Design and Analysis of Experiments Exploring the Main Effects of Preanalytical Variables on Molecular Research

2nd Annual Biospecimen Research Network (BRN) Symposium:
Advancing Cancer Research Through Biospecimen Science

The experiments

Think blood. Others tissues should come later.

I'll illustrate the principles using a hypothetical experiment with blood. A real experiment has just been carried out.

Statistical principles of experimental design

General

Randomization
Replication
Local control
The factorial principle

Special

Designs with confounding
Two-phases
External validity

References

R.A. Fisher

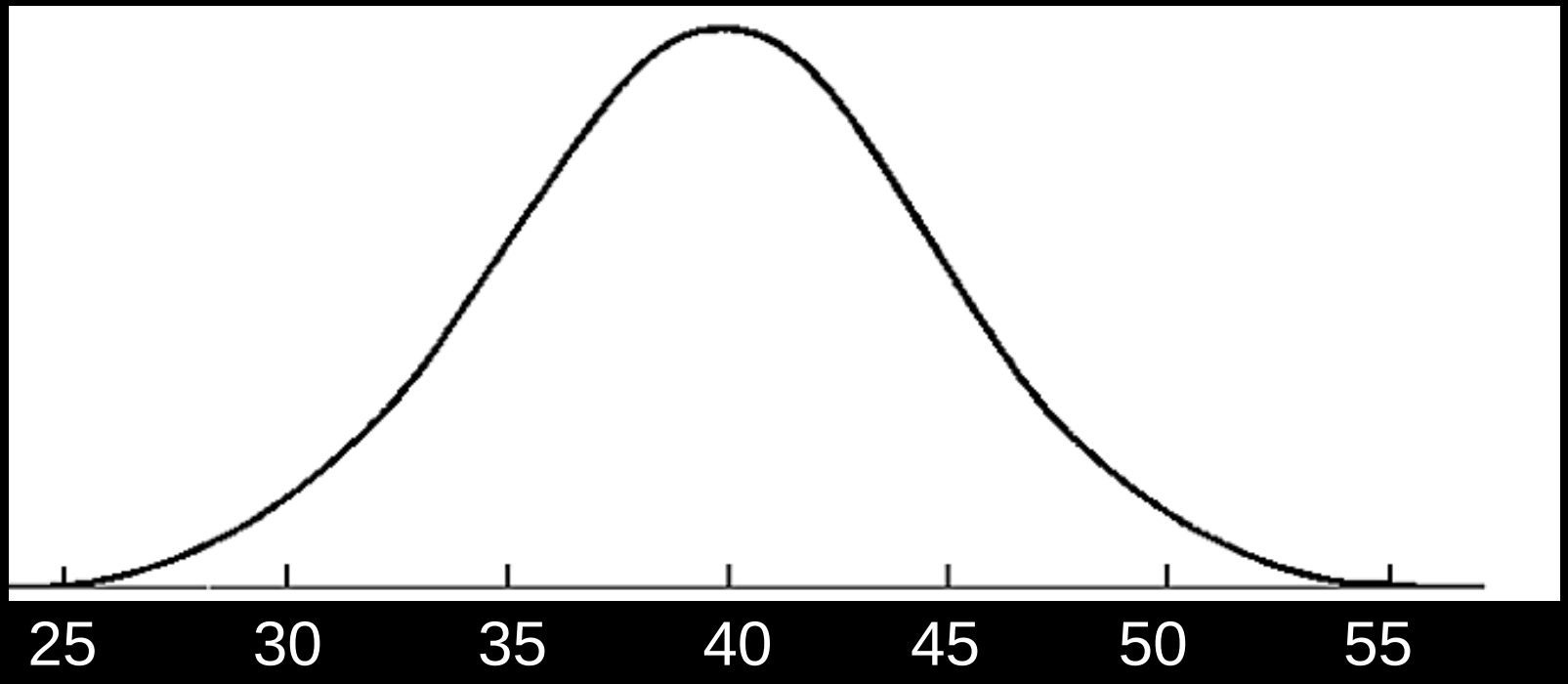
The arrangement of field experiments. *Journal of the Ministry of Agriculture of Great Britain* **33**: 503-513 (1926)

Reprinted in *Breakthroughs in Statistics Volume II*, S Kotz and NL Johnson, eds, Springer 1992.

