

# Experimental Design



Roslin Russell  
Cambridge Research Institute (CRI) CRUK



UNIVERSITY OF  
CAMBRIDGE



CANCER  
RESEARCH  
UK

# Experimental Design

- **Importance**

- Planning
- Well-designed Experiment
- What Type of Experiment
- Define Variables & Good Hypotheses
- Standardise
- Experimental Controls
- Experimental Units
- Bias & Confounding Factors
- Replicates & Sample Size
- Common Experimental Set-ups



## Ronald A. Fisher (1890-1962)

- "*a genius who almost single-handedly created the foundations for modern statistical science*" *Anders Hald*
- "*the greatest biologist since Darwin*" *Richard Dawkins*
- Innovative book ***The Design of Experiments*** (1935).
- Developed approach to design and analyse experiments:
  - Ensures that errors and variance in the measurement of variables are controlled.
  - True causes of changes in variables more clearly distinguished.



## Ronald A. Fisher (1890-1962)

Fisher's fundamental experimental design principles:

- **Replication:**
  - Having more samples ***reduces variability of the estimates*** and increases confidence in the results.
- **Randomisation:**
  - Rule used to assign the samples to treatments and helps to *protect from bias* in results.
- **Blocking:**
  - a homogenous group of **experimental units** e.g litter or cage.



## Ronald A. Fisher (1890-1962)

*“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.” (1938)*

# Experimental Validity

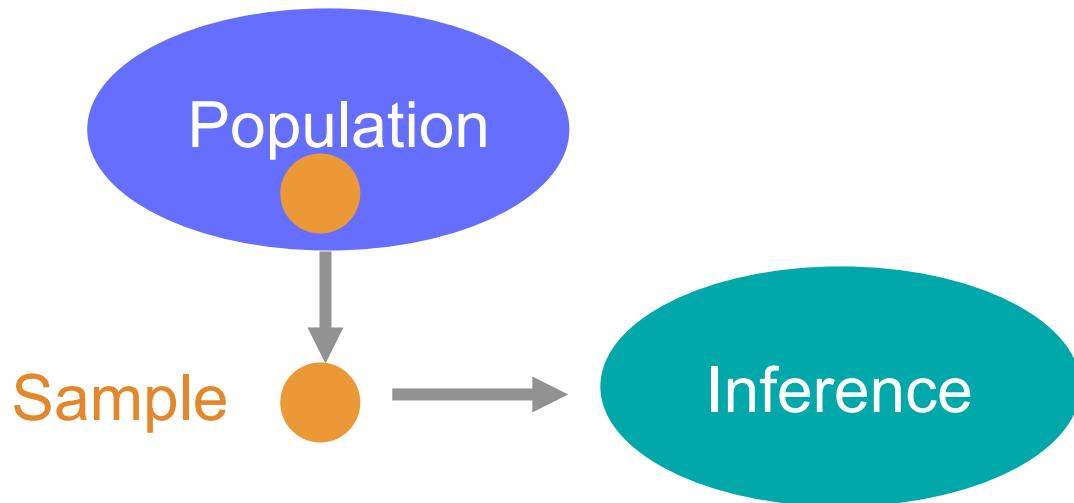
- ED increases the **validity** of your experiment.
- **Internal Validity:** study's ability to determine if a **causal relationship** exists between variables.
- **External Validity:** refers to the **generalisability** of a study.
- Threats to validity are **chance**, **bias** and **confounding**.



Researchers at MIT prove that rolling shopping carts will almost invariably hit the most expensive car in their vicinity.

*Without a valid design, valid scientific conclusions cannot be drawn.*

# Generalisability



- **Generalisability** is ensuring that your results apply to the population that you want to apply your results to.
- This requires a random **sample** of the population to ensure that the sample is representative of the population of interest.
- e.g. the results of a multicentre study will be more generalisable than a single centre study.

# More important now than ever!

Attention to the fundamentals of experimental design is more important now than ever before:

- **Cost** of experimentation.
- **Limited & Precious** material.
- **Immortalization** of data sets in public databases.
- **Ethical concerns** of experimentation.

# Animal Studies & the 3Rs

- The National Centre for Replacement Refinement & Reduction of Animals in research (**NC3Rs**)
  - UK Government, 2004
  - <http://www.nc3rs.org.uk/>
  - good scientific practice.
- Know the 3Rs:
  - **Replacement**: wherever possible live animals should be replaced by alternatives such as cell cultures, invertebrates or mathematical models.
  - **Refinement**: pain distress or lasting harm should be minimised.
  - **Reduction**: use the minimum number of animals consistent with achieving the objectives of the study.



National Centre for the Replacement, Refinement  
and Reduction of Animals in Research

# Animal Studies & the 3Rs

- the NC3R assessed the scope for improved experimental design, statistical analysis and reporting, and to further the implementation of the 3Rs.
- “*For scientific, ethical and economic reasons, experiments involving animals should be appropriately designed, correctly analyzed and transparently reported. This increases the scientific validity of the results and maximizes the knowledge gained from each experiment*”  
Kilkenny  
et al. PlosOne, 2009
- Yet a survey of 271 manuscript found **60% did not achieve this goal!**
- None justified the numbers they used, and in many cases the design of the experiments and/or the statistical analysis were inadequate or even incorrect.

# Animal Studies & the 3Rs

- 92% did not provide **raw data**.
- 87% did not report **random allocation** of subjects to treatments
- 86% did not report “**blinding**” where it seemed to be appropriate
- 100% failed to justify the **sample sizes** used
- 5% did not clearly state the **purpose** of the study
- 6% did not indicate how many **separate experiments** were done
- 13% did not identify the **experimental unit**
- 26% failed to state the **sex** of the animals
- 24% reported neither **age** nor **weight** of animals
- 4% did not mention the **number of animals** used
- 35% which reported numbers used, these **differed** in the materials and methods and the results sections

## Good Experimental Design

- Improves **quality & validity** of your science
- Saves **time & money**
- Obtains **meaningful** results
- Gets published in **higher impact journals**

## Good Science

- *Be clear about the shortcomings of your own experiment.*
- *If you need to take short-cuts, be clear about the caveats.*

# Experimental Design

- Importance
- **Planning**
- Well-designed Experiment
- What Type of Experiment
- Define Variables & Good Hypotheses
- Standardise
- Experimental Controls
- Experimental Units
- Bias & Confounding Factors
- Replicates & Sample Size
- Common Experimental Set-ups at CRI

# Value of Planning

Rushing into experiments without thoughtful planning invites failure.

**“Seventy percent of whether your experiment will work is determined before you touch the first test tube”**

Tung-Tien Sun (2004).

Excessive trust in authorities and its influence on experimental design.

*Nature Reviews Molecular Cell Biology*

# Value of Strategic Planning

- Strategic planning leads to '**best**' scientific practice.
- Have **clearly stated objectives**.
- Involve both the **Analyst & Statistician** while you are planning.
- Involve a **technology expert** (e.g. Genomics/ Proteomics Core Facility)
- Plan the experimental design and analysis methods **before you collect data** / process samples.
- You are going to make **research decisions** based on the results!

# Strategic Planning

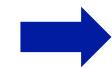
- Specify the experimental **objectives**.
- Determine **type of experiment**.
- Define **variables** & develop **hypotheses**.
- **Standardise** & identify the **experimental units**.
- Plan a **powerful & unbiased** experiment.
- Specify **statistical analysis**.



**Conduct Experiment and Collect Data**



**Analyse & Interpret Results**

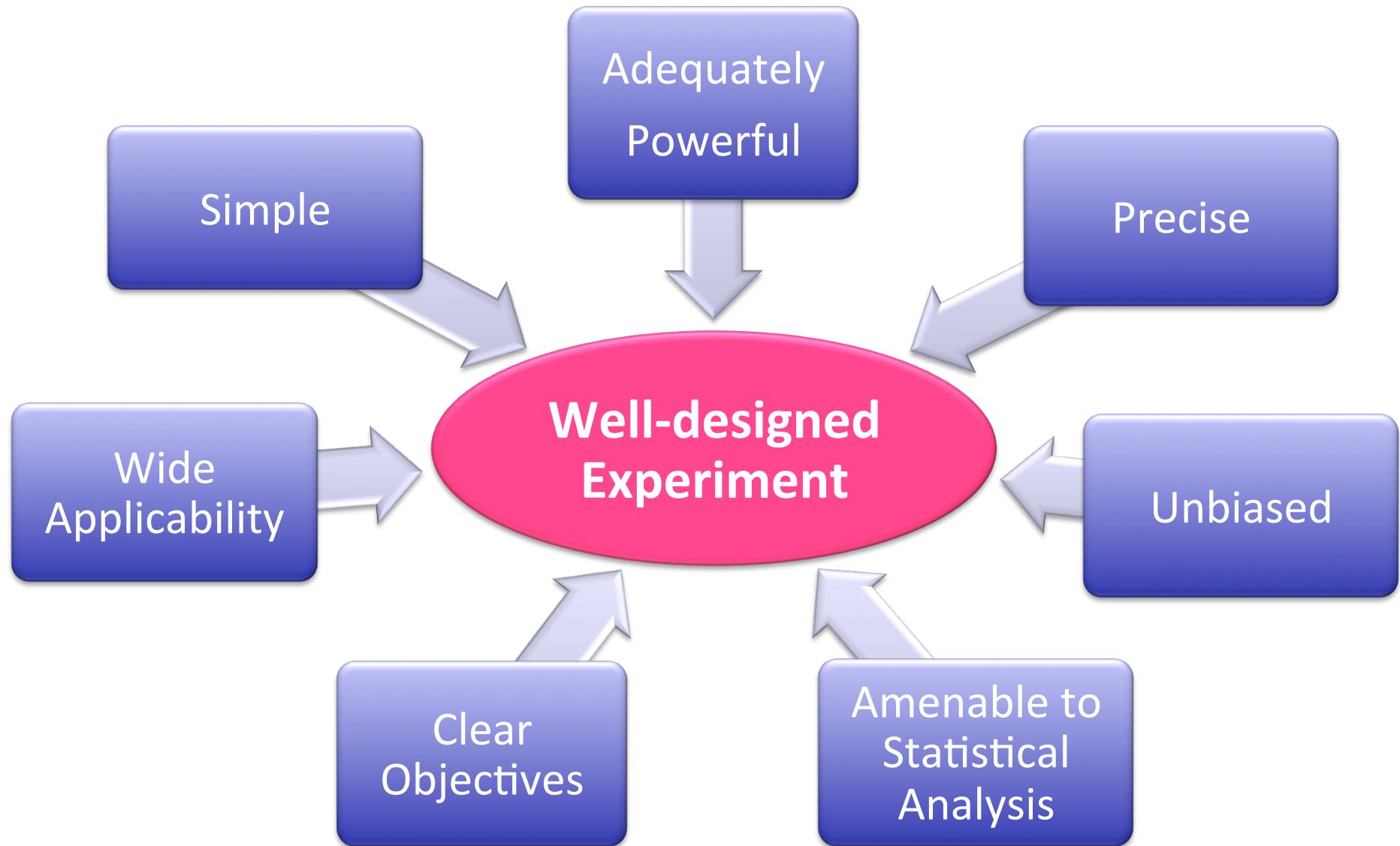


**Publish**

# Experimental Design

- Importance
- Planning
- **Well-designed Experiment**
- What Type of Experiment
- Define Variables & Good Hypotheses
- Standardise
- Experimental Units
- Experimental Controls
- Bias & Confounding Factors
- Replicates & Sample Size
- Common Experimental Set-ups at CRI

# Well-designed Experiment



# Clear Objectives

- **Information gathering:**
  - Why are we doing this?
  - What do we already know?
  - What are the biological objectives?
  - What is the treatment?
  - How many treatment levels?
  - What and how are you going to measure?
  - What technology?
- **What you are assessing mathematically?**
  - ie are you comparing the means of two groups, proportions or variance, or are you studying a relationship?
- Define **variables** and develop a **good hypothesis**.



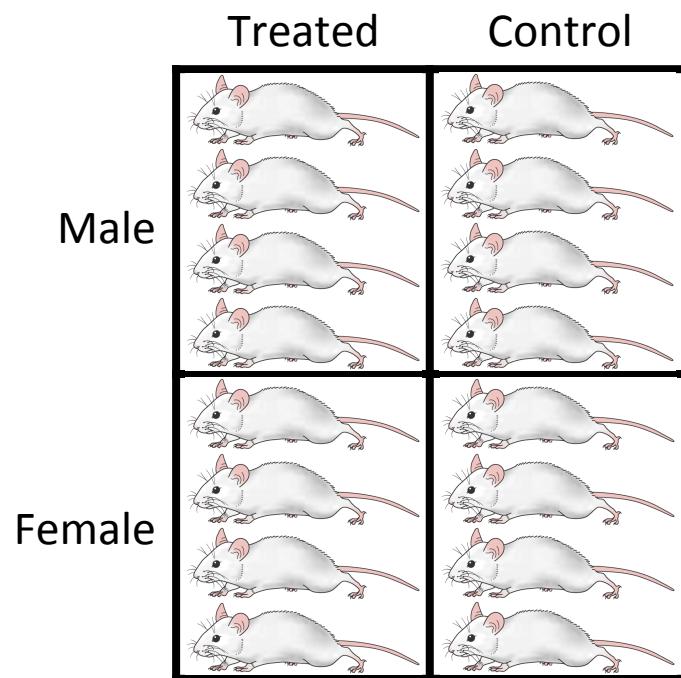
# Applicability

- What is the **scope/breadth** of your experiment?  
Should you broaden it?
- It is often useful to find out whether similar results are obtained in males and females, in different strains, or as a result of different diets or environments etc.
- These effects can be studied efficiently using **factorial experimental designs**:
  - used to investigate the effect of **multiple factors** (2 or more) simultaneously.

