



## Basics of Neurostatistics for Research, Appraisal and Application of Scientific Information to Patient Care.

Kengo Nathan Ezie<sup>1</sup>, Tatsadjieu Ngoune Léopoldine Sybile<sup>1</sup>, Misté Zourmba Ines<sup>1</sup>, Aminatou Dadda<sup>1</sup>, Diele Modeste<sup>1</sup>, Tsilla Nsegue Marie Jeanne Annick<sup>1</sup>, Yada Dirdi Gilbert<sup>1</sup>, Djamilatou Oumarou<sup>1</sup>, Berjo Dongmo Takoutsing<sup>2</sup>, MD, Ignatius N Esene<sup>3</sup>, MD, M.Sc, PhD, MPH

1. Faculty of Medicine and Biomedical Sciences, University of Garoua, Cameroon
2. Department of Research, Association of Future African Neurosurgeons, Yaounde, Cameroon
3. Neurosurgery Division, Faculty of Health Sciences, University of Bamenda, Bambili, Cameroon

**Corresponding Author: Ignatius N Esene**, MD, M.Sc, PhD, MPH, Neurosurgery Division, Faculty of Health Sciences, University of Bamenda, Bambili, Cameroon, P.Box 812, Bamenda, NWR, Cameroon.

**Copy Right:** © 2022 Ignatius N Esene, This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Received Date: August 20, 2022**

**Published Date: September 01, 2022**

**Abstract**

*Every neurosurgeon ought to be equipped with basic notions of statistics to comprehend the research process and implement evidence-based patient care. In this succinct review, we present an overview of frequently encountered notions in neurostatistic with illustrative examples related to neurosurgery where applicable. Herein we have succinctly defined the notions of population, sample, probability, sampling, statistical inference, study power, sample size and summarize statistical tests commonly encountered in neurosurgery literature.*

*Neurosurgeons reading articles or conducting research ought to know the fundamentals provided in this review for better appraisal and application of scientific information to patient care.*

**Key Words:** *Data, Probability, Neurostatistics, Neurosurgery, Research methods, Statistical Inference.*

**Introduction**

The recent rise of complex statistics in research continues to drive the demand for neurosurgeons to be competent and confident in core statistical concepts for critical appraisal and understanding of medical literature. Furthermore, those committed to research in neurosurgery require a basic statistics understanding for effective translation of research into evidence-based practice (2). In this topic review, we outline foundational concepts in neurostatistics with illustrative examples related to neurosurgery where applicable, including case studies and an example database of patients with lumbar disc herniation (Table 1)

**Basic Concepts in Neurostatistics (2, 4, 7-10, 12, 14, 16, 18)**

Neurostatistics is defined as the science of collecting, summarizing, presenting and interpreting data to draw conclusions about neurological diseases. Great variability exists between individuals with neurologic diseases and same cause does not always produce the same effect, thus conclusions drawn from studies are often uncertain. Neurostatistics seeks to explain these variabilities and uncertainties and make inference of sample results to a wider population.

An inseparable link exists between statistics and epidemiology and an understanding of the basics of research methods forms the foundation for which these notions are built. Commonly encountered research designs in neurosurgery might be descriptive and/or analytic and/or integrative (11). The importance stems from the fact that different statistical tests are used to determine effect sizes measured for each research design.

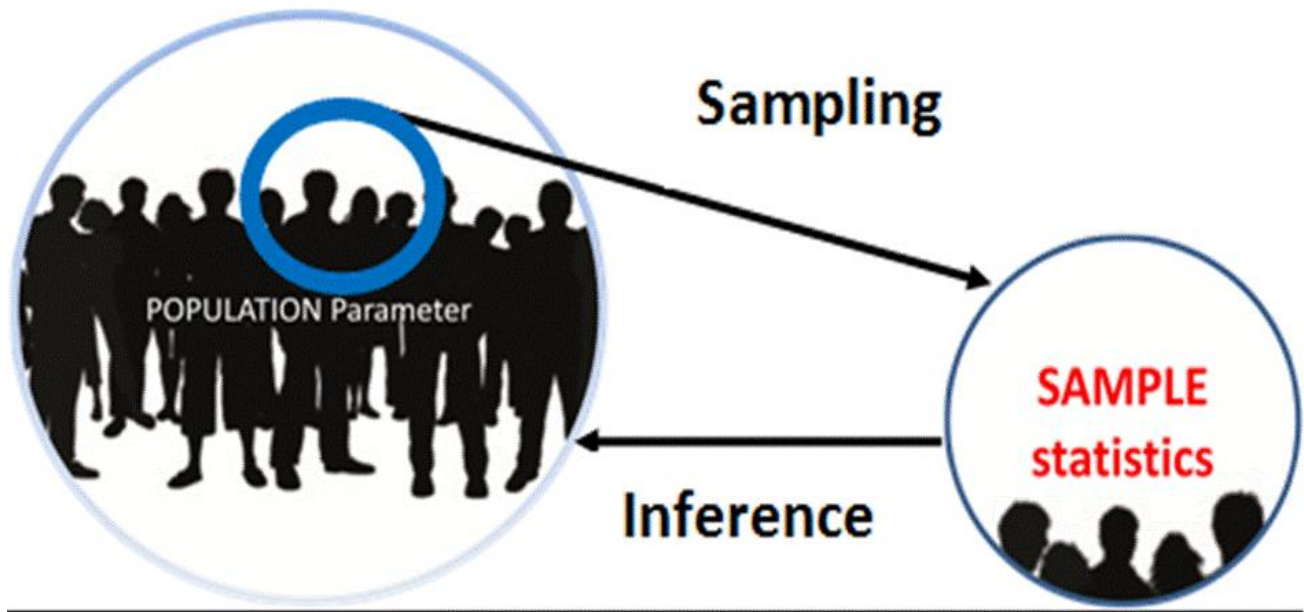
**Basics of Neurostatistics for Research, Appraisal and Application of Scientific Information to Patient Care.**

