Bioequivalence for Highly Variable Drugs – When Replicate/SABE Designs Misbehave or Can’t Be Used: Issues and Solutions

Charles E. DiLiberti, President
Montclair Bioequivalence Services, LLC
charlie@montclairbe.com
Disclaimer

• The approaches described in this presentation are proposals only, and do not reflect established policies of FDA or other regulatory bodies.

• All information provided here, in any form, represents the opinions of Montclair Bioequivalence Services, LLC and may not, in any way, be construed as legal, regulatory, investment, medical, or pharmacological advice.

• Montclair Bioequivalence Services, LLC and Charles E. DiLiberti make no representation or warranty, express or implied, as to the accuracy or suitability for any purpose of any of the information presented here.

• Montclair Bioequivalence Services, LLC and Charles E. DiLiberti may not be held liable for any losses incurred as a result of any use of any of the information presented here.
Highly Variable Drugs (HVDs)

• Traditional definition of highly variable drugs:
  - Within-subject coefficient of variation (intra-subject CV) ≥ 30%
  - Historically, bioequivalence (BE) studies for HVDs required large numbers of subjects

• Scaled Average Bioequivalence (SABE) has been a tremendous success
  - In effect, SABE scales BE criteria to variability of reference (R) drug
  - Enabled generics where there were none
  - Dramatically reduced unnecessary human testing
  - Dramatically reduced sponsors’ costs
Complications with SABE studies
- SABE requires replication – each subject must receive at least 2 doses of R and 1 or 2 doses of test (T) drug
- Incomplete data can pose problems

Study designs that cannot benefit from SABE:
- Scaling is based on within-subject variability of R
- Parallel designs:
  - Each subject receives only one dose of one product
  - No measurement of within-subject variability of R
  - No benefit from scaling
- Parallel studies are still often large, sometimes involving hundreds of subjects
Incomplete data in SABE studies

Partial replicate design

<table>
<thead>
<tr>
<th>Per 1</th>
<th>Per 2</th>
<th>Per 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>×</td>
<td>R</td>
<td>T</td>
</tr>
<tr>
<td>R</td>
<td>×</td>
<td>R</td>
</tr>
<tr>
<td>T</td>
<td>R</td>
<td>R</td>
</tr>
</tbody>
</table>

Patterns of missing data can affect:
- Whether subject can be used
- How remaining data can be used
  - Within R variability
  - T/R ratio

Fully replicate design

<table>
<thead>
<tr>
<th>Per 1</th>
<th>Per 2</th>
<th>Per 3</th>
<th>Per 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>×</td>
<td>T</td>
<td>×</td>
<td>T</td>
</tr>
<tr>
<td>T</td>
<td>R</td>
<td>×</td>
<td>R</td>
</tr>
</tbody>
</table>
Handling incomplete data

• Model (progesterone capsules) BE guidance not clear on handling incomplete data

• Inconsistent practices among statisticians:
  ❖ Some require complete data to include subject
  ❖ Others include subjects with some incomplete data
  ❖ Differences in patterns of missing data that are allowable, and how calculations are done

• Depending on patterns of missing data and practice followed:
  ❖ Could dramatically reduce power
  ❖ Might be questionable statistically
Consequences and solutions

• Most often affects $AUC_\infty$ although sometimes affects $C_{\text{max}}$ and $AUC_t$ due to adverse events, bad weather, etc.

• Not a huge industry issue yet:
  ✓ Creates uncertainty in powering studies
  ✓ Will eventually become important to a sponsor whose borderline study is affected – best to resolve now

• Possible solution: use bootstrapping to calculate upper bound of required SABE metric
  ✓ Should maximize the value of the partial data available/minimize data “waste”
  ✓ Use SAS PROC MIXED at each cycle to address missing data
  ✓ Could be applied routinely for full or incomplete data sets – avoids two-step approach with decision-making
  ✓ Requires further research to confirm
  ✓ Not current FDA policy but if adopted, could revise progesterone guidance accordingly
Parallel design (PK) BE studies

• Used for:
  ❖ Long half-life drugs, including depot injectables
  ❖ To shorten study timelines (e.g., first-to-file)
• Parallel design assesses T and R in different subjects → **between**-subject CV drives study size
  ❖ Between-subject variability can be >> within subject variability
• Can still pose serious practical issues (enormous study sizes) even if strict definition of HVD not met
Improving S/N ratio

• Signal = effects of T/R formulation differences on PK

• Noise = variation in PK response due to factors unrelated to formulation, e.g., between-subject differences in:
  - Age, body mass, metabolism
  - Other unidentified human factors causing differing PK response to the same drug
  - Generally controlled in crossover studies but not in parallel studies
Reducing noise in BE measurements

