

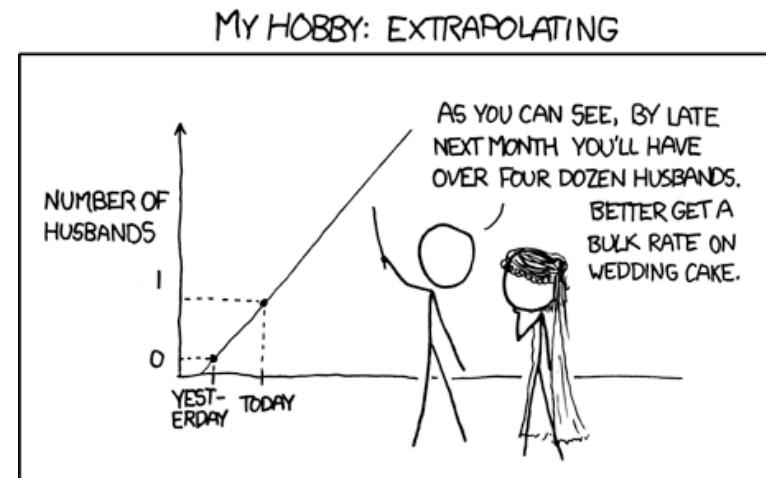
Design and Analysis of Experiments

Steven Forrester, PhD



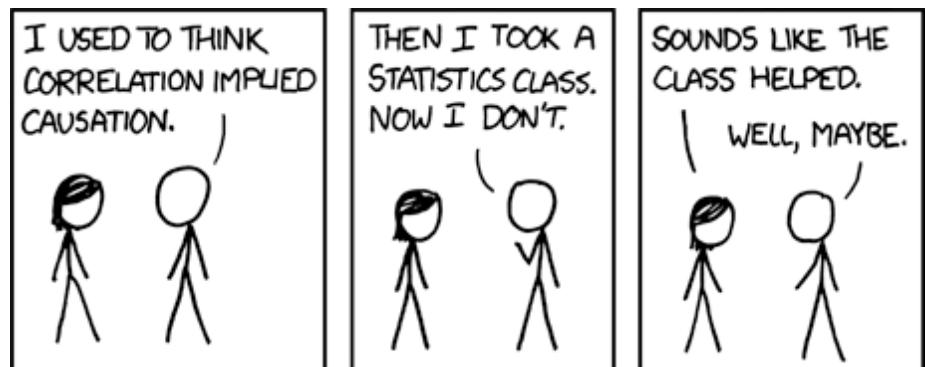
Why Stats and Design Matter

- Help make strong conclusions from limited amounts of data.
- Help distinguish between real differences and random variation.
- Mitigates human inclination to overgeneralize.
“9 out of 10 people that I know are dumb. Therefore, 90% of the population must be dumb.”
- Stats allows extrapolation from sample to population (if exp are designed correctly).
- Hypothesis testing – determine whether an observed difference is likely to be caused by chance (P value).
- Modeling – how well does data fit a mathematical model (regression).



Outline and Goals

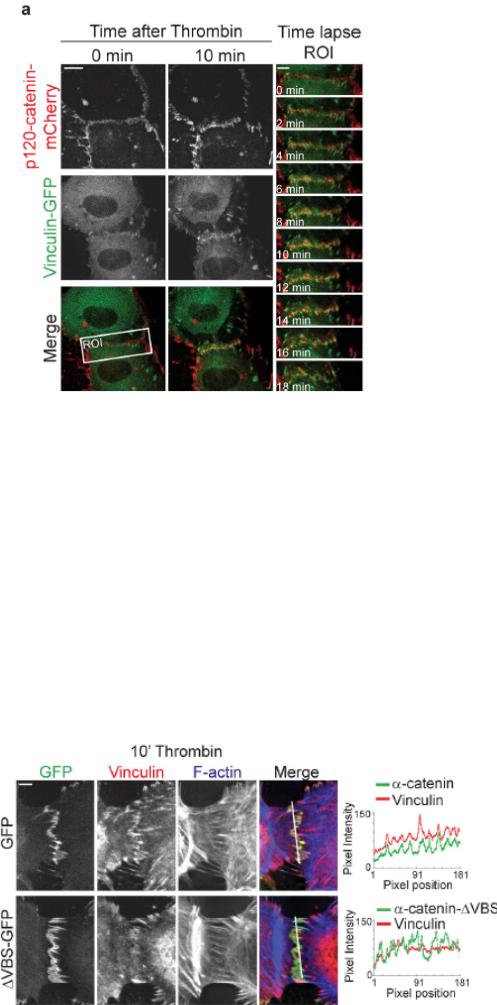
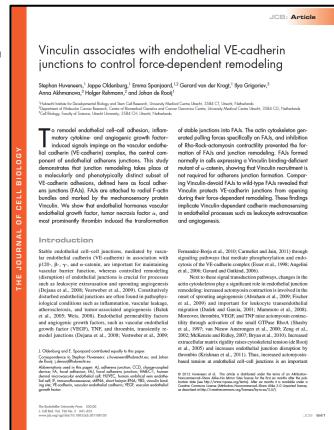
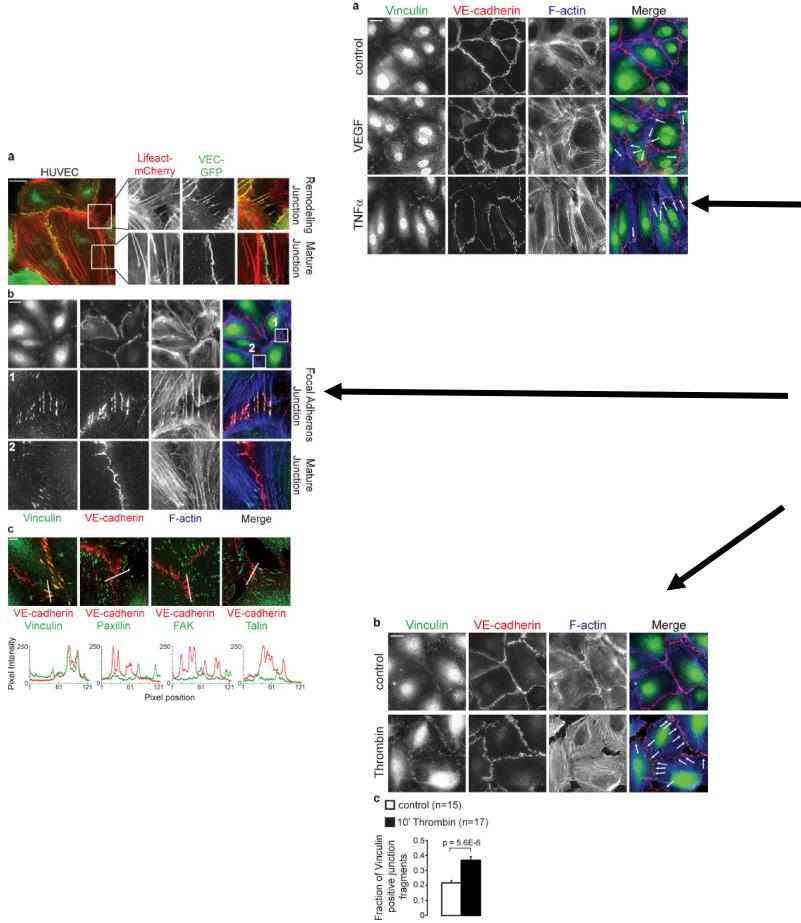
1. Introduce general considerations in the design of biomedical experiments.
2. Understand common statistical tests and when/how to use them.
 - Minimal use of statistical formulas
3. Case studies
4. General considerations to improve quality of your research and reproducibility.
 - Work standardization
 - Process efficiency





