Introduction to experimental design

30 August 2017

Aaron Lun
(based on slides by Rory Stark and Sarah Vowler)

Why perform experiments?

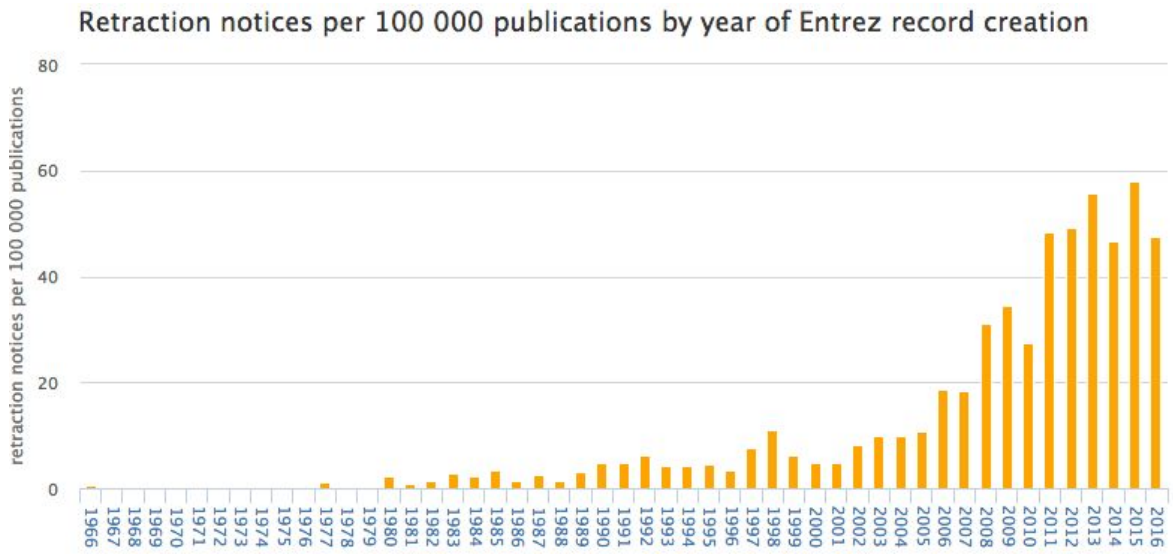
Scientific method:

1. Form a hypothesis about a phenomenon
2. Set up an experiment to test the hypothesis
3. Do the data support/refute hypothesis?
4. Repeat 1 with new hypothesis/phenomenon

Domain knowledge => **1**
Experimental design => **2**
Statistical analysis => **3**

Why think about experimental design?

- Does your experiment have the capability to answer your scientific question?
- Can the results of your experiment be reproduced by yourself and others?



http://rpubs.com/neilfws/65778

47 of 53 high-profile cancer studies were not reproducible (Begley and Ellis, Nature, 483, 531–533)

Why think about experimental design?

*"To consult the statistician **after** an experiment is finished is often merely to ask him to conduct a **post mortem** examination. He can perhaps say what the experiment died of." Ronald A. Fisher, 1938.*

Poor experimental design wastes:

- **Your money & time**, in experimentation
- **Limited & precious** material, esp. clinical samples
- **Other people's time**, due to immortalization of data sets in public databases and literature

What is a well-designed experiment?

- Clear objectives: focused and simple
  - *What are your experimental factors?*
- Accurate
  - *What are your biases?*
- Precise and powerful
  - *What are your sources of variability?*

Experimental Factors

- Independent and Dependent variables
  - Independent variable (IV): what you change (e.g., genotype by KO)
  - Dependent variable (DV): what changes due to IV (e.g., cell division rate)
  - nominal, ordinal, discrete, continuous - to be discussed tomorrow!

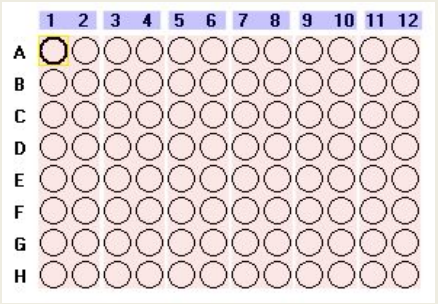
- Experimental factors: IVs that could change the DVs of interest
  - some of interest and to be varied by the experimenter
    - e.g., genotype, time, drug treatment
  - others not of interest and **should be controlled**
    - e.g., plate, batch

Figure out what your experimental factors are in advance!

