

**Analysis of Bioassays:**  
***Personal reflections on the approach of  
the European Pharmacopoeia***

**Rose Gaines Das**

**Consultant in Biostatistics, UK**

***Formerly* Head of Biostatistics, National Institute for Biological  
Standardization and Control, UK**

**[gainesdasre@yahoo.co.uk](mailto:gainesdasre@yahoo.co.uk)**

## ***Brief outline***

- 1. Purpose and scope of the European Pharmacopoeia (EP), noting its emphasis on basic principles of design**
- 2. Brief summary of basic principles of design**
- 3. Comments on the EP approach to ‘non-parallelism’**
- 4. Comments on *in vitro* assays on micro-titre plates, with some examples**

## European Pharmacopoeia Chapter 5.3

# Purpose and Scope

- **Provides guidance: for design and analysis of bioassays prescribed in the EP**
- **Is not intended for statisticians ‘intended for use by those whose primary training and responsibilities are not statistics but who have responsibility for analysis or interpretation of the results of these assays, often without the help and advice of a statistician.’**
- **Is not mandatory ‘Alternative methods can be used and may be accepted by the competent authorities, provided that they are supported by relevant data and justified during the assay validation process’**

## Purpose and Scope

- **Covers a limited range of designs**
  - **3.1.3 Calculations and Restrictions: same number of dilutions for each preparation; equal number of experimental units for each treatment**
- **Recommends expert professional advice: for designs outside the covered range; for designs needed for research and development; for full validation; at various points in the text, special cases are noted and professional advice recommended**



# European Pharmacopoeia Chapter 5.3

## Contents

- 1. Introduction**
- 2. Randomisation and independence of individual treatments**
3. Assays depending upon quantitative responses
4. Assays depending upon quantal responses
5. Examples
6. Combination of assay results
7. Beyond this Annex
8. Tables and generating procedures
9. Glossary of symbols
10. Literature

## European Pharmacopoeia Chapter 5.3

### Applicability of statistical method

#### Section 1. Introduction

#### Section 1.1. General Design and Precision

***“In every case, before a statistical method is adopted, a preliminary test is to be carried out, with an appropriate number of assays, in order to ascertain the applicability of the method.”***

## European Pharmacopoeia Chapter 5.3

### Design and randomization

#### Section 2. Randomisation and independence of individual treatments

#### Failure of randomization and independence:

***inherent variability not fully represented in the experimental error variance, resulting in***

***unjustified increase in the stringency of the tests in analysis of variance***

***under-estimation of the true confidence limits***

## **European Pharmacopoeia Chapter 5.3**

### **Design and randomization**

#### **Section 3. Assays depending upon quantitative responses**

##### **Section 3.1.1. General principles**

*Assays are 'dilution assays' (condition of similarity)*

*Applicability of statistical methods requires*

- 1) Random assignment of treatments to experimental units*
- 2) Normal distribution of responses*
- 3) Homogeneity*

## European Pharmacopoeia Chapter 5.3

### Applicability of statistical method

**Section 3. Assays depending upon quantitative responses**

**Section 3.1.2. Routine assays**

***“The applicability of the statistical model needs to be questioned only if a series of assays shows doubtful validity.”***

***...in which case, a new series of preliminary investigations....***

# **Approach taken by the European Pharmacopoeia**

(adapted from presentation by A. Daas, EDQM)