Designing a Study vs Experiment

- Most basic science studies reported in journals are not fully designed upfront
- Combination of many experiments



Designing a Study vs Experiment

- Most basic science studies reported in journals are not fully designed upfront
- Combination of many experiments

Study – General research topic
(How does LPS affect the vascular system?)

Experiments – Specifics of the topic

Exp Logic

Question

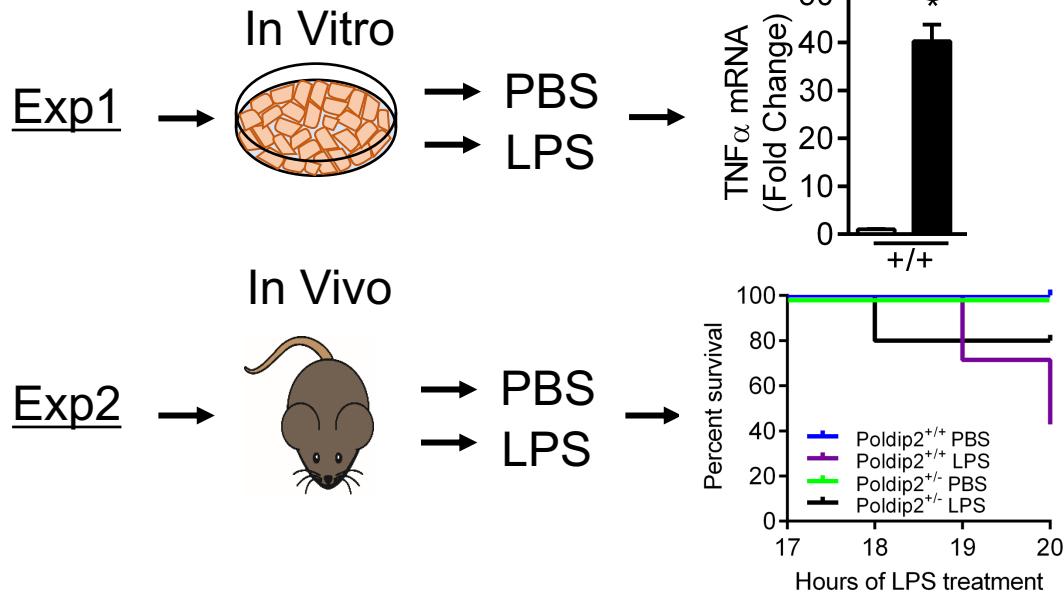
Hypothesis

Experiment

Observation

Analysis

Conclusion



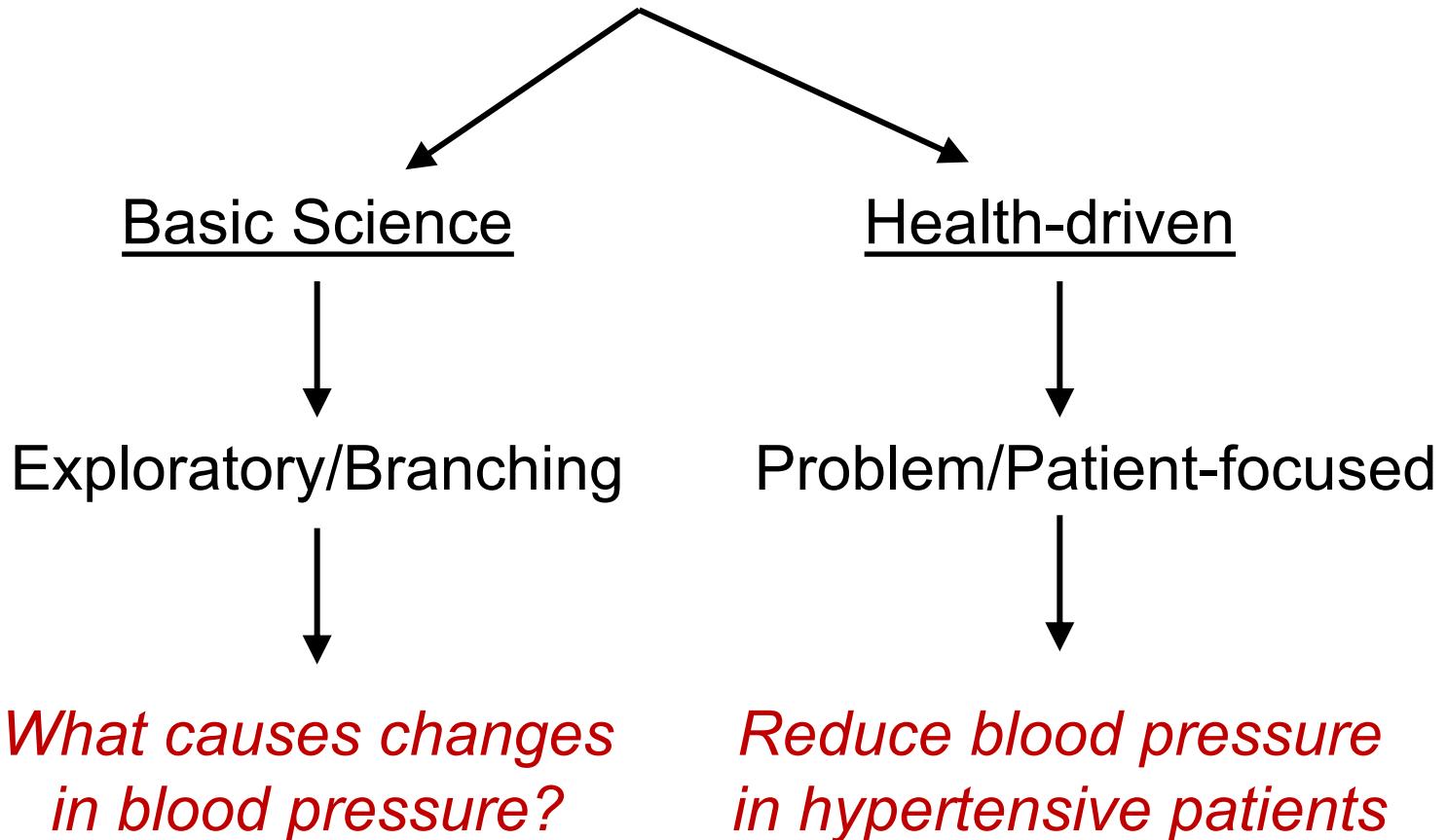
**What do you think makes a
good experiment?**



**So how do you design and conduct
an experiment?**

What is your goal?

Biomedical Research Areas





Principles of Experimental Design

Question

What scientific question are you trying to address?

How does TNF α affect adhesion molecule expression in endothelial cells?

- General but specific
- Actionable

Experiment

Observation

Analysis

Conclusion

What do you think is true/will happen?

I hypothesize TNF α increases adhesion molecule expression?

- Relates to question
- Something that can be disproven
- Provides framework on how to design experiment



Principles of Experimental Design

Pre-Experimental Planning

- **DO NOT RUSH THIS PHASE!!!**
- Often ignored or rushed in order to get to the experiment.
- Large influencer of end-product quality
- Poor planning = poor performance
- Helps you figure out what you need to do



Principles of Experimental Design

Design the process you will use to test your hypothesis.

Question

Hypothesis: *TNF α increases adhesion molecule expression?*

Hypothesis

Pre-Experiment



- What should my experiment be set up to do?

Answer: *Probably test adhesion molecule expression in response to TNF α*

Experiment

- How?

Observation

- What is your end-point?

What are you trying to observe?

Indirect vs direct – may depend on cost

Marker vs physiological

Analysis

Conclusion

Principles of Experimental Design

Design the process you will use to test your hypothesis.