Confounding Factors

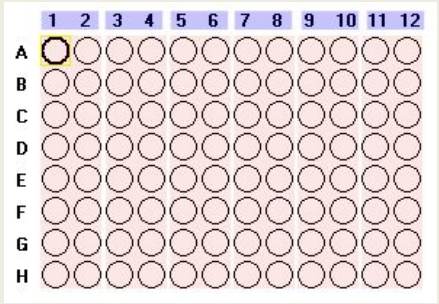
- Uninteresting factors that are *not* controlled
- May mask an actual association or **falsely** demonstrate an apparent association between the factor of interest and the dependent variables.
- Simple example: test for change in weight between WT and KO mice
  - but all WT mice are male, and all KO mice are female
  - is the difference due to sex or gene knockout?

Day1, Plate 1



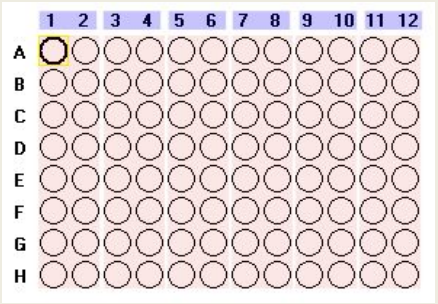
Control

Day2, Plate 2



Treatment 1

Day3, Plate 3



Treatment 2

*Differences between Control, Treatment 1 and Treatment 2 are confounded by **day** and **plate.***

Confounding Factors

- Experimental data becomes difficult to interpret if it's not just the the factor of interest that varies.
- Confounding factors introduce biases into the results, potentially resulting in misleading conclusions.
  - *Bias results in inaccurate estimates of your effects*
  - *Even worse, they may be reproducibly inaccurate!*
- If a study does not consider confounding factors, **don't believe it!**

Make sure you know what your confounding factors are!

AN EXPRESSION OF CONCERN HAS BEEN PUBLISHED ON THIS PAPER; SEE LAST PAGE

Sciencexpress Report

Genetic Signatures of Exceptional Longevity in Humans

Paola Sebastiani,[1]* Nadia Solovieff,[1] Annibale Puca,[2] Stephen W. Hartley,[1] Efthymia Melista,[3] Stacy Andersen,[4] Daniel A. Dworkis,[3] Jemma B. Wilk,[5] Richard H. Myers,[5] Martin H. Steinberg,[6] Monty Montano,[3] Clinton T. Baldwin,[6,7] Thomas T. Perls[4]*

[1]Department of Biostatistics, Boston University School of Public Health, Boston, MA 02118, USA. [2]IRCCS Multimedica, Milano, Italy; Istituto di Tecnologie Biomediche, Consiglio Nazionale delle Ricerche, Segrate, 20122, Italy. [3]Department of Medicine, Boston University School of Medicine, Boston, MA 02118, USA. [4]Section of Geriatrics, Department of Medicine, Boston University School of Medicine and Boston Medical Center, Boston, MA 02118, USA. [5]Department of Neurology, Boston University School of Medicine, Boston, MA 02118, USA. [6]Departments of Medicine and Pediatrics, Boston University School of Medicine and Boston Medical Center, Boston, MA 02118, USA. [7]Center for Human Genetics, Boston University School of Medicine, Boston, MA 02118, USA.

- GWAS study: 800 centenarians vs. controls
- Found 150 SNPs predicting centenarians with 77 % accuracy
- Problem: they used **different SNP chips** for centenarians and controls
- Retracted in 2011 following independent review and QC of data

http://www.the-scientist.com/blog/display/57558/

Confounding factors = bias

Type of Bias	Description
Selection bias	Systematic differences between **baseline characteristics** or **treatment groups** that are being compared.
Performance bias	Systematic differences between groups in exposure to factors other than the interventions of interest (aka ***confounding*** **or** *extraneous* **factors**).
Attrition bias	Systematic differences between groups due to samples being **withdrawn** from the study or **excluded** from the analyses.
Detection or Measurement bias	Systematic differences between groups in how outcomes are assessed or determined, e.g. **measurement errors** and inefficient use of data.
Reporting bias	Systematic differences between reported and unreported findings due to manipulation in the reporting of findings such as **selective or distorted reporting**, e.g. papers with more 'interesting results ' are more likely to be submitted and accepted for publication.