# Factorial Experimental Design

e.g. the effect of a drug on both males and females without doing two separate experiments or using twice as many animals.

- This would show whether or not the two sexes will respond in the same way or not.
- This is not possible if the two sexes are used in different experiments.



2 x 2

# Simple



- Complicated experiments can lead to **mistakes** or the statistical analysis becomes unduly complex.
- e.g. If you do need to process a lot of samples, would it be logically feasible to do this all in the time that you have without introducing any **error**?

# Adequately Powered

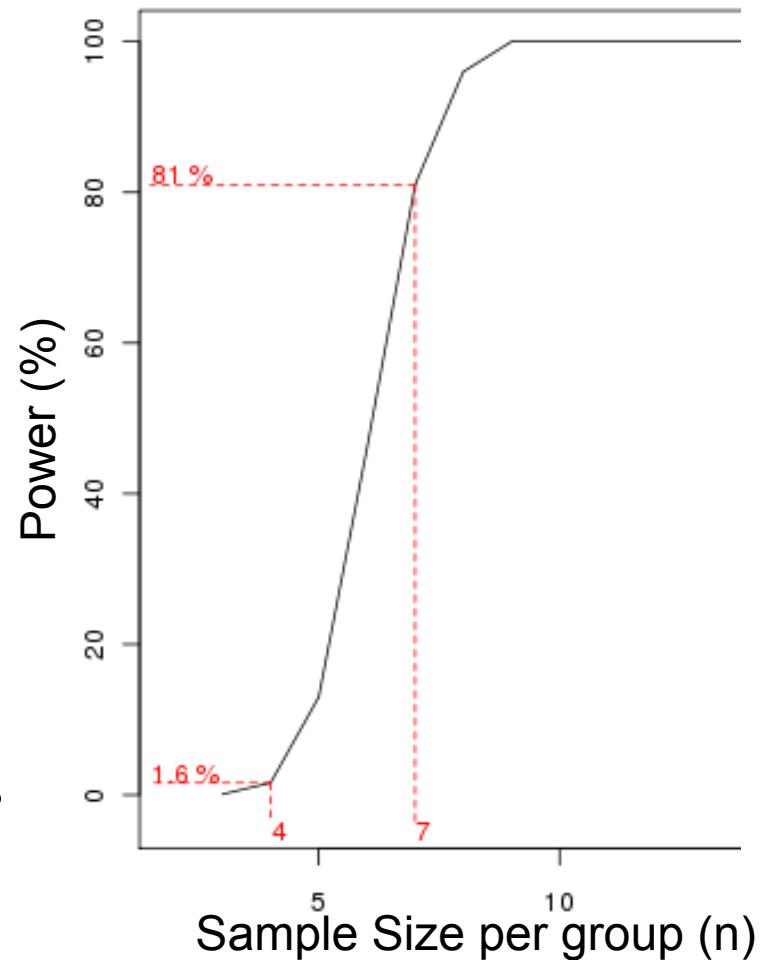
*The power of an experiment is the **probability** that it can detect an effect, if it is present.*

- Power is often **overlooked**.
- A **probability** - any value between 0% and 100%.
- Achieved by:
  - Using **appropriate numbers** of animals  
(sample size)
  - Controlling **sources of variation**  
(e.g. by standardisation)

If you increase the **variability** when you increase the size then it won't necessarily have more power.

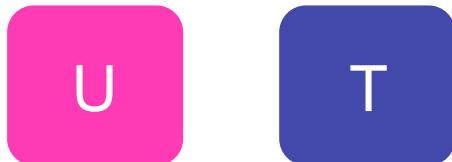
# Adequately Powered

- **Too many samples:**
  - wastes resources e.g. animals, money, time and effort, and is **unethical**.
- **Too few samples:**
  - may lack power and miss a scientifically important effect.
  - Also wastes resources and is **unethical**.



# Which experiment?

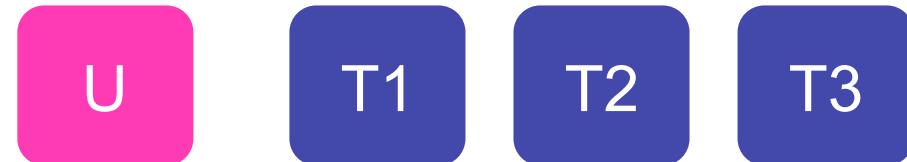
**Expt A**



$n = 12$  / group  
Total = 24

Power to detect 2 fold change =  
**0.70** ie **70%** of the time  
**(most of the time)**

**Expt B**



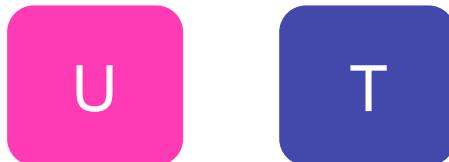
$n = 6$  / group  
Total = 24

Power to detect 2 fold change =  
**0.35** ie **35%** of the time  
**(some of the time)**

Options: **decrease scope** (increase n) or **increase scope** (decrease n)  
ie focus or expand?

# Which experiment?

**Expt A**



$n = 12 / \text{group}$   
Total = 24

Power to detect 2 fold change =  
**0.70** ie **70%** of the time  
(**most** of the time)

**Expt B**



$n = 6 / \text{group}$   
Total = 24

Power to detect 2 fold change =  
**0.35** ie **35%** of the time  
(**some** of the time)

But if **increase scope**, should really get  
**more resources** to maintain power!

# Precise & Unbiased

## PRECISION:

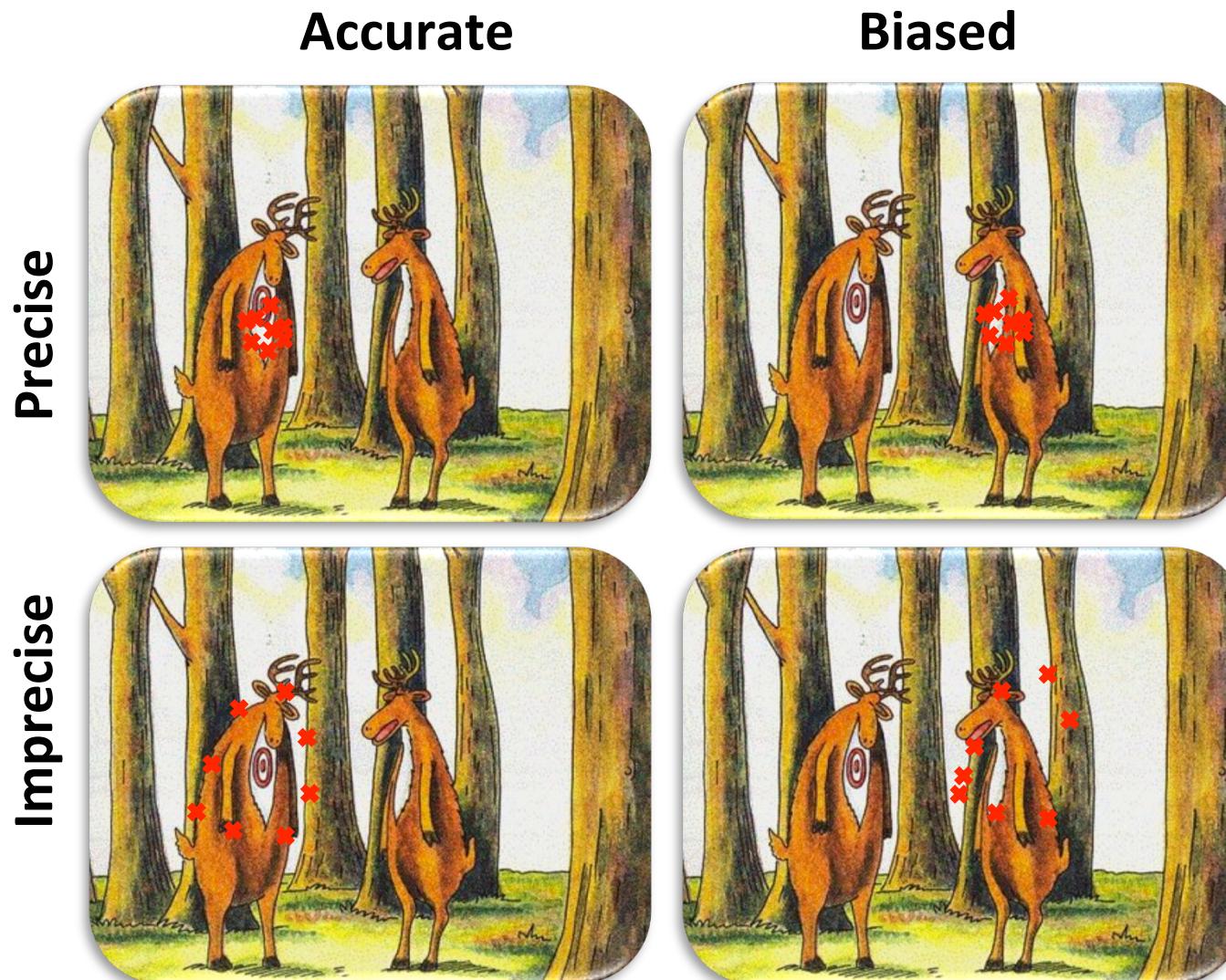
- **Reproducibility** of repeated measurements.
- **Precise estimation** of the quantity of interest.
- *Random variation (chance)* leads to results being **imprecise**.



## UNBIASED:

- Bias can affect **accuracy**.
- Should control for **systematic differences** between the measure and some “true” value (target).
- **Doesn’t confound** that estimate with a technical effect.
- *Systematic variation (bias)* leads to results being **inaccurate**.

# Accuracy & Precision



Very difficult to get both accuracy & precision - normally it's a trade-off.

# Amenable to Statistical Analysis

*An investigator should never start an experiment without knowing how it is going to be analysed.*

Decide on the following **before** you start the experiment:

- The **type of data** to be collected (e.g. measurement, qualitative, binary etc.)
- The **method of statistical analysis** of the resulting data.

Involve **the analyst and statistician**.



# Experimental Design

- Importance
- Planning
- Well-designed Experiment
- **What Type of Experiment**
- Define Variables & Good Hypotheses
- Standardise
- Experimental Controls
- Experimental Units
- Bias & Confounding Factors
- Replicates & Sample Size
- Common Experimental Set-ups at CRI

# What Type of Experiment?

- **Pilot?**
  - Small study to test the logistics of proposed larger study,
  - Gain some information on variability
- **Exploratory?**
  - Used to generate data with which to develop hypotheses for future testing.
  - No clearly stated hypothesis: may “work” or “not work”.
- **Confirmatory?**
  - Used to test some relatively simple hypothesis stated *a priori*.

# Confirmatory Experiments

- **True experimental design** is regarded as the most accurate form of experimental research.
- Involve **comparisons** between two or more sample groups.
- Their aim is to *test a “null hypothesis”*.
  - eg “there is no difference in tumour growth when comparing drug treatment with control”.
- Only type of research that is accepted by all disciplines as **statistically provable**.

Remember **Fisher's fundamental experimental design principles!!!**  
**(Replication, Randomisation, Blocking)**

# Experimental Design

- Importance
- Planning
- Well-designed Experiment
- What Type of Experiment
- **Define Variables & Good Hypotheses**
- Standardise
- Experimental Controls
- Experimental Units
- Bias & Confounding Factors
- Replicates & Sample Size
- Common Experimental Set-ups at CRI

# Variables & Types

- Any factor that can take on different values is a scientific variable and **influences the outcome** of experimental research.
  - eg time, weight, drug, gender, ethnic group, country, plate, cage etc
- If the values of a variable are in distinct categories they are called **levels**.
  - eg the varying doses of each drug
  - eg gender can be either M or F

# Variables & Types

What type of measurement?

- Categorical (**nominal**) , e.g. gender
- Categorical with ordering (**ordinal**), e.g. tumour grade
- **Discrete**, e.g. shoe size
- **Continuous**, e.g. body weight in kg, height in cm

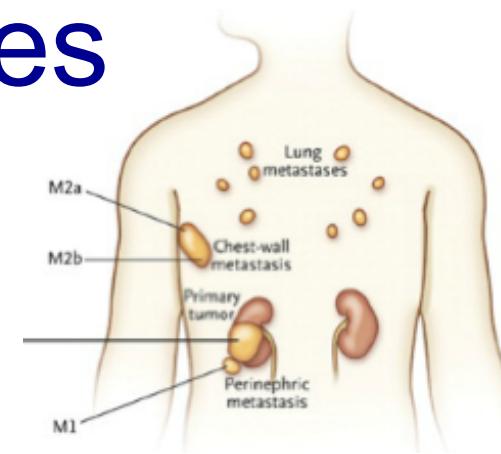
# Related Variables

Are the variables **related**?