Question

Hypothesis

Pre-Experiment



Experiment

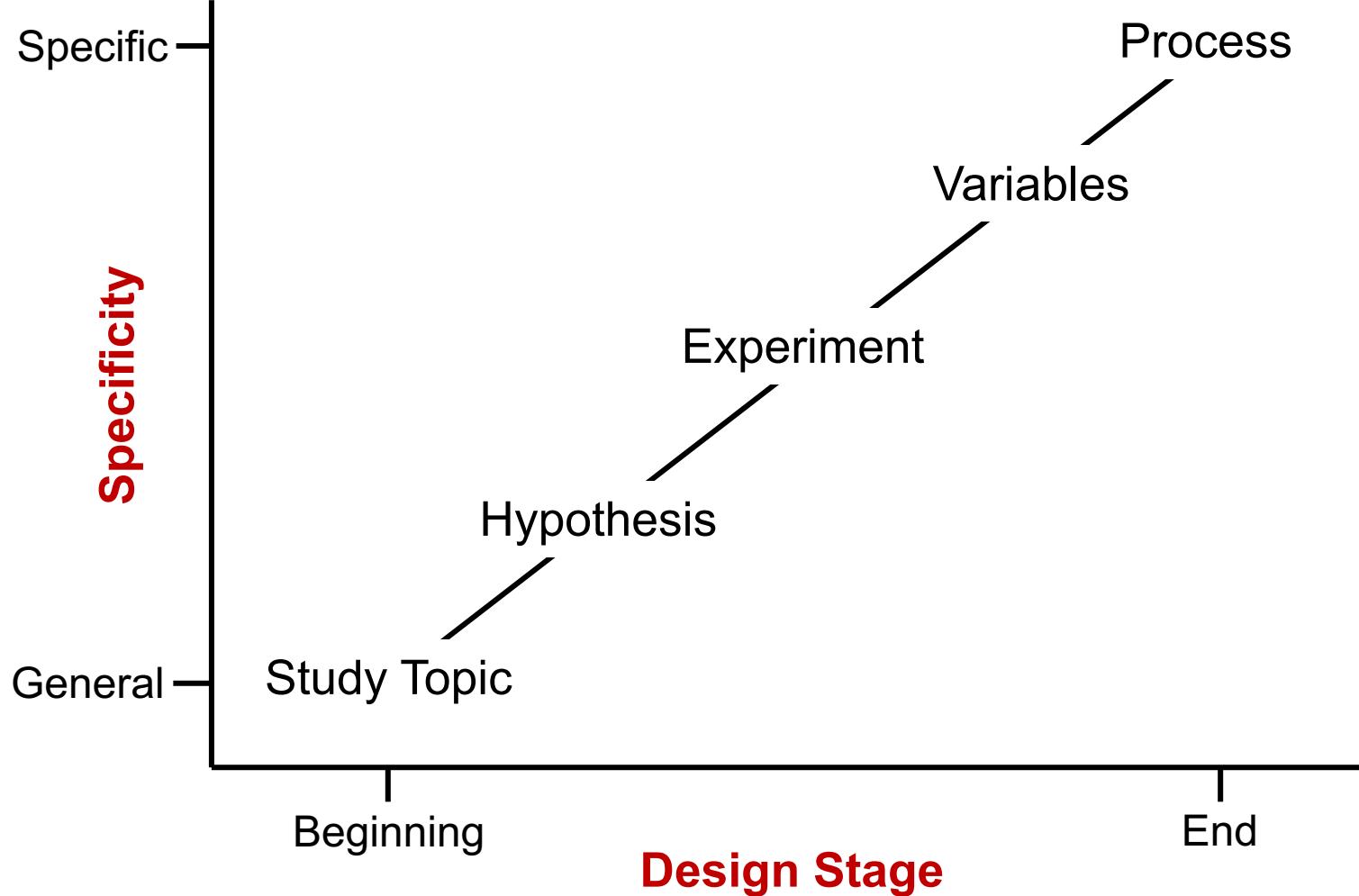
Observation

Analysis

Conclusion

- Consult literature
 - *What is the current knowledge in the field? Is this prior research reproducible?*
 - *What are the best techniques to test your endpoint?*
 - *Ex: Western, PCR*
- What are your variables?
 - Independent – X – study groups
 - Dependent – Y (*Value of Y depends on what happens with X*)
- What is the best experiment you could do ...*considering resources and cost?*
- How will you get there?
 - *Materials, equipment, study population*

Principles of Experimental Design



Example Case

Haploinsufficiency of gene X in humans has been correlated with protection against hypertension and other inflammatory conditions. Your lab is curious as to how gene X may affect the vascular system and blood pressure.

Question: What is an initial experiment that you could conduct?

Materials on hand: cells (any vascular cell type – ECs, VSMCs, CMs, Fibs), transgenic and KO mice for gene X, standard laboratory equipment and materials, inflammatory agonists, AngII, equipment for measuring blood pressure.

Define:

- Study topic
- Specific hypothesis
- X and Y – control and experimental
- Experiment

Example Experiments

- **Study topic:** The affect of gene X on the vascular system and inflammation