Removing bias with controls

Question: *Is the observed effect caused by my factor of interest?*

Assume we have "treatment" samples, affected by our factor.

- Construct a set of control samples, where **only the factor of interest differs** between control and treatment.
- Effect of our factor is determined by comparing the DV values between the control and treatment groups.

Set up appropriate controls for each factor of interest!
(not always obvious, depends on knowing the other factors)

Types of experimental controls

- Negative controls = no expected effect:
  - e.g., a WT control when studying a KO phenotype
  - avoid “type I” errors, i.e., false positives
- Positive controls = expected known effect
  - e.g., a qPCR reaction with known template
  - avoid “type II” errors, i.e., false negatives
- Technical controls
  - e.g., standard curve for qPCR, spike-ins
  - Detect/correct technical biases, normalization

Examples of experimental controls

- Wild-type organism (knockouts)
- Inactive siRNA (silencing)
- Vehicle (treatments)
- Input: fragmented chromatin (ChIP)
- Spike-ins (quantification/normalisation)
- “Gold standard” datapoints
- Multi-level controls
  - e.g. ChIP-seq: Vehicle/Input vs. Treatment/Input

Removing bias with randomization

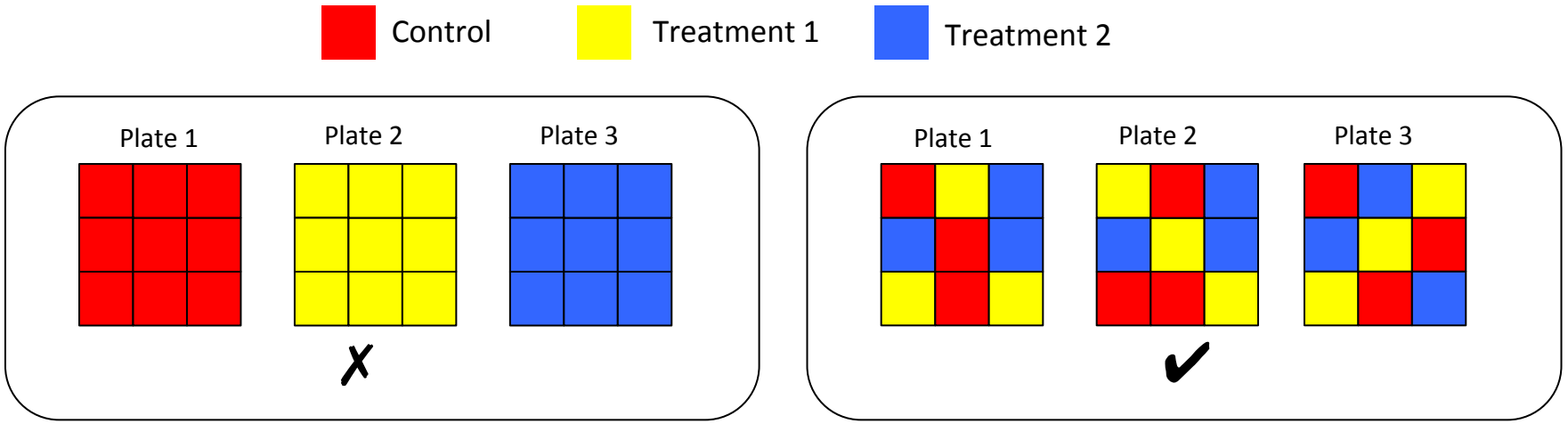
- Some variables cannot be easily controlled
  - e.g., random fluctuations in measuring devices
  - logistics are not feasible

- Randomize to eliminate systematic biases
  - e.g., can't process all samples in a single day
  - randomize the samples to be processed across days, avoid biases due to time effect

If you can't control a factor, randomize across it!

Randomised block design

Blocking is the arranging of *experimental units* in groups (blocks) that are similar to one another.



Each plate contains spatially randomised equal proportions of control and treatments 1 and 2, controlling plate effects.