Case N°	A G E	SE X	Pre-op Pain (sciatica) Severity (VAS)	Abnormal Medial Hamstring Reflex	Disc Level	Treatment	Outcome (sciatica) (VAS)	Complicati on
1	31	m	8	Yes	L45	microsurgery	1	yes
2	35	m	6	no	L34	microsurgery	7	no
3	38	m	9	yes	L45	microsurgery	1	yes
4	31	m	10	yes	L45	microsurgery	1	yes
5	31	f	10	yes	L5S1	microsurgery	1	yes
6	41	f	10	no	L5S1	microsurgery	0	yes
7	42	f	8	yes	L45	microsurgery	1	no
8	42	f	7	yes	L45	microsurgery	3	no
9	30	m	9	yes	L45	microsurgery	1	no
10	40	m	7	yes	L45	microsurgery	1	yes
11	37	m	9	yes	L45	microsurgery	5	yes
12	35	m	8	yes	L5S1	microsurgery	2	no
13	40	m	9	yes	L45	conservative	1	no
14	21	m	10	no	L12	conservative	6	no
15	45	m	7	yes	L5S1	conservative	1	no
16	57	f	10	no	L34	conservative	0	yes
17	48	m	8	no	L45	conservative	1	no
18	75	m	7	yes	L45	conservative	0	yes
19	54	m	8	no	L5S1	conservative	1	no
20	26	m	9	no	L45	conservative	0	yes

**Table 1** Illustrative database of patients with herniated lumbar disc

**Concept of Population, Sample and Sample Design (1, 4, 9, 10, 17, 17)**

The Population is the entire group of individuals that we want information about. The Sample is part of the population that is studied. Conclusions about the population are drawn based on data from the sample (Figure 1).