- Adjust for human factors unrelated to formulation, e.g., elimination rate constant ($k_{el}$)
- Could adjust $C_{max}$ and AUC by $k_{el}$
  - Theoretically, $AUC \propto 1/k_{el}$
  - $AUC*k_e = (\text{fraction absorbed})*(\text{dose})/(\text{volume of distribution})$
  - Eliminates effect of clearance
  - $C_{max}$ usually strongly correlated with AUC
- Because PK analyses are done on ln-transformed data:
  - $\ln(AUC * k_{el}) = \ln(AUC) + \ln(k_{el})$
  - Perform analysis of covariance (ANCOVA) on $\ln(AUC)$ and $\ln(C_{max})$ with $\ln(k_{el})$ as covariate
  - ANCOVA useful to remove influence of nuisance variables (e.g., $k_{el}$) correlated with desired response variables [e.g., $\ln(C_{max})$, $\ln(AUC)$]
Solid oral dosage form example
$\ln(C_{\text{max}})$ vs. $\ln(k_{el})$
ln(AUC) vs. ln($k_{el}$)

InAUCt vs Inkel

Treatment

○ X
△ Y
**ln(k_{el}) properties in this example**

- Regression between desired PK variables \([\ln(C_{max}), \ln(AUC)]\) and \(\ln(k_{el})\):
  - As expected, strong negative correlation
  - Slopes of regression lines similar for T and R products (parallelism)

- ANOVA on \(\ln(k_{el})\):
  - Treatment effect is not significant \((p = 0.64)\)
  - No evidence that \(\ln(k_{el})\) is different for T and R products
  - \(\ln(k_{el})\) is unrelated to formulation \((in this case)\)
**Effect of ln($k_{el}$) as covariate for ln($C_{max}$)**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>ANOVA</th>
<th>ANCOVA w/ln($k_{el}$) as covariate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual CV</td>
<td>41.8%</td>
<td>24.5%</td>
</tr>
<tr>
<td>N required*</td>
<td>142</td>
<td>~52</td>
</tr>
<tr>
<td>90% CI</td>
<td>91.0 – 132.2%</td>
<td>102.3 – 128.1%</td>
</tr>
<tr>
<td>p (treatment effect)</td>
<td>0.41 (NS)</td>
<td>0.05</td>
</tr>
<tr>
<td>p (ln($k_{el}$))</td>
<td></td>
<td>&lt; 0.000001</td>
</tr>
</tbody>
</table>

*Total (both treatment groups), assuming equal allocation, T/R ratio = 0.95, power = 80%
# Effect of \( \ln(k_{el}) \) as covariate for \( \ln(AUC) \)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>ANOVA</th>
<th>ANCOVA w/( \ln(k_{el}) ) as covariate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual CV</td>
<td>68.0%</td>
<td>22.2%</td>
</tr>
<tr>
<td>N required*</td>
<td>328</td>
<td>~44</td>
</tr>
<tr>
<td>90% CI</td>
<td>73.1 – 129.6%</td>
<td>94.8 – 116.3%</td>
</tr>
<tr>
<td>( p ) (treatment effect)</td>
<td>0.87 (NS)</td>
<td>0.43 (NS)</td>
</tr>
<tr>
<td>( p ) (( \ln(k_{el}) ))</td>
<td></td>
<td>&lt; 0.000001</td>
</tr>
</tbody>
</table>

*Total (both treatment groups), assuming equal allocation, T/R ratio = 0.95, power = 80%
Long-acting injectable example
$\ln(C_{\text{max}})$ vs. $\ln(k_{\text{el}})$
ln(AUC) vs. ln($k_{el}$)

lnAUCt vs lnkel

Treatment
- Q
- Z
What’s going on here?

- Expect $\ln(C_{\text{max}})$, $\ln(\text{AUC})$ to be negative correlated with $\ln(k_{\text{el}})$, but are positively correlated.
- Flip-flop kinetics:
  - True absorption rate $k_{\text{abs}} \ll \text{true } k_{\text{el}}$
  - $k_{\text{abs}}$ masquerades as apparent $k_{\text{el}}$
- Makes sense that true absorption rate (masquerading as apparent $k_{\text{el}}$) is positively correlated with $C_{\text{max}}$ and AUC.
- But absorption rate should be determined by formulation, so adjusting by it would be “cheating,” right?
In this case, \( \ln(k_{el}) \) linked to subject factors, not formulation