Randomised Block Design

Good design example: Alzheimer's study from GlaxoSmithKline

Plate effects by ***plate***

Left PCA plot show *large plate effects*.
Each colour corresponds to a different plate

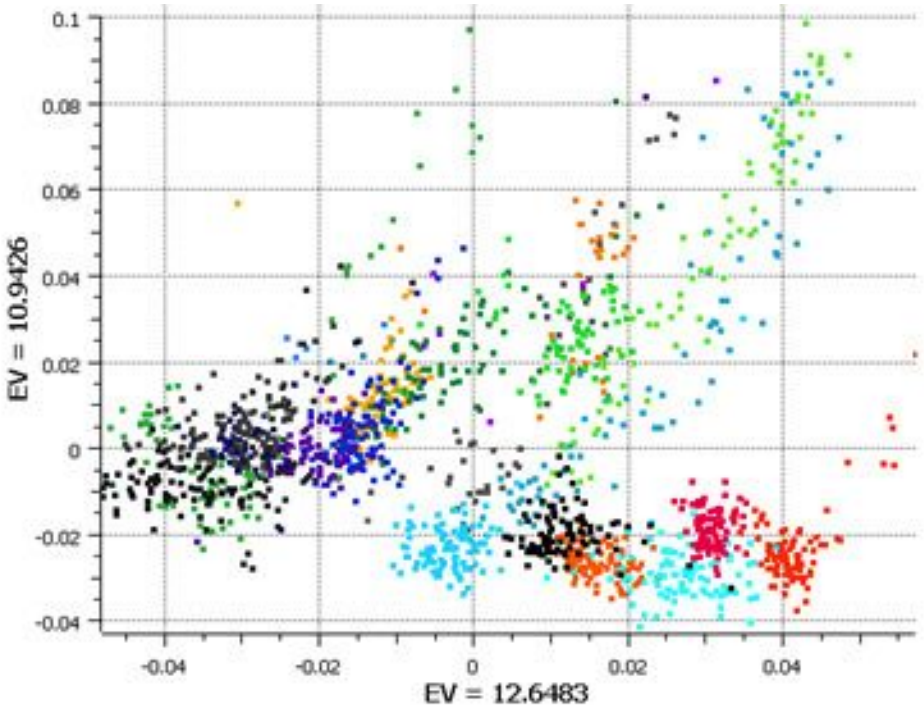
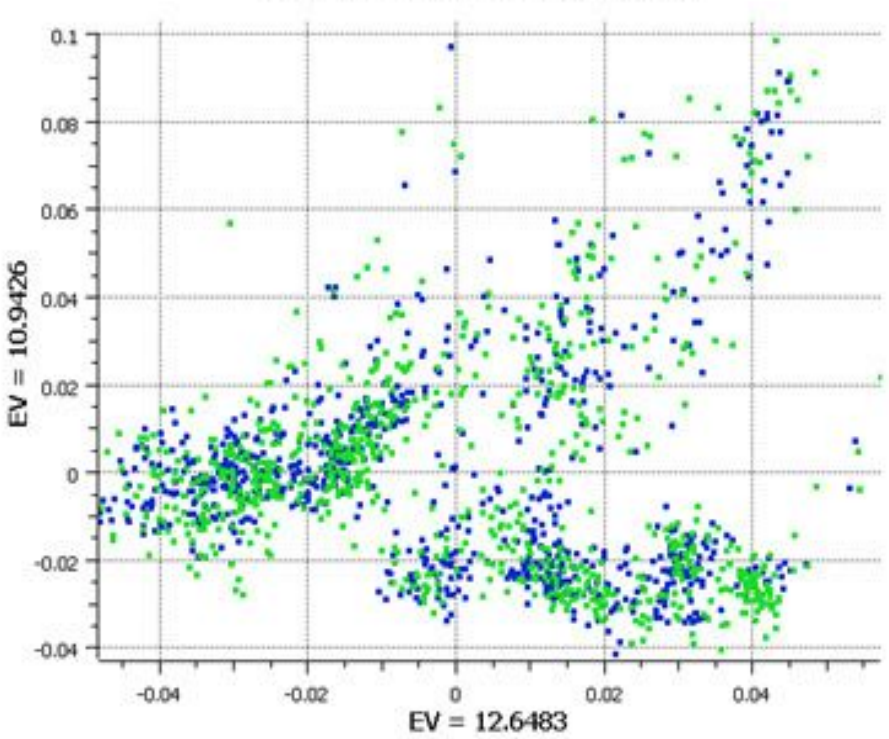


