

Effective Use of Data in Incomplete Crossover Designs: Theory and Application

Yonghua Wang
Global Biometric Sciences
Bristol-Myers Squibb

Outline

- Introduction
- The Design
- Statistical Methods
- Understanding of the Analyses
- Summary

Disclaimer

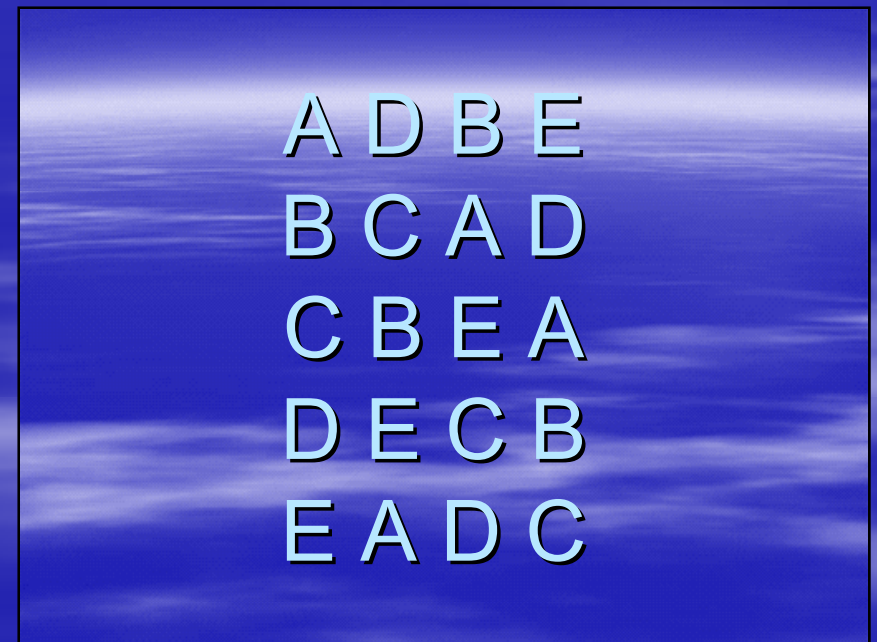
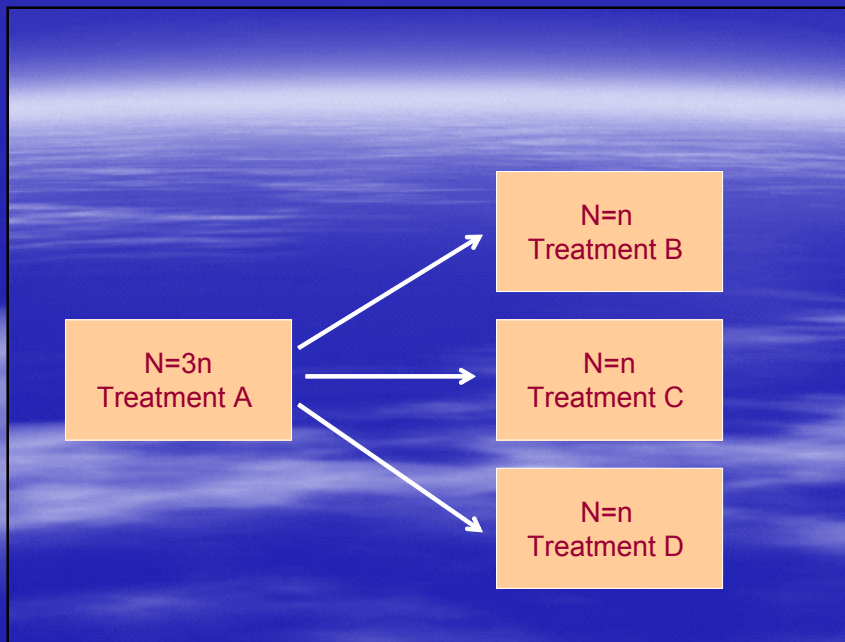
- This is not about new statistical theories or methods even though I went through serious formulation generation.
- This is about statistical understanding of the theory when it is used in application
- This is NOT based on the FDA guidance (because I could not find one).