- **Specific hypothesis:** Gene X suppresses *TNF α -induced NF- kB nuclear translocation*
- **X and Y:**
 - *X variables*
 - Groups (2) – *siControl and siGeneX*
 - Within groups (2) – *PBS and TNF α*
 - *Y variable*
 - *Outcome - NF- kB translocation*
- **Experiment:** *Stimulation of cells followed by fluorescent staining and imaging via confocal microscopy*

Example Experiments

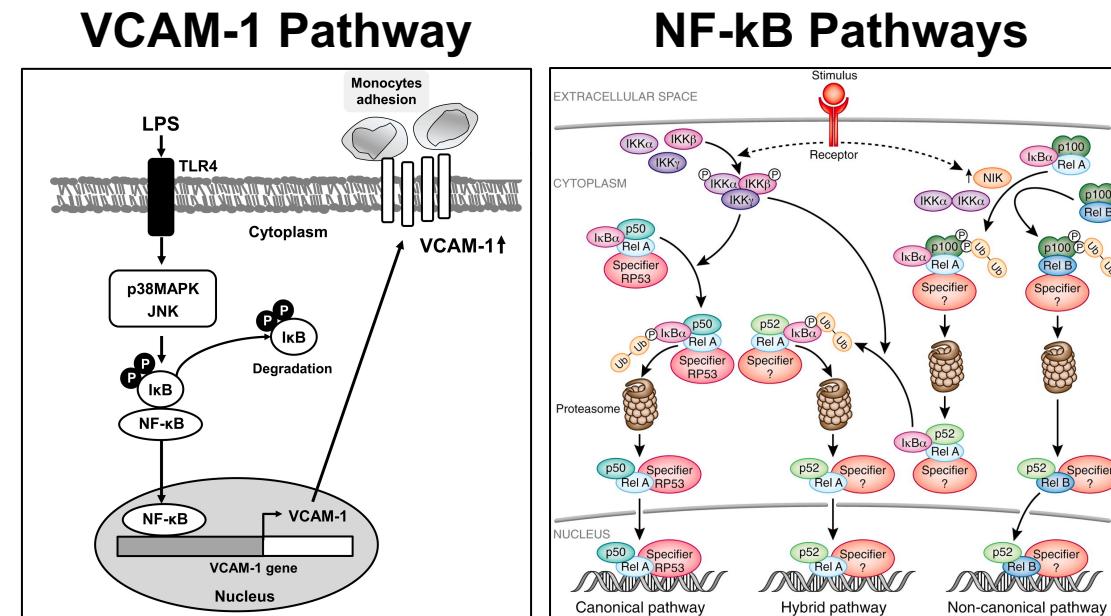
- **Study topic:** The affect of gene X on the vascular system and inflammation
- **Specific hypothesis:** *Gene X reduces blood pressure in mice.*
- **X and Y:**
 - *X variables*
 - Groups (2) – *Wild-type and Gene X KO mice*
 - *Y variable*
 - *Outcome – blood pressure*
- **Experiment:** *Telemetry or tail-cuff experiments on mice to measure blood pressure.*

Experiments

- Order of experiments matters (think general to specific)
 - Progression of a study requires interpretation of previous findings and sense of direction.
 - Limits re-work and waste

If I'm studying GeneX and NF- κ B activity, does it make more sense to investigate specific NF- κ B pathways or general VCAM-1 induction first?

- Drop-dead experiment
 - Experiment where failure to reject H_0 changes project direction.
 - Conclusive experiment



Principles of Experimental Design

Design the process you will use to test your hypothesis.