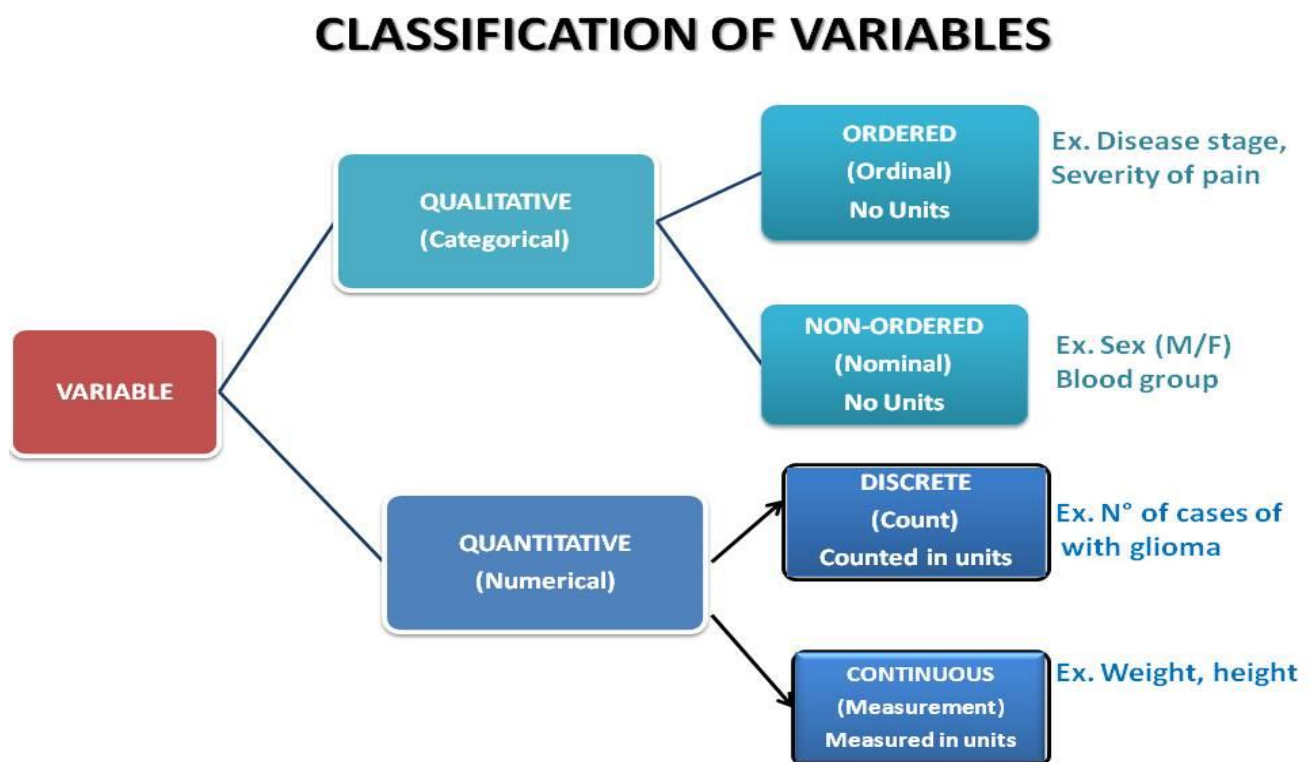


**Figure 1:** Relationship between population and sample

**Sample Design:** Methods used to choose a sample from the population. Bad sample designs (e.g.: Voluntary response sampling and convenience sampling) lead to bias. Bias can occur from the manner in which subjects are selected into the study and the way information is obtained, reported or interpreted. A biased sample systematically favors certain outcomes and is not representative of the entire population. For example, if for the patients in Table 1, the doctor decides without randomization which patients go for either conservative treatment or surgery, then there will be selection bias. Information bias another category of bias might occur if the measurement of information (e.g pain severity) on exposure or outcome is different amongst the two treatment groups. Bias unlike chance and confounding cannot be quantified and should be minimized during study design and conduct as this cannot be corrected during the analytic phase of research(12). Most common in neurosurgery is “expertise bias” a major problem with analytic studies where surgeons involved in the study might have a high level of expertise with one procedure (usually the standard or old procedure), but only limited experience with the other procedure investigated – usually the novel procedure (3).

Simple Randomized Sampling (SRS) is the standard for which every set of individuals has an equal chance to be selected from the population. SRS is unbiased and independent. For example, if the patients are randomly allocated to either conservative treatment or microdiscectomy then each patient has an equal chance of getting either treatment modality. Three simple steps to choose a SRS are: Sampling Frame (an exhaustive list of all individuals in the population), Numerical Labeling of individuals in the population and Use of Tables of Random Digits to select labels at random. There exists computer softwares for choosing SRS.

Data Analysis are methods for organizing and describing data to help reveal the information they contain and include Picturing distributions with Graphs and describing distributions with Numbers (called Statistics). Individuals are objects described by a set of data such as patients with gliomas. Variables are any characteristics of an individual such as age and sex. Variables are classified as shown in Figure 2.



**Figure 2:** Types of Variables (Modified from David Bowers,2008) (5).

**Frequency Distribution.** It tells us about variables and how many times (Count=Frequency) they occur. It might include the Relative Frequency (a proportion of individuals having a given value of a variable) and the Cumulative Frequency and Cumulative Relative Frequency. Categorical variables

are often divided into class intervals. As a rule: if the sample size is less than 100 then divide into 5-10 classes and if greater than 100 select 10-20 classes.

### **Picturing Distributions with Graphs (4, 9, 15).**

Diagrams are powerful tools for conveying information about the data, for providing simple summary pictures, and for spotting outliers and trends before any formal analyses are performed. The type of graph used depends on the kind of variable (Quantitative or Qualitative).

For Qualitative variables, use Bar charts and Pie Charts (Area of sector proportional to the frequency).

For Quantitative variables, use Stem-and-leaf plots (if no available computer) or Histograms (Box 1).

Other types of graphs include Dot plot (very simple to draw, but can be cumbersome with large data sets), Scatter plot and Box plot (see Figure 3). Tables too are often used to summarize data.

### **Five Key elements to be looked for in interpreting a histogram:**

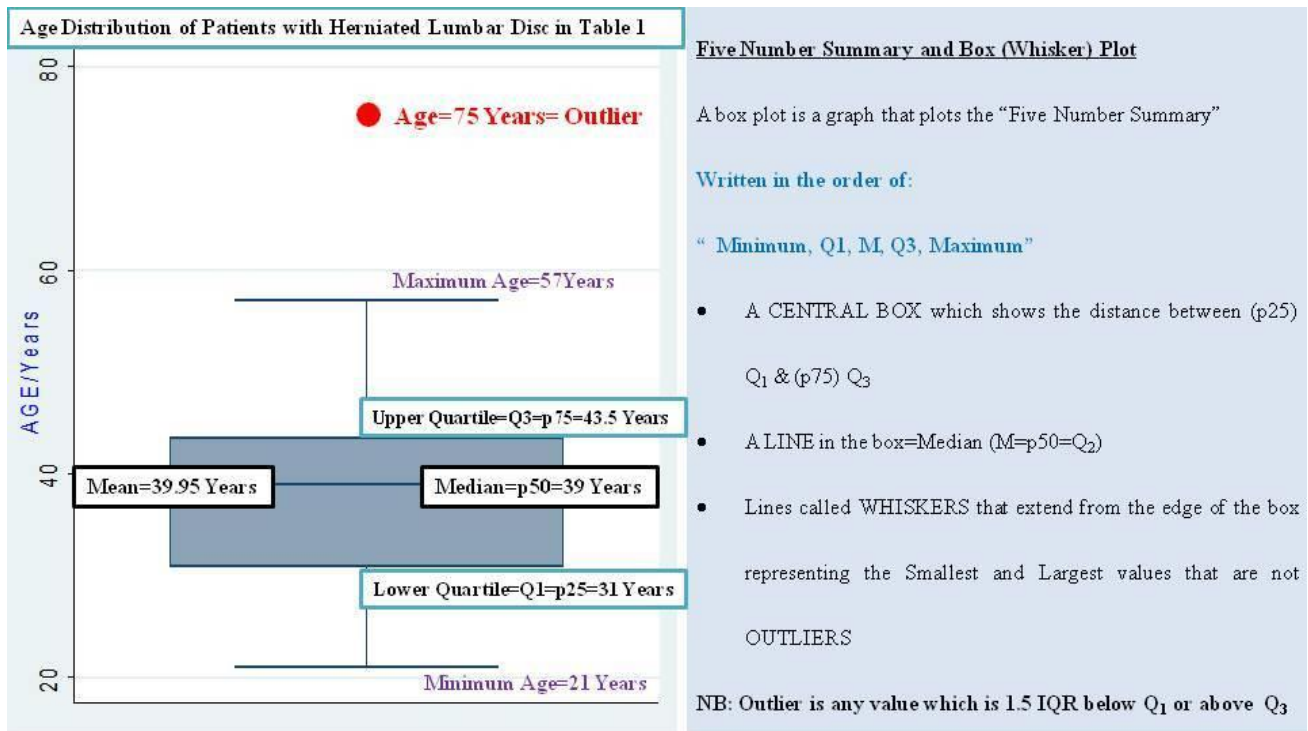
1. Shape: Symmetrical or asymmetrical (skewed),
2. Center (Midpoint of distribution),
3. Peak (Uniform, Uni, Bi or Multimodal),
4. Spread (Minimum↔Maximum values), and
5. Outliers.