- ANOVA on \( \ln(k_{el}) \): \( p \) (treatment) = 0.68 (NS), GMR = 96.3%
- ANOVA (or ANCOVA) on \( \ln(k_{el}) \) with formulation and subject-related factors:

<table>
<thead>
<tr>
<th>Effect tested</th>
<th>( p ) effect tested</th>
<th>( p ) ( k_{el} ) treatment effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race*</td>
<td>0.0006</td>
<td>0.90 (NS)</td>
</tr>
<tr>
<td>Ethnicity*</td>
<td>0.70 (NS)</td>
<td>0.65 (NS)</td>
</tr>
<tr>
<td>Tobacco*</td>
<td>0.45 (NS)</td>
<td>0.71 (NS)</td>
</tr>
<tr>
<td>Age**</td>
<td>0.066 (NS)</td>
<td>0.68 (NS)</td>
</tr>
<tr>
<td>Weight**</td>
<td>&lt; 0.000001</td>
<td>0.55 (NS)</td>
</tr>
<tr>
<td>Height**</td>
<td>0.43 (NS)</td>
<td>0.64 (NS)</td>
</tr>
<tr>
<td>BMI**</td>
<td>&lt; 0.000001</td>
<td>0.71 (NS)</td>
</tr>
<tr>
<td>BSA**</td>
<td>0.00001</td>
<td>0.53 (NS)</td>
</tr>
</tbody>
</table>

*Categorical variable \( \rightarrow \) ANOVA  
**Continuous variable \( \rightarrow \) ANCOVA
Really, what is going on here?

• Consider the rate-limiting step controlling absorption of drug from injection site:
  ❖ Appears to be controlled by between-subject factors related to injection site anatomy/physiology:
    • Depth of injection/thickness of fat layer
    • Vascular perfusion of injection site
    • Muscle density
  ❖ Probable contribution from within-subject factors related to variability in injection process:
    • Specific site within muscle
    • Tissue damage
  ❖ Clearly independent of formulation (in this case)
  ❖ Formulations have the potential to release drug faster than it can leach out of injection site into systemic circulation
**Effect of \( \ln(k_{el}) \) as covariate for \( \ln(C_{max}) \)**

<table>
<thead>
<tr>
<th>Parameter:</th>
<th>ANOVA</th>
<th>ANCOVA w/( \ln(k_{el}) ) as covariate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual CV</td>
<td>67.3%</td>
<td>34.0%</td>
</tr>
<tr>
<td>N required*</td>
<td>324</td>
<td>~96</td>
</tr>
<tr>
<td>90% CI</td>
<td>99.2 – 129.4%</td>
<td>103.3 – 119.3%</td>
</tr>
<tr>
<td>( p (\text{treatment effect}) )</td>
<td>0.12 (NS)</td>
<td>0.017</td>
</tr>
<tr>
<td>( p (\ln(k_{el})) )</td>
<td></td>
<td>&lt; 0.000001</td>
</tr>
</tbody>
</table>

*Total (both treatment groups), assuming equal allocation, T/R ratio = 0.95, power = 80%*
**Effect of ln($k_{el}$) as covariate for ln(AUC)**

<table>
<thead>
<tr>
<th>Parameter:</th>
<th>ANOVA</th>
<th>ANCOVA w/ln($k_{el}$) as covariate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual CV</td>
<td>26.2%</td>
<td>24.7%</td>
</tr>
<tr>
<td>N required*</td>
<td>60</td>
<td>~54</td>
</tr>
<tr>
<td>90% CI</td>
<td>96.7 – 108.2%</td>
<td>96.7 – 107.5%</td>
</tr>
<tr>
<td>p (treatment effect)</td>
<td>0.50 (NS)</td>
<td>0.55 (NS)</td>
</tr>
<tr>
<td>p (ln($k_{el}$))</td>
<td></td>
<td>&lt; 0.000001</td>
</tr>
</tbody>
</table>

*Total (both treatment groups), assuming equal allocation, T/R ratio = 0.95, power = 80%
Testing multiple covariates

Significance (p) of multiple covariates

<table>
<thead>
<tr>
<th>Covariate</th>
<th>$\ln(C_{\text{max}})$</th>
<th>$\ln(\text{AUC})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.56 (NS)</td>
<td>0.0008</td>
</tr>
<tr>
<td>Weight</td>
<td>0.048</td>
<td>0.79 (NS)</td>
</tr>
<tr>
<td>BMI</td>
<td>0.67 (NS)</td>
<td>0.38 (NS)</td>
</tr>
<tr>
<td>Height</td>
<td>0.33 (NS)</td>
<td>0.36 (NS)</td>
</tr>
<tr>
<td>BSA</td>
<td>0.14 (NS)</td>
<td>0.44 (NS)</td>
</tr>
<tr>
<td>$\ln(k_{el})$</td>
<td>&lt;0.000001</td>
<td>0.0007</td>
</tr>
</tbody>
</table>
### Effect of ln(k_{el}) and weight as covariates for ln(C_{max})