Introduction

- The causes of an Incomplete and/or Unbalanced Designs can be:
 - Unintentional
 - Intentional
- There is a need to fully understand the statistical analysis methods for this kind of data

Introduction

- Unintentional Incomplete Designs are caused by Dropouts
- Intentional Incomplete and/or Unbalanced Designs



Introduction

- The research initiated with ClinPharm studies because of the incomplete designs for different reason and the small sample sizes used by each treatment group.
- Easily extend to Phase II/III large studies when there drop outs and changing from baseline at the later time point is the interest of the study.
- While the approach is applicable in any situation, the examples will be concentrated on ClinPharm type of studies

Introduction

The ClinPharm examples used in this presentation:

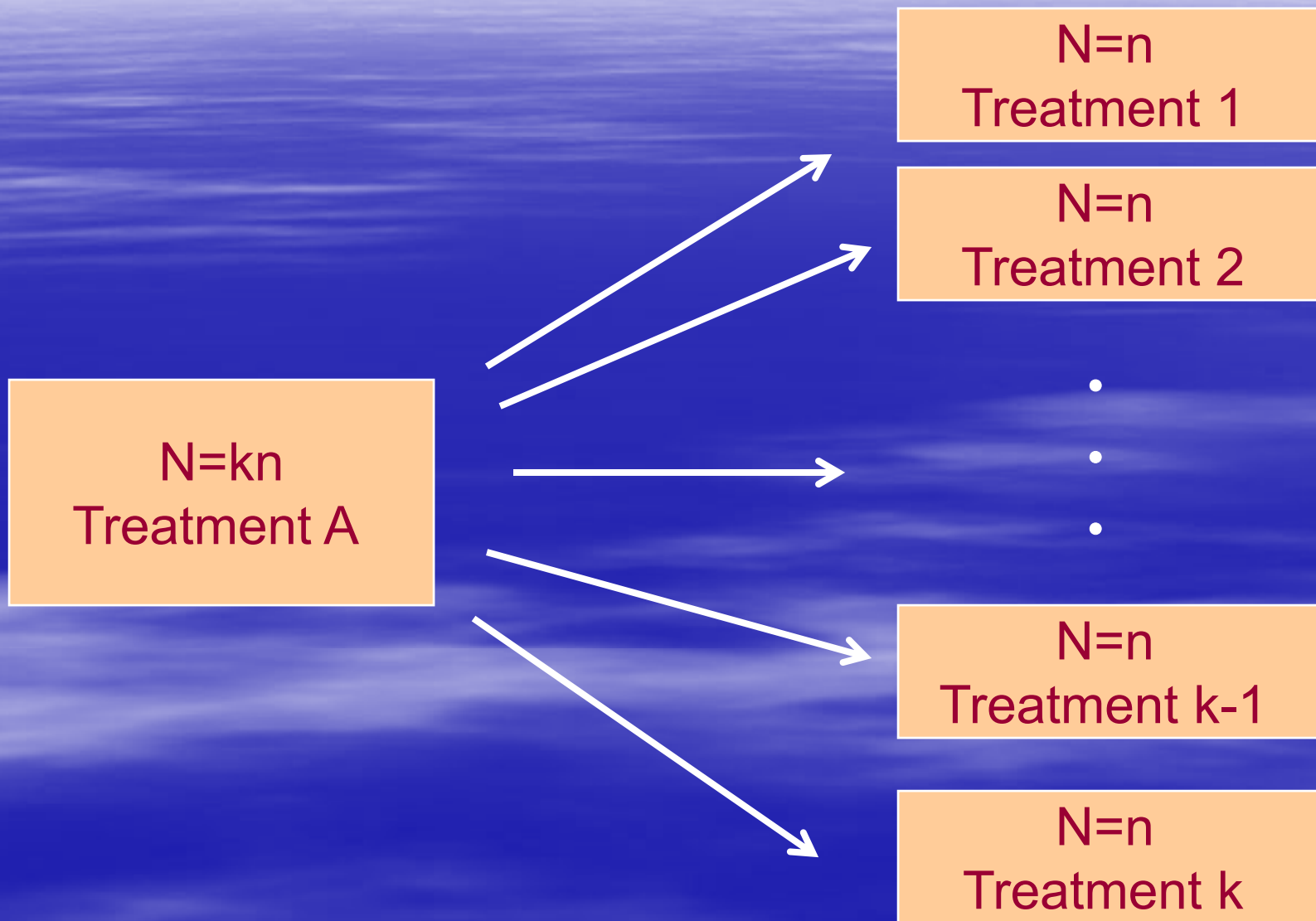
- The bioavailability of a promising drug can be reduced when it must be co-administered with another drug.
- In some disease areas, the co-administration of several different drugs is necessary (e.g. HIV infection).
- There is a need to identify one or more schemes for the co-administration of two drugs that keeps both drugs' bioavailability (efficacy).

