029
Bootstrap30	-27.0	12.56	-2.15	0.029
Bootstrap31	-25.7	12.65	-2.03	0.035
Bootstrap32	-19.1	12.06	-1.59	0.072
Bootstrap33	-25.5	14.16	-1.80	0.051
Bootstrap34	-32.1	14.51	-2.21	0.026
Bootstrap35	-24.7	11.74	-2.10	0.031
Bootstrap36	-25.6	13.31	-1.93	0.042
Bootstrap37	-24.6	12.71	-1.94	0.041
Bootstrap38	-27.7	12.89	-2.15	0.029
Bootstrap39	-22.5	13.08	-1.72	0.058
Bootstrap40	-26.4	12.98	-2.04	0.035
Bootstrap41	-31.2	12.69	-2.46	0.017
Bootstrap42	-18.9	14.43	-1.31	0.110
Bootstrap43	-29.2	12.04	-2.43	0.018
Bootstrap44	-29.0	14.91	-1.94	0.040
Bootstrap45	-26.0	14.10	-1.84	0.048
Bootstrap46	-20.7	12.21	-1.69	0.061
Bootstrap47	-27.0	13.08	-2.06	0.033
Bootstrap48	-26.7	11.15	-2.40	0.019
Bootstrap49	-27.8	13.06	-2.13	0.030
Bootstrap50	-29.9	15.49	-1.93	0.041

Note: The degree of the t statistics is 10 and the p-value is based on a one-sided testing.

BIBLIOGRAPHY

Baddeley, A., and Vedel Jensen, E. (2005), *Stereology for Statisticians*, London: Chapman & Hall/CRC.

Cochran, W. (1977), *Sampling Techniques, Third Edition*, New York: John Wiley and Sons.

Correa, C. (2001), "Statistical Analysis of Stereological Estimators," , .

Cruz-Orive, L. M. (1980), "Best linear unbiased estimators for stereology," *Biometrics*, 36, 595–605.

Cruz-Orive, L. M. (1999), "Precision of Cavalieri sections and slices with local errors," *Journal of Microscopy*, 193(3), 182–198.

Cruz-Orive, L. M., and Geiser, M. (2004), "Estimation of Particle Number by Stereology: An Update," *Journal of Aerosol Medicine*, 17(3), 197–212.

Cruz-Orive, L. M., and Weibel, E. (1981), "Sampling designs for stereology," *Journal of Microscopy*, 122, 235–257.

Dorph-Petersen, K.-A., Nyengaard, J. R., and Gundersen, H. J. G. (2001), "Tissue Shrinkage and Unbiased Stereological Estimation of Particle Number and Size," *Journal of Microscopy*, 204(3), 232–246.

Dorph-Petersen, K.-A., Pierri, J. N., Wu, Q., Sampson, A. R., and Lewis, D. A. (2007), "Primary visual cortex volume and total neuron number are reduced in schizophrenia," *The Journal of Comparative Neurology*, 501(2), 290–301.

Gould, A. L., and Shih, W. J. (1998), "Modifying the design of ongoing trials without unblinding," *Statistics in Medicine*, 17, 89–100.

Gould, A., and Shih, W. (1991), "Sample size re-estimation without unblinding for normally distributed outcomes with unknown variance," *Communications in Statistics-Theory and Methods*, 21, 2833–2853.

Grandell, J. (1997), *Mixed Poisson Proecess*, London: Chapman & Hall.

Gundersen, H., and Jensen, E. (1987), "The efficiency of systematic sampling in stereology and its prediction," *Journal of Microscopy*, 147, 229–263.

Gundersen, H., Jensen, E., Kiêu, K., and Nielsen, J. (1999), "The efficiency of systematic sampling in stereology - reconsidered," *Journal of Microscopy*, 193, 199–211.

Gundersen, H., and Osterby, R. (1981), "Optimizing sampling efficiency of stereological studies in biology: or 'do more less well!'," *Journal of Microscopy*, 121(1), 65–73.

Jensen, E. (1987), "Design and model-based stereological analysis of arbitrarily shaped particles," *Scandinacian Journal of Statistics*, 14, 161–180.

Jensen, E., and Sunberg, R. (1986), "Statistical models for stereological inference about spatial structures: on the applicability of best linear unbiased estimators in stereology," *Biometrika*, 42, 735–751.

Kieser, M., and Friede, T. (2001), "A comparision of methods for estimating sample size adjustment," *Statistics in Medicine*, 20, 3861–3873.

Konopaske, G. T., Dorph-Petersen, K.-A., Pierri, J. N., Wu, Q., Sampson, A. R., and Lewis, D. A. (2007), "Effect of chronic exposure to antipsychotic medication on cell numbers in the parietal cortex of macaque monkeys," *Neuropsychopharmacology*, 32, 1216–1223.

Konopaske, G. T., Dorph-Petersen, K.-A., Sweet, R., Pierri, J., Zhang, W., Sampson, A., and Lewis, D. (2008), "Effect of chronic antipsychotic exposure on astrocyte and oligodendrocyte numbers in macaque monkeys," *Biological Psychiatry*, 63(8), 759–765.

Liu, Q., and Chi, G. Y. H. (2001), "On sample size and inference for two-stage adaptive designs," *Biometrics*, 57, 172–177.

Liu, X., and Liang, K. (1992), "Efficacy of Repeated Measures in Regression Models with Measurement Error," *Biometrics*, 48(2), 645–654.

Mentre, F., Mallet, A., and Baccar, D. (1997), "Optimal design in random-effects regression models," *Biometrika*, 84(2), 429–442.

Miles, R., and Davy, P. (1976), "Precse and general conditions for the validity of a comprehensive set of stereological fundamental formulae," *Journal of Microscopy*, 107, 211–226.

Proschan, M. A., and Hunsberger, S. A. (1995), "Designed extension of studies based on conditional power," *Biometrics*, 51, 315–1324.

Schmitz, C., and Hof, P. (2007), *Design-Based Stereology in Brain Aging Research in Brain Aging: Models, Methods, and Mechanisms* Boca Raton, FL: CRC Press, Taylor & Francis Group, LLC.

Shun, Z. (2001), "Sample size reestimation in clinical trials," *Drug Information Journal*, 35, 1409–1422.

Stein, C. (1945), "A two-sample test for a linear hypothesis whose power is independent of the variance," *Annual of Nathematical Statistics*, 16, 243–258.

Vickers, A. J. (2003), "How many repeated measures in repeated measures designs? Statistical issues for comparative trials," *BMC Medical Research Methodology*, 3(1), 22.

West, M.J., S. L., and Gundersen, H. . (1991), "Unbiased stereological estimation of the total number of neurons in the subdivisions of the rat hippocampus using the optical fractionator," *The Anatomical Record*, 231, 482–497.