Introduction to Experimental Design

BIM 283

Advanced Design of Experiments for Biomedical Engineers

Basic Principles of Experimental Investigation

- Sequential Experimentation
- Comparison
- Manipulation
- Replication
- Randomization
- Blocking
- Simultaneous variation of factors
- Main effects and interactions
- Sources of variability

Blocking

- If some factor may interfere with the experimental results by introducing unwanted variability, one can *block* on that factor
- In agricultural field trials, soil and other location effects can be important, so plots of land are subdivided to test the different treatments. This is the origin of the idea

- If we are comparing treatments, the more alike the units are to which we apply the treatment, the more sensitive the comparison.
- Within blocks, treatments should be randomized
- Paired comparisons are a simple example of randomized blocks as in the tomato plant example
- If 12 RNA samples are distributed to 4 lanes, with 3 labeled samples per lane, then the lanes are blocks
- Sometimes these are called batch effects and can cause trouble if correlated with an important factor.

"Comparison of the transcriptional landscapes between human and mouse tissues," Shin Lin, Yiing Lin, Joseph R. Nery, Mark A. Urich, Alessandra Breschi, Carrie A. Davis, Alexander Dobin, Christopher Zaleski, Michael A. Beer, William C. Chapman, Thomas R. Gingeras, Joseph R. Ecker, and Michael P. Snyder, *PNAS*, **111**, 2014.

To date, various studies have found similarities between humans and mice on a molecular level, and indeed, the murine model serves as an important experimental system for biomedical science. In this study of a broad number of tissues between humans and mice, high-throughput sequencing assays on the transcriptome and epigenome reveal that, in general, differences dominate similarities between the two species. These findings provide the basis for understanding the differences in phenotypes and responses to conditions in humans and mice.

Claims in Lin et al. 2014

- This article claims that difference in gene expression between species is larger than the difference in expression across organs, casting doubt on the use of the mouse model.
- Heart, kidney, liver, small bowel, spleen, testis, adipose, adrenal, sigmoid colon, lung, ovary, brain, pancreas.
- One sample per species per organ, in itself a problem of lack of denominator.
- The use of clustering does not fix this problem.

"A reanalysis of mouse ENCODE comparative gene expression data," Yoav Gilad and Orna Mizrahi-Man, *F1000Research*, 2015; **4**: 121.

D87PMJN1 (run 253, flow cell D2GUAACXX, lane 7)	D87PMJN1 (run 253, flow cell D2GUAACXX , lane 8)	D4LHBFN1 (run 276, flow cell C2HKJACXX , lane 4)	MONK (run 312, flow cell C2GR3ACXX , lane 6)	HWI-ST373 (run 375, flow cell C3172ACXX , lane 7)
heart	adipose	adipose	heart	brain
kidney	adrenal	adrenal	kidney	pancreas
liver	sigmoid colon	sigmoid colon	liver	brain
small bowel	lung	lung	small bowel	spleen
spleen	ovary	ovary	testis	● Human
testis		pancreas		● Mouse

D87PMJN1 (run 253, flow cell D2GUAACXX, lane 7)	D87PMJN1 (run 253, flow cell D2GUAACXX , lane 8)	D4LHBFN1 (run 276, flow cell C2HKJACXX , lane 4)	MONK (run 312, flow cell C2GR3ACXX , lane 6)	HWI-ST373 (run 375, flow cell C3172ACXX , lane 7)
heart	adipose	adipose	heart	brain
kidney	adrenal	adrenal	kidney	pancreas
liver	sigmoid colon	sigmoid colon	liver	brain
small bowel	lung	lung	small bowel	spleen
spleen	ovary	ovary	testis	Human
testis		pancreas		Mouse

- Most of the human samples are in two lanes in the same run (253) on one sequencer.
- Most of the mouse samples are in two lanes in two runs in two different sequencers.
- One run on a different sequencer had both human and mouse samples.
- Batch effects: machine, run in machine, lane in run and machine; these are known to have effects on the results.