The Design

Two-Period Sequential Crossover

- Each period can take either several days or several weeks to reach a steady state in each period, depending on the drug.
- Randomization occurs at the beginning of the second period.
- PK samples are collected at either the end (steady state) or the beginning (single dose) of the each period.

A General Design



Design: An example

Treatment A = Drug X

Treatment B = Drug X + Drug Y *at the same time*

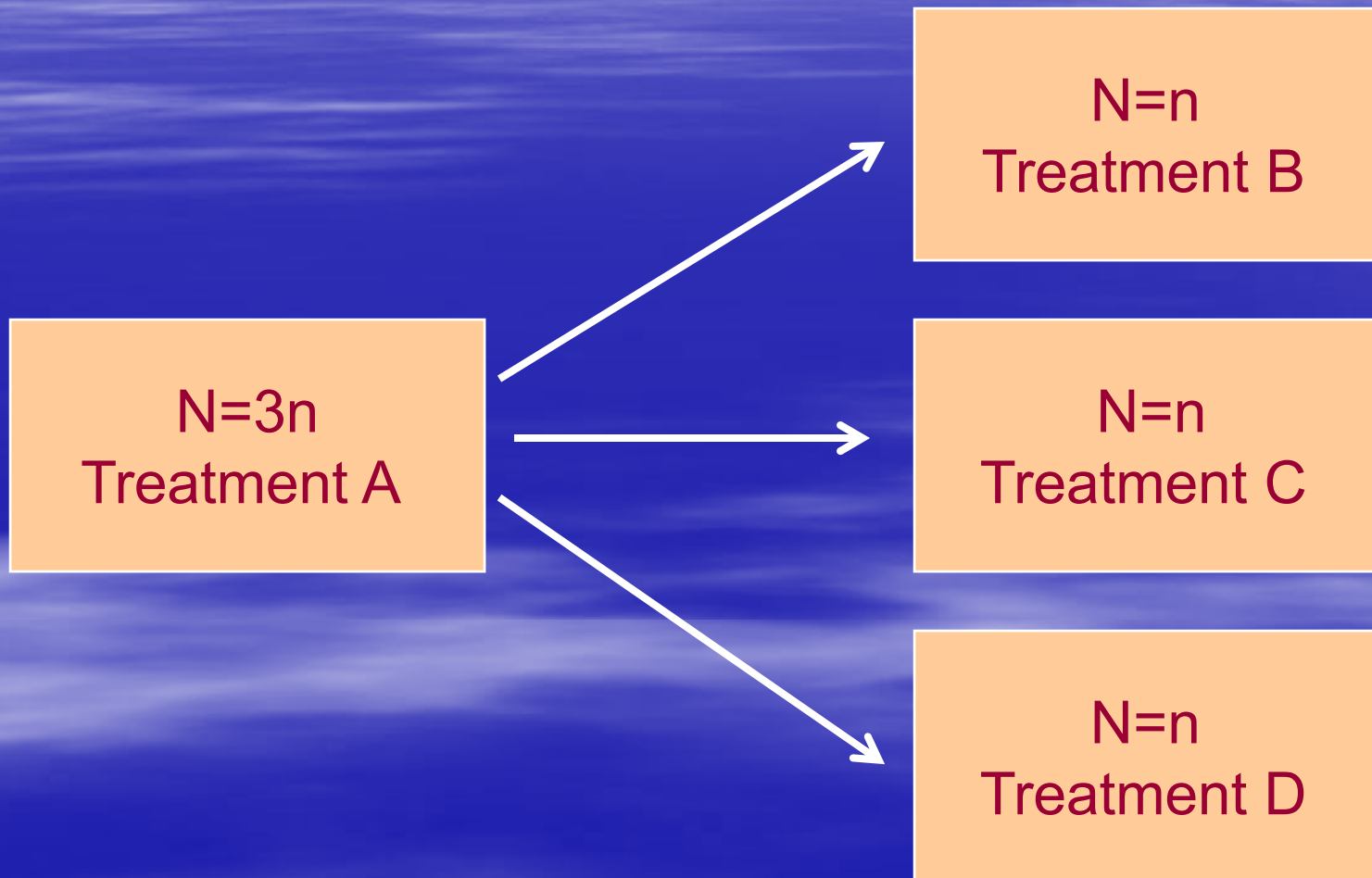
Treatment C = Drug X + Drug Y *2 hours apart*

