THE UNIVERSITY OF ALABAMA AT BIRMINGHAM

Section ON Statistical Genetics

# Design of Experiment in Metabolomics

Hemant K. Tiwari, Ph.D.
Professor and Head

Section on Statistical Genetics
Department of Biostatistics
School of Public Health

## Metabolomics: Bench to Bedside

| Workflow | Considerations | Choices |
|---|---|---|
| Study design | Population | For example, Caucasian, Asian, African |
| | Study type | For example, population-based, twins study, clinical studies |
| | Sample type | For example, blood, urine, saliva |
| Sample collection | Standard operating protocols | Compatibility between study centres |
| | Fasting state | For example, fasting, non-fasting, controlled nutritional challenges |
| | Sample quantities | Serum, plasma, small volumes to avoid thawing |
| Sample storage | Temperature | –80°C, liquid nitrogen |
| | Aliquoting | 200 µl for mass spectrometry, 1 ml for NMR, avoid thawing cycles |
| | Biobanking | Manual, automated |
| Sample preparation | Metabolite extraction | For example, polar, charged |
| | Derivatization | Changing biochemical properties for better measurement |
| Sample analysis | Method | $^1$H NMR, LC–MS/MS, GC–MS/MS |
| | Identification | Targeted, non-targeted, quantitative |
| | Provider | Proprietary, core facility, fee-for-service |
| Data analysis | Covariates | Age, gender, body mass index, medication, lifestyle |
| | Statistical analysis | For example, linear model, using ratios, advanced statistics |
| | Initial data processing | Log-normal scaling, principal-component transformation |
| Data interpretation | Functional | For example, GRAIL, overlay with eQTL data |
| | Biochemical | KEGG, HMDB |
| | Medical | GWAS catalogue, pharmacogenomics database |

Suhre and Gieger (Nature Review Genetics, Vol 13, Nov 2012)

## Design of Experiments/ Experimental Design

- A controlled experiment to either test a hypothesis or generate hypotheses
- In the design of experiments, the experimenter is usually interested in the effect of some process or intervention on some subjects

## History of Experimental Design

- In 1747, James Lind (a Scottish Physician) developed the theory that citrus fruits cured scurvy, while serving as surgeon on HMS Salisbury Ship of the Royal Navy. This was the first ever clinical trial conducted.
- Scurvy is disease resulting from a deficiency of Vitamin C.

## Lind's Experiment

- Lind selected 12 men from the ship suffering from scurvy. He divided them into six pairs, giving each pair different supplements to their basic diet for two weeks. The treatments were all remedies that had been proposed:
  - A quart of cider every day
  - Twenty five drops of *elixir vitriol* (sulphuric acid) three times a day upon an empty stomach
  - One half-pint of seawater every day
  - A mixture of garlic, mustard, and horseradish in a lump the size of a nutmeg
  - Two spoonful of vinegar three times a day
  - Two oranges and one lemon every day

Result: The men given citrus fruits recovered dramatically within a week.

Source: Wikipedia

## Aims of experimental design

- To provide answers to research hypothesis or generate hypothesis
- Minimize the biological variance and technical or experimental variance since metabolome can change very rapidly in response to subtle changes in environment

## How to avoid technical variance

- Randomization
- Normalization
- Selection of covariates in the analyses
- Selection of statistical method most appropriate for the research question

## Basic Steps in Experimental Design

- Formalize the question and study design
- The number of replicates to be used for an experiment
- Randomization
- Blocking (stratification)

## Study Design for human study

- Pedigree data
- Twins
- Population data
- Case-control
- Case only

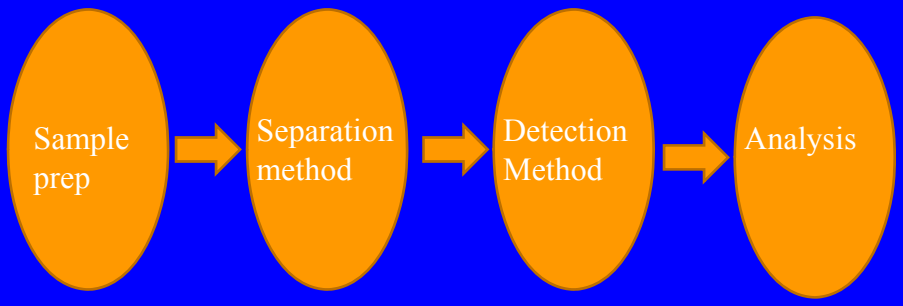
## Choice of Platform

- GC/MS
- LC/MS
- NMR

## Choosing which metabolites to study

- Targeted vs. non-targeted
- Tissue Type
  - Blood
  - Urine
  - Buckle Swab (?)

