1. Why the hell do we need statistics?

"There are three kind of lies: lies, damned lies, and statistics", British Prime Minister Benjamin Disraeli (as credited by Mark Twain):

"It is easy to lie with statistics, but it is easier to lie without them", Fredrick Mosteller (Professor of Biostatistics at Harvard School of Public Health)

When should statistics be employed?
Every stage of the experimental process, from formulation of the problem (hypothesis) through data collection to drawing conclusions.

- Project planning (PowerAnalysis)
- Experimental design
- Data analysis
- Data interpretation

Statistical experimental design good is for:

- Eliminating known sources of bias
- Guarding against unknown sources of bias
- Ensuring the experiment provides precise information about the subject of interest
- Ethical considerations (preventing the waste of resources and lives)

So, that means that experienced investigators use proper statistics?
Several studies on the use of statistics in medical journal articles have consistently observed that roughly half of all journal articles are published with incorrect statistics. This is amongst professionals, presumably trained in the art of statistics. Misuse of statistics has caused both harm to patients bodies and to their pocketbooks, and it is important that we not perpetuate this misuse.

2. Data … what data?

- **Ratio Scale**: constant size interval between any adjacent units; a 39cm – 40cm step is equivalent to a 61cm – 62cm step. Zero point exists and there is a physical meaning to it (true zero point). Examples: weights, volumes, lengths of time etc.

- **Interval Scale**: Constant interval but not a true zero. Examples: temperatures scales (C and Fahrenheit; Kelvin is a ratio scale), time of day and year (circular interval scale).

- **Ordinal Scale**: Data dealing with relative rather than quantitative differences. *Ordinal* being from the Latin word for "order". Examples: shorter, faster, or sizes (e.g., 1, 2, 3, 4, and 5). Often, interval or ratio scale data are changed to ordinal scale data. By doing so, information is lost but analyses can be simplified.

- **Nominal Scale**: *Nominal* is from the Latin name for "name". Thus, data are qualitatively rather than quantitatively expressed by a specific attribute. Examples: dead or alive, types of trees, etc.

- **Continuous and discrete data**: When any conceivable value could be observed, then the variable is continuous. Discrete variables are usually consecutive integers but exceptions exist (e.g., ratio of number of wings to legs of insects). Ratio, interval, and ordinal scale data can be either continuous or discrete; nominal scale data are always discrete.

3. Ok, we have data … but why not simply calculate means and SD?

Variability is inherent to all experimental data:

Sources for variability:
Where do variations in measurements arise?

- Fixed differences among machines (e.g. instruments with different calibrations.)
- Random differences due to ambient conditions (e.g. temperature, time of day, etc)
- Measurement error in instrument readings (e.g. reading the number incorrectly)
- **Biological influences including genetics and environmental influences**

Experimental errors - when test conditions change
Measurement errors - even when test conditions are not changed

After collecting data, always plot your data to determine the underlying distribution

How to Summarize Distributions

- **Normal (Gaussian):** mean, standard deviation are appropriate
- **Non-normal:** mean, standard deviation are typically not appropriate. You could use the median, 25th and 75th percentile instead (which do not rely on the underlying distribution)

How can we describe the shapes of distributions?

- Skew
- Kurtosis
- Modality
- Range
- Percentiles

Some Important Term when Summarizing Data
Population - all possible measurements (N) of a certain type (e.g., height of all New Yorkers. Note a population may exist only conceptually). A statistical population consists of all possible items or units possessing one or more common characteristics under specified experimental or observational conditions.

Sample - a subset (n) of the population which you actually measured given financial and temporal constraints. Uses N-1 instead of N to calculate the variance and the SD. In practice, you hardly ever have the whole population.

Sample Size - should be carefully considered not only prior to the experiment (power calculations) but will also need to be considered during data analysis.

The central limit theorem describes the relation between SD and SEM

- Standard errors should not be used to describe data variability in a data sample as they provide information on the *accuracy* of the measured mean, rather than on its *precision*.

Coefficient of variation

CV = SD/Mean * 100% - allows direct comparison of variations independent of the magnitude.
Use CV only for ratio scale data

Z-Score

Z is known as a normal deviate or a standard score as it tells us how many standard deviations a particular value is located from the mean.

4. The use and misuse of Student's t-test…

The t-test is the most commonly used (and misused) statistical test in the biomedical sciences.

- For comparisons of 2 groups only
- This is a parametric test

There are different types of t-tests:

- Unpaired or paired t-test
- One-tailed or two-tailed
- Equal vs unequal variance

5. Of false positives and negatives….does my study have enough power?

	Ho is true	**Ho is false**
Ho is rejected	Type I error	No error
Ho is NOT rejected	No error	Type II error

Statistical power is the likelihood of not committing a type II error, 1-β

What variables affect type II error (power)?

- α – the risk of type II error increases with decreasing alpha
- Size of desired difference
- Variability
- Sample size

When should we calculate power?

- *Calculating power after data have been collected (post hoc):* Determination of post-doc power is typically **not** useful because the conclusions are trivial. Instead, calculate power for the difference that you were **originally** interested in and/or use the data for the design of a more appropriately powered (future) experiment.

- *During the design phase (critical for all grant applications):* Estimate the required sample size based on your willingness to accept a certain β-level.

6. Now that I know the t-test … do I really have to learn another statistical test? Practical advice for using ANOVA.

- **Analysis of variance (ANOVA)** is a powerful tool that can deal with a number of sophisticated data analysis questions
- This is a parametric test
- Student t test is a special case of the ANOVA for two groups
- A **single-factor** or **one-way** analysis of variance is used to determine whether 3 or more experimental groups are different from each other
- ANOVA itself will not tell you which groups are different from each other
- A **multi-factorial** or **multi-way** analysis of variance is used to determine the effects and interactions of a number of *factors*.

Post-Hoc Tests

- Bonferroni Test: Applies t-tests but corrects the α-value. If k comparisons are done, alpha becomes α_T/k. Works OK for very few groups but becomes overly conservative with increasing sample size and should be avoided.
- Tukey HSD Test (Honestly Significant Difference Test): it is commonly used and robust from departures of normality and homogeneity of variance; on the downside, it is somewhat conservative.
- SNK (Student Newman Keul test): very powerful, less conservative, popular, awkward to use with SPSS.
- Dunnet t-test: somewhat similar to the SNK test; used for making comparisons to only one control group

7. My sample size is really small. Should I still do statistics on my data? The beauty of non-parametric tests.

- Use a non-parametric test whenever you do not have a normal distribution or you are unable to assess the data distribution. Non-parametric tests are ***ranked*** tests.
- Can also be used for ordinal data (although other, parametric, methods are available)
- Non-parametric tests can always be used but a parametric test will ALWAYS be more powerful if it is justified. The power of a parametric test can be very low under conditions that warrant a non-parametric test.

Parametric Test	***Equivalent Non-Parametric Test***
Student t-test (unpaired)	Mann Whitney test
Student t-test (paired)	Wilcoxon test
ANOVA	Kruskal Wallis test

8. Regression? Correlation? What's the difference anyways?

Correlation versus Regression

Regression implies dependency but both analyses relate 2 variables to each other (even in the same way)

Regression

- Independent Variable
- Dependent Variable