Treatment D = *Higher doses of* Drug X + Drug Y

Assumption: Drug Y reduces bioavailability of Drug X

Objective: Identify a scheme that both drugs can be co-administered safely and effectively

The Design



Analyses Objectives

The Comparisons:

The bioavailability of Drug X in
Treatment B, C, or D
to that in Treatment A

Challenges

It is an **unbalanced and incomplete** design with partial **correlated data**. (this is the case when wk 48 value comparing to baseline value in Phase II/III)

In sequence AB, A and B are correlated.

In sequence AC, A and C are correlated.

In sequence AD, A and D are correlated.

Statistical Methods

- **Easy ways to handle the incompleteness:**
 - Avoid incomplete designs and always use complete balanced crossover
 - Ignore data from incomplete part
 - Assuming parallel, ignore correlation:
ANOVA or ANCOVA
- **Harder way to deal with incomplete balanced/unbalanced crossover:**
general linear model with measurements within subject as repeated measures

A Illustration on When Incomplete Part Ignored

Only use subj. in TrT A that crossed to TrT B
to compare TrT A and TrT B:

$$(1) \text{Var}(\bar{A}_{A(B)} - \bar{A}_B) = 2 (1 - \rho) \sigma^2 / n$$

Use all subj. in TrT A:

$$(2) \text{Var}(\bar{A}_A - \bar{A}_B) = [1 + (1 - 2\rho)/k] \sigma^2 / n$$

⇒ $\rho = 1/2$, the two are the same

⇒ $\rho < 1/2$, (1) > (2)

⇒ $\rho > 1/2$, (1) < (2)

Statistical Methods

- **Easy ways to handle the incompleteness:**
 - Avoid incomplete designs and always use complete balanced crossover
 - Ignore data from incomplete part:
 - Assuming parallel, ignore correlation: **inefficient**
- **ANOVA or ANCOVA**
- **Harder way to deal with incomplete balanced/unbalanced crossover:**
 - general linear model with measurements within subject as repeated measures

Statistical Methods

- **Easy ways to handle the incompleteness:**
 - Avoid incomplete designs and always use complete balanced crossover
 - Ignore data from incomplete part:
 - Assuming parallel, ignore correlation:
ANOVA or ANCOVA
- **Harder way to deal with incomplete balanced/unbalanced crossover:**
general linear model with measurements within subject as repeated measures
PROC MIXED – are you using it correctly?

Statistics Modeling

To illustrate, we will work with the reduced design of treatments A, B, and C.

$$\begin{pmatrix} Y_{AB} \\ Y_B \\ Y_{AC} \\ Y_C \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mu \\ \beta \\ \eta \end{pmatrix} + \varepsilon$$

Commonly Used Method – Random Effect

PROC MIXED;

class treatment subject sequence;

model y = treatment;

random subject /group=sequence;

$$\begin{pmatrix} \sigma_0^2 + \sigma^2 & \sigma_0^2 & & & & \\ \sigma_0^2 & \sigma_0^2 + \sigma^2 & & & & \\ & & \sigma_1^2 + \sigma^2 & \sigma_1^2 & & \\ & & \sigma_1^2 & \sigma_1^2 + \sigma^2 & & \\ & & & & & \end{pmatrix}$$

Statistics Methods for the Analysis

PROC MIXED;

class treatment subject sequence;

model y = treatment;

repeat /subject=subject group=sequence type=un;

$$\begin{pmatrix} \sigma_1^2 & \rho_b \sigma_1 \sigma_2 & & & & \\ \rho_b \sigma_1 \sigma_2 & \sigma_2^2 & & & & \\ & & \sigma_4^2 & \rho_c \sigma_4 \sigma_3 & & \\ & & \rho_c \sigma_4 \sigma_3 & \sigma_3^2 & & \end{pmatrix}$$

The Sequence Effect was not Used in the Model Statement!

- Unbalanced design—
 - all subjects take Treatment A in the first period
 - in the second period, the sequence overlaps with the treatments.
- The sequence effect is not of interest.
 - The adjustment of the model was carried out by individual subjects.
- The sequence effect in the model will cost efficiency in estimating treatment difference.