G.E.P. Box, J.S. Hunter & W.G. Hunter

Statistics for Experimenters. 2nd edition John Wiley & Sons, Inc., 2005.

Population distribution of a hypothetical molecular variable



Population mean = 40, standard deviation (SD) = 5
Mean of females = 42, mean of males = 38
Within person replicate SD = 2
Assay replicate SD = 1.

Hypothetical experiments with blood measuring our molecular variable

Three pre-analytical variables (factors), with their baseline and alternative levels

Patient exercise prior to phlebotomy:
Levels: None, Moderate

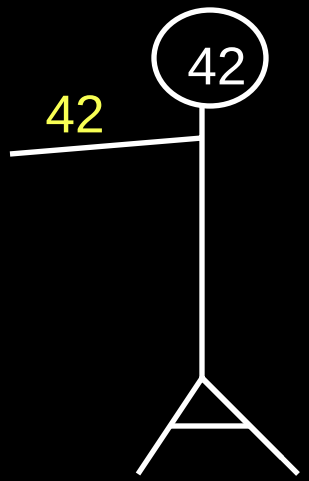
Patient position during phlebotomy:
Levels: Sitting, Standing

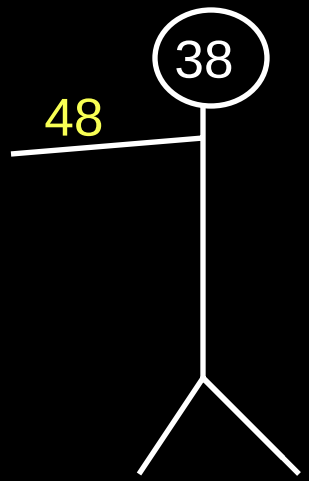
Temperature during blood processing:
Levels: Room temperature (RT), 4ºC

Let's suppose that in all cases, the alternative adds 10.

Comparison 1: Between subjects

Assign exercise treatment *at random*





……

No exercise
(baseline)

Moderate exercise
(+10)