D87PMJN1 (run 253, flow cell D2GUAACXX, lane 7)	D87PMJN1 (run 253, flow cell D2GUAACXX , lane 8)	D4LHBFN1 (run 276, flow cell C2HKJACXX , lane 4)	MONK (run 312, flow cell C2GR3ACXX , lane 6)	HWI-ST373 (run 375, flow cell C3172ACXX , lane 7)
heart	adipose	adipose	heart	brain
kidney	adrenal	adrenal	kidney	pancreas
liver	sigmoid colon	sigmoid colon	liver	brain
small bowel	lung	lung	small bowel	spleen
spleen	ovary	ovary	testis	Human
testis		pancreas		Mouse

- Species effect is almost completely confounded with machine/run/lane, so that similarity of results from sample of different organs from the same species are indistinguishable from similarity of results from samples in the same lane, same run, and same sequencer.
- The blocks are confounded with species, making a good analysis impossible.
- Also, no replicates from different animals/subjects from the same organ.

Simultaneous Variation of Factors

- The simplistic idea of "science" is to hold all things constant except for one experimental factor, and then vary that one thing
- This misses interactions and can be statistically inefficient
- Multi-factor designs are often preferable

Interactions

- Sometimes (often) the effect of one variable depends on the levels of another one
- This cannot be detected by one-factor-at-a-time experiments
- These interactions are often scientifically the most important

- **Experiment 1**. I compare the room before and after I drop a liter of gasoline on the desk. **Result**: we all leave because of the odor.

- **Experiment 1**. I compare the room before and after I drop a liter of gasoline on the desk. **Result**: we all leave because of the odor.
- **Experiment 2**. I compare the room before and after I drop a lighted match on the desk. **Result**: no effect other than a small scorch mark.

- **Experiment 1**. I compare the room before and after I drop a liter of gasoline on the desk. **Result**: we all leave because of the odor.
- **Experiment 2**. I compare the room before and after I drop a lighted match on the desk. **Result**: no effect other than a small scorch mark.
- **Experiment 3**. I compare all four of ±gasoline and ±match. **Result**: we are all killed.
- Large interaction effect
- The effect of the match is small in the –gasoline condition and large in the +gasoline condition
- **An interaction is when the effect of one variable depends on the level of another**

Statistical Efficiency

- Suppose I compare the expression of a gene in a cell culture of either
  - keratinocytes or fibroblasts,
  - confluent and nonconfluent,
  - with or without a possibly stimulating hormone,
  - with 2 cultures in each condition,
- requiring 16 culture dishes

- I can compare the cell types as an average of 8 cultures vs. 8 cultures
- I can do the same with the other two factors
- This is more efficient than 3 separate experiments with the same controls, using 48 cultures
- Can also see if cell types react differently to hormone application (interaction)

Fractional Factorial Designs

- When it is not known which of many factors may be important, fractional factorial designs can be helpful
- With 7 factors each at 2 levels, ordinarily this would require $2^7 = 128$ experiments
- This can be done in 8 experiments instead!
- Each two factors form a replicated two-by-two
- Some sets of three factors form an unreplicated two-by-two-by-two

	F1	F2	F3	F4	F5	F6	F7
1	H	H	H	H	H	H	H
2	H	H	L	H	L	L	L
3	H	L	H	L	H	L	L
4	H	L	L	L	L	H	H
5	L	H	H	L	L	H	L
6	L	H	L	L	H	L	H
7	L	L	H	H	L	L	H
8	L	L	L	H	H	H	L

	F1	F2	F3	F4	F5	F6	F7
1	H	H	H	H	H	H	H
2	H	H	L	H	L	L	L
3	H	L	H	L	H	L	L
4	H	L	L	L	L	H	H
5	L	H	H	L	L	H	L
6	L	H	L	L	H	L	H
7	L	L	H	H	L	L	H
8	L	L	L	H	H	H	L

	F1	F2	F3	F4	F5	F6	F7
1	H	H	H	H	H	H	H
2	H	H	L	H	L	L	L
3	H	L	H	L	H	L	L
4	H	L	L	L	L	H	H
5	L	H	H	L	L	H	L
6	L	H	L	L	H	L	H
7	L	L	H	H	L	L	H
8	L	L	L	H	H	H	L

	F1	F2	F3	F4	F5	F6	F7
1	H	H	H	H	H	H	H
2	H	H	L	H	L	L	L
3	H	L	H	L	H	L	L
4	H	L	L	L	L	H	H
5	L	H	H	L	L	H	L
6	L	H	L	L	H	L	H
7	L	L	H	H	L	L	H
8	L	L	L	H	H	H	L

Main Effects and Interactions

- Factors Cell Type (C), State (S), Hormone (H)
- Response is expression of a gene
- The main effect C of cell type is the difference in average gene expression level between cell types

- For the interaction between cell type and state, compute the difference in average gene expression between cell types separately for confluent and nonconfluent cultures. The difference of these differences is the interaction.
- The three-way interaction CSH is the difference in the two way interactions with and without the hormone stimulant.