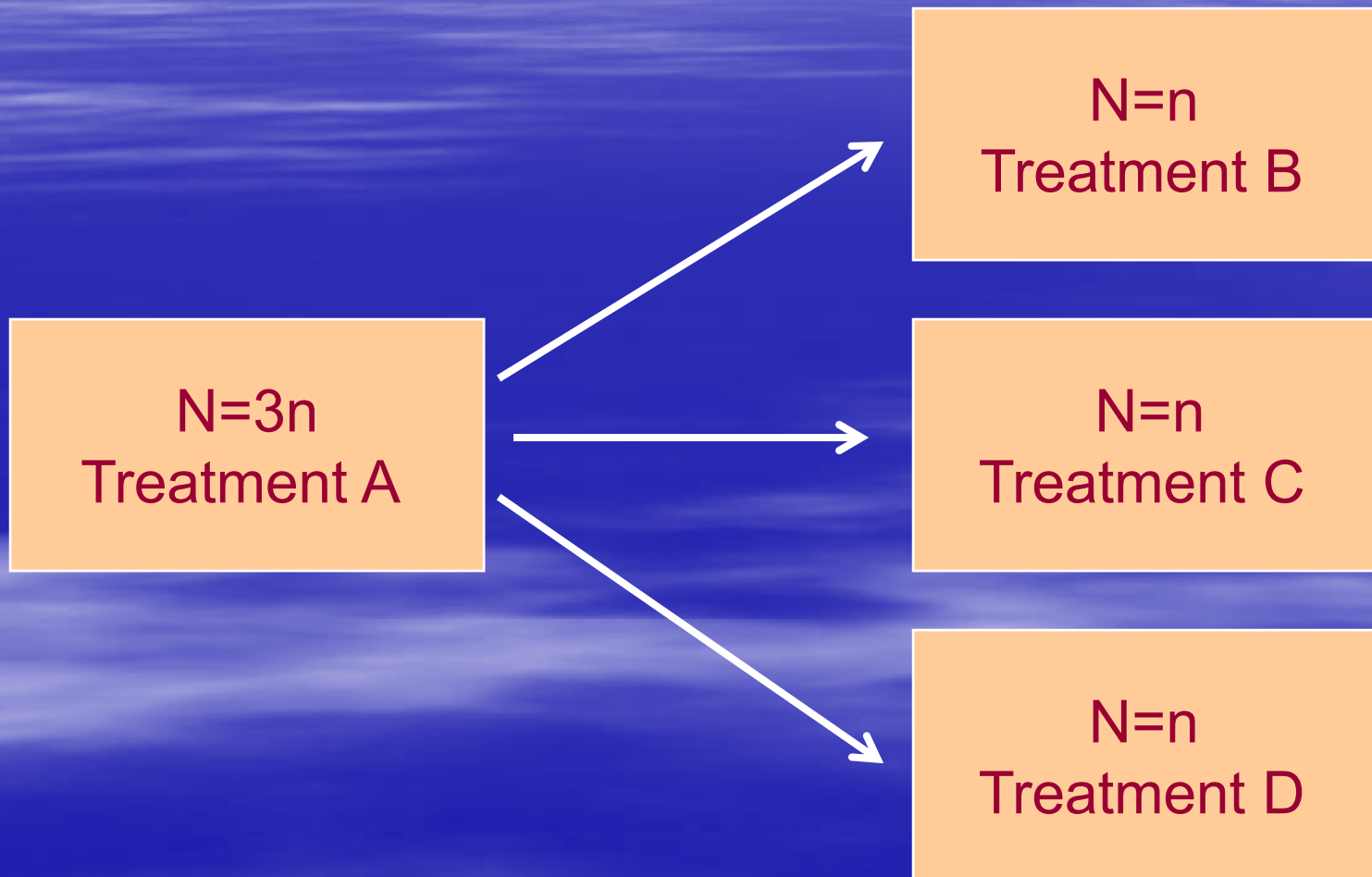
Different Statistical Methods

1. Repeated measurement model including only subjects with both treatments for the comparison
2. Repeated measurement model with the sequence effect using all subjects
3. Repeated measurement model without the sequence effect but including all subjects
4. Mixed effect model with subject as random effect
5. Repeated measurement model with all observations of Treatment A comparing to those of Treatment B (do not use Treatment C&D)

Comparing the Efficiency Among the Methods

- Step 1: derive the theoretical standard deviation for LSM estimate
- Step 2: simulate data
- Step 3: summarize simulated data for different methods

Using This Design in the Simulation



Simulation Assumptions

A and B, A and C, A and D are from the bivariate normal distribution.