We wouldn’t want confounding of treatment with gender or…

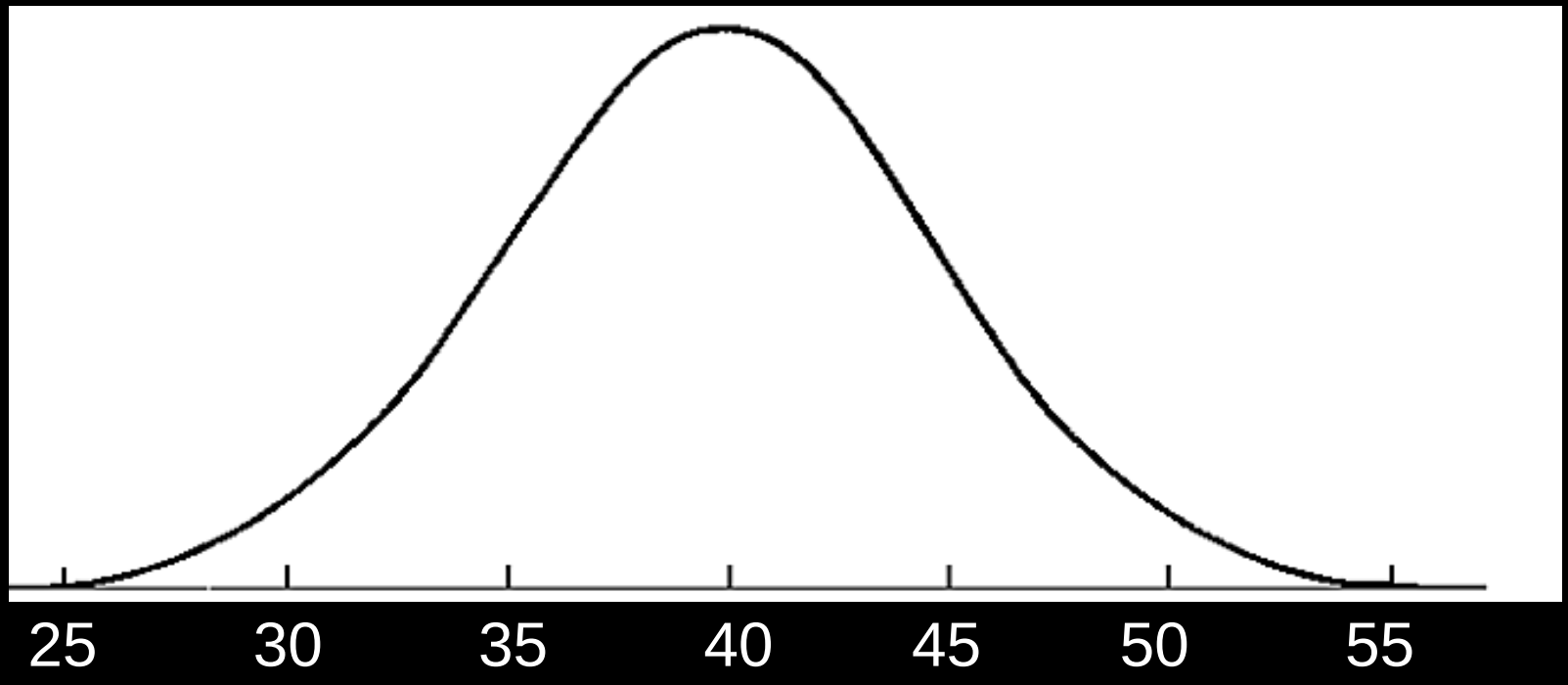
Randomization

Any assignment that would otherwise be haphazard should be done at random using a chance device.

Randomization

- Achieves a rough balancing, leading to a lower chance of biasing the outcome by known or unknown variables
- Permits reliable statistical analyses (error estimation or randomization tests)

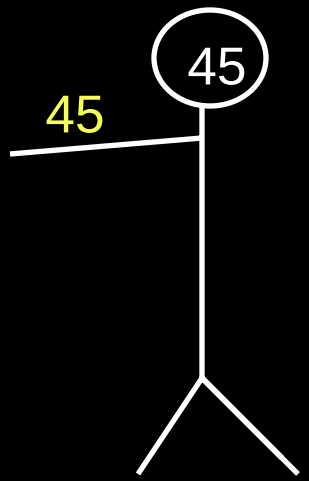
Population distribution of a hypothetical molecular variable

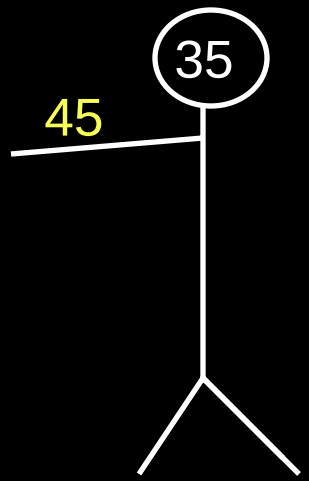


Population mean = 40, standard deviation (SD) = 5
Mean of females = 42, mean of males = 38
Within person replicate SD = 2
Assay replicate SD = 1.

Comparison: Between subjects

Assign exercise treatment *at random*





……

No exercise
(baseline)

Moderate exercise
(+10)

Here we'll need replication to see the effect of exercise.

Replication

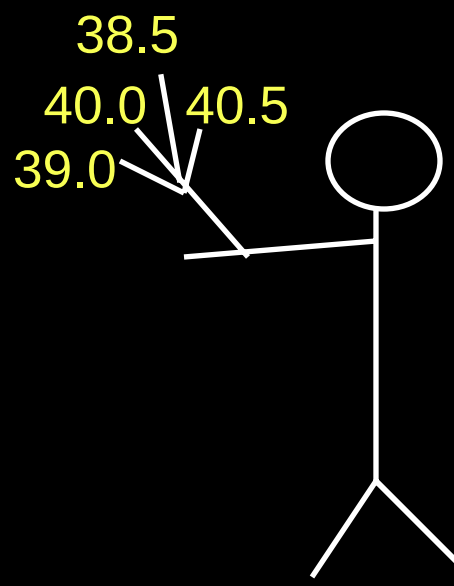
Is having more than one measurement with the experimental conditions held constant.