### **Box 1: Interpretation of a Histogram**

#### **Summary Statistics.**

Every set of measurements has two important characteristics: Position & Spread.

Position (Central tendency statistics): Include three main statistics viz.: Mode (most frequent value amongst the observations), Mean (average) and Median (middle observation in ordered data=50th centile= 2nd quartile = 5th decile). From Table 1; for age, mode is 31 years (occurs in three patients), mean is 39.95 and median age is 39 years. Notice the equality of mean and median because of the symmetrical distribution of age in our sample.



(Data from Esene IN et al. Diagnostic performance of the medial hamstring reflex in L5 radiculopathy. Surg Neurol Int 2012; 3:104)

**Figure 3: Five Summary Statistics**

Spread or Scatter (Dispersion statistics): Include the Range, Interquartile range, Variance, standard deviation and coefficient of variation.

Range: Difference between the highest (max) and lowest (min) values in a data set. The range is often presented as the two extreme values rather than their difference i.e. Range= [Min Max].E.g. from Table 1 age range= [min=21 and max. = 57 years]. Can be misleading if outliers such as the patient with age 75 years in Figure 3.

Interquartile Range (IQR): Difference between the 1st (Q1) and 3rd (Q3) quartile. IQR measures spread about the median. IQR=Q3-Q1. IQR is not affected by outliers!

From Table 1, for age of patient, Q1(p25)=31 years and Q3(p75)= 43.5 years. IQR=Q3-Q1=12.5 years.

Also known as midsread or middle fifty, the IQR is often used during statistical analysis to find outliers in data. Outliers are observations that fall below Q1 - 1.5(IQR) or above Q3 + 1.5(IQR). In Figure 3 above, the patient with age 75 years is an outlier.

**Variance (S<sup>2</sup>):** Average distance of the values from the mean. Less used in practice.

**Standard Deviation (SD):** Square root of variance.  $SD = \sqrt{S^2}$ . In routine statistics, SD are usually presented with means. An Empirical Rule: For nearly Symmetrical, Unimodal distributions,  $\approx 68\%$  of the data are within one SD of the mean and  $95\%$  within 2 SD of the mean.

**Coefficient of Variation (CV):** It is used to compare variability of values measured in different units as illustrated in Example 1.

### **Example 1**

In the data below CSF protein content is measured in mg/dl and BMI in  $\text{Kg/m}^2$

If for CSF Protein: Mean=187.6 mg/dl and  $S=21.9$  mg/dl and

For BMI: Mean=26.0  $\text{Kg/m}^2$  and  $S=4.1$   $\text{Kg/m}^2$

$CV = S/\bar{x}$  For CSF Protein=0.12 and BMI=0.16.

Therefore, the variability of BMI is greater than that of CSF protein for that sample.

### **Relationship between Mean, Median and Mode:**

For symmetrical distributions: Mean( $\bar{x}$ ) = Median( $\check{X}$ ) = Mode (M). The mean is the most used because more statistical methods have been developed for it. Statistical tests for symmetrical distributions are called parametric tests.

### **Example 2**

Consider the data below (mass in kilograms of 5 patients from Table 1):

Data 1: 49 55 63 70 75 Mean=62.4 kg Median= 63 Kg

Data 2: 49 55 63 70 100 Mean=67.4 Kg Median=63 Kg

For asymmetrical distributions: Mean is influenced by large values towards the tail (skewed side). Median almost the same and the mode the same (Example 2). The median is therefore used because it is less sensitive to extreme values. Statistical tests for skewed distributions are called non-parametric tests.

**Probability and Probability Distribution(1, 4).**

Probability and Probability Distribution: The theory of probability enables us to link “samples” and “populations” and to draw conclusions about the population from the samples. Every neurosurgeon is recommended to be acquainted with the theory and laws of probability as they form the bases of real-life experiments. Just as for frequency distribution of sample data, probability distributions can also be graphed and summarized in terms of midpoint and spread. For the sake of simplicity, there exist three important probability distributions: Binomial and Poisson distribution for discrete data and Gaussian (Normal) distribution for random continuous data.

In statistical inference, the Binomial distribution is the model for counts and proportions and the Poisson distribution the mathematical model for rates. The reader is not obliged to know the mathematical intricacies involved in these distributions but should have an idea of their applications. For example statistical tests used to calculate the number of, or proportion of cases with/without a disease are based on binomial models whereas those that compute say the monthly rate of new cases of gliomas are done on a background of the Poisson distribution.