Treatment	Mean	Standard Deviation	Correlation
A	1.5	0.8	
B	1.2	0.2	0.5
C	1.2	0.7	0.3
D	1.2	0.4	0.3

Simulation Results

100 Simulations, $k=3$, $n=15$, $N=kn=45$,

Estimating:

- difference between A and B
- the standard error of the estimate

Methods	Mean (SD)	Standard Deviation (SD)
1. Just AB sequence	0.3059 (0.1029)	0.1904 (0.0357)
2. With sequence effect	0.3059 (0.1029)	0.1904 (0.0357)
3. No sequence effect	0.2990 (0.0908)	0.1123 (0.0116)
4. Subject as random effect	0.2986 (0.0897)	0.1797 (0.0151)
5. Use all observations from A	0.2983 (0.0904)	0.1150 (0.0115)

Theoretical Standard Error

- Methods 3 and 5 give unbiased estimates of the actual standard deviation of the LSM estimate of the treatment difference (standard error for the actual mean) when the covariate matrix structure is build by following block as:

$$\begin{pmatrix} \sigma_1^2 & \rho_b \sigma_1 \sigma_2 \\ \rho_b \sigma_1 \sigma_2 & \sigma_2^2 \\ & & \sigma_1^2 \end{pmatrix}$$

The theoretical standard deviation for LSM estimate is:

$$\{[k-(k-1) \rho_b] \sigma_2^2 + \sigma_1^2 - 2 \rho_b \sigma_1 \sigma_2\} / kn \quad (=0.1135, k=3)$$

Simulation Results

- Bigger differences between Models 3 and 5 maybe seen if the correlation between A and C (D) increases.
- Methods 1, 2, and 4 gives biased estimation of the theoretical standard deviation of the LSM estimate of the difference.

Real Data Analyses

Methods	Mean	Standard Error	P-value
1. Just AB sequence	-0.2622	0.1283	0.0603
2. With sequence Effect	-0.2622	0.1283	0.0476
3. No sequence effect	-0.2626	0.1197	0.0340
4. Subject as random effect	-0.2631	0.0955	0.0087
5. Use all observations from A	-0.2651	0.1200	0.0324

Conclusion on the Methods

- The repeated measurement model should be used instead of the mixed effect model with the subject as the random effect if the correlation structure is unknown.
- The sequence effect negatively affects the efficiency of the estimate. It should not be included in the model.
- The full model gives results similar to those of the model without treatments C and D when comparing A to B using all data observed from A.
 - For an unbalanced and incomplete design with partial subjects crossed over to the next treatment, all data should be included in the analysis.

Understanding of the Analyses

Perfect world:

within subject comparison in a balanced and complete crossover study and no dropouts

Statistical understanding:

The subjects in this study are a random sample of the general population. The treatment difference calculated from these subjects representing the treatment difference of the general population.

Understanding of the Analyses

Real world:

A study can be an incomplete crossover study due to design or dropout.

Statistical understanding:

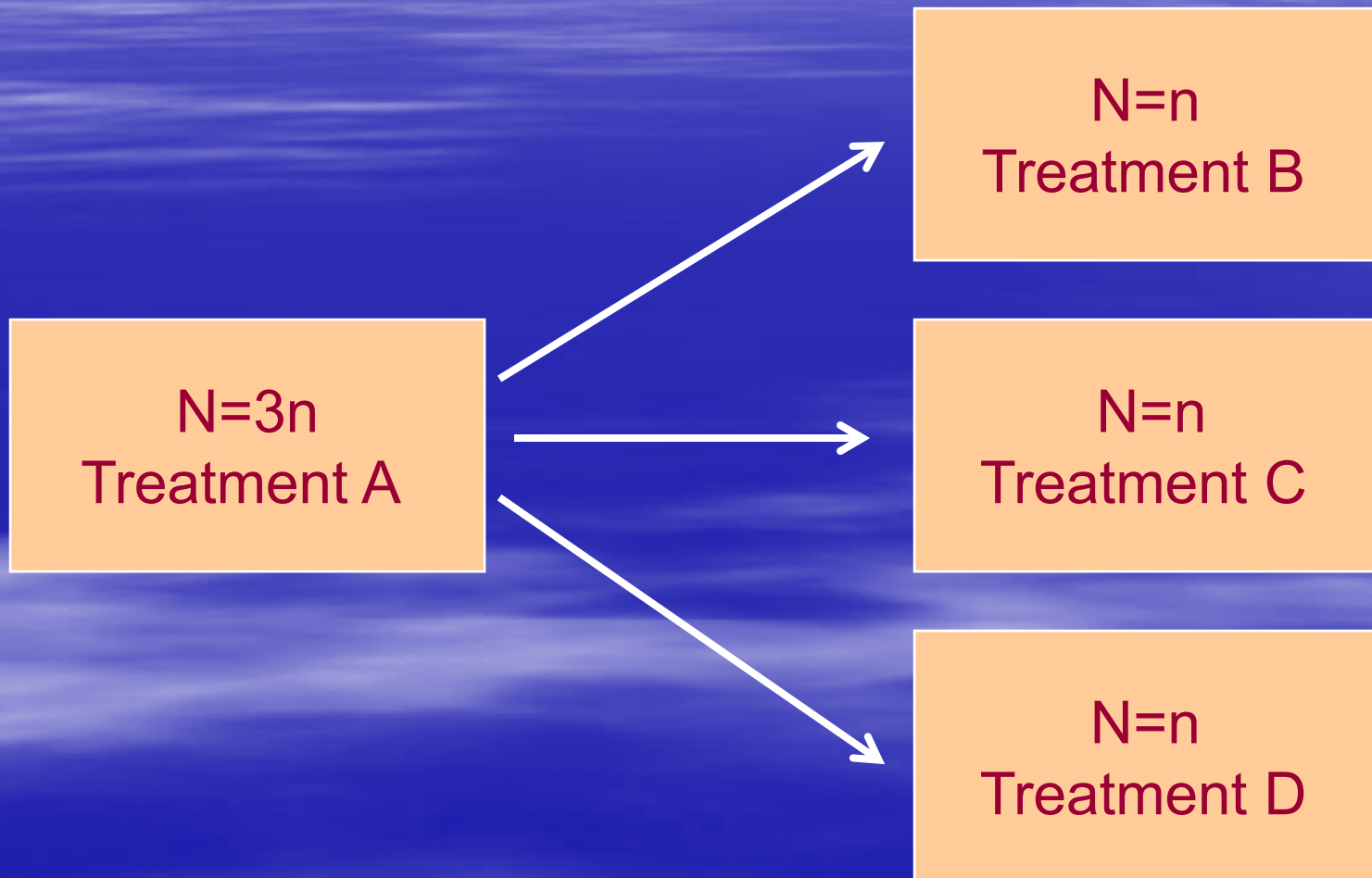
- The subjects in Treatment A are a random sample of the general population with Treatment A.
- The subjects in Treatment B are a random sample of the general population with Treatment B.
- The treatment different in the general population is based on the inference from the two-sample comparison