Replication:

- Increases precision (the law of averages)
- Permits estimation of variability (including error)

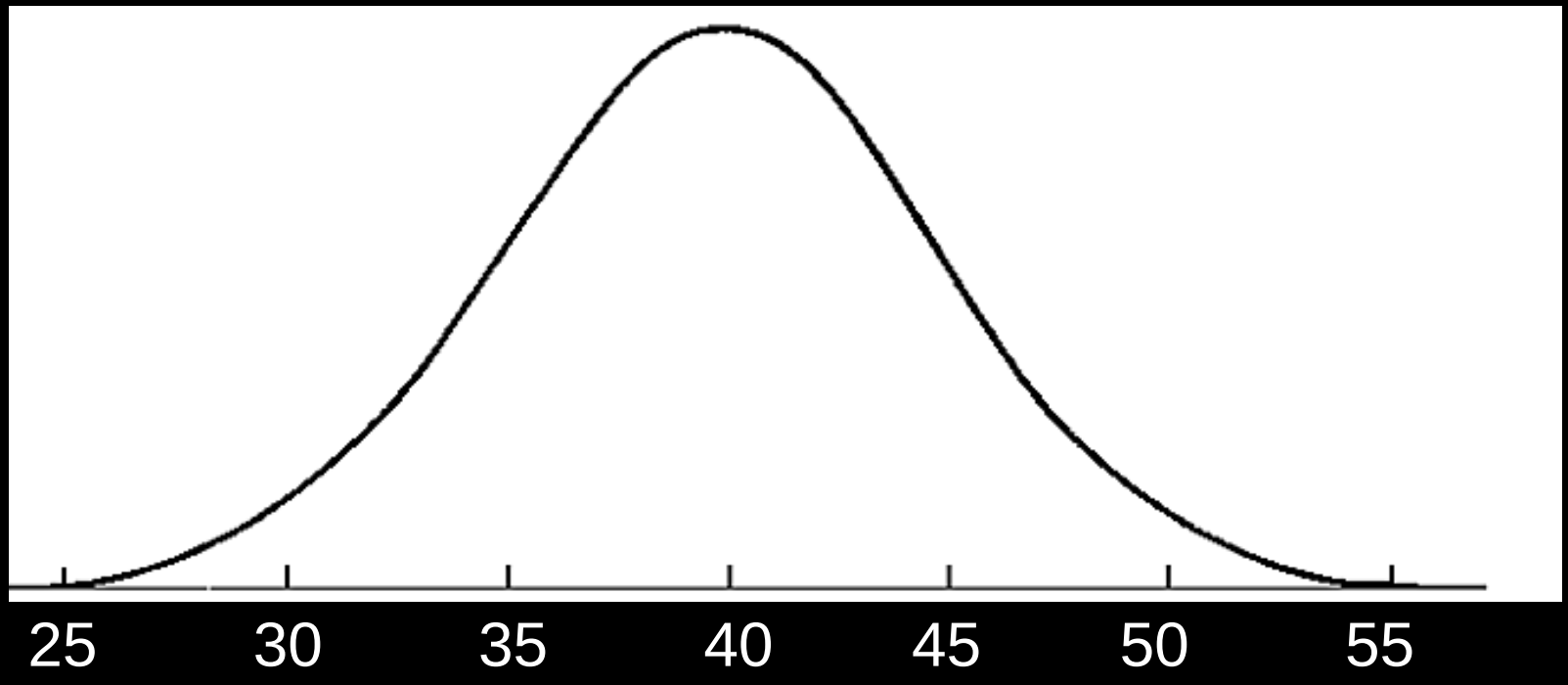
It has to be in the right place.

Replication here not relevant to estimation of the effect of exercise



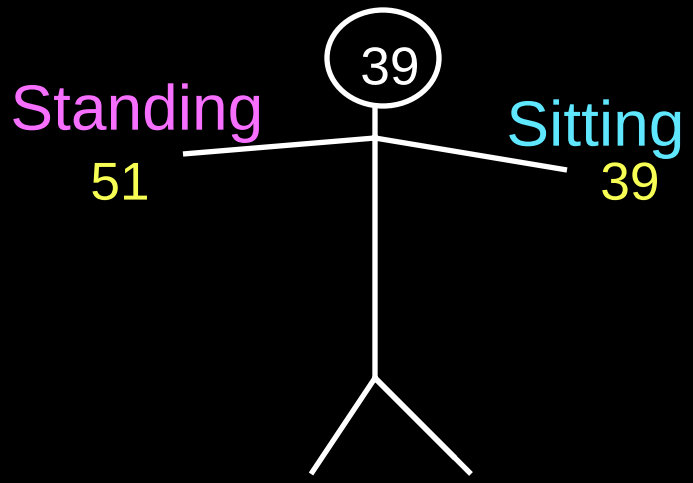
The real difficulty estimating the effect of exercise is the variability between people (“biological” vs “technical”).