The Gaussian distribution is the fundamental probability distribution in statistics as it is the standard model that fits many observed frequency distributions and because it occupies the central place in sampling theory thus statistical inference.

**Statistical Inference: Confidence Interval and Statistical Significance(4, 8, 9, 15)**

Statistical Inference is the process of drawing conclusions from a sample to a population. Statistics is a numerical quantity, which describes some characteristic of a Sample. A Parameter is a numerical quantity that describes some characteristic of a population. Figure 1 illustrates the relationship between a sample and population, and the notions of sampling and inference.

When a study is conducted, the population “parameter” is not known. Statistics is used to estimate the unknown parameter through the process of inference (Figure 4).

Two common types of statistical inferences are used: Confidence Interval (CI) [used to estimate the population parameter e.g. Mean ( $\mu$ ).] and Test of Significance [Used to assess whether a hypothesis made about a population parameter is true or not].

In statistics, confidence interval (CI) is a type of “interval estimate” of a population parameter and is used to indicate the reliability of an estimate. CIs consist of a range of values (interval) that act as good

estimates of the unknown population parameter. When a study is conducted, the sample data gives us the “Point Estimate” whose precision is the “Confidence Interval”. That is; the mean of a sample can be used to estimate the mean of the population from which it was drawn, with a known degree of precision and confidence(10).

Level of Significance ( $\alpha$ )	10%=0.1	5%=0.05	1%=0.01	0.1%=0.001
Confidence Level ( $1-\alpha$ )	90%	95%	99%	99.9%
Critical Value [ $Z_{(1-\alpha)}$ ]	1.645	1.96	2.575	3.29

$(1-\alpha) \% CI = \bar{x} \pm Z_{(1-\alpha)} (\delta/\sqrt{n})$

e.g. for 95%  $CI = \bar{x} \pm 1.96 (\delta/\sqrt{n})$

$\alpha$ =Level of Significance e.g. 0.05, 0.01

$(1-\alpha)$ = Confidence Level e.g. 95%

$\bar{x}$  = Point Estimate

$\pm Z_{(1-\alpha)}$  = Critical Value

$(\delta/\sqrt{n})$  = Standard Error

$Z_{(1-\alpha)} (\delta/\sqrt{n})$  = Margin of Error

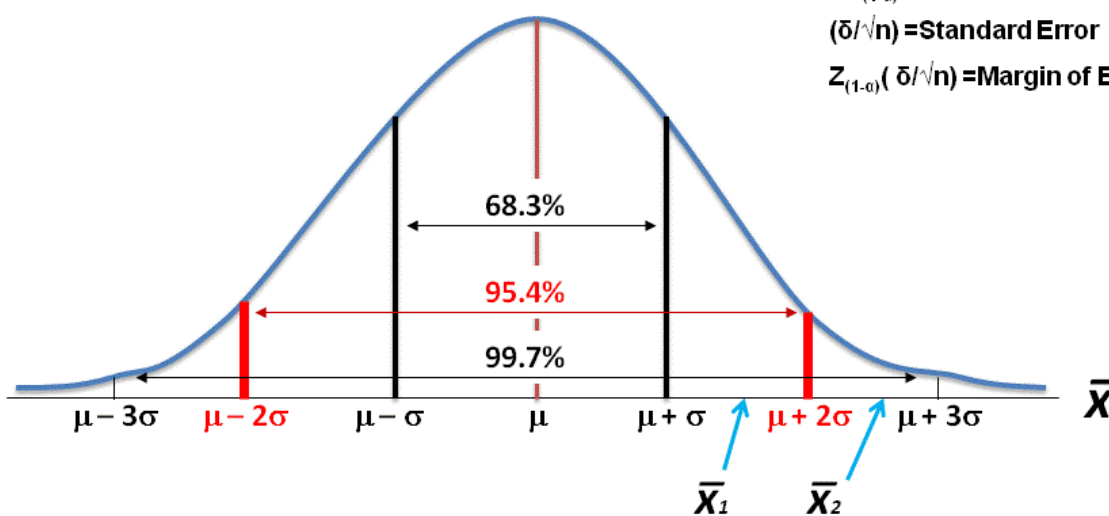


Figure 4: Normal Distribution illustrating Notions of Confidence Interval and P-Value