*ie paired, repeated or matched variables*

Examples:

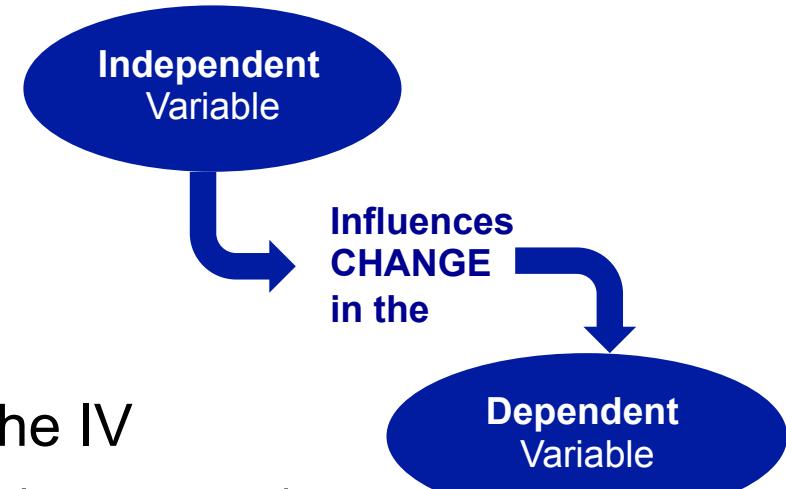
- Multiple samples taken from the **same cell-line**.
- Multiple samples from the **same patient**:
  - Multiple metastatic and primary tumours from the same patient.
  - Matched normal and tumour samples from the same patient.
  - Before and after treatment samples from the same patient.
- Paired mice from the **same litter**, one given the drug and the other the vehicle control.



**Paired experimental design** is more powerful than unpaired experimental design because the differences between individuals are factored out in the analysis.

# Independent & Dependent Variables

- **Variables:**
  - Independent Variable (IV)
    - What you change
  - **Dependent** variable (DV):
    - What changes because of the IV  
(ie what you are measuring/observing)
- **Hypothesis:**
  - A proposed explanation for a phenomenon e.g.:
    - “Smoking, increases the risk of lung cancer”.
    - “Interleukin (IL)-6 gene knock-out, increases tumour growth”.
  - “**If (independent** variable), **then (dependent** variable)”



# Experimental Design

- Importance
- Planning
- Well-designed Experiment
- What Type of Experiment
- Define Variables & Good Hypotheses
- **Standardise**
- Experimental Units
- Experimental Controls
- Bias & Confounding Factors
- Replicates & Sample Size
- Common Experimental Set-ups at CRI

# Standardise



- It is essential to '**compare apples to apples**' in order to arrive at legitimate conclusions.
- **Controlling variation** is vital:  
the more **uniform** the subjects are within a treatment group or sub-group, the fewer of them will be needed, or the greater the **power** of the experiment.

# Standardise

Standardise **subjects & procedures** and to achieve uniformity:

- Sample groups must contain **similar subjects**:
  - eg similar age and weight of your mice in both the control wild-type and treated groups.
- Each subject in each sample group must receive the **same treatment under the same conditions**:
  - e.g a drug must be delivered in the same way to both your control and treated groups.

# Experimental Design

- Importance
- Planning
- Well-designed Experiment
- What Type of Experiment
- Define Variables & Good Hypotheses
- Standardise
- **Experimental Controls**
- Experimental Units
- Bias & Confounding Factors
- Replicates & Sample Size
- Common Experimental Set-ups at CRI

# Experiment Controls

- A very important part in your experiment, because it's very difficult to eliminate all of the possible confounding variables & bias.
- Designing the experiment with controls in mind is often more crucial than determining the independent variable.
- Increases the statistical validity of your data.
- Two types: **positive** and **negative**.

# Experiment Controls

- **Positive controls**

- Check the procedure is effective in observing the effect ie the set-up is capable of producing results.
- If the control fails, then there is something wrong with the experimental set-up.
- Reduces the chances of false negatives.
- e.g. use an established antibody that's known to work.

- **Negative controls**

- Make sure that no confounding variable has affected the results or take into account any likely source of bias.
- Uses a sample not expected to work.
- Control for false positives
- e.g. mock siRNA as a control for silencing a gene etc

# Experimental Controls

- **Sham control or *placebo***

**Mimic** a procedure or treatment without the actual use of the procedure or test substance.

- e.g. same surgical procedure but without X implanted



- **Vehicle control**

Used in studies where a substance is used to deliver an experimental compound.

e.g. apply EtOH to cell lines on its own as a control since it's used as a vehicle for delivering the Tamoxifen drug.

# Experimental Design

- Importance
- Planning
- Well-designed Experiment
- What Type of Experiment
- Define Variables & Good Hypotheses
- Standardise
- Experimental Controls
- **Experimental Units**
- Bias & Confounding Factors
- Replicates & Sample Size
- Common Experimental Set-ups at CRI

# Experimental Units

- You are experimenting on **units** or subjects (e.g. humans, mice, cages of mice, litter of mice, cell lines).
- The **experimental unit** is the physical entity which can be assigned, at random, to a an experimental condition or treatment.
- The **unit of randomisation** and the **unit of statistical analysis** when comparing groups.
- Commonly it is an individual animal but this is not always the case.

# Experimental Units

The Experimental Unit could be:

- The individual subject/animal
- The cage
- An animal for a period of time
- The breeding female & litter
- Part of animal

See **NC3Rs** for more details:

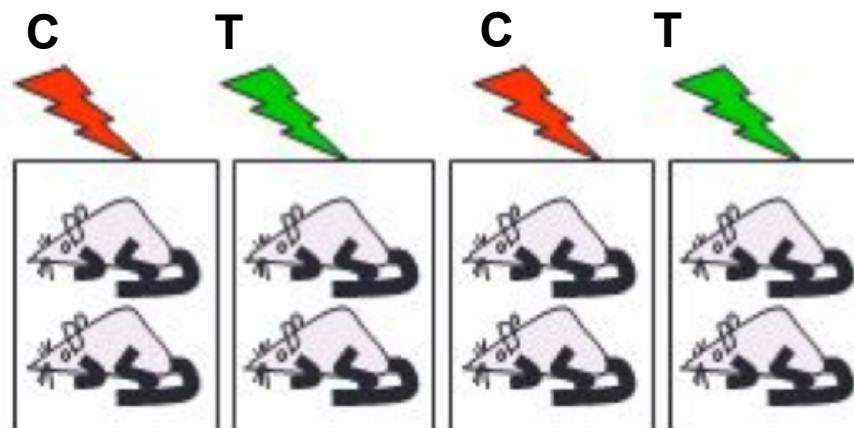
[http://isogenic.info/html/6\\_experimental\\_unit.html](http://isogenic.info/html/6_experimental_unit.html)

# Experimental Units

Example 1:

Animals are housed two per cage and the treatment (C or T) is given in the food or water.

What do you think is “N”, the total number of experimental units in this case?

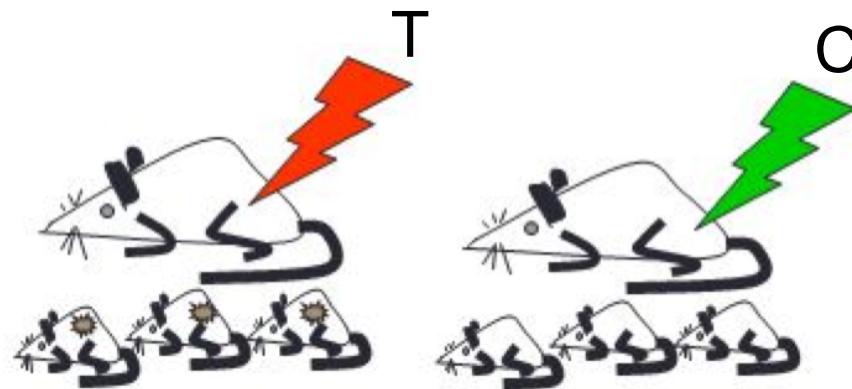


# Experimental Units

Example 2:

Pregnant female is treated with the test compound (T) or a placebo (C).

The pregnant females are killed at about mid-gestation and the pups are weighed, measured and studied for abnormalities.



How many experimental units are there in the experiment?

# Experimental Design

- Importance
- Planning
- Well-designed Experiment
- What Type of Experiment
- Define Variables & Good Hypotheses
- Standardise
- Experimental Controls
- Experimental Units
- **Bias & Confounding Factors**
- Replicates & Sample Size
- Common Experimental Set-ups at CRI

***“A biased scientific result is no different from a useless one”***

Daniel Sarewitz

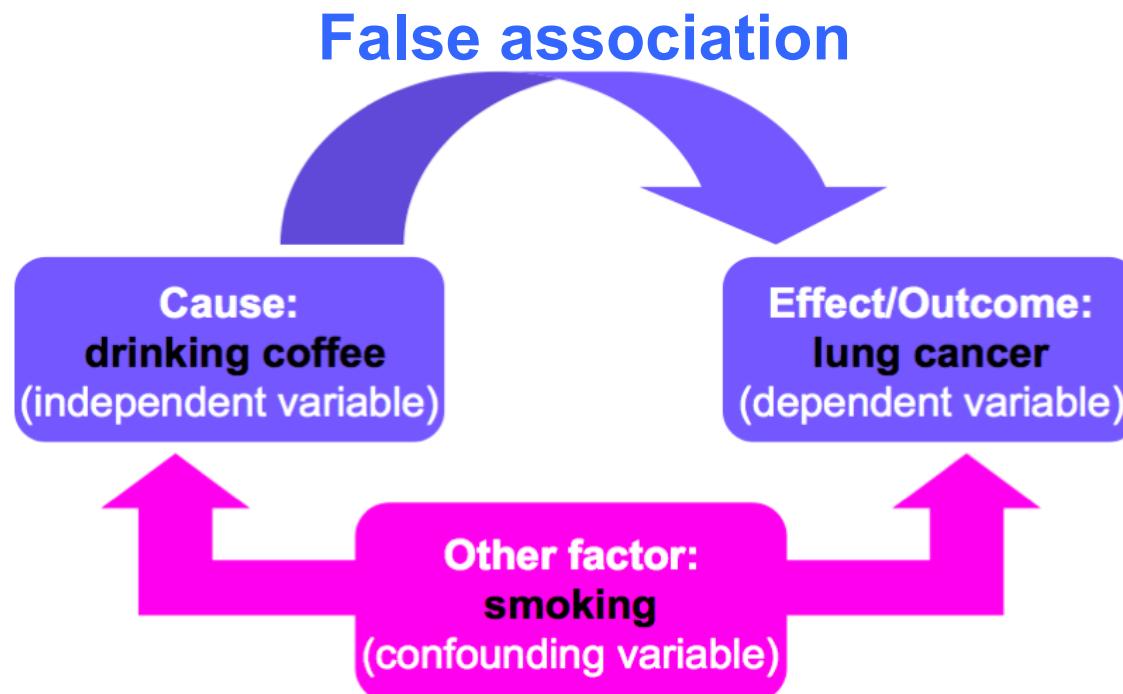
(Beware the creeping cracks of bias. *Nature*, 2012)

# Various Forms of Bias

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

# Confounding factors

- Also known as **extraneous**, **hidden**, **lurking** or **masking** factors, or the **third variable** or **mediator variable**.
- May mask an actual association or **falsely** demonstrate an apparent association between the independent & dependent variables.
- Hypothetical Example would be a study of coffee drinking and lung cancer.



# Confounding factors

- Other examples:
  - Democrats were less satisfied with their sex lives than Republicans.  
(ABC poll report).
  - Slightly overweight people live longer than thin people  
(US Centre for Disease Control).
- **Inadequate management and monitoring of confounding factors**
  - one of the most common causes of researchers wrongly assuming that a correlation leads to a causality.
- Confounding factors - obvious with hind-sight but often missed through lack of thought or money etc.

# Identify what the other Influences are

- You only want the the **independent variable** to influence the results.
- Identify what **other factors/variables** could influence the dependent variable.
- Comes from the **investigator** and their knowledge about the data or process.
  - e.g. gender, location, time, instrument, experimenter, batch etc
- Plan how this is going to be **managed**:
  - remember **Fisher's** fundamental experimental design principles!!!