***Maintain the classical approach whenever suitable***

***Do not explicitly describe alternative approaches because that is likely to confuse***

***Briefly stipulate that alternative approaches are allowed with references to relevant literature***

***Make clear that the chapter is not intended as a 'strait-jacket'***

***Recommend consultation with expert statisticians whenever the classical approach is found to be not suitable***

# Basic principles of assay design

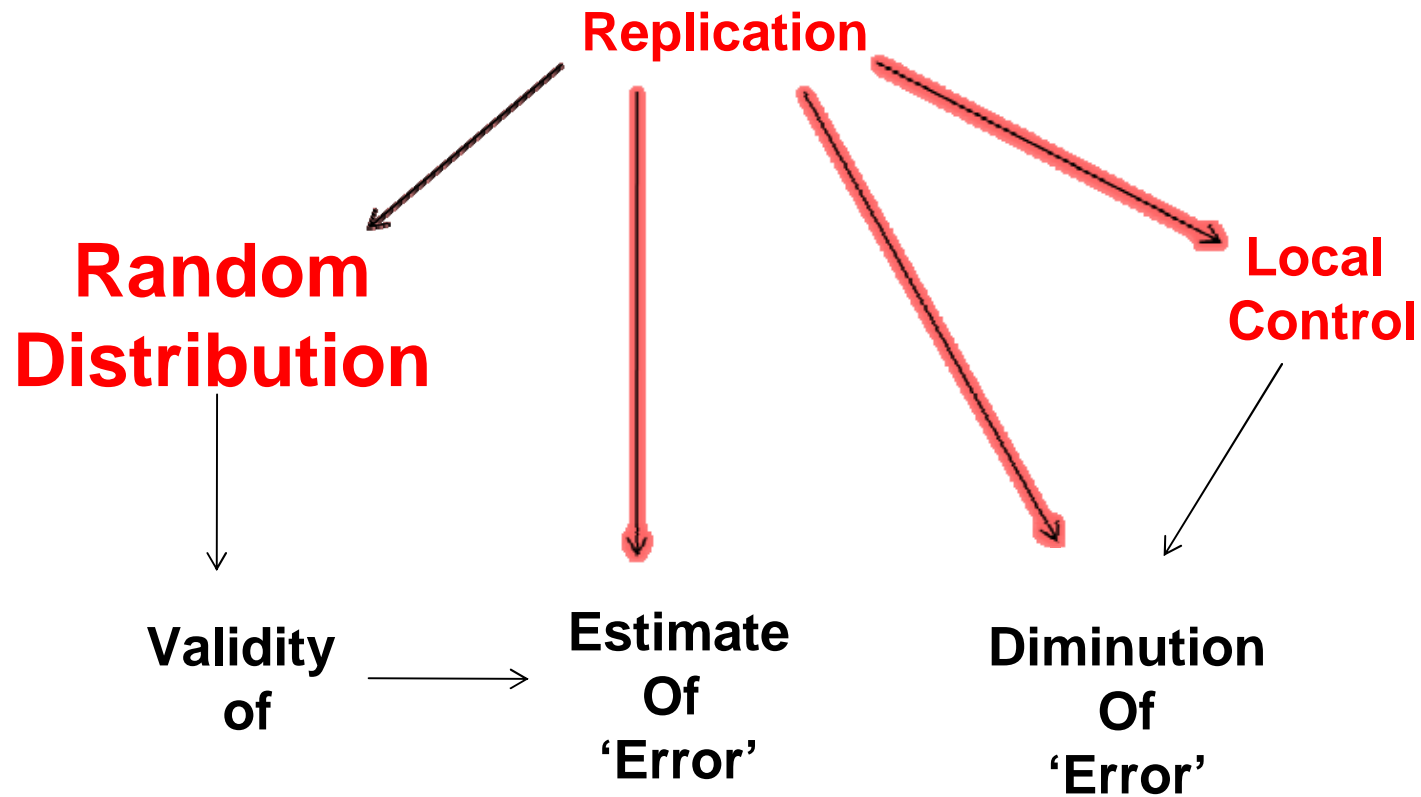


**R. A. Fisher**

**“To consult a 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.”**

*R. A. Fisher, 1938*

# Three fundamental principles of experimental design (R.A. Fisher 1931)



- **The three fundamental principles**
  - **Apply to the 'experimental unit'**
    - **Experimental unit** must be clearly **defined** and identified
    - **Experimental unit** can be **assigned at random** to any treatment; different units must be capable of receiving different treatments.
    - **Experimental units** must be **independent**. The treatment applied to one unit should not affect the treatment applied to any other unit.
    - The population from which experimental units are selected and the way they are selected should be defined and identified.
    - Failure to correctly identify the experimental unit is one of the most common mistakes in design and analysis of bioassays. A common error is to assume the experimental unit is the individual animal when all animals in a cage are treated identically

**Statistical methods used for analysis of assays are typically based on the assumptions that responses are:**

**Normally distributed**

*No? Consider transformation of responses / analysis not dependent on normality*

**Homogeneous**

*No? Consider transformation of responses / weighted analysis to take account of heterogeneity*

**Independent**

**No? Fatal flaw!**

# Failure of assumption of Independence

- **Unpredictably affects**
  - **significance level of tests**
  - **sensitivity of tests**
  - **accuracy of estimates**



# **Consequences of failure of Independence Unknown!**

- **Generalization is not possible**
- **In many biological situations**
  - **True variability is frequently larger than that estimated (too many significant results)**
  - **Statistical / numerical technique may not be invalidated but interpretation of results is uncertain and significance levels are ‘approximate’**
  - **(Cochran and Cox, 1957)**

- **“Randomization is one of the essential features of most experiments. The investigator who declines to randomize is digging a hole for himself, and he cannot expect the statistician to provide the ladder that will help him out.”**
- D. J. Finney, 1970

## EP 5.3

# Non-parallelism of dose – response curves

- ❖ **In 2005 a *National* delegation to the EP Commission requested revision of Chapter 5.3**
- ❖ **Summary of reasons for request:**
  - **Validity test for non-significant non-parallelism said to be unreasonable**
  - **Validity criterion for test frequently needs to be modified as a consequence of ‘overly precise’ data**
  - **Simplify approval of alternative statistical analyses (seemingly not in line with EP)**
- ❖ **Specific request: Commission should work towards solving acknowledged problem with the parallelism test**

## EP 5.3

# Non-parallelism of dose – response curves

Request for revision referred to Expert Group STA

- In response the Expert Group
  - Recognized issues raised
  - Questioned description of test as ‘unreasonable’
  - Noted that methods open for use are not restricted provided they are supported by relevant data
  - Proposed changes to text to emphasize ability to use alternative methods
  - Proposed additional section to discuss non-parallelism specifically
- Proposed revision published, adopted by EP Commission in 2006