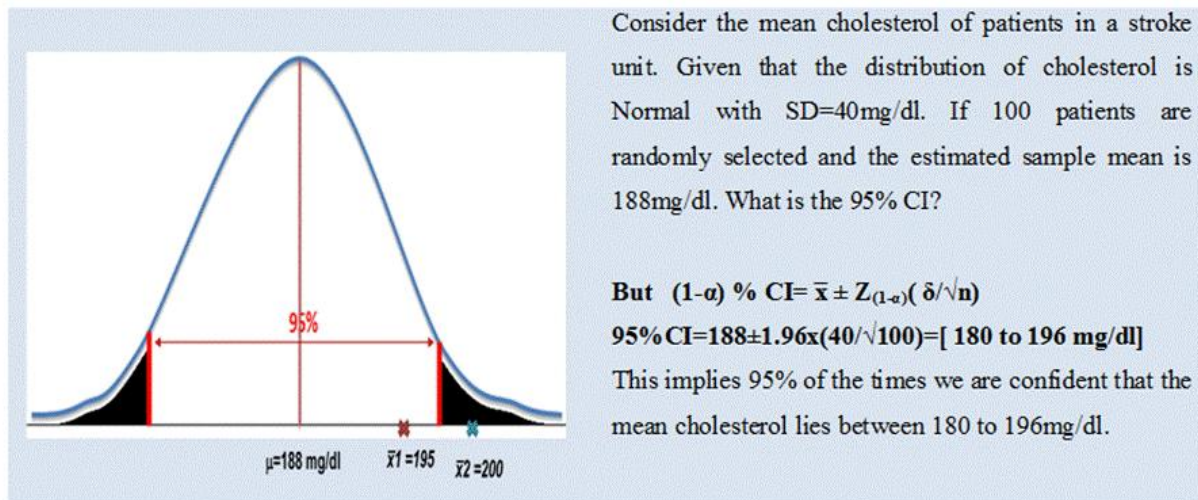
The normal distribution in Figure 4 describes entirely the number of standard deviations (□) that values are away from the mean (□). Values can lie  $\pm 1SD$ ,  $\pm 2SD$ ,  $\pm 3SD$  etc from the mean. The most commonly used in scientific literature is the 95% CI but this Confidence Level ( $1-\alpha$ ) can be changed by changing the critical values ( $Z(1-\alpha)$ ). So we can also have a 68% CI (Means between 1 SD from Point Estimate) and 99% CI (Means will lie between 3SD from point estimate) (Figure 4).

“95% CI” (read as the 95% confidence interval) means 95% of all possible samples will have a mean ( $\bar{x}$ ) lying between 2 SD from the point estimate. It means there is a 95% chance that these limits will capture the true population mean ( $\mu$ ) i.e. we are 95% confident that the right (true) population means lies within these limits. But this also means there is a 5% chance of not having the mean within this interval.

From Table 7 the 95% CI of the mean age for patients treated conservatively [95% CI (31.3 60.2)] is wider (less precision) than for patients treated surgically [95% CI (33.2 40)].

The margin of error indicates how accurate the estimate or how close the CI pins down the true value of the parameter. It varies inversely with the sample size (n).

In order to have High Confidence (Narrow CI) and small margin of error, the sample size must be adequately large. The sample size must be planned before the data is collected. Figure 5 is an illustrative example of the concept of confidence interval.



**Figure 5:** Illustrative example on Confidence Interval (and Example 3)

**Basics of Test of Significance and P-Value.** Every research seeks to answer a question. This starts with a hypothesis called the Null Hypothesis (H0) which is a statement of “No Difference” or “No Effect” and refers to the population parameters. It is the supposition that the effect which we are checking for does not exist. e.g. when comparing treatment to no-treatment, H0 = treatment has “no effect”

H0:  $\mu = \mu_0 \rightarrow$  “no difference” between the population mean ( $\mu$ ) and a given sample value ( $\mu_0$ ). The Alternative hypothesis will be: Ha:  $\mu$  “not” equal to  $\mu_0$  (“there is a difference”). The idea of significance test is to ask if the data gives any evidence against the H0.

Summarily, the null hypothesis (H0) is always a claim about a population parameter we seek evidence “against” and the alternative or research hypothesis (Ha) is the claim about the population parameter we seek evidence “for”.

“If the sample mean falls outside the confidence limits (e.g. outside the 95% CI), in the area of rejection (i.e.  $\alpha=5\%$ ), the null hypothesis is rejected, and the alternative hypothesis is accepted”(10).

**P-Value= Probability Value:** The p-value of a statistical hypothesis test is the probability of getting a value of the test statistic as extreme as or more extreme than that observed by chance alone, if the null hypothesis  $H_0$ , is true. It is the probability of wrongly rejecting the null hypothesis if it is in fact true. It is equal to the significance level of the test for which we would just reject the null hypothesis. In other words, it is the probability of having observed our data (or more extreme data) by chance when  $H_0$  is true ( $H_0: \mu = \mu_0$ ).

The p-value is compared with the actual significance level of our test and, if it is smaller, the result is “statistically” significant. That is, if the null hypothesis were to be rejected at the 5% significance level, this would be reported as " $p < 0.05$ " and as  $p < 0.01$  if the significance level is 1% (Figure 4).

Small p-values suggest that the null hypothesis is unlikely to be true. The smaller it is, the more convincing is the rejection of the null hypothesis. It indicates the strength of evidence for say, rejecting the null hypothesis  $H_0$ , rather than simply concluding "Reject  $H_0$ " or "Do not reject  $H_0$ ".

#### Example 4

Consider two samples of stroke patients with Mean cholesterol  $\bar{x}_1 = 195$  mg/dl and  $\bar{x}_2 = 200$ mg/dl

Our Null Hypothesis  $H_0: \mu = \mu_0$

If we set our Level of Significance at  $\alpha=5\%$

Then our Confidence level  $1-\alpha=100-5=95\%$

Using Z-statistics,  $Z = \frac{\bar{x} - \mu}{\delta/\sqrt{n}}$  where Z= Normal Standard Deviate

- For  $\bar{x}_1 = 195$  mg/dl,  $Z = \frac{188 - 195}{40/\sqrt{100}} = \pm 1.75$  & P-Value (from standard Z-tables)= 0.08
- $H_0: \mu = \mu_0 \leftrightarrow H_0: \mu = 195$  mg/dl. P-Value=0.08 is  $> 0.05$  (Not Statistically significant) thus we DO NOT reject the Null Hypothesis (NRH0)
- For  $\bar{x}_2 = 200$ mg/dl,  $Z = \frac{188 - 200}{40/\sqrt{100}} = \pm 3$  and P-Value= 0.0026
- $H_0: \mu = \mu_0 \leftrightarrow H_0: \mu = 200$  mg/dl P-Value=0.00026 which is  $< 0.05$  (Statistically significant) thus we reject the Null Hypothesis that the population mean equals 200mg/dl (RH0)

Note that the 95%CI= [180 to 196 mg/dl]. (from Example 3)

$\bar{x}_1 = 195$  lies within the confidence limits but  $\bar{x}_2 = 200$  lies out of the 95% confidence limit.