Sources of Variability in Laboratory Analysis

- Intentional sources of variability are treatments and blocks
- There are many other sources of variability
- Biological variability between organisms or within an organism
- Technical variability of procedures like RNA extraction, labeling, hybridization, chips, etc.

Replication

- Almost always, biological variability is larger than technical variability, so most replicates should be biologically different, not just replicate analyses of the same samples (technical replicates)
- However, this can depend on the cost of the experiment vs. the cost of the sample
- 2D gels are so variable that replication is required
- Expression arrays, PCR, RNA-Seq, Mass Spec and others do not usually require technical replication, but biological replication is essential.

Levels of Analysis

- I have 10 mice each of two strains and two RNA samples per mouse obtained by splitting a single sample. Correct comparison is 10 vs. 10 and the results from the RNA samples for each mouse should be averaged.
- I have 10 mice each of two strains and two RNA samples per mouse one obtained at baseline and one after a treatment. Correct comparison is a paired t-test, 10 vs. 10, or equivalently, a two-way ANOVA of strain and time.

Levels of Analysis

- I am examining keratinocyte mobility under two conditions; I have four glass slides in each condition and have measured the distance traveled by 30 cells on each slide. Correct analysis is obtained by averaging the 30 cells on each slide, so we have a 4 vs. 4 comparison. The sample size is not 240!
- There are advanced methods in which we model the cell effect nested in the slide, but the results will be similar to averaging the cells.

Quality Control

- It is usually a good idea to identify factors that contribute to unwanted variability
- A study can be done in a given lab that examines the effects of day, time of day, operator, reagents, etc.
- This is almost always useful in starting with a new technology or in a new lab

Possible QC Design

- Possible factors: day, time of day, operator, reagent batch
- At two levels each, this is 16 experiments to be done over two days, with 4 each in morning and afternoon, with two operators and two reagent batches
- Analysis determines contributions to overall variability from each factor

Glycomic Analysis of Ovarian Cancer Samples vs. Controls

- Cancer samples vs. controls in cells obtained from an ovarian cancer archive.
- Measured sugar species adducted to proteins by mass spectrometry (glycoproteins are important in human physiology).
- In the first set of samples, all the controls were under age 45 and all the patients were over age 45. This is problematic since there are substantial possible differences in physiology between pre- and post-menopausal women
- Next set of samples were more balanced by age

Allowing for Confounders

- Samples were grouped in sets of 12 (the group size was determined by the pre-processing machinery).
- Six patient samples and six controls per set
- Approximately balanced by age within each set of 12.
- Samples coded and blinded to the mass spec lab.
- There were some detected differences between patients and controls
- But the largest differences were variation from group to group!
- Might be due to operator, temperature in the lab, or just random

References

- *Statistics for Experimenters*, Box, Hunter, and Hunter.
- ISwR = *Introductory Statistics with R*, Peter Dalgaard (on the web site).
- *Fundamentals of Biostatistics*, Bernard Rosner, or other introductory biostatistics text.

Exercise 1

- You have a clinical study in which 12 patients will either get the standard treatment or a new treatment
  - Randomize which 6 of the 12 get the new treatment so that all possible combinations can result. Use Excel or R or another formal randomization method.
  - Instead, randomize so that in each pair of patients entered by date, one has the standard and one the new treatment (blocked randomization).
- What are the advantages of each method?
- Why is randomization important?

Course Website

- http://dmrocke.ucdavis.edu/Class/BIM283.2025.Winter/BST283-Winter-2025.html
- Or go to dmrocke.ucdavis.edu, choose Courses, then BIM 283.