# Managing Confounding Factors

- **Bias is avoided by:**
  - Correct selection of experimental units.
  - Randomisation of the experimental units.
  - Randomisation of the order in which measurements are made.
  - “Blinding” and the use of coded samples.
- **Failure to randomise and blind can lead to false positive & negative results!!!**

# “Food colouring affects childrens’ behaviour”



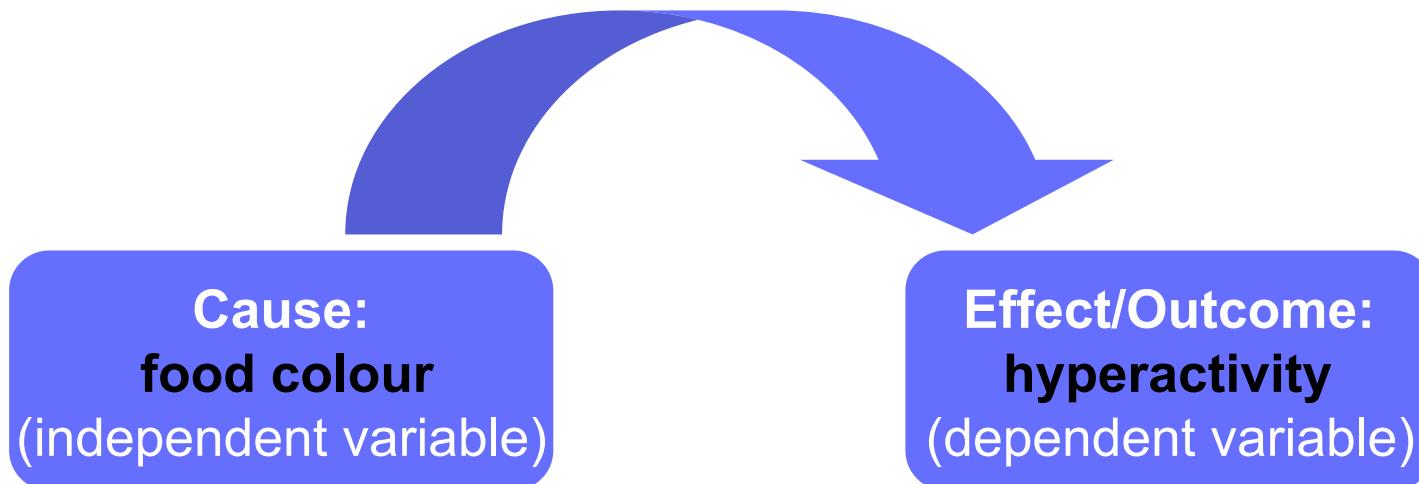
Food with artificial  
colouring (junk food)



Food with no artificial  
colouring (healthy food)

# Confounding factors:

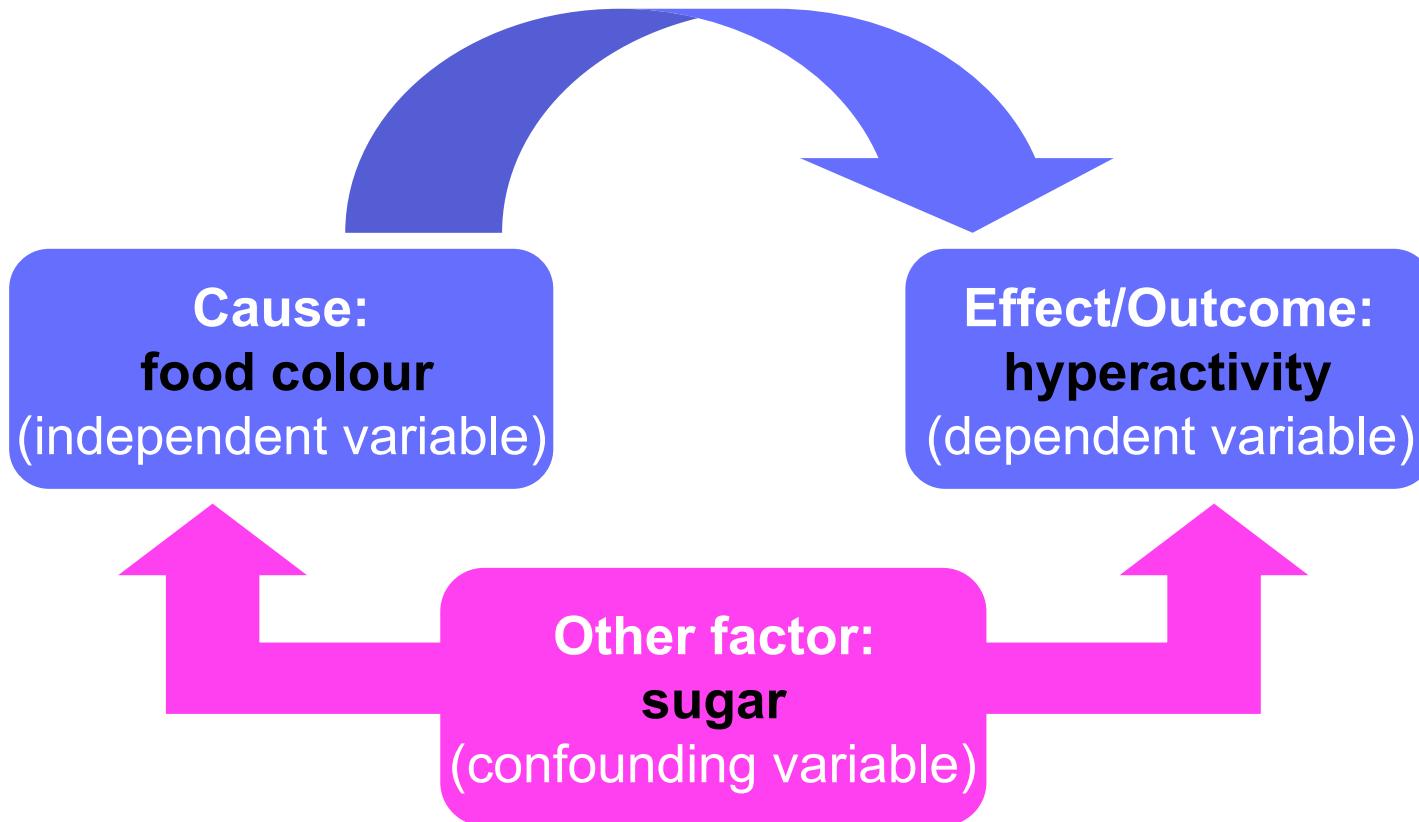
example 1



# Confounding factors:

example 1

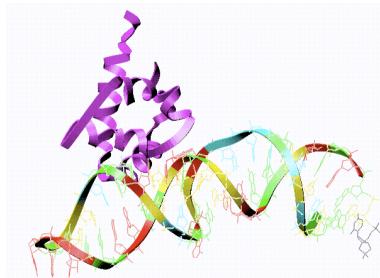
## False association



Confounding variables are variables that the researcher **failed to control**, or eliminate, damaging the internal validity of an experiment.

# Confounding factors:

example 2



RNA Extraction

Plate1

	1	2	3	4	5	6	7	8	9	10	11	12
A	○	○	○	○	○	○	○	○	○	○	○	○
B	○	○	○	○	○	○	○	○	○	○	○	○
C	○	○	○	○	○	○	○	○	○	○	○	○
D	○	○	○	○	○	○	○	○	○	○	○	○
E	○	○	○	○	○	○	○	○	○	○	○	○
F	○	○	○	○	○	○	○	○	○	○	○	○
G	○	○	○	○	○	○	○	○	○	○	○	○
H	○	○	○	○	○	○	○	○	○	○	○	○

Control

Plate2

	1	2	3	4	5	6	7	8	9	10	11	12
A	○	○	○	○	○	○	○	○	○	○	○	○
B	○	○	○	○	○	○	○	○	○	○	○	○
C	○	○	○	○	○	○	○	○	○	○	○	○
D	○	○	○	○	○	○	○	○	○	○	○	○
E	○	○	○	○	○	○	○	○	○	○	○	○
F	○	○	○	○	○	○	○	○	○	○	○	○
G	○	○	○	○	○	○	○	○	○	○	○	○
H	○	○	○	○	○	○	○	○	○	○	○	○

Treatment 1

Plate3

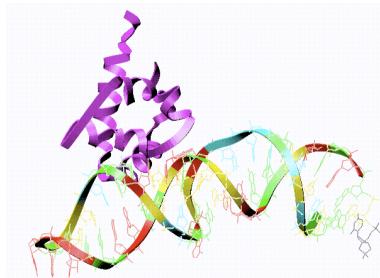
	1	2	3	4	5	6	7	8	9	10	11	12
A	○	○	○	○	○	○	○	○	○	○	○	○
B	○	○	○	○	○	○	○	○	○	○	○	○
C	○	○	○	○	○	○	○	○	○	○	○	○
D	○	○	○	○	○	○	○	○	○	○	○	○
E	○	○	○	○	○	○	○	○	○	○	○	○
F	○	○	○	○	○	○	○	○	○	○	○	○
G	○	○	○	○	○	○	○	○	○	○	○	○
H	○	○	○	○	○	○	○	○	○	○	○	○

Treatment 2

The difference between Control, Treatment 1  
and Treatment 2 is confounded by **Plate**

# Confounding factors:

example 2



RNA Extraction

Day1, Plate 1

	1	2	3	4	5	6	7	8	9	10	11	12
A	○											
B												
C												
D												
E												
F												
G												
H												

Control

Day2, Plate 2

	1	2	3	4	5	6	7	8	9	10	11	12
A	○											
B												
C												
D												
E												
F												
G												
H												

Treatment 1

Day3, Plate 3

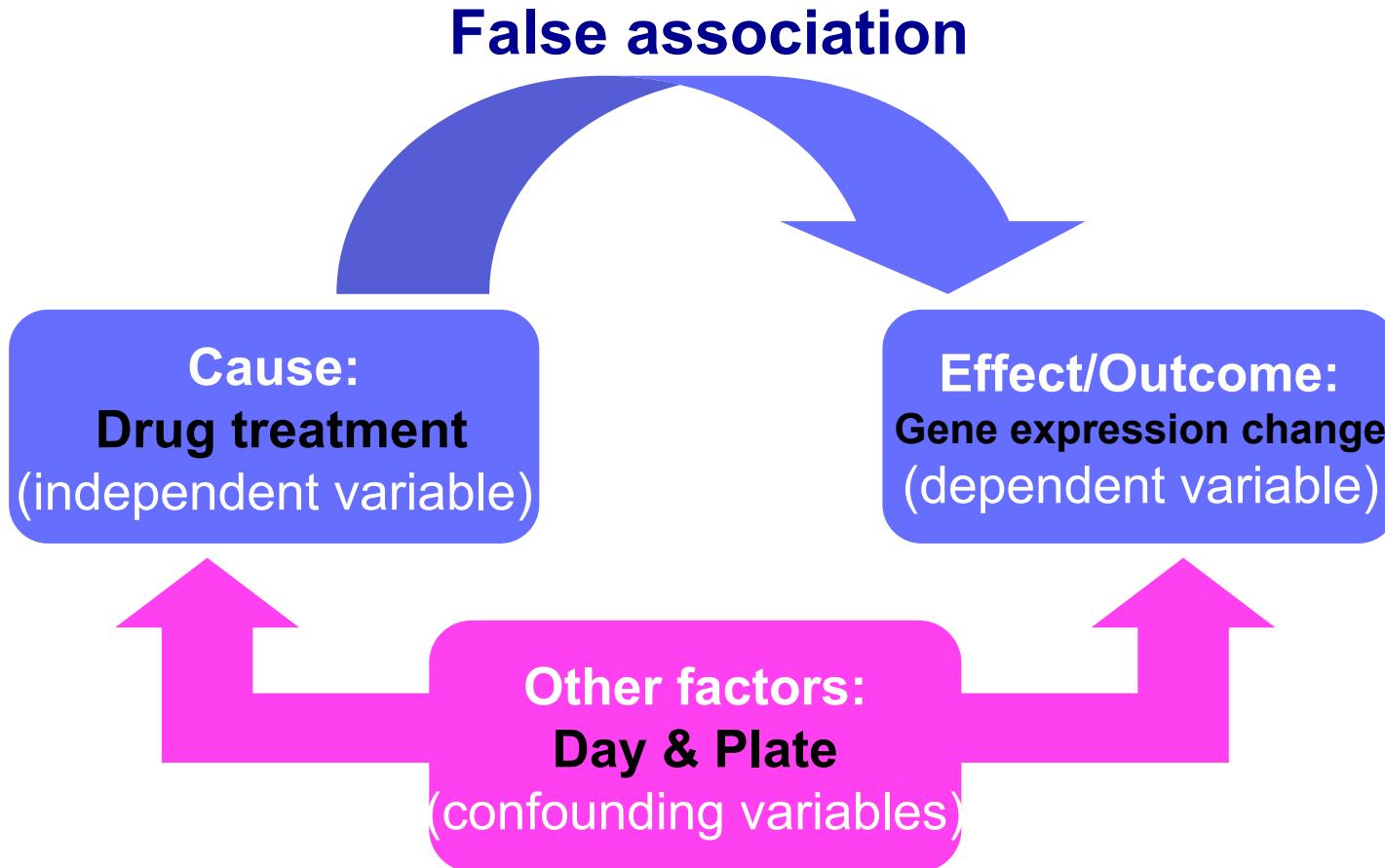
	1	2	3	4	5	6	7	8	9	10	11	12
A	○											
B												
C												
D												
E												
F												
G												
H												

Treatment 2

The difference between Control, Treatment 1  
and Treatment 2 is confounded by **day** and **plate**.