Question

Hypothesis

Pre-Experiment

Experiment

Observation

Analysis

Conclusion

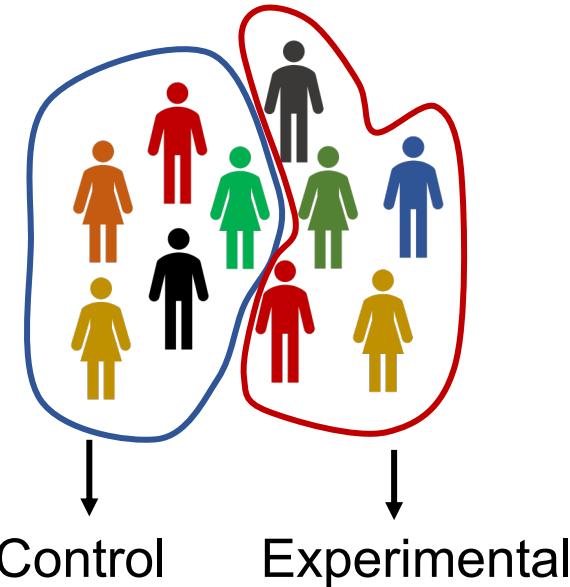


- How big of a difference are you expecting to see?
(Is there preliminary data?)
- Calculate sample size
- Experimental methodology
 - *Independent samples – most common*
 - *Paired samples*
 - *Repeated measures*
- How will you clean and analyze data?
 - *Will you get rid of outliers? – justification*
 - *Analyses: T-tests, ANOVA, regression, etc.*

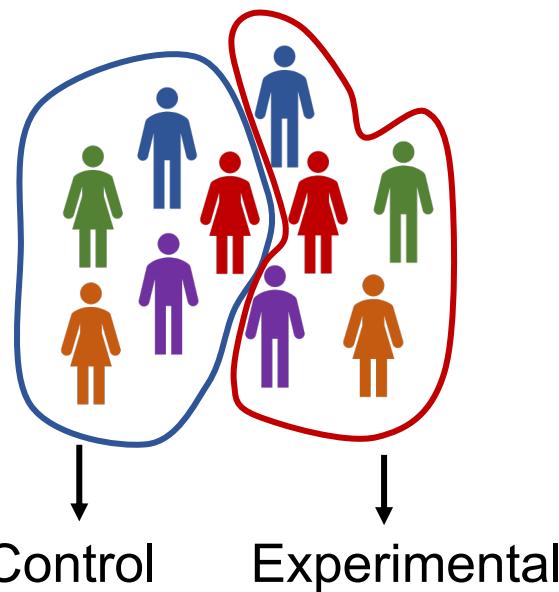


Principles of Experimental Design

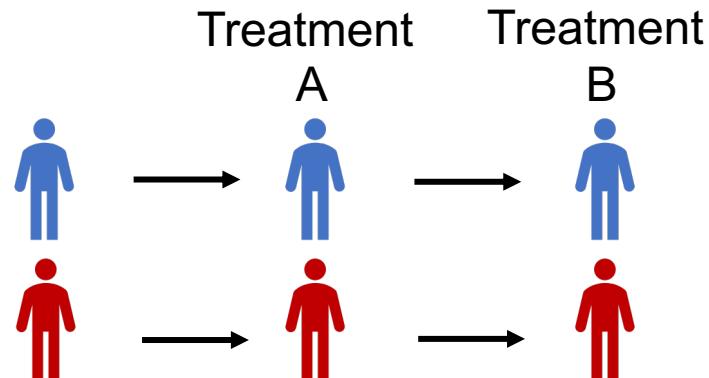
Independent Samples



Paired Samples



Repeated Measures



Sample Size

Sample Size Calculation

$$N = \frac{2 \cdot SD^2 \cdot (Z_\alpha - Z_B)^2}{\Delta^2}$$

Question

Hypothesis

Pre-Experiment

Experiment

Observation

Analysis

Conclusion

- $(Z_\alpha - Z_B)^2$ - power index
- Δ^2 - minimum difference

α	Power Index					
	$\beta = .01$	$\beta = 0.05$	$\beta = 0.10$	$\beta = .20$	$\beta = .50$	
1-sided	2-sided	Power = 99%	Power = 95%	Power = 90%	Power = 80%	Power = 50%
0.05	0.1	15.8	10.9	8.6	6.2	2.7
0.025	0.05	18.3	13	10.5	7.9	3.8
0.005	0.01	23.9	17.8	14.9	11.7	6.6

- α - probability of rejecting the null hypothesis when it is true
- β – probability of accepting the null hypothesis when it is false

α and β Error

α and β – Brief Review

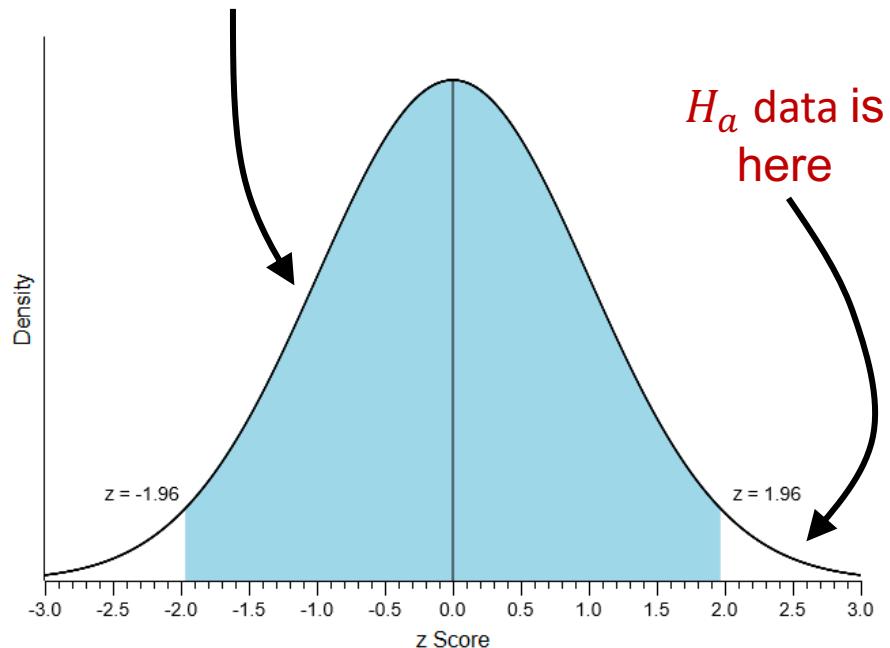
α - Also known as the significance level and is the probability of rejecting the null when it is true.