Plate effects by ***case/control***

Right PCA plot shows each plate cluster contains *equal proportions* of cases (blue) and controls (green).



http://blog.goldenhelix.com/?p=322

Removing bias with blinding

Especially important where subjective measurements are taken

Unconscious (or conscious!) biases of the experimenter may affect the measurements

Blind the labels so that person doing the measuring can only use the data

- e.g., wine tasting, scoring phenotypes...

Experimental variability

- Biological variability (a.k.a. “noise”)
  - Biological processes are inherently stochastic
  - Single cells, cell populations, organs, individuals...
  - Timepoints, cell cycle, stresses, etc.
- Technical variability
  - Reagents, antibodies, temperatures, pollution
  - Platforms, runs, operators

Some of these are obvious factors to be controlled.

Others are not; how do you ask someone to be “less variable”?

Why use replicates?

Question: *How can I be sure that my effect is real and reproducible, and not just due to random variability?*

Do it again, and again, and again…

Always include multiple replicates in your design!

(not always obvious: what type? How many?)

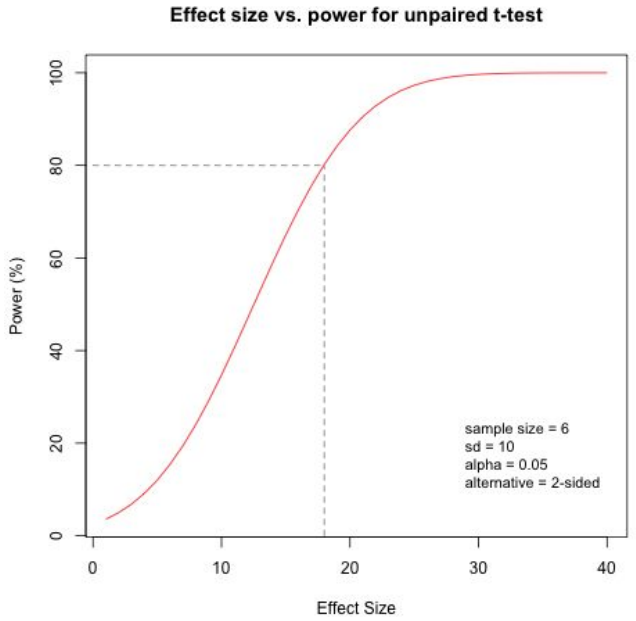
Types of Replication

- Biological replication = measurements of *independent biological samples.*
  - accounts for variability in the biological system (plus variability in the measurements)
  - *In vivo:* Patients, mice
  - *In vitro:* Different cell lines, passages
- Technical replication = repeated measurements of the *same biological sample.*
  - only accounts for variability in experimental protocol, measurement platform (e.g., sequencer, Nanodrop)
  - easier to generate, but much less useful

How many replicates/samples?

- Depends on variability and effect size
  - smaller effects = need more samples
  - greater variance = need more samples
  - (assuming that adding more samples doesn't increase the variability)