## EP 5.3

# Non-parallelism of dose – response curves

## Revisions

- 1) **More explicit reference to use of alternative methods in Section 1, Introduction**

**‘Alternative methods can be used and may be accepted by the competent authorities, provided that they are supported by relevant data and justified during the assay validation process.’**

*Previously ‘Alternative methods may be used, provided that they are not less reliable than those described here.’ (EP 2005)*

- 2) **New Section added to Section 7, Beyond this annex**  
**7.6 Non-parallelism of dose – response curves**

## EP 5.3

# Non-parallelism of dose – response curves

## Key points of new section 7.6

(Part 1)

- ❖ **Similarity of dose – response curves is a fundamental criterion for valid potency estimation**
- ❖ **Similarity criterion frequently met by non-significant statistical test for deviations from parallelism**
- ❖ **Sources of statistically significant non-parallelism considered**
  - 1) **Failure of fundamental assumption of similarity**
  - 2) **Failure of statistical assumptions**

## EP 5.3

# Non-parallelism of dose – response curves

## Key points of new section 7.6

(Part 2)

### ❖ Sources of statistically significant non-parallelism

#### 1) Failure of fundamental assumption of similarity

Possible solutions: suitable standard, more specific assay system, involvement of regulatory authorities

#### 2) Failure of statistical assumptions

Possible solutions: modifications of assay design, more appropriate analysis, more correct estimate of residual error

### ❖ Section 7.6 concludes

**“No simple, generally applicable statistical solution exists to overcome these fundamental problems. The appropriate action has to be decided on a case-by-case basis “**

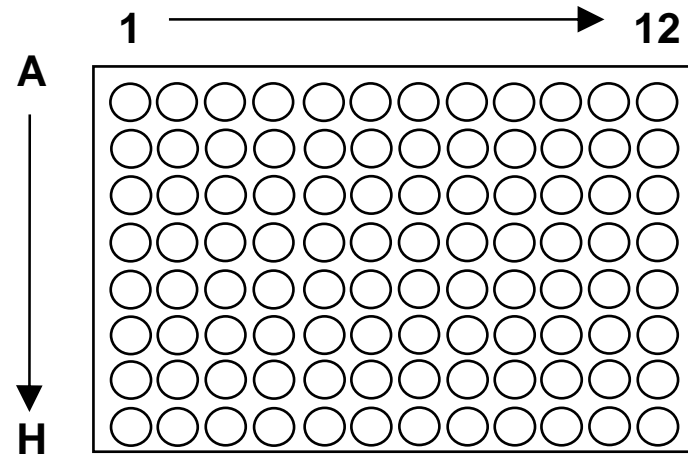
## ***In vitro* assays**

### **Cell based bioassays carried out on micro titre plates**

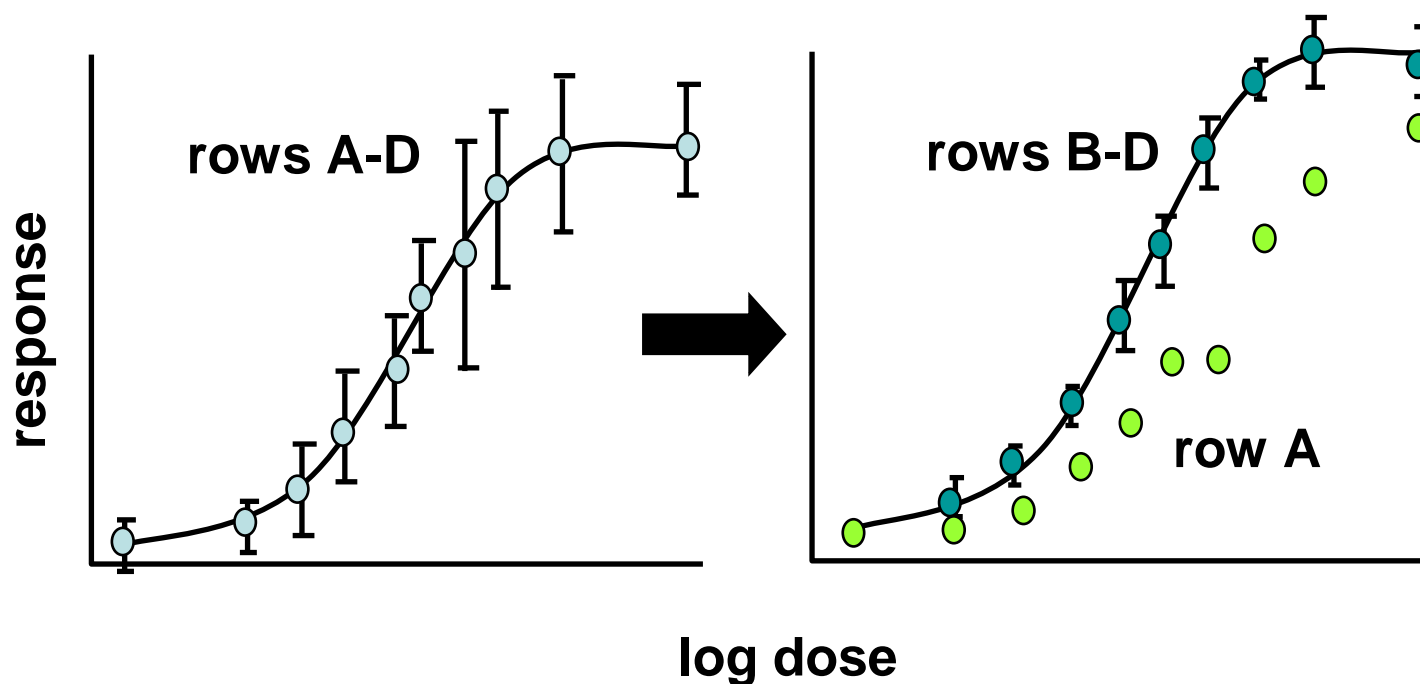
- **Assays on micro titre plates seldom ‘randomized’**
  - **Time**
  - **Feasibility**
  - **Possibility of errors**
  - **Ease of serial dilutions**
  - **Coding link between treatment and position difficult**
  