- p of 0.05 would indicate a 5% risk of concluding that a difference exists when there is no actual difference (Type I error)
- $p < 0.05$ is a scientific standard

Why?

- Because someone said so
- α is determined beforehand and set in context of the specific experiment.
- Does a patient really care if $p = 0.07$?

Distribution of data
when H_0 = True



α and β Error

α and β – Brief Review

β – Also known as the Type II error and occurs when you conclude that there is no significant difference between means, when in fact H_a is true.

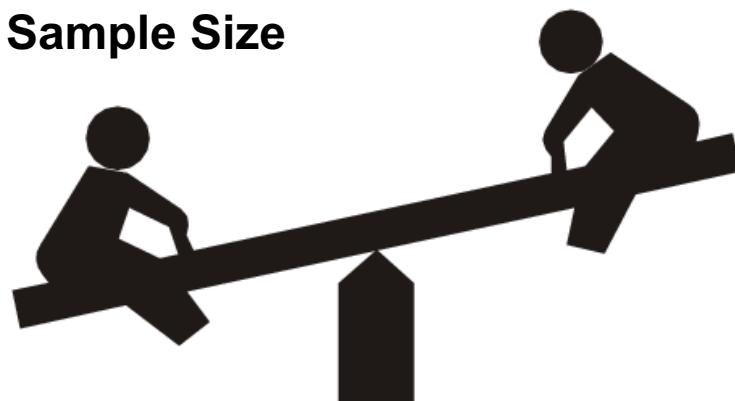
- *Power* – defined as $1 - \beta$ and is the probability that an experiment will detect a difference, when a difference exists.

$$\text{power} = \Pr(\text{reject } H_0 \mid H_1 \text{ is true})$$

- *Power* is regulated by both sample size and effect size within an experiment.

Effect Size

Sample Size



Sample Size

Sample Size Calculation Examples

$$N = \frac{2 \cdot SD^2 \cdot (Z_\alpha - Z_B)^2}{\Delta^2}$$

α	Power Index					
	$\beta = .01$	$\beta = 0.05$	$\beta = 0.10$	$\beta = .20$	$\beta = .50$	
1-sided	2-sided	Power = 99%	Power = 95%	Power = 90%	Power = 80%	Power = 50%
0.05	0.1	15.8	10.9	8.6	6.2	2.7
0.025	0.05	18.3	13	10.5	7.9	3.8
0.005	0.01	23.9	17.8	14.9	11.7	6.6

Question: Does deletion of AT1R in C57BL6 mice reduce systolic blood pressure?

Experiment: Telemetry in AngII-infused AT1R +/+ and -/- mice.

SD: 10 mmHg, Δ : 15 mmHg, Power: 80%

$$\frac{2 \cdot 10^2 \cdot (7.9)}{15^2} = 7 \text{ mice per group}$$

Question

Hypothesis

Pre-Experiment

Experiment

Observation

Analysis

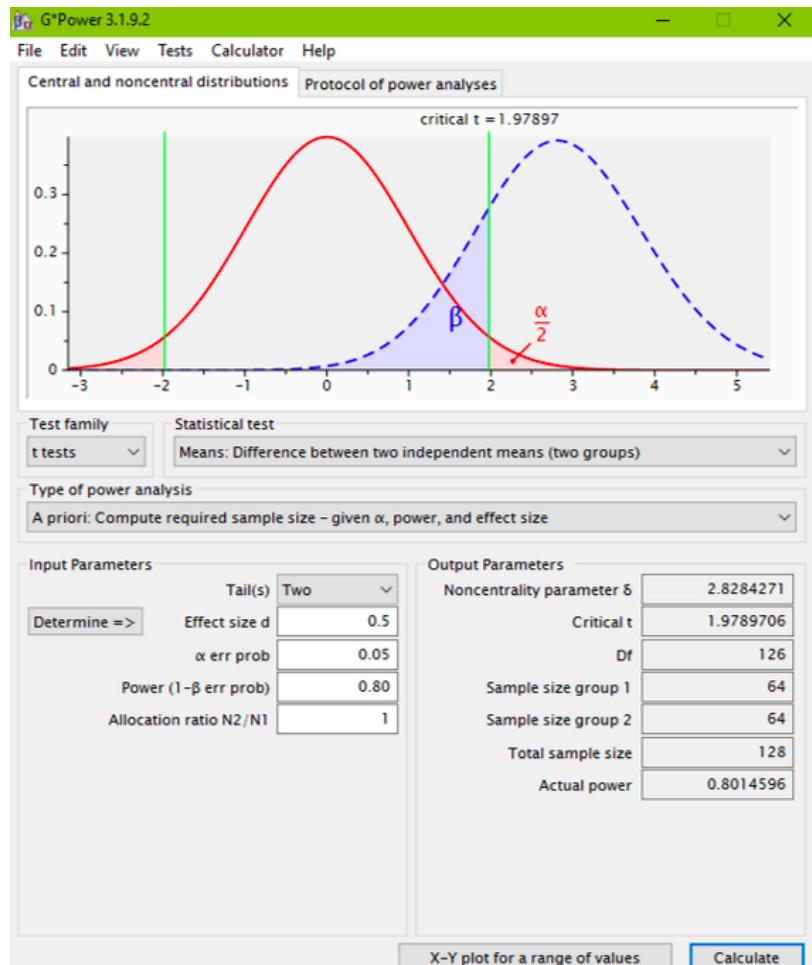
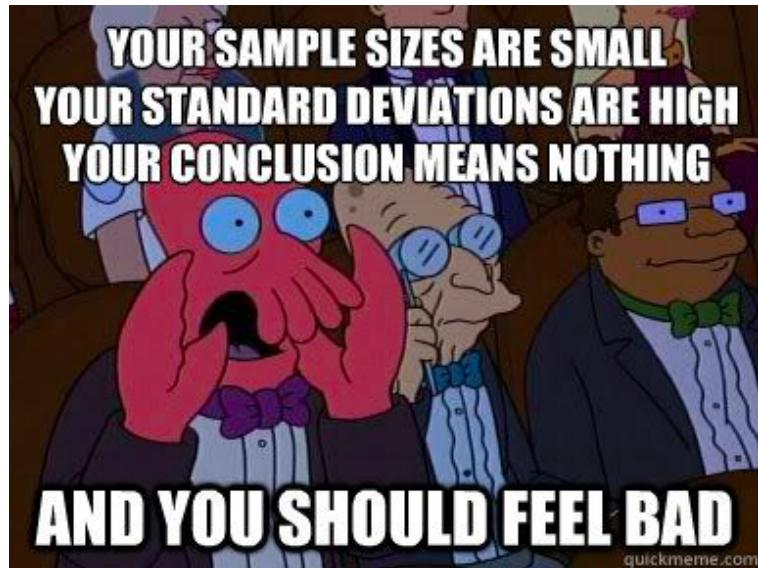
Conclusion



