Add-On and Sequential Designs

Helmut Schütz
BEBAC

Biostatistics: Basic concepts & applicable principles for various designs in bioequivalence studies and data analysis | Mumbai, 29 – 30 January 2011
Add-On Designs

- Were extensively discussed at the Bio-International ’92 Conference, Bad Homburg, Germany
  - Sometimes (rarely) accepted in the US and EU
  - Only if planned, same protocol, batches, clinical and analytical method.
    If first part not BE, recalculation of sample size, initiation of second part.
  - Pooling of both parts under certain conditions.
Add-On Designs

- Canada
  Second part in at least 12 subjects. Pooling is only allowed if two consistency tests not significant ($p>0.05$):
    - Equality of residual mean squares ($F$-test) of the two parts. Smaller MSE must be used as the denominator.
  Example:
  - $0.01321$ ($1^{st}$ part: $n=55$, df 53)
  - $0.01718$ ($2^{nd}$ part: $n=14$, df 12)
### Add-On Designs

#### Sum of Squares

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>DF</th>
<th>SSE</th>
<th>MSE</th>
<th>F_stat</th>
<th>P_value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence</td>
<td>1</td>
<td>0.0271658</td>
<td>0.0271658</td>
<td>0.173639</td>
<td>0.6786</td>
</tr>
<tr>
<td>Sequence*Subject</td>
<td>53</td>
<td>8.29185</td>
<td>0.15645</td>
<td>11.844</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Treatment</td>
<td>1</td>
<td>0.0349142</td>
<td>0.0349142</td>
<td>2.03203</td>
<td>0.1795</td>
</tr>
<tr>
<td>Period</td>
<td>1</td>
<td>0.0271536</td>
<td>0.0271536</td>
<td>2.05565</td>
<td>0.1575</td>
</tr>
<tr>
<td>Error</td>
<td>53</td>
<td>0.700088</td>
<td>0.0132092</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Sum of Squares

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>DF</th>
<th>SSE</th>
<th>MSE</th>
<th>F_stat</th>
<th>P_value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence</td>
<td>1</td>
<td>0.0489527</td>
<td>0.0489527</td>
<td>0.419204</td>
<td>0.6786</td>
</tr>
<tr>
<td>Sequence*Subject</td>
<td>53</td>
<td>8.29185</td>
<td>0.15645</td>
<td>11.844</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Treatment</td>
<td>1</td>
<td>0.211196</td>
<td>0.211196</td>
<td>15.9885</td>
<td>0.0002</td>
</tr>
<tr>
<td>Period</td>
<td>1</td>
<td>0.0271536</td>
<td>0.0271536</td>
<td>2.05565</td>
<td>0.1575</td>
</tr>
<tr>
<td>Error</td>
<td>53</td>
<td>0.700088</td>
<td>0.0132092</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
\hat{F} = \frac{\text{MSE}_{\text{large}}}{\text{MSE}_{\text{small}}} = \frac{0.0171819}{0.0132092} = 1.30075
\]

\[
F_{1=0.05,12,53} = 1.940
\]

\[
p(\hat{F}) = 0.24595 > 0.05 \quad \checkmark
\]
Add-On Designs

- Canada cont’d
  Second part in at least 12 subjects. Pooling is only allowed if two consistency tests not significant ($p>0.05$):
  - Since first test not significant ($p > 0.246$), pool studies.
    Now test for study by formulation interaction.
    ANOVA model:
    Fixed: Study + Treatment + Treatment × Study
    Random: Subject(Study) + Period(Study)
Add-On and Sequential Designs

Add-On Designs

Tests of Model Effects

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Numer_DF</th>
<th>Denom_DF</th>
<th>F_stat</th>
<th>P_value</th>
</tr>
</thead>
<tbody>
<tr>
<td>int</td>
<td>1</td>
<td>56.5</td>
<td>2144.16</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Study</td>
<td>1</td>
<td>56.5</td>
<td>0.0007</td>
<td>0.9784</td>
</tr>
<tr>
<td>Treatment</td>
<td>1</td>
<td>64.6</td>
<td>9.9949</td>
<td>0.0024</td>
</tr>
<tr>
<td>Treatment*Study</td>
<td>1</td>
<td>64.6</td>
<td>0.1156</td>
<td>0.7349</td>
</tr>
</tbody>
</table>