- **Wells on a plate are not ‘identical’ although frequently they are treated as such**

# The 96-well plate format used for many in vitro assays

- 8 x 12 well : common format
- supporting instrumentation, software
- edge, row, column, plate effects



# Position may affect dose – response curve

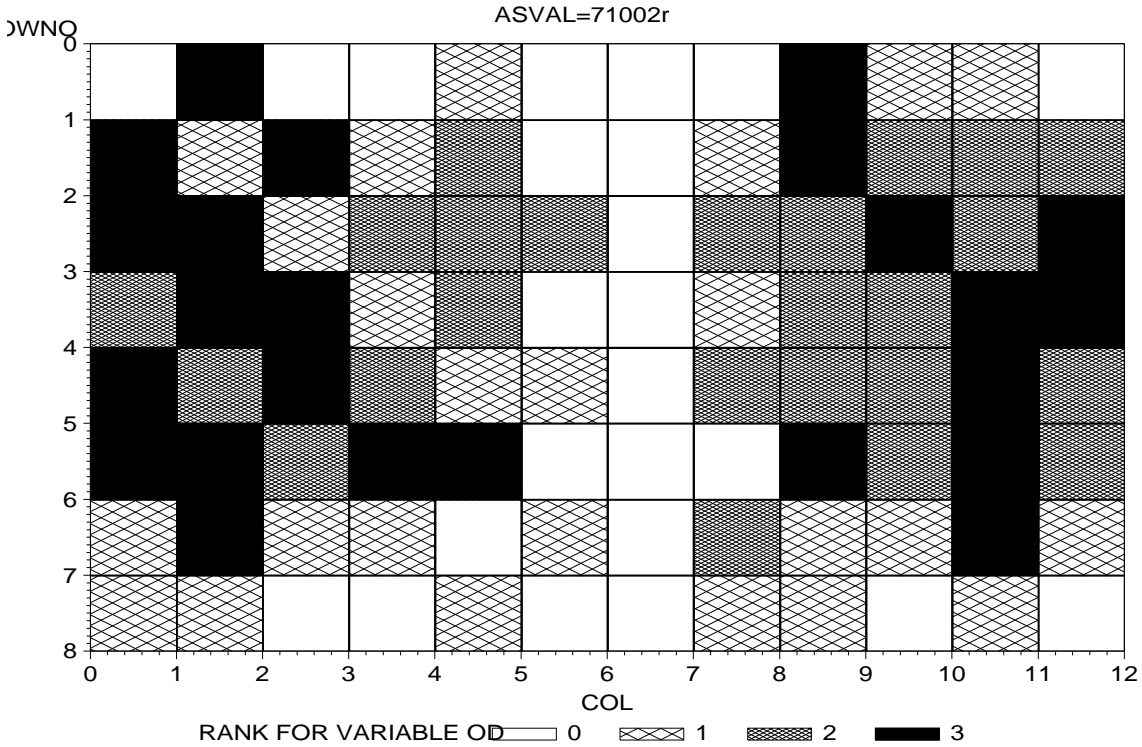


Row A is anomalous. Note that it was necessary to have an appropriate graphical presentation of data for this effect to be detected. Possible causes: edge effect, serial dilution effect

## **Cell-based bioassays: Uniformity plates**

- **Wells on a plate are not ‘independent’ and ‘identical’ although frequently they are treated as such**
- **Uniformity plates: All wells on a plate are treated with the same reagents and same dose of analyte**
- **Aim: To assess ‘plate effects’**
- **The following slides show uniformity plates with responses ranked for size into quartiles**