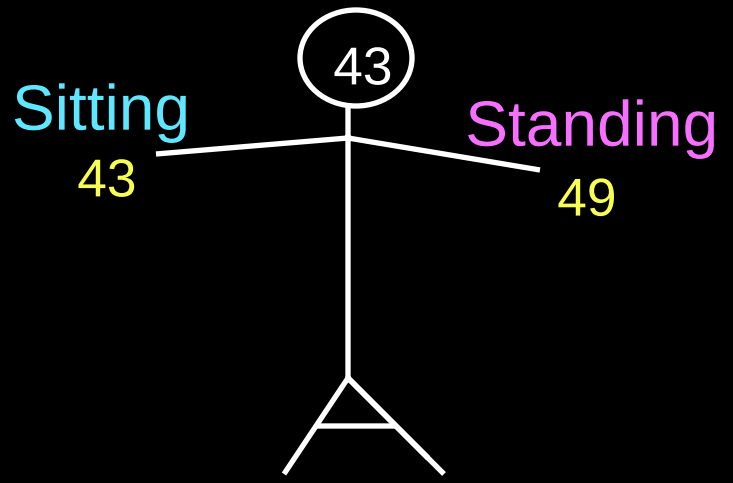
Population distribution of a hypothetical molecular variable



Population mean = 40, standard deviation (SD) = 5
Mean of females = 42, mean of males = 38
Within person replicate SD = 2
Assay replicate SD = 1.

Comparison 2: Between arms Within subjects

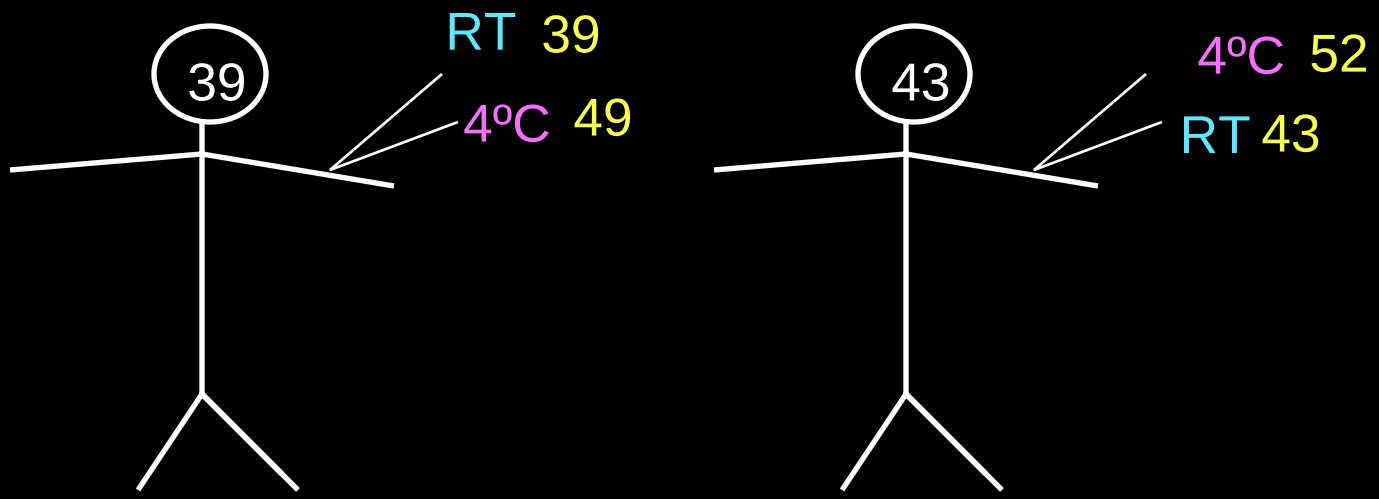




……

Sitting = baseline, Standing = +10

Comparison 3: Between sub-samples Within samples (arms)



RT = baseline, 4ºC = +10

Local control (of variability)

Dividing the available experimental units - sub-samples within arms within subjects - into groups of similar units, and assigning treatments to them, at random, as appropriate for the experimental comparisons.