# Confounding factors:

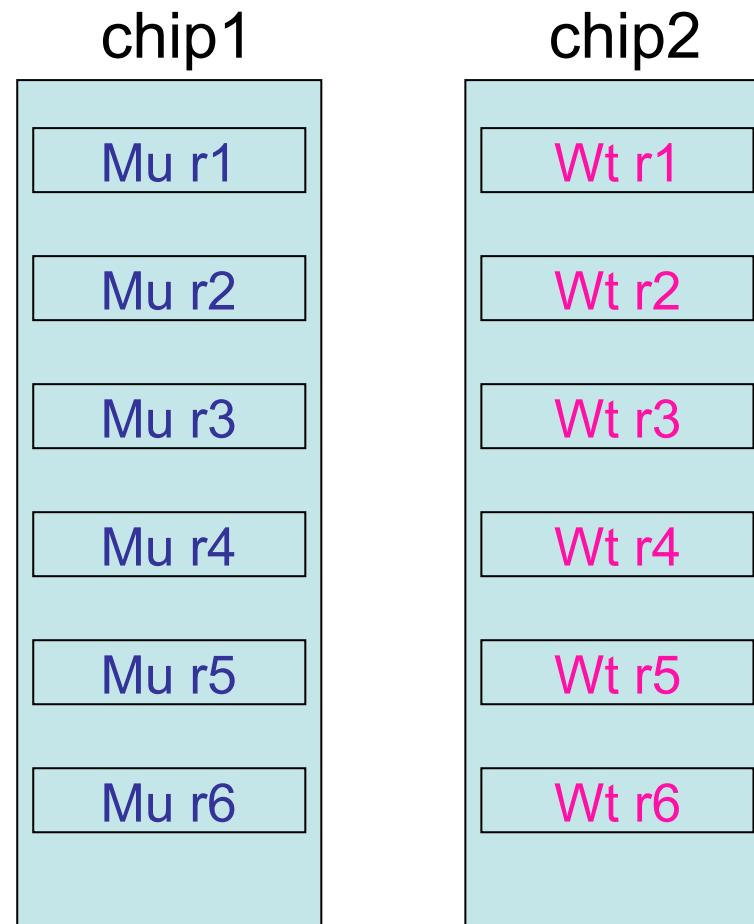
example 2



Confounding variables are variables that the researcher **failed to control**, or eliminate, damaging the internal validity of an experiment.

# Confounding factors:

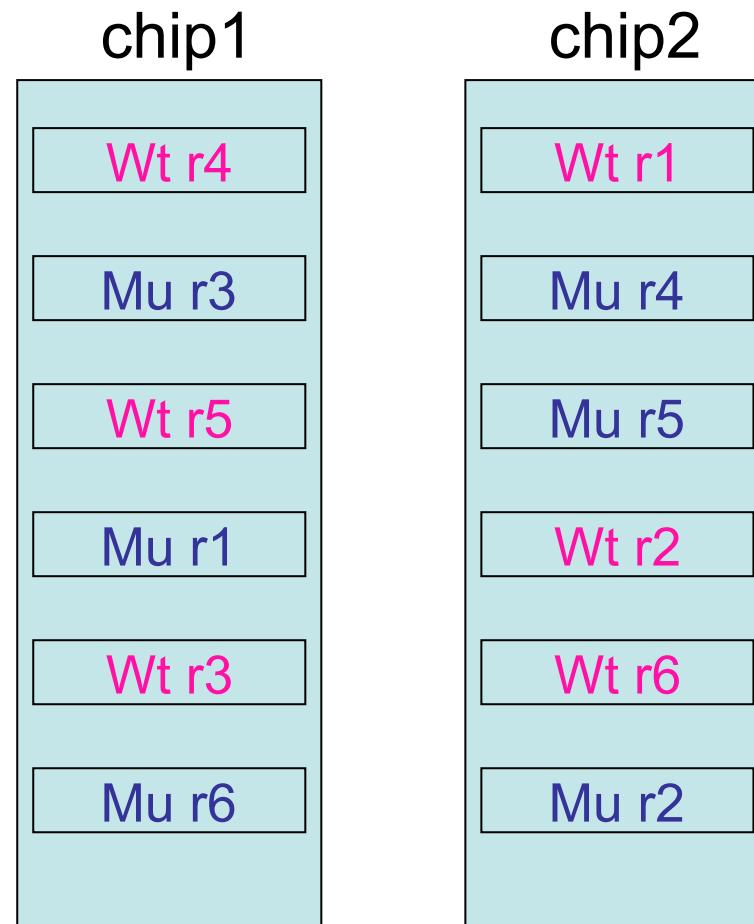
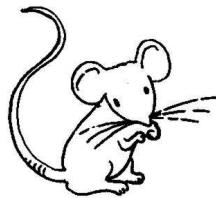
example 3



The difference between the two sample groups:  
'mutant' and 'wild type' is confounded by **chip**

# Confounding factors:

example 3



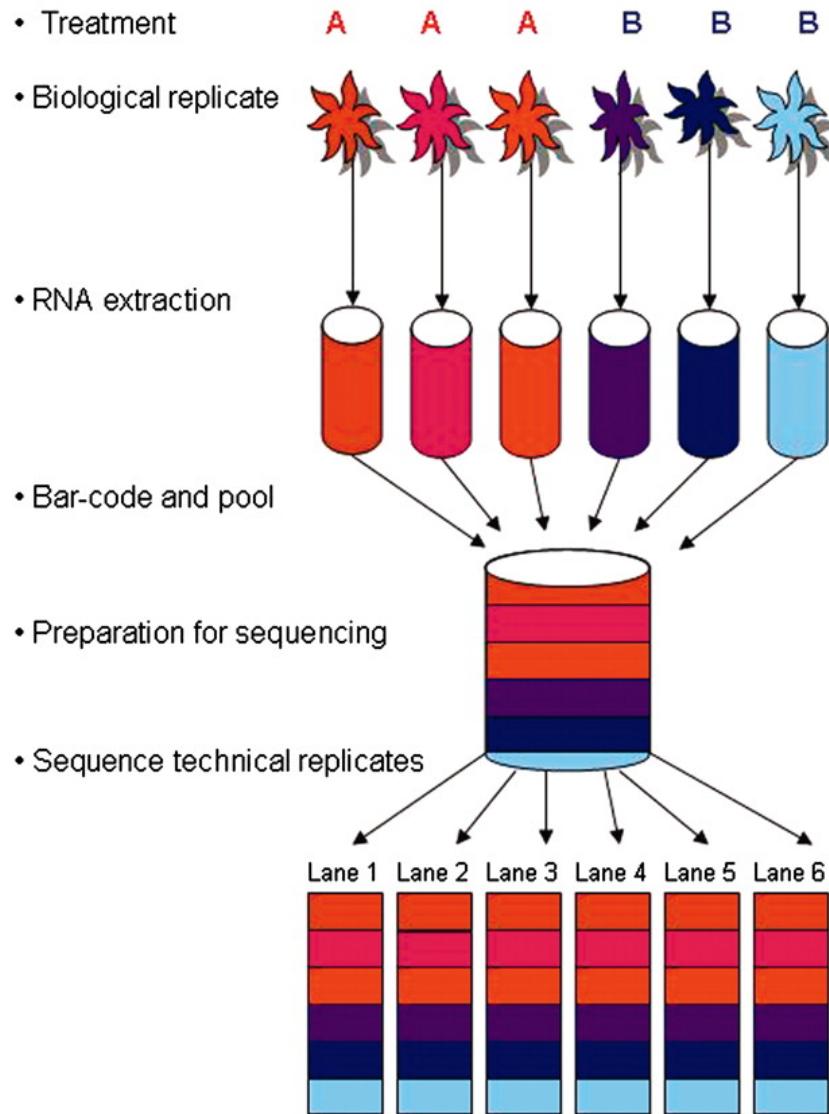
Balanced &  
Randomised

Randomly assign mice from each sample group into two chip groups,  
and then randomly assign the order on each chip

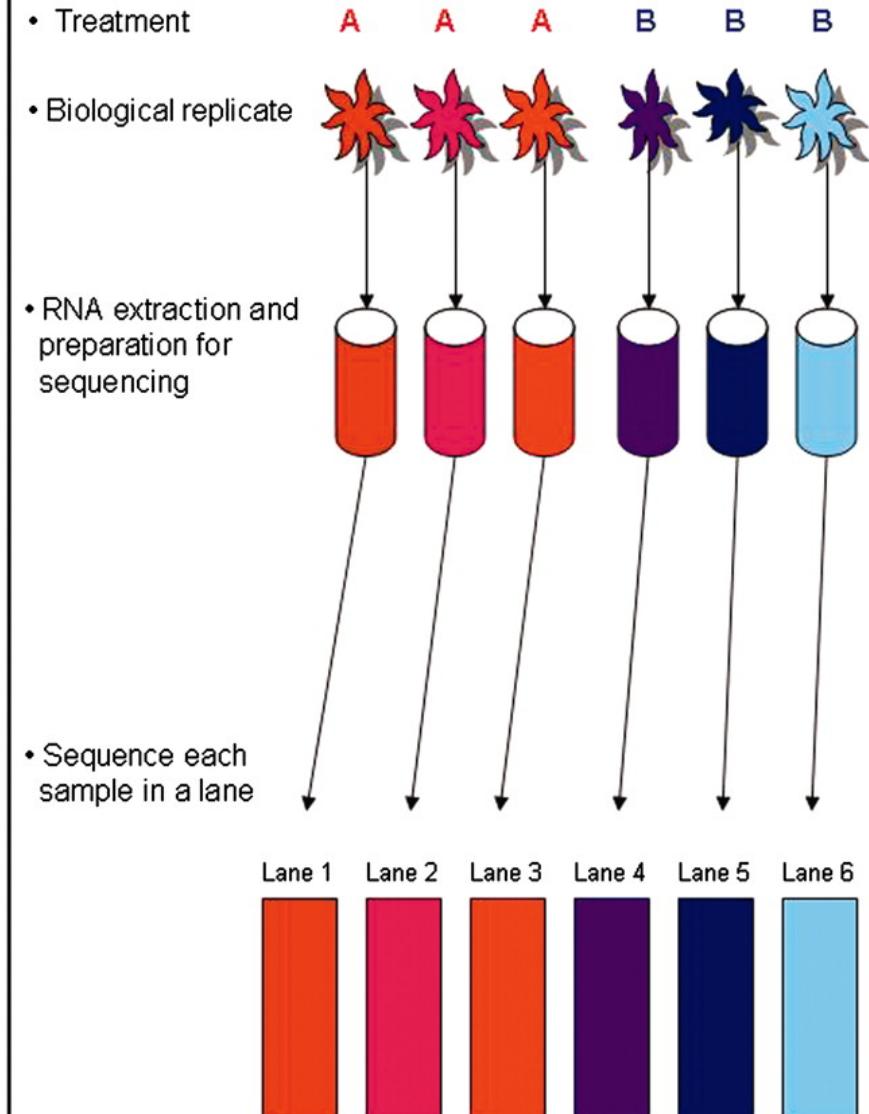
# Confounding factors

example 4: RNA-seq

## Balanced Blocked Design



## Confounded Design



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



Report

## Genetic Signatures of Exceptional Longevity in Humans

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

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

- A GWAS study of 800 centenarians against controls found 150 SNPs which can predict if a person is a centenarian with 77 % accuracy.
- Problem: they used **different SNP chips** for centenarian vs control.
- Retracted 2011 following an independent lab reviewed the data and QC applied.

# Managing potential confounders

## Randomisation:

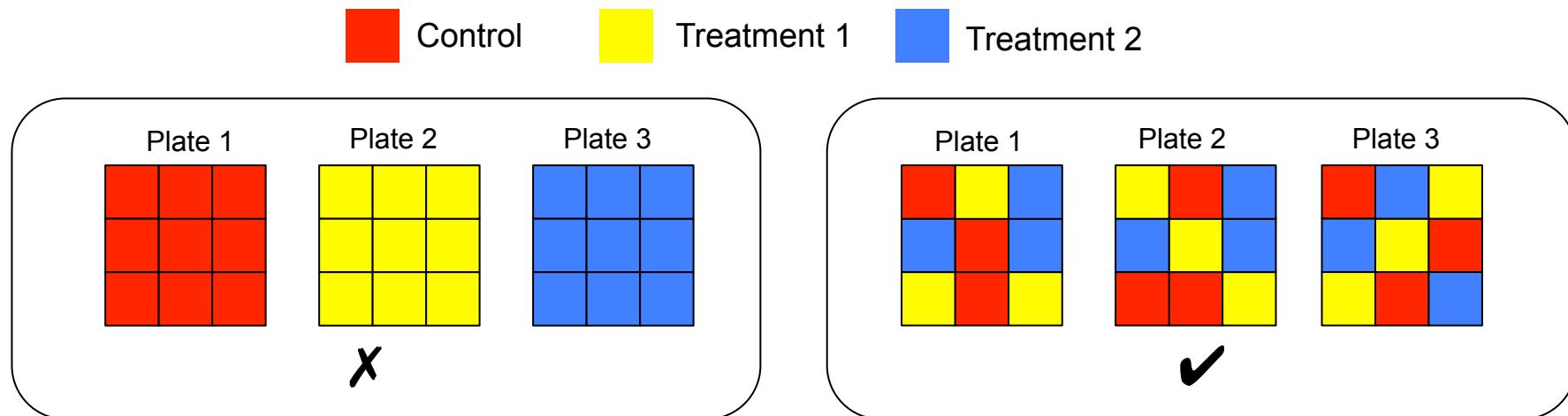
- ensures that each experimental unit has an equal probability of receiving a particular treatment.
- It reduces the chance of systematic differences between the treatment groups.

## Which Method?

- ✗ – alternate allocation
- ✗ – day of week
- ✗ – odd or even last digit of sample number.
- ✗ – Flip a coin
- ✓ – Random number generator program

# Randomised Block Design

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



- RBD across plates so that each plate contains **equal proportions** of:
  - Control
  - Treatment 1
  - Treatment 2eliminating spurious associations due to plate effects.