Understanding of the Analyses

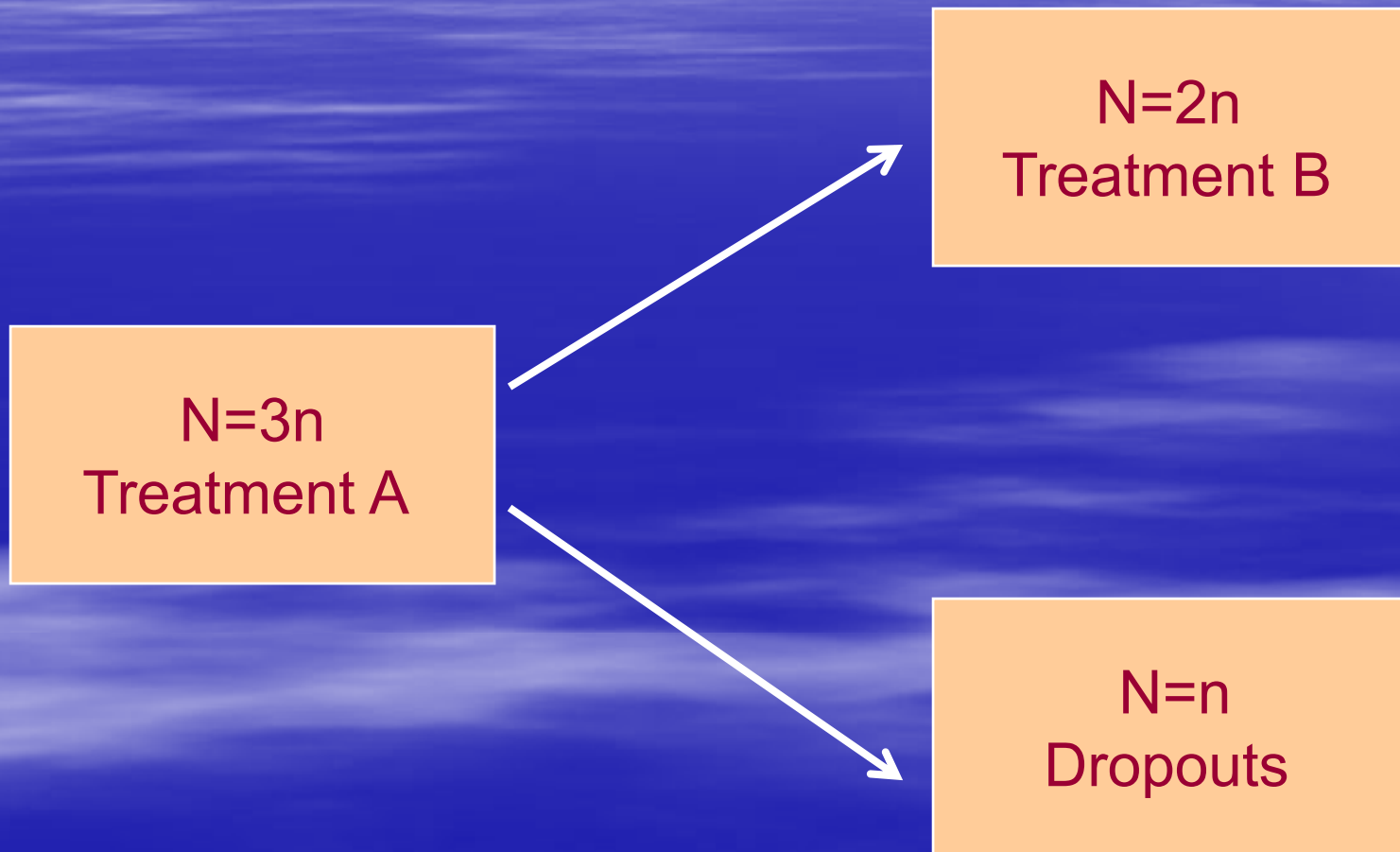
- Easy to accept: parallel treatment comparisons or historical comparisons.
- Difficult to understand: all data should be used instead of just the completers.

The confusion is actually caused by the crossover design and the subjects who were treated by both treatments.

Back to This Design



Or Incomplete Due to Dropouts



Understanding of the Analyses

The clinical trial's conclusion: Drug Y decreases Drug X's AUC by $x\%$.

The result: geo. mean ratio of TrT A to TrT B in AUC is $1+x\%$.

- This result should only be concluded by using all data from TrT A.
- Only use completers or subjects crossed over to the other treatment:
 - The estimate of the geo. mean of TrT A would be poor in the two examples.
 - Different geo. means of TrT A would be used when comparing to TrT B, TrT C, TrT D.

Understanding of the Analyses

The reasons of confusion:

- The subjects were treated by both treatments or have values at both time points in the study.
- It gives people false understanding that the difference should only be within these subjects and forgot the basic meaning of the comparison in general sense.
- It is difficult to explain to medical doctors and sometimes, PK scientists.

An Example of Misunderstand Statistics in DDI, BA, BE Studies

One of the proposed regulatory guidance includes (outside USA):

Present PK data by:

Geo. Mean Ratio and the range of the ratios

- No “statistics” is needed.
- No parallel studies or historical comparisons.
- No use of incomplete subjects

- The reason: the MDs can understand the limit of the impact!
- The danger: the MDs use the sample range as the real range in practice.
- Don't believe that this is proposed by a statistician

Summary

Statistical analysis in CT should be:

- Always find the best ways to efficiently use data
- Methods driven by statistical theory
- Analysis methods should be decided by statisticians!

Acknowledgement

Jingdong Xie
Forest Lab ??

My former colleague who validated
the formula.

Thank you very much for your
attention!