<table>
<thead>
<tr>
<th>Parameter:</th>
<th>ANOVA</th>
<th>ANCOVA w/ln(k_{el}) as covariate</th>
<th>ANCOVA w/ln(k_{el}) and weight as covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual CV</td>
<td>67.3%</td>
<td>34.0%</td>
<td>32.0%</td>
</tr>
<tr>
<td>N required*</td>
<td>324</td>
<td>~96</td>
<td>~86</td>
</tr>
<tr>
<td>90% CI</td>
<td>99.2 – 129.4%</td>
<td>103.3 – 119.3%</td>
<td>104.6 – 119.8%</td>
</tr>
<tr>
<td>p (treatment effect)</td>
<td>0.12 (NS)</td>
<td>0.017</td>
<td>0.0064</td>
</tr>
<tr>
<td>p (ln(k_{el}))</td>
<td>&lt; 0.000001</td>
<td>&lt; 0.000001</td>
<td>&lt; 0.000001</td>
</tr>
<tr>
<td>P (weight)</td>
<td>&lt; 0.000001</td>
<td>&lt; 0.000001</td>
<td>&lt; 0.000001</td>
</tr>
</tbody>
</table>

*Total (both treatment groups), assuming equal allocation, T/R ratio = 0.95, power = 80%
## Effect of ln($k_{el}$) and age as covariates for ln(AUC)

<table>
<thead>
<tr>
<th>Parameter:</th>
<th>ANOVA</th>
<th>ANCOVA w/ln($k_{el}$) as covariate</th>
<th>ANCOVA w/ln($k_{el}$) and age as covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual CV</td>
<td>26.2%</td>
<td>24.7%</td>
<td>24.3%</td>
</tr>
<tr>
<td>N required*</td>
<td>60</td>
<td>~54</td>
<td>~52</td>
</tr>
<tr>
<td>90% CI</td>
<td>96.7 – 108.2%</td>
<td>96.7 – 107.5%</td>
<td>96.8 – 107.4%</td>
</tr>
<tr>
<td>p (treatment effect)</td>
<td>0.50 (NS)</td>
<td>0.55 (NS)</td>
<td>0.53 (NS)</td>
</tr>
<tr>
<td>p (ln($k_{el}$))</td>
<td>&lt; 0.000001</td>
<td>0.000001</td>
<td></td>
</tr>
<tr>
<td>p (age)</td>
<td></td>
<td></td>
<td>0.0036</td>
</tr>
</tbody>
</table>

*Total (both treatment groups), assuming equal allocation, T/R ratio = 0.95, power = 80%
ANCOVA with $\ln(k_{el})$ as covariate

• Potential to dramatically reduce residual variability and sample size for parallel BE studies on some products
  - May not be feasible for all products – case by case approach advisable

• Other subject-related covariates (age, weight, BMI, BSA, etc.) could be included

• Before using ANCOVA, should verify:
  - $\ln(k_{el})$ does not have a significant formulation effect
  - $\ln(k_{el})$ is significant in ANCOVA model
  - Slopes of PK variables [$\ln(C_{max})$ or $\ln(AUC)$] vs. covariates [$\ln(k_{el})$, weight, age, etc.] are comparable for test and reference products

• Requires further research
• Not current FDA policy
Conventional SABE

- Based on **switching** a patient from one formulation to another formulation
- High **within**-subject variability provides evidence that desired safety and efficacy properties are insensitive towards dose
- Justifies “relaxing” (scaling) BE criteria based on **within**-subject variability of reference product
A parting thought: SABE for parallel designs?

- Some drugs (e.g., LA injectables) are dosed once or are expected to work on the first dose in a treatment-naïve patient
  - Switching and therefore, within-subject variability are irrelevant
  - High between-subject variability in PK provides evidence that desired safety and efficacy properties are insensitive towards dose
  - Justifies “relaxing” (scaling) BE criteria based on between-subject variability of reference product
- Adapting the current SABE procedure to scale for between-subject variability of reference product
  - Appears to be feasible
  - Requires more research
  - Not current FDA policy
Many thanks to...

- GPhA
- Anonymous sponsors
- AAPS Generic Product Focus Group Steering Committee
- Industry statisticians