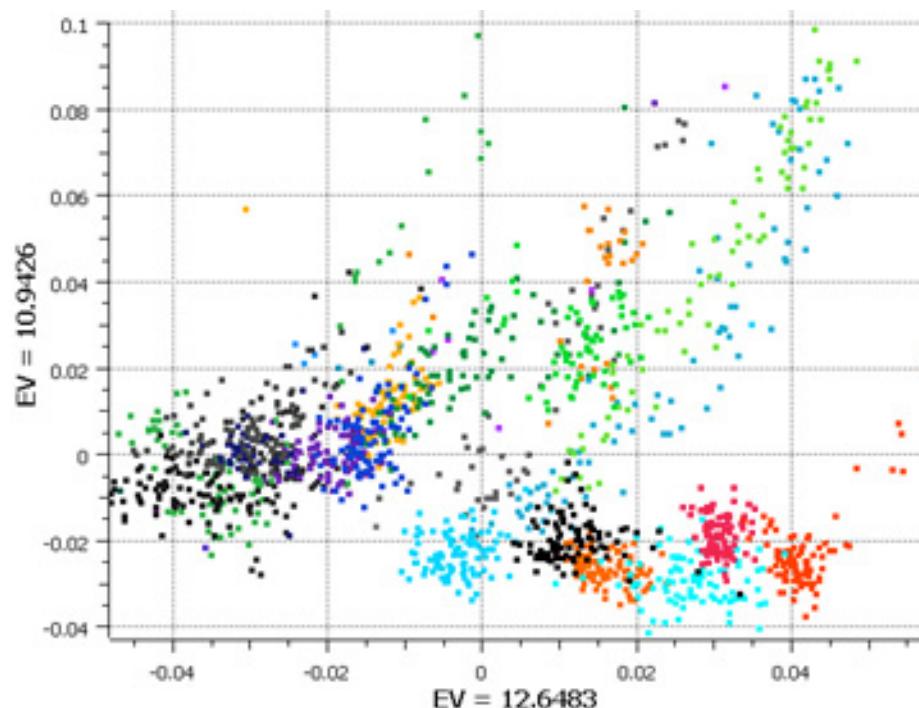
# Randomised Block Design

**Good** design example: Alzheimer' s study from GlaxoSmithKline

## Plate effects by *plate*

Left PCA plots show *large plate effects*.

Each colour corresponds to a different plate

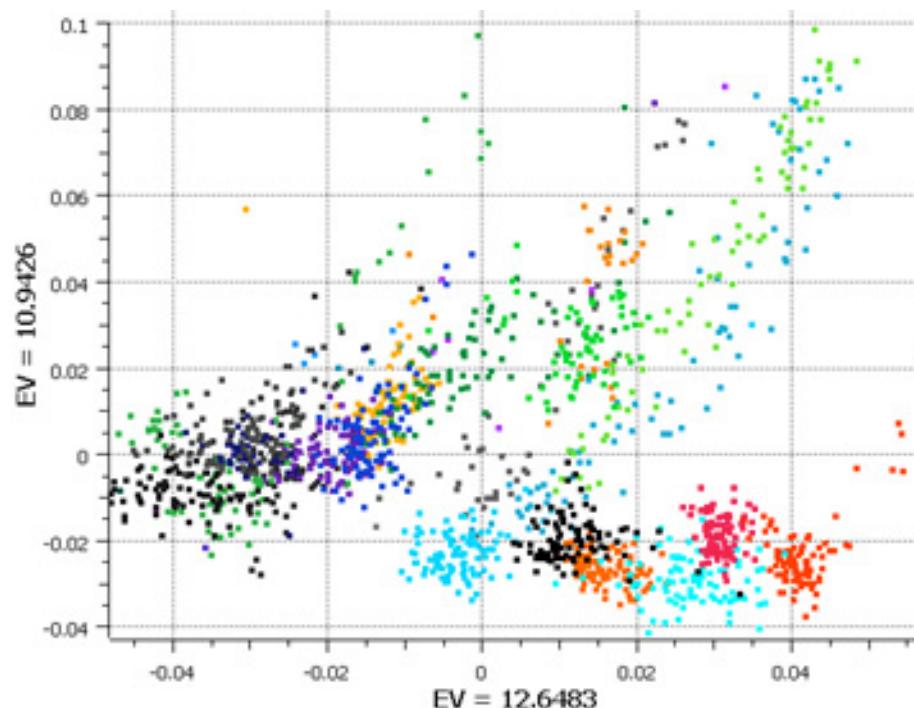


# Randomised Block Design

**Good** design example: Alzheimer' s study from GlaxoSmithKline

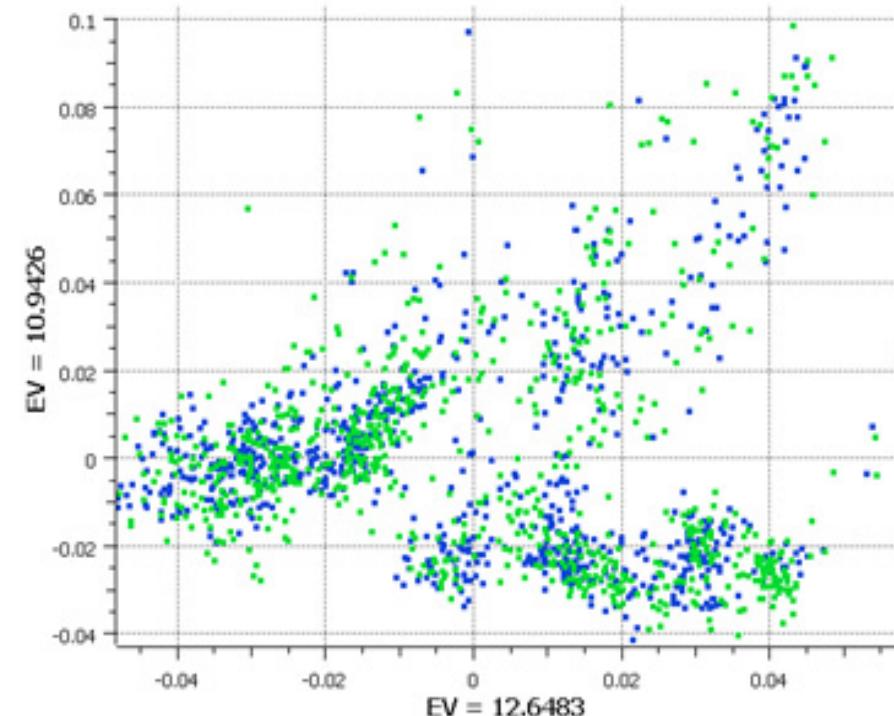
## Plate effects by plate

Left PCA plots 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).



# Randomised Block Design

- Example plate layout for a ~4,000-subject *Parkinson's genotyping study*: ~2,000 cases and ~2,000 controls arranged in 45 different plates.
- Note: it is key that the **randomisation is controlled** so that the blocks are almost perfectly balanced.

Block	Case Status		Site				DNA Extraction Method			Number
	Case	Control	1	2	3	4	1	2	3	
1	X		X				X			407
2	X		X					X		42
3	X		X						X	61
4	X			X				X		854
5	X				X			X		417
6	X					X		X		219
7		X	X				X			191
8		X	X					X		684
9		X	x					x		28
10		X		X				X		684
11		X			X			X		300
12		x				x		X		113

-----> 407 samples from EU 1 are evenly divided at Random among the 45 plates, resulting in either nine or ten samples per plate.



# Blinding

*When humans have to make observations there is always the possibility of bias!*

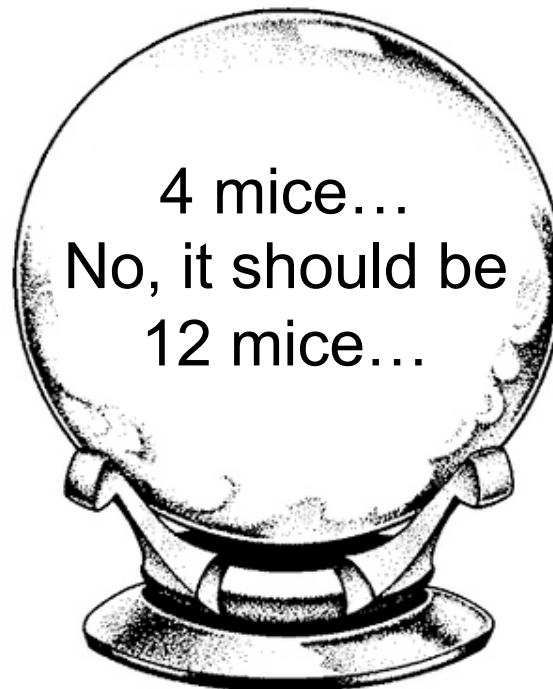
- It is possible, **unintentionally**, to bias the results towards that favoured outcome, leading to incorrect conclusions.
- Ideally, the **researcher should not know** which treatment the animal/patient has received.
- Reduces the risk of bias from a **placebo effect**,
  - where people believe they are receiving medicine and they show signs of improvement in health.
- There are some situations where blinding is **impossible**.
  - In this case reliance needs to be placed in **randomisation** of the order in which the animals are tested.

# Experimental Design

- Importance
- Planning
- Well-designed Experiment
- What Type of Experiment
- Define Variables & Good Hypotheses
- Standardise
- Experimental Controls
- Experimental Units
- Bias & Confounding Factors
- **Replicates & Sample Size**
- Common Experimental Set-ups at CRI

# How many replicates?

Look into my crystal ball.....



The sensitivity, ability or **power** to detect changes depends on the **sample size**.

# How many replicates?

Depends on the resources, the goals of the study, and the reliability of the technology:

- How much **money** do you have?
- Can you **handle** all these samples without problem?
- What size of differences (**effect size**) to detect?
- It's an **accurate representation** of the population?
- Large enough to achieve **meaningful** results?
- What variability (**noise**) in your system?

# How many replicates?

Larger Samples are needed when...

- Large number of **uncontrolled variables** are interacting unpredictably.
- Interested in also studying **subgroups** within the sample.
- Population is made up of a **wide range of variables** and characteristics.
- Differences in the results (**effect size**) are expected to be small.
- **High attrition** of subjects is expected.

# Biological or technical replicates?

Microarray Processing Workflow



Choose Samples

Biological Replication

Extract RNA

Technical Replication

Quality Control

Convert to cRNA

Technical Replication

Quality Control

Hybridization

Technical Replication

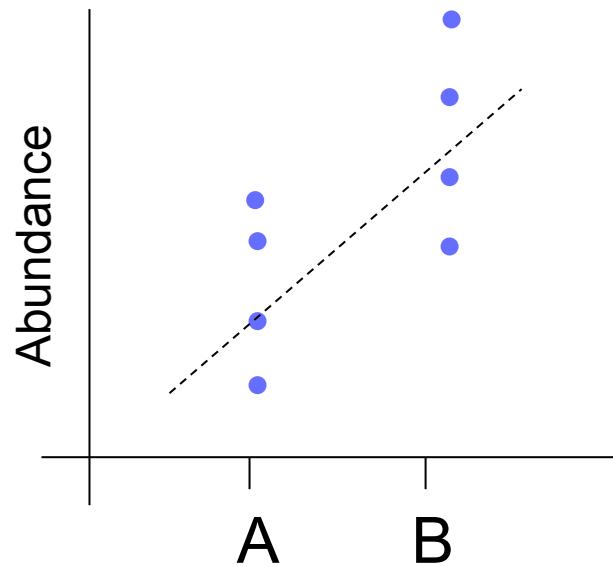
Washing

Scanning

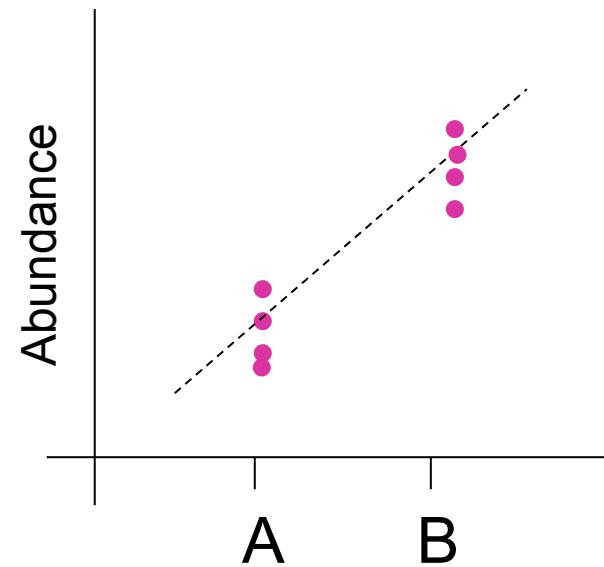
Technical Replication

Analysis

# Biological or technical replicates?



Biological Replicates

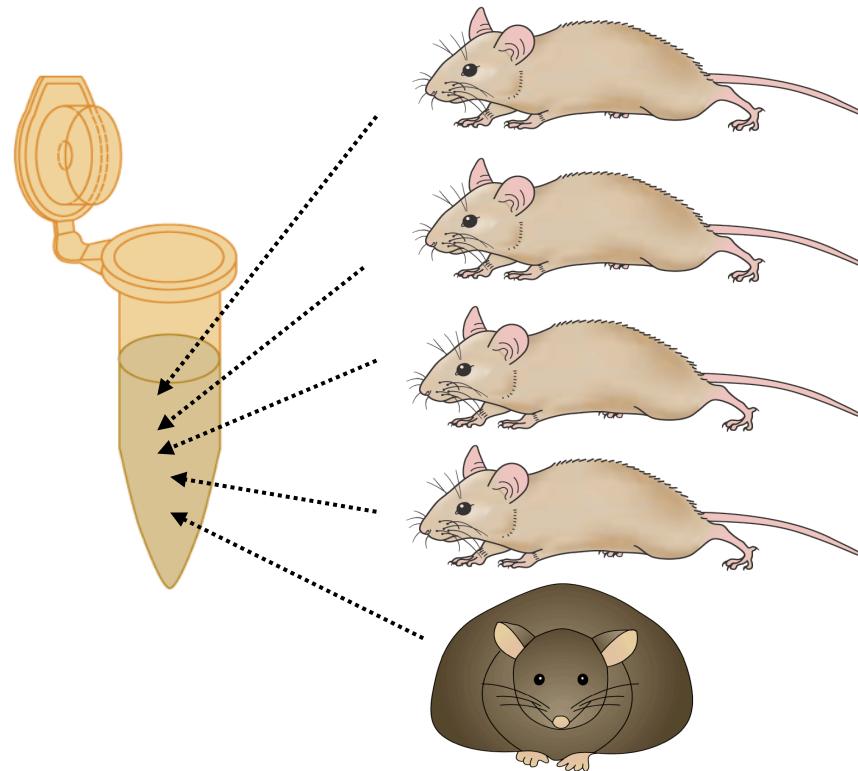


Technical Replicates

Don't do technical replicates unless you are planning a technical rather than a biological study.