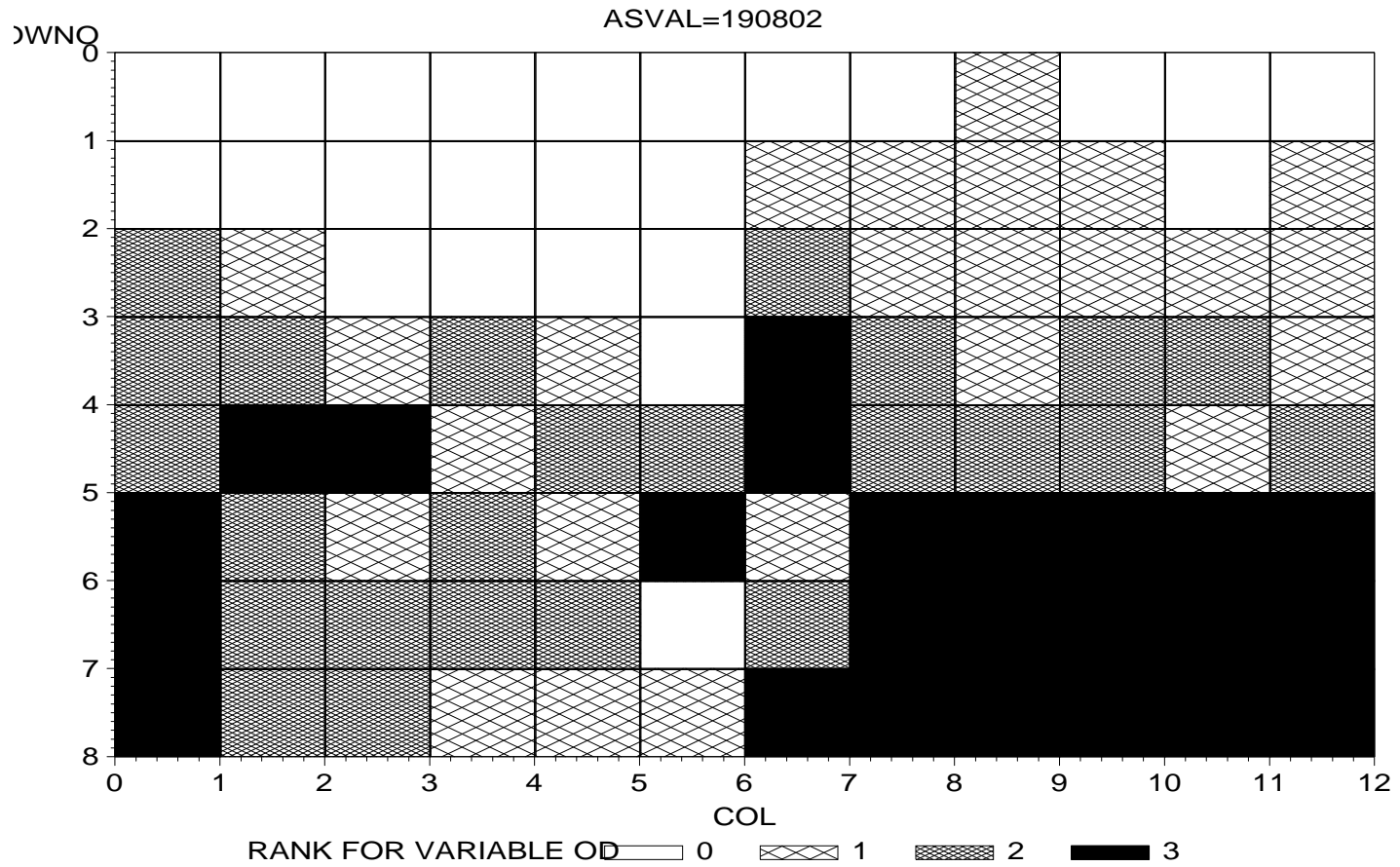
# Uniformity plates from same assay system: Plate 1



Shading shows magnitude of responses with larger responses more darkly shaded

- Smallest 25% of responses
- Largest 25% of responses

# Uniformity plates from same assay system: Plate 2



Shading shows magnitude of responses with larger responses more darkly shaded

- Smallest 25% of responses
- Largest 25% of responses

# Uniformity plates show row / column effects

## Causes??

- **Wells within plate affected by**
  - **Position on plate**
    - **Effect of neighbouring wells**
  - **Position in time**
    - **For each added reagent**
    - **Interaction of time with 'mixture'**
  
- **Effect of 'environment' on plate as a whole**



Photograph courtesy of  
C.J. Robinson

# What happens when there is non-random assignment?

*i.e. what is the effect of non-random assignment on the statistical analysis / interpretation?*

- Effects not known **unpredictable**
- Depend on how treatments are assigned and how this relates to the pattern of variation in responses across the plates
- An example using uniformity plates

# Randomization of treatments on 96-well microtitre plate

- **Comments from Hooten JWL and Paetkau V (1986). Random assignment of treatments in a 96-well (8 x 12) microtiter plate. J. Immunological Methods 94: 81-89.**
- **“The difficulties in placing different treatment aliquots without error into the wells of a microtiter plate when there is no predictable sequence are considerable.”**
- **Uniformity plates were used to evaluate effect of different designs by assigning ‘hypothetical’ treatments;**
- **No treatment and therefore expect variability between and within treatments to be the same; Expect F ratio to be 1**
- **Three ‘designs’**
  - **Adjacent placement**
  - **Quadrant placement**
  - **Random placement**