Sample Size

Modern Sample Size Calculation

- Variety of software available to calculate sample size (Graphpad/Prism, SPSS, *R*, Python, GPower).
- Use them

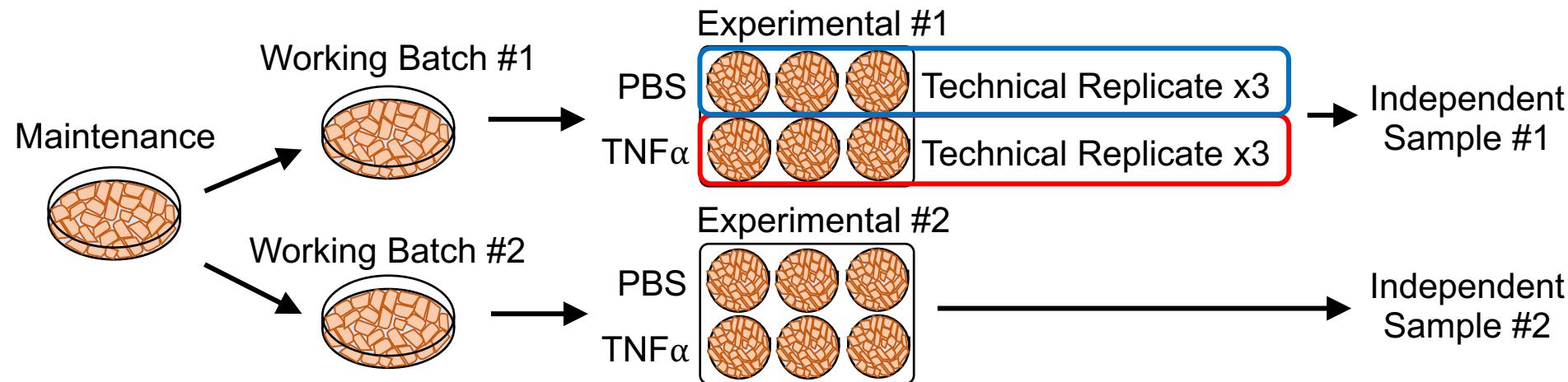




Replicates

Sample Size Calculation – Additional Considerations

- Previous tools apply to all facets of basic science; however, *in vitro* experiments are tricky because sometimes differences of over 100% are observed – what do we do?
- General guidelines for *in vitro* work
 - Perform 3-4 independent experiments with technical replicates (2-3) at minimum



Instilling Quality into Experiments

Reproducibility and Scientific Rigor

Question

Hypothesis

Pre-Experiment

Experiment

Observation

Analysis

Conclusion

- The most important aspect of an experiment is designing and conducting it in a way so that it's reproducible by others and minimizes bias.
- Patients don't care about innovation. Improved therapies require standards of quality.
- Design quality into your experiments.
- Common approaches:
 - Record and report all details of an experiment!
 - Create robust experiments
 - Experimental blinding
 - Using proper controls
 - Using proper analytic methods
 - TRY AND DISPROVE YOUR HYPOTHESIS!

Experimental Process

Question

Hypothesis

Pre-Experiment

Experiment

Observation

Analysis

Conclusion

Carry out the process to test your hypothesis

Think of every step of your process

Endpoint: VCAM-1 protein expression

Steps:

1. Grow cells – Human Endothelial cells
2. Stimulate cells with TNF α
3. Collect cells in buffer suitable for Western
4. Run Western blot
5. Transfer bands from gel to membrane
6. Probe with antibody for adhesion molecule
7. Image
8. Put in report – commonly powerpoint
9. Discuss with peers, labmates, PI
10. Interpret
- 11. Repeat**

Experimental Process

Question

Hypothesis

Pre-Experiment

Experiment

Observation

Analysis

Conclusion

Record observations for others

- Record data and observations so that others can follow along.
- Details:
 - Note experimental conditions
 - Animal behavior
 - Reagents used (lot #, dose, expiration, storage)
 - Equipment used
 - Who helped you
- Keeping detailed records saves time when writing pubs/grants and will help others measure the quality of the experiments.