# Should I pool samples?

- Insufficient amount of sample for hybridisation.



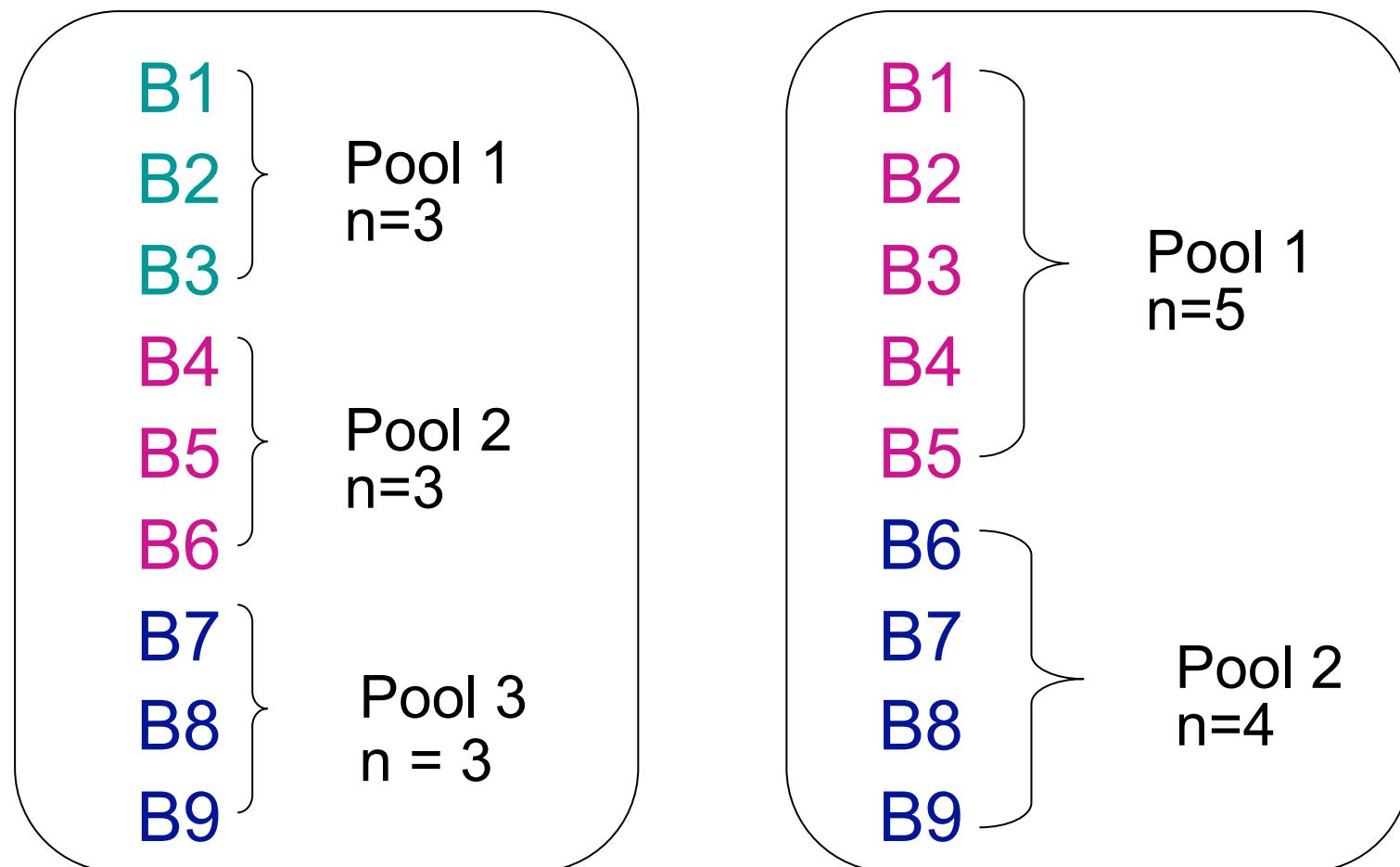
# Should I pool samples?

*When working with very small amounts of tissue sample, pooling is a good way of reducing the noise while keeping the number of  $n$  reasonably small.*

*e.g. where  $n$  = no. of array hybridizations.*

- Pool the **biological material** (tissue, cells), not the purified RNA or labeled cDNA! In this way, problems are far easier to spot.
- **No** way to estimate **variation between individuals** in a pool, which is sometimes important and often interesting.
- **Beware of outliers!** Don't include any sample that looks suspicious!
  - In some studies (pooled vs unpooled designs) the majority of DEGs turn out extreme in only one individual.

# How should I pool samples?

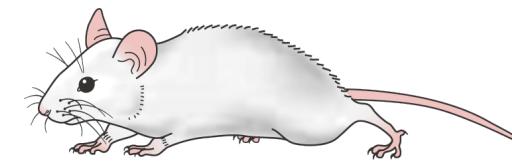
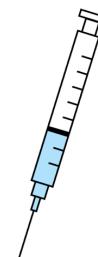


# Experimental Design

- Importance
- Planning
- Well-designed Experiment
- What Type of Experiment
- Define Variables & Good Hypotheses
- Standardise
- Experimental Controls
- Experimental Units
- Bias & Confounding Factors
- Replicates & Sample Size
- **Common Experimental Set-ups**

# Experimental Set-ups

- Patient/Tissue comparisons
- Drug Treatment
- Knock out mice
- Transfected cell lines





# Patient/Tissue Comparisons

- Look at **baseline characteristics** and identify any potential biases:
  - eg Breast cancer study on treatment groups: age, ethnic group, tumour size, ER status, geographical location etc
- **Randomisation** addresses baseline inequality for characteristics that can or can not be identified.
- **Thorough reporting:**
  - more important in observational research, where biases may not be addressed by powerful methods such as randomisation.
- **Pure tissue samples:**
  - try to obtain tissue samples devoid of adjacent tissue, or use microdissection. Collect **cellularity scores**.
- For large-scale studies, include **technical replicates** to look at batch effects.

**Table 1.** Base-Line Characteristics of the Patients and Tumors and Primary Treatment.\*

Variable	Exemestane (N=2362)	Tamoxifen (N=2380)
Demographic characteristics		
Age — yr	64.3±8.1	64.2±8.2
White race — no. (%)	2308 (97.7)	2325 (97.7)
Nodal status — no. (%)		
Negative	1211 (51.3)	1211 (50.9)
1–3 Positive nodes	715 (30.3)	706 (29.7)
≥4 Positive nodes	321 (13.6)	330 (13.9)
Positive, but no. of nodes missing	5 (0.2)	9 (0.4)
Unknown	84 (3.6)	96 (4.0)
Missing data	26 (1.1)	28 (1.2)
Histologic type — no. (%)		
Infiltrating ductal	1814 (76.8)	1871 (78.6)
Infiltrating lobular	346 (14.6)	327 (13.7)
Other	172 (7.3)	156 (6.6)
Unknown	3 (0.1)	1 (<0.1)
Missing data	27 (1.1)	25 (1.1)
Estrogen-receptor status — no. (%)†		
Positive	1917 (81.2)	1936 (81.3)
Progesterone-receptor positive	1312 (55.6)	1307 (54.9)
Progesterone-receptor negative	351 (14.9)	384 (16.1)
Progesterone-receptor status unknown or missing	254 (10.8)	245 (10.3)
Negative	26 (1.1)	33 (1.4)
Unknown	398 (16.9)	392 (16.5)
Missing data	21 (0.9)	19 (0.8)
Progesterone-receptor status — no. (%)		
Positive	1320 (55.9)	1313 (55.2)
Negative	360 (15.2)	395 (16.6)
Unknown	659 (27.9)	653 (27.4)
Missing data	23 (1.0)	19 (0.8)
Type of surgery — no. (%)		
Mastectomy	1222 (51.7)	1235 (51.9)
Breast-conserving	1116 (47.2)	1123 (47.2)
Unknown	3 (0.1)	2 (0.1)
Missing data	21 (0.9)	20 (0.8)
Previous chemotherapy — no. (%)		
Yes	766 (32.4)	765 (32.1)
No	1575 (66.7)	1596 (67.1)
Missing data	21 (0.9)	19 (0.8)
Previous hormone-replacement therapy — no. (%)		
Yes	567 (24.0)	557 (23.4)
No	1723 (72.9)	1747 (73.4)
Unknown	51 (2.2)	54 (2.3)
Missing data	21 (0.9)	22 (0.9)
Duration of tamoxifen therapy at randomization — yr		
Median	2.4	2.4
Interquartile range	2.1–2.7	2.1–2.7
Tamoxifen dose — no. (%)		
20 mg	2243 (95.0)	2270 (95.4)
30 mg	77 (3.3)	76 (3.2)
Missing data	42 (1.8)	34 (1.4)

\* Plus-minus values are means ±SD. Patients with missing data had no value reported for a given variable; for patients in the "unknown" category, data were reported as unknown.

† Data for positive and negative estrogen-receptor status include retrospectively ascertained status for some patients whose status was unknown at randomization.

Coombes, RC et al (2004)

## A randomised trial of exemestane after two to three years of tamoxifen therapy in postmenopausal women with primary breast cancer.

*N. Engl. J. Med.* 350, 1081–1092.

Age

Sex

Ethnic group

Nodal status

Histological type

ER status

PR status

Type of surgery

Previous chemotherapy

Previous hormone therapy

Duration of tamoxifen therapy

Tamoxifen Dose

Check clinical data for:

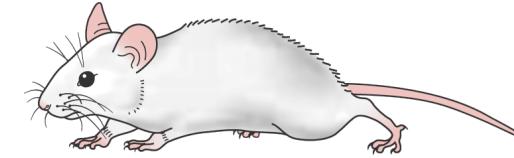
- biases
- missing data
- errors
- data standardisation

# Drug Treatments



- How many **time-points** or just '**before vs. after**'?
- Include a **time 0** baseline for time-series?
- If you are interested in a drug effect, try to examine the **earliest time point** possible.
- This minimizes **secondary effects** and focuses on the drug-specific changes.
- Include '**vehicle controls**' .
- **Blinding** can address any unintentional bias.

# Knock-out mice



- **Increase scope?**
  - e.g. sex, litter, age, strain, weight etc
- **Any potential confounding factors?**
  - e.g. cage location?
- **Random sampling:**
  - Each subject in the population has an equal chance of being selected
  - Avoid any bias in assigning the animals to the treatment groups
  - Randomisation should also extend to cage placement within rooms in the animal house.
- **Blinding:**
  - can address any unintentional bias.