- **Adjacent placement: Error Mean Square was seriously underestimated in all plates giving in all cases apparently significant treatment effects where there are none**

Source of Variability	Trial 1		Trial 2	
	Mean Square	F	Mean Square	F
Treatments	13.06	2.96* <b>&gt;&gt;1</b>	9.46	2.18* <b>&gt;&gt;1</b>
Residual Error	4.41		4.33	

	1	2	3	4	5	6	7	8	9	10	11	12
A	1	2	3	4	5	6	7	8	9	10	11	12
B	1	2	3	4	5	6	7	8	9	10	11	12
C	13	14	15	16	17	18	19	20	21	22	23	24
D	13	14	15	16	17	18	19	20	21	22	23	24
E												
F												
G												
H												

- **Quadrant placement: Error Mean Square was consistently overestimated and likely to obscure treatment effects if there were any ('true' treatment effects may not appear significant)**

Source of Variability	Trial 1		Trial 2	
	Mean Square	F	Mean Square	F
Treatments	5.93	0.52* <<1	6.55	0.91 </~1
Residual Error	11.39		7.18	

	1	2	3	4	5	6	7	8	9	10	11	12
A	1	5	9	13	17	21						
B	2	6	10	14	18	22						
C	3	7	11	15	19	23						
D	4	8	12	16	20	24						
E							1	5	9	13	17	21
F							2	6	10	14	18	22
G							3	7	11	15	19	23
H							4	8	12	16	20	24

- Random placement: Error Mean Square consistently provided a reasonable estimate of real error

Source of Variability	Trial 1		Trial 2	
	Mean Square	F	Mean Square	F
Treatments	8.99	1.07 ~ 1	7.22	1.11 ~ 1
Residual Error	8.40		6.52	




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

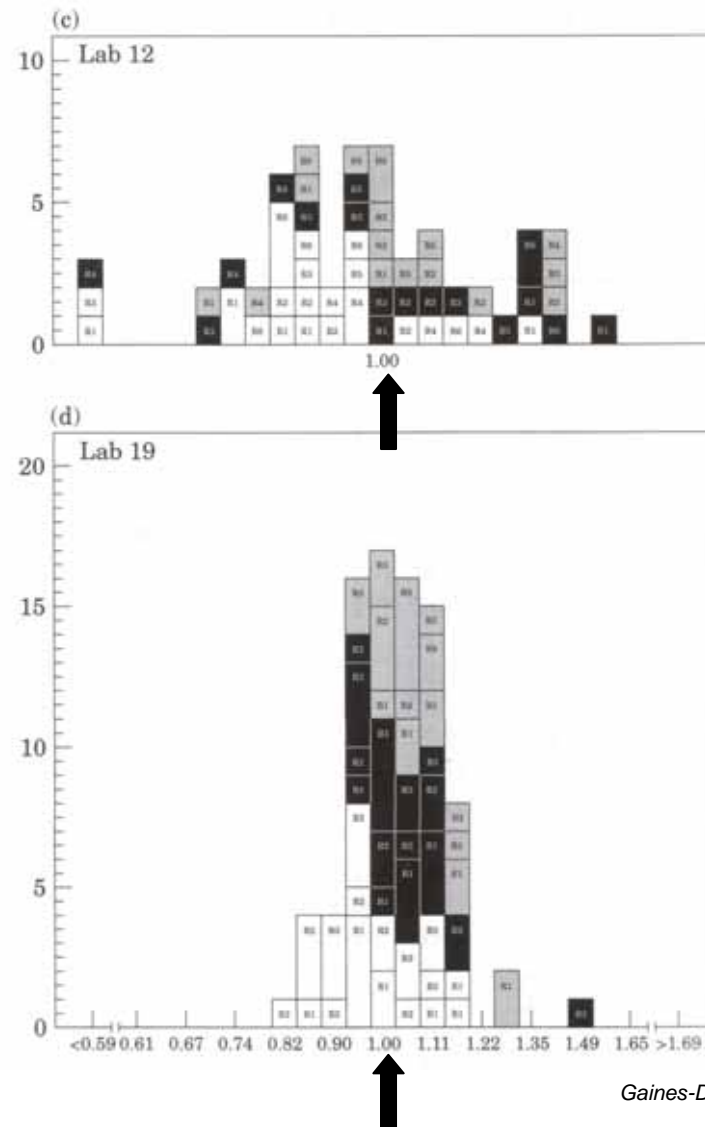
# What happens when there is non-random assignment

- Effect on analysis / interpretation can be unpredictable
- Depend on how treatments are assigned and how this relates to the pattern of variation in responses across the plates
- *Adjacent placement is common – leads to overestimation of ‘treatment’ effects / underestimation of residual error and hence occurrence of ‘too many significant effects’*
- The following examples use relative potency estimates for duplicate dilution series within plates and for coded duplicate preparations in many assays