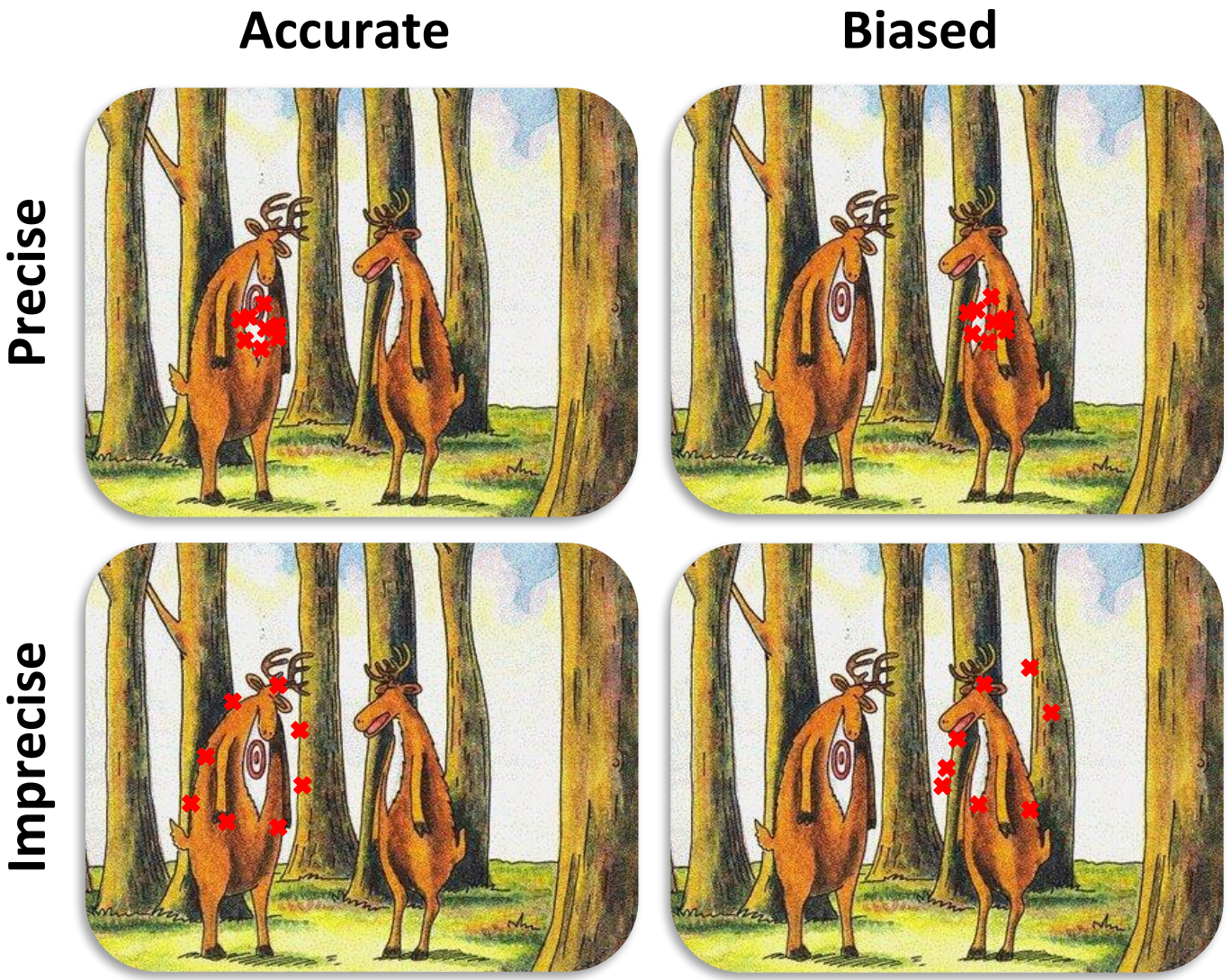
- Calculating appropriate sample sizes
  - Power calculations
  - Resource equation



- Power: the **probability** of detecting an **effect** of a specified size if present.
  - Calculation requires knowledge of variability; some expectation of the effect; and knowledge of the statistical analysis and acceptable error rate
  - determine **appropriate numbers** of samples (sample size/replicates)

- Mead's resource equation: get enough samples to quantify the error well
  - usually around 10-20 "degrees of freedom" for simple experiments
  - i.e., 10-20 more samples than groups (WT/KO/treated/untreated, etc.)

Precision, Accuracy & Bias



Types of Experimental Designs

- Block designs: randomisation
- Matched: tumour/normal
- Factorial/multifactorial designs
- Nested designs
- Time series

https://rawgit.com/bioinformatics-core-shared-training/experimental-design/master/ExperimentalDesignManual.pdf

Can you design a good experiment?

1. Know what your experimental factors are:
   - Which ones are interesting and to be varied?
   - Which ones should be controlled?
2. Deal with the biases in your system:
   - Controls, randomization and blinding
3. Deal with the variability in your system:
   - Replicates, and enough of them!

Make life easier for your statistical analysis!

Specific issues for sequencing

- Platforms (MiSeq, HiSeq, etc.)
- Library preps
- Multiplexing and pooling strategies
- Single-end vs paired end
- Sequencing depth
  - Coverage
  - Lanes
- Validation