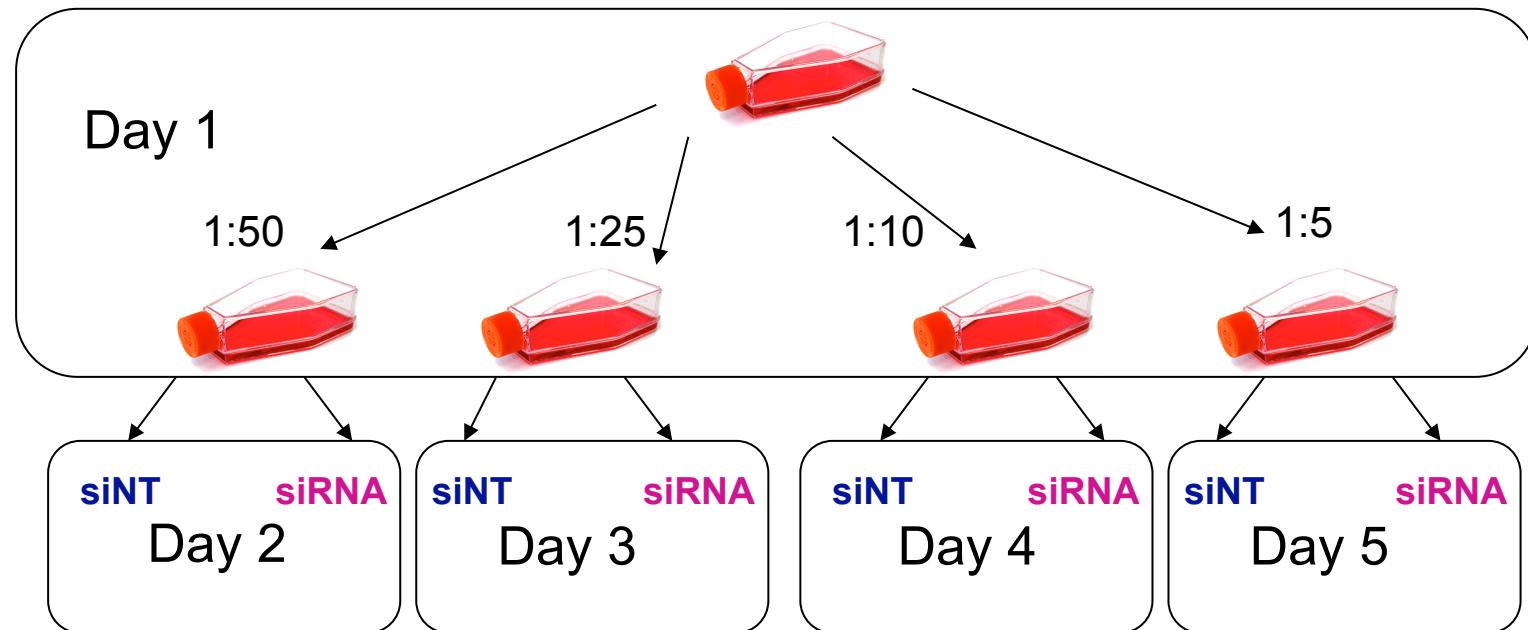
# Transfected cell-lines



- For example, cell lines treated with **siRNA**.
- Check for consistent and optimal **cell density** and confluence.
- Use a **mock transfected cell line** as a control.
- Use **independent transfectants** for each replicate.

# Transfected cell-lines

Are there any issues with this experimental set up below?



# Summary

- **Understand** the basic statistical and philosophical issues involved in any successful experimental design.
- **Design & Plan** carefully before you collect!
- Involve the **Genomics/Proteomics** and the **Analyst/Statistician** while you plan.
- **Record & Collect** as much information as possible about the samples and how they were collected/processed.
- **Randomisation & Blinding** can address potential confounding effects and biases.
- **Sample Size Calculations** – power!

Remember Fisher's fundamental principles of ED:

**Replication, Randomisation, Blocking**

This leads to:

**Robustness, Increased Validity, Power, Reproducibility and Meaningful Results**

# Useful Resources

**“Understanding sample size: what determines the required number of microarrays for an experiment?”** Jorstad et al (2007), Trends Plant Sci. 12(2):46-50

**“A Discussion of Statistical Methods for Design and Analysis of Microarray Experiments for Plant Scientists”** Dan Nettleton (2006), *The Plant Cell* 18:2112-2121 (2006)

“Statistical Design and Analysis of RNA Sequencing Data”. *Genetics* 185(2):405-416 (2010)

“Design and validation issues in RNA-seq experiments.” *Brief Bioinform* (2011) 12(3): 280-287

“Efficient experimental design and analysis strategies for the detection of differential expression using RNA-Sequencing”. *BMC Genomics* 2012, 13:484

“15 steps in the design and statistical analysis of experiments involving laboratory animals.” <http://www.isogenic.info/index.html>

“Practical aspects of experimental design in animal research” Johnson & Besselsen (2002) ILAR Journal

“The Design of Animal Experiments” Festing et al (2010)

“Deriving chemosensitivity from cell lines: forensic bioinformatics and reproducible research in high through-put biology”. Baggerly & Coombes (2009) Ann Appl Stat 3(4):1309-1334

“Tackling the widespread and critical impact of batch effects in high-throughput data” Leek et al (2010) Nature Reviews Genetics 11, 733–739

# Useful Resources

“Understanding sample size: what determines the required number of microarrays for an experiment?” Jorstad et al (2007), Trends Plant Sci. 12(2):46-50

“A Discussion of Statistical Methods for Design and Analysis of Microarray Experiments for Plant Scientists” Dan Nettleton (2006), *The Plant Cell* 18:2112-2121 (2006)

“Statistical Design and Analysis of RNA Sequencing Data”. *Genetics* 185(2):405-416 (2010)

“Design and validation issues in RNA-seq experiments.” *Brief Bioinform* (2011) 12(3): 280-287

“Efficient experimental design and analysis strategies for the detection of differential expression using RNA-Sequencing”. *BMC Genomics* 2012, 13:484

“15 steps in the design and statistical analysis of experiments involving laboratory animals.” <http://www.isogenic.info/index.html>

“Practical aspects of experimental design in animal research” Johnson & Besselsen (2002) ILAR Journal

“The Design of Animal Experiments” Festing et al (2010)

“Deriving chemosensitivity from cell lines: forensic bioinformatics and reproducible research in high through-put biology”. Baggerly & Coombes (2009) Ann Appl Stat 3(4):1309-1334

“Tackling the widespread and critical impact of batch effects in high-throughput data” Leek et al (2010) Nature Reviews Genetics 11, 733–739

# Useful Resources

“Understanding sample size: what determines the required number of microarrays for an experiment?” Jorstad et al (2007), Trends Plant Sci. 12(2):46-50

“A Discussion of Statistical Methods for Design and Analysis of Microarray Experiments for Plant Scientists” Dan Nettleton (2006), *The Plant Cell* 18:2112-2121 (2006)

“Statistical Design and Analysis of RNA Sequencing Data”. *Genetics* 185(2):405-416 (2010)

“Design and validation issues in RNA-seq experiments.” *Brief Bioinform* (2011) 12(3): 280-287

“Efficient experimental design and analysis strategies for the detection of differential expression using RNA-Sequencing”. *BMC Genomics* 2012, 13:484

**“15 steps in the design and statistical analysis of experiments involving laboratory animals.”** <http://www.isogenic.info/index.html>

**“Practical aspects of experimental design in animal research”** Johnson & Besselsen (2002) ILAR Journal

**“The Design of Animal Experiments”** Festing et al (2010)

“Deriving chemosensitivity from cell lines: forensic bioinformatics and reproducible research in high through-put biology”. Baggerly & Coombes (2009) Ann Appl Stat 3(4):1309-1334

**“Tackling the widespread and critical impact of batch effects in high-throughput data”**  
Leek et al (2010) Nature Reviews Genetics 11, 733–739

# Useful Resources

“Understanding sample size: what determines the required number of microarrays for an experiment?” Jorstad et al (2007), Trends Plant Sci. 12(2):46-50

“A Discussion of Statistical Methods for Design and Analysis of Microarray Experiments for Plant Scientists” Dan Nettleton (2006), *The Plant Cell* 18:2112-2121 (2006)

“Statistical Design and Analysis of RNA Sequencing Data”. *Genetics* 185(2):405-416 (2010)

“Design and validation issues in RNA-seq experiments.” *Brief Bioinform* (2011) 12(3): 280-287

“Efficient experimental design and analysis strategies for the detection of differential expression using RNA-Sequencing”. *BMC Genomics* 2012, 13:484

“15 steps in the design and statistical analysis of experiments involving laboratory animals.” <http://www.isogenic.info/index.html>

“Practical aspects of experimental design in animal research” Johnson & Besselsen (2002) ILAR Journal

“The Design of Animal Experiments” Festing et al (2010)

“Deriving chemosensitivity from cell lines: **forensic bioinformatics and reproducible research in high through-put biology**”. Baggerly & Coombes (2009) Ann Appl Stat 3(4):1309-1334

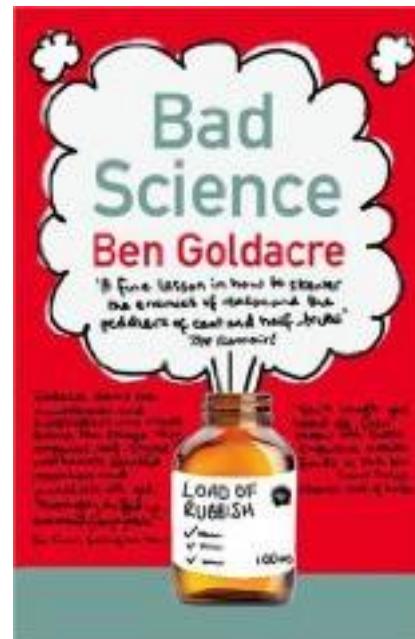
“**Tackling the widespread and critical impact of batch effects in high-throughput data**” Leek et al (2010) Nature Reviews Genetics 11, 733–739

# Books

Free online book on experimental design

<http://www.stat.cmu.edu/~hseltman/309/Book/Book.pdf>

Ben Goldacre's "Bad Science"



# Acknowledgements

## Bioinformatics Core Facility (Cambridge)

Thomas Carroll  
Sarah Dawson  
Mark Dunning  
Silvia Halim  
Suraj Menon  
Rory Stark  
Sarah Vowler  
**Matthew Eldridge**

## Computational Biology Group (Cambridge)

Nuno Barbosa-Morais  
Natalie Thorne  
Matt Ritchie  
Andy Lynch  
Christina Curtis  
**Benilton Carvalho**  
**Oscar Rueda**

## Genomics Core Facility (Cambridge)

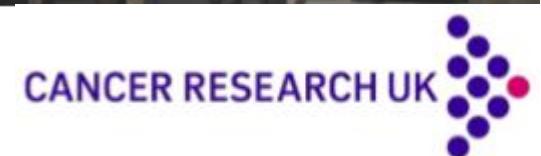
Michelle Osborne  
Sarah Leigh-Brown  
Hannah Haydon  
Claire Fielding  
Fatimah Madni  
**James Hadfield**

## Sanger Institute (Cambridge)

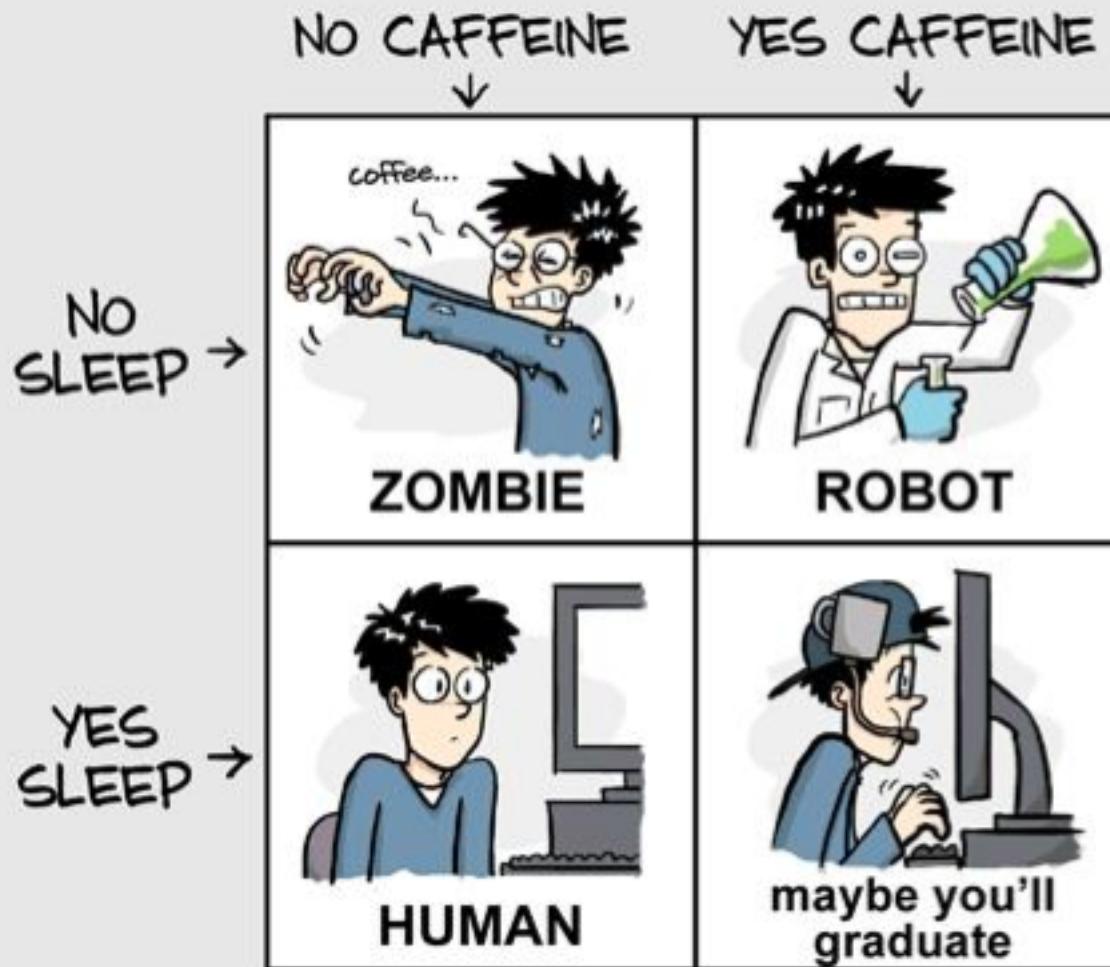
Natasha Karp



UNIVERSITY OF  
CAMBRIDGE



# GRAD SCHOOL ENERGY LEVELS



JORGE CHAM © 2011