- **Consequences of 'row effects' on estimates of relative activity based on comparisons between adjacent nominally identical rows in two different laboratories**

# Estimates of relative activity of one row in terms of the adjacent nominally identical row

- Each square represents one estimate
- Expected value for each estimate is 1 (↑), because rows are nominally identically treated
- The shading shows which rows are compared
- C and B open squares 
- E and D solid squares 
- F and G shaded squares 
- **Note: Estimates of relative activity show a consistent bias across multiple independent plates**
- (Figure 3 from Gaines-Das and Meager in *Biologicals* 1995; 23: 285-297)

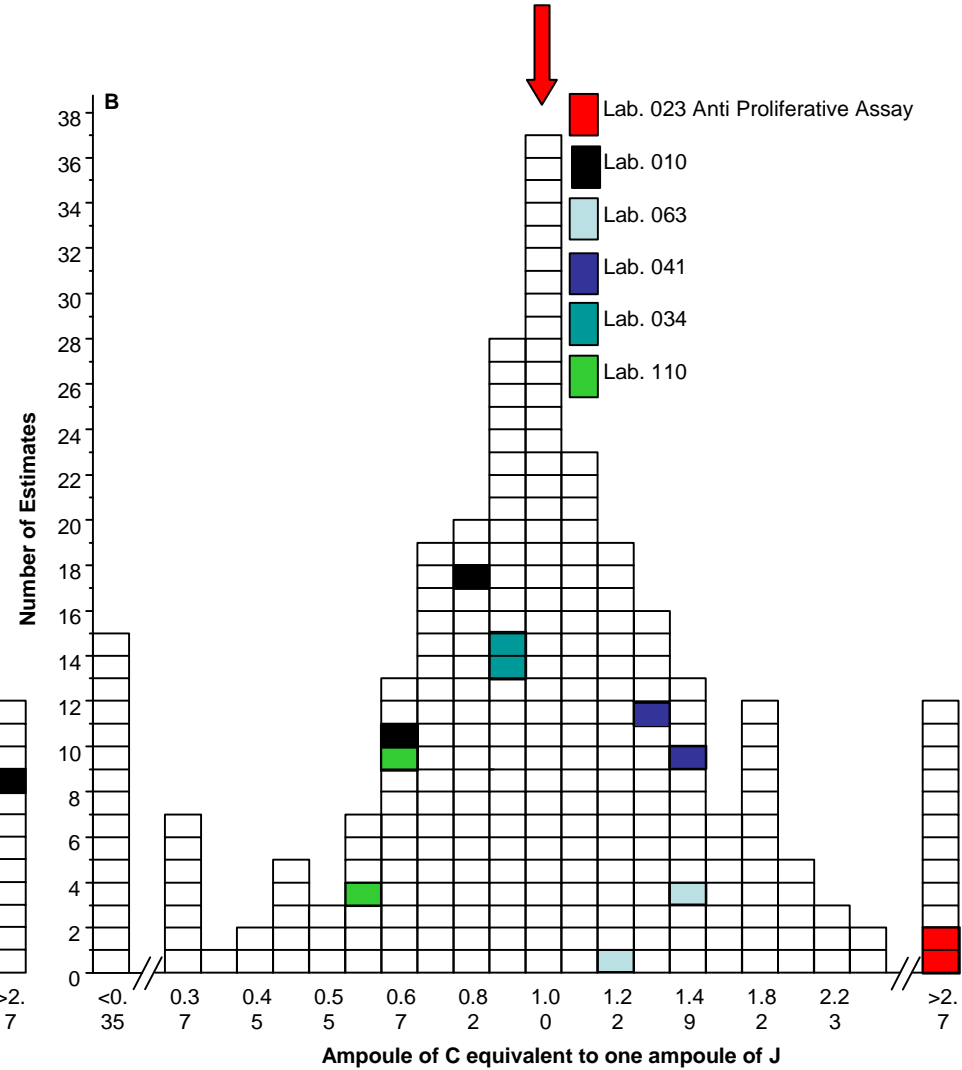
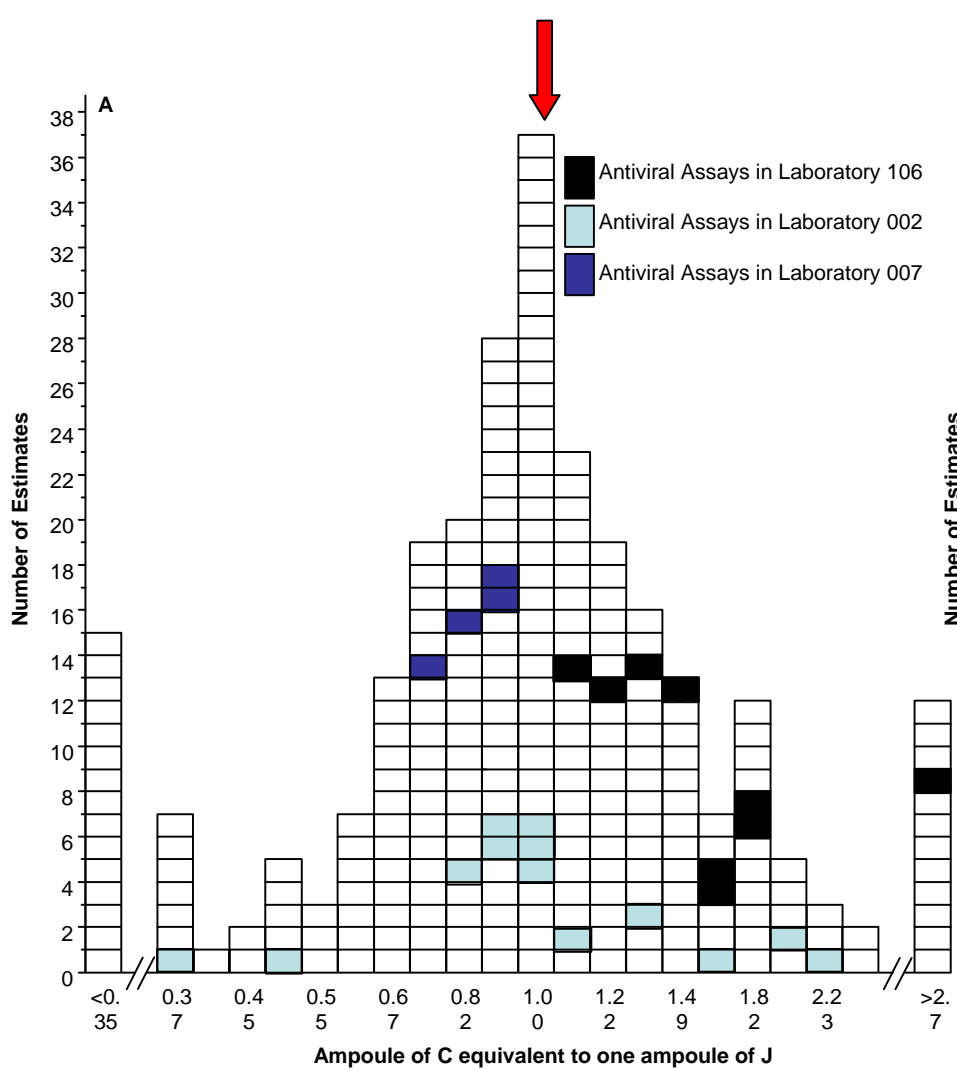