We've already seen three examples:

Exercise, assigned to whole subjects (must be so)

Position, assigned to arm, within subject (design choice)

Temperature, assigned to sub-sample, within arm (ditto)

The science and art of statistical experimental design

Lies in exploiting the structure of the experimental units when making the assignment of the levels of treatment factors to the units. The general aim is:

- To assign different levels of a treatment to units that are as similar as possible, to maximize the precision of the estimate of the difference, but
- Bearing in mind the nature of the treatment factor.

Actual assignments should be done using a chance device.

The factorial principle (RA Fisher, 1926)

No aphorism is more frequently repeated in connection with field trials, than that we must ask Nature few questions, or, ideally, one question, at a time. The writer is convinced that this view is wholly mistaken. Nature, he suggests, will best respond to a logical and carefully thought out questionnaire; indeed, if we ask her a single question, she will often refuse to answer until some other topic has been discussed.

Suppose that we are interested in exploring the joint impact of Position and Temperature

		Position: Sitting	Position: Standing
Temperature	RT	Baseline	+10
	4°C	+10	?

Suppose that we are interested in exploring the impact of Position and Temperature

		Position: Sitting	Position: Standing
Temperature	RT	Baseline	+10
	4°C	+10	+20

Additive: no interaction

Suppose that we are interested in exploring the impact of Position and Temperature

		Position: Sitting	Position: Standing
Temperature	RT	Baseline	+10
	4°C	+10	+10

Non-additive: Interaction!

Advantages of factorial experiments

Permitting more than one variable to change in the experiment gives

- The possibility of observing interactions;
- Increased efficiency - getting more information per observation - even if no interactions are present

Features of experiments with blood

Designs with confounding principally of interest

There are two phases

Designs with confounding

These are efficient designs (i.e. ones with fewer runs) for estimating the main effects of factors, and their first-order interactions, assuming that higher-order interactions are absent, or relatively small. In jargon, they are confounded.

They can be amazingly efficient, i.e. lots of factors can be explored with comparatively few runs. Box, Hunter & Hunter give an introduction to these ideas. (Assumptions must be made.)

Two-phase experiments

Experiments to determine the effect of pre-analytical variables on molecular measurements on serum or plasma need to explore many variables, including patient, acquisition, processing and storage factors.

Phase one consists of the part leading to the blood, and so includes patient and acquisition variables.

Phase two concerns what happens to the blood, and so includes processing and storage factors.

For greater efficiency, blood experiments should explore factors from both phases simultaneously. Addressing the design issues here is a relatively recent research topic.

External validity

We want the conclusions of our experiment to apply more widely, ideally universally. No single experiment can guarantee this.

The conditions of the experiment may restrict the applicability of its conclusions. The subjects, the reagents, equipment and facilities, the phlebotomists, etc. may be atypical. Experiments don't always fully reflect day to day practice.

One way to work towards external validity is to conduct many experiments, under a wide range of conditions.

A second way is to carry out observational studies, and compare the results with those of the experiments.

From many things not mentioned

Deciding upon the right scale for the measurements in the experiment, e.g. log?

Determining which error is appropriate for any given treatment comparison. (Different errors can be appropriate for different comparisons.)

Describing a suitable two-phase design with confounding for the blood experiments. We have several.

Further references

Bailey, R.A. (2008) Design of two-phase experiments.
http://www.newton.ac.uk/webseminars/pg+ws/2008/doe/doew02/0813/bailey/

Vivacqua, C. A. and Bisgaard, S. (2009). Post-Fractionated Strip-Block Designs *Technometrics* ***51****: 47-55.*

Acknowledgements

Carla Vivacqua, UFRN, Brazil

Rosemary Bailey, QMUL, London

Helen Moore & Mark Lim, OBBR