## Steps in metabolomics experiment




At each step, the errors can occur
Aim of the experiment is to minimize the errors

## Aims of experimental design

- Minimize the variance of extraneous variable variance that may have impact on outcome variables
  - Examples: Metabolites are sensitive to environmental influences such as sample storage
    - Temperature
    - Aliquoting
    - Biobanking

## Measurement Issues in metabolomics

- To get the sample ready
  - The sample is prepped and put onto wells on a silicon plate
  - Each well's aliquot is subjected to gas and/or liquid chromatography
  - After separation, the sample goes to a mass spectrometry

## Measurement Issues

- Sources of errors at the prep stage
  - Within subject variation
  - Within tissue variation
  - Contamination by cleaning solvents
  - Calibrate uncertainty
  - Evaporation of volatiles

## Errors at separation method

- Gas chromatography (GC) (GC creates ionized aerosol, each droplet evaporates to a single ion and this is separated by mass in the column, then ejected to the spectrometer):
  - Imperfect evaporation
  - Adhesion in the column
  - Ion fragmentation

## Measurement Issues at Detection

- Mass spectrometry (MS) is used to identify and quantify metabolites after separation by GC, HPLC (LC-MS)
  - Baseline removal
  - denoising
  - Peak detection (the process of distinguishing peptide and noise peaks)
  - Normalization

## Sample Size and Power

- Purpose
  - Planning a study: number of individuals to recruit to test a research hypothesis
  - Understand sample size implications of alternative study designs
  - Sample was already collected and wants study using new technology
    - GWAS was done, but wants to do metabolomics GWAS

## Sample Size and Power Calculation

- Often the number of samples to be used for the experiments dictated by the reality of resources available, not science.
  - How much money is available for the experiment
  - What is the cost per sample
  - Thus, sample size = $ available through NIH/ cost per sample

## Hypothesis Testing

- Power calculations are based on the principles of hypothesis testing
- A hypothesis is a statement about population parameter
- The two complemetary hypotheses in a hypothesis testing problem are called the null hypothesis ($H_0$) and alternate hypothesis ($H_1$)

# Two types of errors in hypothesis testing

$H_0$: $\theta=0$ versus $H_1$: $\theta \neq 0$

| | | Decision | |
|---|---|---|---|
| | | Accept $H_0$ | Reject $H_0$ |
| Truth | $H_0$ | Correct Decision | Type I error ($\alpha$) |
| | $H_1$ | Type II error ($\beta$) | Correct decision |

Reject $H_0$ when we should not

Don't reject $H_0$ when we should

- Type I Error: Probability of finding a statistically significant effect when the truth is that there is no effect
- Type II Error: Probability of not finding a statistically significant effect when there is none.
- Power = 1- $\beta$, Significance level = $\alpha$
- Goal is to minimize both types of errors

# Power depends on …

- Design
- The method of analyzing the data
- The effect size
- Standard deviation of the effect of interest
- Measurement variability
- The chosen significance level ($\alpha$)
- The sample size

We usually use significance level of 5% and 80% power to estimate the sample size

## Factors Affecting the sample size

| | |
|---|---|
| Effect size ↑ | Required sample size ↓ |
| Variation of data ↑ | Required sample size ↑ |
| Type I error rate ↓ | Required sample size ↑ |
| Power ↑ | Required sample size ↑ |

- Type I error rate ($\alpha$) is kept fixed and becomes smaller as number of tests increase
- Effect size and variation of the data ($\sigma^2$) is either obtained through pilot study or vary to calculate different sample sizes.

## To calculate Sample Size

- Need to know level of significance ($\alpha$)
- Statistical power ($1- \beta$)
- Effect size (expected difference)
- Standard deviation
- What statistical test we are going to use

# Sample Size Formula for difference in means

- A sample size formula to test difference of means between two groups
  - $n_1 = ((r + 1)\sigma^2 (Z_\beta + Z_{\alpha/2})^2)/ (r \Delta^2)$

*where,*

- $n_1$= size of the smaller group
- r = ratio of larger group to smaller group
- $Z_\beta$= standard normal deviate corresponds to β
- $Z_{\alpha/2}$= standard normal deviate corresponds to two-tailed significance level
- Δ= clinically meaningful difference in means of the outcome
- $\sigma^2$=Variance of the trait or characteristic

# Common standard normal deviates $Z_\alpha$ and $Z_\beta$

| α or β | One-sided test $Z_\alpha$ or $Z_\beta$ | Two-sided test $Z_\alpha$ |
|---|---|---|
| 0.001 | 3.09 | 3.29 |
| 0.005 | 2.58 | 2.81 |
| 0.01 | 2.33 | 2.58 |
| 0.025 | 1.96 | 2.24 |
| 0.05 | 1.64 | 1.96 |
| 0.10 | 1.28 | 1.64 |
| 0.20 | 0.84 | 1.28 |

## Simple Example

- How many people would you need to sample in each group (assuming both groups of equal size) to achieve power of 80% if SD $=\sigma=10$, difference in mean is 5 with fixed $\alpha=0.05$. So $Z_{\alpha/2} = Z_{0.025} = 1.96$, $Z_{\beta}=0.84$, and
- $n_1 = ((r + 1)\sigma^2 (Z_{\beta} + Z_{\alpha/2})^2)/ (r \Delta^2)$

  $= 2 (100) (.84+1.96)^2/ (5^2)$

  $= 62.72 \sim 63$

  63 per group implies 126 altogether.

## Sample Size

- Two-tailed test: When the investigator is interested in determining whether a treatment A is different from a treatment B
- Usually a 2 tailed test is performed with the risk of making a Type-1 error set at $\alpha / 2$ in each tail.
- For a 2 tailed test at $\alpha = .05$ and equal allocation of type-1 error to each tail $Z_{\alpha} = 1.96$

## Sample Size

Sometimes an investigator is only interested in a difference between treatments in one direction.

This is appropriate when The hypothesis is to test whether a treatment A is better than treatment B

For a 1 tailed test at $\alpha = .05$ $Z_\alpha = 1.65$

## Conclusions

- Experiment should be designed with consultation with the statistician and metabolomics assays provider
- Good design and good analytic methods can lead to reduced sample size and also lead to valid meaningful results