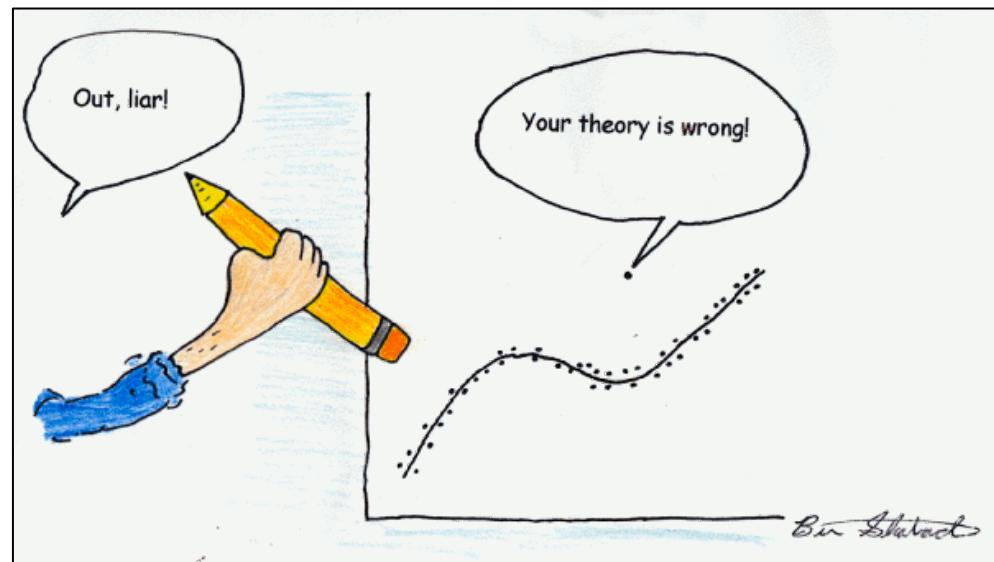


Analysis



Analyzing results

- Which statistical test do I use?
 - Depends on experimental design
 - Parametric vs non-parametric
- What about outliers?



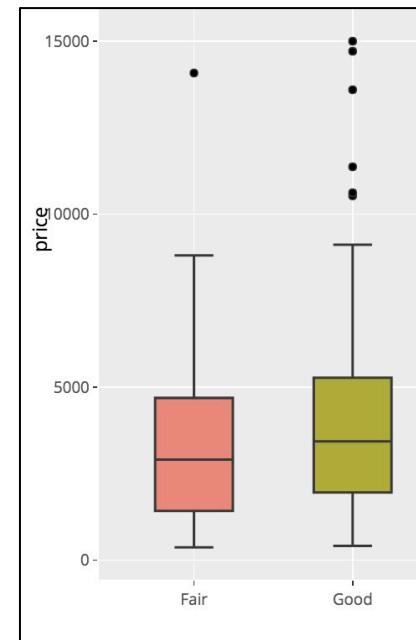
Outliers



Outliers and Data Exclusion

- When is it OK to remove a data point?
 - If an experiment is done incorrectly

Use the wrong dose of TNF α
Animal dies during experiment
Data is coded incorrectly
Failure to blind individual performing analysis
- When should you keep outliers?
 - Usually always unless you can prove the data point is due to bad data.
- How do you test for outliers?
 - Grubb's test – Graphpad/Prism
 - Quartiles and IQR rule

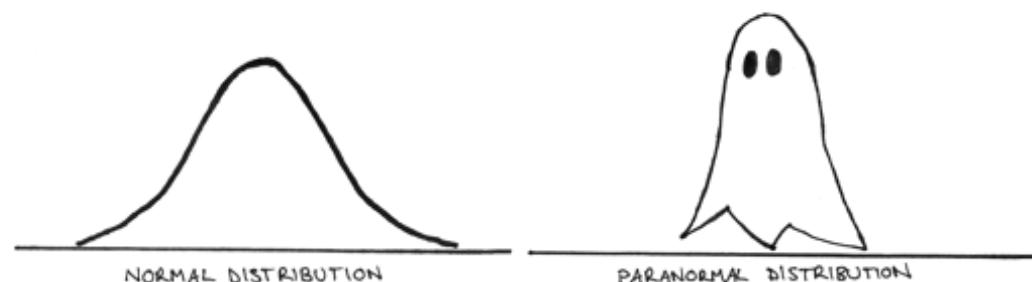


Statistical Tests



Types of Analyses (Comparisons)

- **Parametric** – compare means
 - Assume normal distribution
 - CLT – given a large sample size, parametric tests will work for non-normal data
 - High statistical power – more likely to detect significant effect
- **Non-parametric** – compare medians
 - Make no assumptions about distribution
 - Used for small sample sizes if non-normal



Statistical Tests

Question

Hypothesis

Pre-Experiment

Experiment

Observation

Analysis

Conclusion

Types of Analyses

- **T-test**

- Test difference between 2 group means
- Uses t statistic
- Can be paired or unpaired
- Non-parametric – Wilcoxon/Mann-Whitney

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

difference between groups

Variability of groups

- **ANOVA**

- Test difference between 3 or more groups means
- Analyzes variance within and between groups

Statistical Tests

Which Group is Different in an ANOVA?

- **Post-hoc tests**
 - Methods for specific comparisons of means after an ANOVA
 - Compare control to all other groups – Dunnett's test
 - Compare only selected means – Bonferroni test
 - Compare all means – Bonferroni, Tukey, SNK
- **Assumptions of ANOVA**
 - Assumption of normality (variables and residuals)
 - Variances of populations are similar
 - Responses are independent



Table of Tests

Goal	Type of Data		
	Measurement (from normal distribution)	Rank, Score, or Measurement (non-normal distribution)	Binomial (Two Possible Outcomes)
Describe one group	Mean, SD	Median, IQR	Proportion
Compare one group to a hypothetical value	One-sample <i>t</i> test	Wilcoxon test	Chi-square or Binomial
Compare two unpaired groups	Unpaired <i>t</i> test	Mann-Whitney	Fisher's test
Compare two paired groups	Paired <i>t</i> test	Wilcoxon test	McNemar's test
Compare three or more unmatched groups	One-way ANOVA	Kruskal-Wallis test	Chi-square
Compare 3 or more matched groups	Repeated-measures ANOVA	Friedman test	
Quantify association	Pearson correlation	Spearman correlation	
Predict value from another variable	Regression	Non-parametric regression	

General Considerations

- **Beware of very large and very small samples**
 - Tiny differences will be statistically significant, even if scientifically trivial
 - Small have little power – important differences may be insignificant
- **Don't focus on averages – outliers may be important**
- **Garbage in, Garbage out**
 - Statistics won't solve bad data or methods
- **Confidence interval are as informative as P values**
- **Statistically significant does no mean scientifically important**
- **$P < 0.05$ is not sacred**
- **Distinguish between studies designed to generate a hypothesis vs studies designed to test one**