- **Consequences of structured non-random designs for estimates of relative potency based on within assay comparison of **coded duplicate preparations****
  
- **Potency estimates for two ampouled materials identical except for code (coded duplicates) obtained in independent assays**
  
- **Therefore expected potency is 1 in all cases**
  
  
- Figure 4 taken from Gaines-Das RE and Meager A (1995). Evaluation of Assay Designs for Assays Using Microtitre Plates: Results of a Study of In Vitro Bioassays and Immunoassays for Tumour Necrosis Factor (TNF). *Biologicals* 23: 285-297



- **Coded duplicates** from an international collaborative study of interferon alpha
- **Potency estimates for two ampouled materials identical except for code (coded duplicates) obtained in independent assays**
- **Therefore expected potency is 1 in all cases**

# Coded Duplicate Preparations



## Estimates of relative potency for **coded duplicates** showed bias in many laboratories

- **Bias may be positive or negative** (Estimates may be consistently larger or smaller than the expected value of 1.)
- **Between assay variability is not consistent between laboratories even when the same reference method is used** (compare for example, TNF estimates in laboratories 15 and 19)
- **Distribution of estimates over all laboratories is consistent with expected value of 1, but estimates in individual laboratories are frequently biased to an unpredictable extent**

# **Design: 96 well micro titre plates**

## **How?**

- **Wells on a plate are not ‘independent’ and ‘identical’ although frequently they are treated as such**
- **‘Awareness’ is the first step**
- **General solution is not possible, but some practical approaches can be considered**
- **Coded duplicate preparations can help to provide a measure of the effects within a specified design**

# Design: 96 well micro titre plates

- complete randomization is often impossible
- for replicates, try to avoid clustering, with the same or “equivalent” positions always used for the same samples
- - assessing sources of bias
  - – ‘uniformity plates’ may help
  - beware of linked causes of variation, for example, sequence of dosing linked to plate position
- analysis may make some allowance for structure of design and position on plate

# Conclusion

**Always remember / be aware of**

- **Principles of assay design**
- **Underlying statistical assumptions**
  - **Independence**

# Three questions

- 1) **Are the experimental units clearly defined and identified?**
  - ❖ **Has pseudo-replication been avoided?**
- 2) **Have the three fundamental principles been correctly applied?**
  - ❖ **Controls? Replication? Randomization?**
- 3) **Are the statistical assumptions met?**
  - ❖ **Normality? Homogeneity? Independence?**



**Yes** (to all)  
**Well done!**  
**Proceed with analysis**



**NO** (to any)  
**Stop and think**  
**Analysis may not be valid**