When we reject the Null hypothesis we accept the alternative hypothesis and vice versa.

In Figure 5, the P-Value represents the area under the tail below and above the 95% CI bringing us to the notion of two-tailed versus one-tailed P-value.

**One-tailed versus two-tailed tests:** When we say the null hypothesis;  $H_0: \mu = \mu_0 \rightarrow$  “no difference” between the population mean ( $\mu$ ) and a given sample value ( $\mu_0$ ). It implies that when we reject the null hypothesis, the population mean ( $\mu$ ) can be either  $\mu > \mu_0$  or  $\mu < \mu_0$  (i.e.  $\mu \neq 0$ ) (P2 in Table 7). This is two-tailed testing and in medicine corresponds to testing for example whether the treatment results in outcomes are different from chance (either better or worse). Often one is generally interested in whether a treatment results in outcomes that are better than “chance”, and thus uses a one-tailed test. In the latter, the p-value represents the area in a single tail of the distribution. The alternative hypothesis says what direction the parameter differs from the null value [i.e.  $H_a: \mu < \mu_0$  implies p-value in the lower tail (P1 in Table 7) and  $H_a: \mu > \mu_0$  implies p-value in the upper tail (P3 in Table 7)].

In general, a two-tailed test is preferred unless one has a specific reason to make a one-tailed test in which case the decision whether a test will be one- or two-sided must be made before data analysis and must always be justified.

If the P-value is “small”, then the data gives evidence against  $H_0 \rightarrow$  Test is statistically significant  $\rightarrow R_{H_0}$ . If the P-value is “large”, the data does NOT give evidence for  $H_0 \rightarrow$  Test is NOT statistically significant  $\rightarrow N_{R_{H_0}}$  ( $H_0$  cannot be rejected). But what is “Small” or “Large” p-value?

Statisticians usually use a Cut-off value of 0.05. If P-Value is  $< 0.05$ , we reject the Null Hypothesis but If the P-Value  $\geq 0.05$ , we do NOT reject the Null Hypothesis.

The cut-off chosen for P-Value in order to  $R_{H_0}$  is called the “Significance Level” ( $\alpha$ ) which is usually 0.05 but can be 0.01 or 0.001. Many formulae exist for the manual calculation of P-values (e.g. Using Z-statistics, t-statistics) but this is usually given automatically from statistical Softwares. The significance level chosen depends on how much evidence we require from our data to reject the null hypothesis.

P-Value  $< 0.05 \rightarrow$  Reasonable Evidence against Null Hypothesis=Significant

P-Value  $< 0.01 \rightarrow$  Strong Evidence against  $H_0$ = Highly Significant

P-Value  $< 0.001 \rightarrow$  Very Strong Evidence against  $H_0$ =Very Highly Significant

**Power of a test (4, 8, 12) and Sample Size (2):**

The power of a statistical hypothesis test measures the test's ability to reject the null hypothesis when it is actually false - that is, to make a correct decision (or the probability of finding an effect when an effect actually exists). In other words, the power of a hypothesis test is the probability of not committing a type II error (Table 8) . It is calculated by subtracting the probability of a type II error from 1. The maximum power a test can have is 1, the minimum is 0. Ideally we want a test to have high power, close to 1. It is usually expected to be  $\geq 80\%$ . The higher the power, the more sensitive the test.

**Sample Size (2):** The power of study and sample size are entities that should be determined before the study is conducted to achieve the goal of the study. The importance of sample size stems from the fact that if the study is too small then it is unlikely that a small but clinically important difference will be detected and if too large, then time and resources will be wasted. The required sample size for a study depends on four criteria(6): The minimum clinically relevant effect that one wishes to detect for the endpoint of interest, the required power of the study (usually set at 80% or 90%), the required significant level (usually set at 5%) and the variance of the endpoint or outcome of interest.

Different formulae exist for sample size calculation depending on the study design and the endpoint under the study. Neurosurgeons are advised to always confirm their sample size calculation from professional statisticians as the consequences of wrong sample size calculation can't be neglected.

**Data Presentation**

To glean any meaningful information from the raw data obtained in any study, they must be meticulously analyzed, interpreted and clearly presented. Common softwares for statistical analysis: Microsoft Excel, Epi Infos, SPSS (Statistical Package for the Social Sciences) and STATA.

Once the raw data are collected, they are usually passed on to a statistician (usually not a neurosurgeon) who does the cleaning and analysis. The neurosurgeon must be thus equipped with basic skills to discern, extract and interpret the results. Table 6 shows some tests commonly encountered in neurosurgery.

A wide variety of tests of statistical significance (1, 4, 8-10, 12, 13, 15) are available which even though may be daunting at first sight can be easily understood if studied systematically. Two-thirds of these tests relate to the comparison of proportions of individuals falling into categories (discrete data)

while 1/3 compare means (continuous data)(12). They are further grouped into specific types depending on sample sizes and then number of samples.(Table 6 ).

The principal research question for discrete data relates to how the proportion of individuals in a particular category compares with one or more other proportions. Depending on the “Sample Size” two groups of statistical procedures are utilized: Exact tests (as a function of the underlying distribution of the variable) and the Chi-Square test (an approximation!)

For continuous data, “testing the difference of means” is the commonplace. Depending on the sample size and knowledge of information about variance of variables, the research question can be addressed using either the standard normal distribution or the t-distribution.

Worth underscoring are the notions of parametric tests (used for symmetrical distributions e.g., age of patients in Table 1) and non-parametric tests (used for skewed distributions e.g., pain severity (Table1)).

Generally, data are presented under two broad categories: Description of Independent/Explanatory variable(s) and Analysis of “Response/Dependent” variable(s). Separate tables should be used for independent and response variables. The table for independent variables should be in the order: Sociodemographic, clinical and paraclinical data.

<b>Summary statistics for age (years)</b>	<b>Age</b>
<b>N (sample size)</b>	20 patients
<b>Min</b>	21
<b>P25(Q1)</b>	31
<b>P75 (Q3)</b>	43.5
<b>Max</b>	75
<b>Mean</b>	39.95
<b>P50 (Median)</b>	39
<b>Range (Q3-Q1)</b>	12.5

**Table 2:** Summary statistics for the illustrative database of patients with herniated lumbar disc

Description	Mean	Standard Deviation	Proportion	Coefficient of Linear correlation
Statistics (Sample)	$\bar{x}$	s	$\hat{p}$	r
Parameter (Population)	$\mu = \text{"mu"}$	$\delta = \text{"sigma"}$	$\Pi = \rho = \text{"pi"}$	$\rho = \text{"rho"}$

**Table 3:** Symbols for Sample Statistics and Population Parameters

Treatment	n	Mean	Standard Error	Standard Deviation	[95% Confidence Interval]	P-Values*
Conservative	8	45.8	6.1	17.3	[31.3 60.2]	<u>P-Value 1 (P1)</u> Ha: diff < 0 Pr (T < t) = 0.9607
Microsurgery	12	36.1	1.3	4.6	[33.2 40.0]	<u>P-Value 2 (P2)</u> Ha: diff! = 0 Pr ( T  >  t ) = 0.0786  <u>P-Value 3 (P3)</u> Ha: diff > 0 Pr(T > t) = 0.0393

\*Two-sample t test with equal variances

**Table 4:** Comparing position and spread of age between two treatment groups

		Truth about the Population	
		H <sub>0</sub> True	H <sub>a</sub> True
Decision Based on Sample	Reject H <sub>0</sub>	Type I Error ( $\alpha$ )	Right decision
	Don't reject H <sub>0</sub>	Right decision	Type II Error( $\beta$ )

**Table 5:** Hypothesis testing and Power

## **Conclusion**

Understanding basic concepts in neurostatistics and neuroepidemiology provides the foundation for the interpretation and appraisal of scientific literature in neurosurgery. It enables neurosurgeons interested in research to appropriately plan and conduct their study, using the correct study designs and appropriately analyze the data. Neurosurgeons reading articles or conducting research ought to know the fundamentals provided in these reviews for better appraisal and application of scientific information to patient care.

## **References**

1. Armitage P, Berry G, Matthews JNS: Statistical Methods in Medical Research. John Wiley & Sons, 2008.
2. Bennett VLF: Handbook Of Clinical Neuroepidemiology. Nova Publishers, 2006.
3. Bhandari M, Joensson A: Clinical Research for Surgeons. Thieme, 2009.
4. Bland M: An introduction to medical statistics. Oxford University Press, 1995.
5. Bowers D: Medical Statistics from Scratch: An Introduction for Health Professionals. John Wiley & Sons, 2008.
6. Feigin VL, Bennett DA: Handbook Of Clinical Neuroepidemiology. Nova Science Publishers, 2006.
7. Feinstein AR: Clinical biostatistics. C. V. Mosby, 1977.
8. Fletcher RH, Fletcher SW: Clinical Epidemiology: The Essentials. Lippincott Williams & Wilkins, 2005.
9. Forthofer RN, Lee ES, Hernandez M: Biostatistics: A Guide to Design, Analysis and Discovery. Academic Press, 2007.
10. Glaser AN: High-Yield™ Biostatistics , 3e. Lippincott Williams & Wilkins, 2005.
11. Grimes DA, Schulz KF: An overview of clinical research: the lay of the land. Lancet 359:57-61, 2002.
12. Hennekens C, Buring J: Epidemiology in Medicine. Lippincott Williams & Wilkins, 1987.
13. Michelson S, Schofield T: The Biostatistics Cookbook: The Most User-Friendly Guide for the Bio/Medical Scientist. Springer, 1996.

14. Norman GR, Streiner DL: Biostatistics: The Bare Essentials. PMPH-USA, 2008.
15. Petrie: Medical Statistics at a Glance Text and Workbook. John Wiley & Sons, Limited, 2013.
16. Porta M: A Dictionary of Epidemiology. Oxford University Press, 2008.
17. Riffenburgh RH: Statistics in Medicine. Academic Press, 2006.
18. Rosner BA: Fundamentals of Biostatistics. Cengage Learning, 2011.