Bioequivalence Statistics

User-Specified Confidence Level for CI's = 95.0000
Percent of Reference to Detect for 2-1 Tests = 20.0%
A.H.Lower = 0.800  A.H.Upper = 1.250

Formulation variable: Treatment
Reference: R  LSMean= 6.088010  SE= 0.132921 GeoLSM= 440.543718

Test: T  LSMean= 6.167145  SE= 0.132921 GeoLSM= 476.822902

Difference = 0.0791, Diff_SE= 0.0250, df= 64.6
Ratio(%Ref) = 108.2351

CI 90% = ( 103.8061, 112.8531)
CI 95% = ( 102.9556, 113.7853)

Average bioequivalence shown for confidence=95.00 and percent=20.0.
Add-On and Sequential Designs

Add-On Designs

- Canada cont’d
  - Formulation by study interaction not significant ($p < 0.735$), pooled analysis acceptable.
  - No $\alpha$-adjustment mentioned in 1992 guideline, but recommended in 2010 draft (Bonferroni: 95% CI).
  - 2010 draft also allows for a group sequential design.
Add-On Designs

- Japan
  - Sample size of first part $\geq 20$ or
    sample size of both parts $\geq 30$
  - Sample size of second part $\geq 50\%$ of first part
  - Study must be added as a source of variation in the pooled analysis
  - PE of the pooled study within 90\% – 111\%
  - Similar dissolution between test and reference
  - No $\alpha$-adjustment (90\% CI).
    I would strongly recommend to use the 95\% CI instead of the conventional 90\% CI.
Add-On Designs

- It was shown in simulations that the patients’ risk in naïve pooling may be inflated up to ~30%!
- Group sequential designs are a well known design in Phase III studies.
- These designs test for a significant difference, i.e., the null hypothesis is formulated differently to the one applied in BE.
  - Led to work on sequential designs in BE; LA Gould (1995) and D Potvin et al. (2008).
Sequential Designs

... have a long and accepted tradition in later phases of clinical research (mainly Phase III).

  - First proposal by LA Gould (1995) in the area of BE did not get regulatory acceptance in Europe, but
  - stated in the current Canadian draft guidance (January 2010).
  - Two-Stage Design acceptable in the EU (GL 2010, Section 4.1.8) and the US (personal communication with Barbara Davit, Ljubljana, May 2010).
Sequential Designs

- Penalty for the interim analysis (94.12% vs. 90% CI)
  - Moderate increase in sample sizes
    - Example: T/R 95%, power 80%
    - ~10–20% increase (sims by Gould 1995, Potvin 2008)
  - Comparison to a fixed sample design is based on a delusion – assuming the CV to be ‘known’!
  - On the long run (many studies) sequential designs will need less subjects.

<table>
<thead>
<tr>
<th>CV%</th>
<th>90% CI</th>
<th>94.12% CI</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>8</td>
<td>8</td>
<td>1.000</td>
</tr>
<tr>
<td>15</td>
<td>12</td>
<td>14</td>
<td>1.167</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>24</td>
<td>1.200</td>
</tr>
<tr>
<td>25</td>
<td>28</td>
<td>34</td>
<td>1.214</td>
</tr>
<tr>
<td>30</td>
<td>40</td>
<td>48</td>
<td>1.200</td>
</tr>
</tbody>
</table>
Sequential Designs

  - Overall $\alpha$-level preserved at $\leq 0.05$. Individual CI level $\sim 93\%$–$94\%$ (calculation rather complicated; dependent on PE and sample size ratio of Stages).
  - Multiple Stages (i.e., $>2$) possible.

- Canada’s draft (2010)
  - Maximum total sample size (Stages 1+2) and size of Stage 1 fixed in protocol.
  - Adjusted $\alpha$-levels stated in the protocol.
Two-Stage Design

- EMA GL on BE (2010)
  - Section 4.1.8
    - Initial group of subjects treated and data analysed.
    - If BE not been demonstrated an additional group can be recruited and the results from both groups combined in a final analysis.
    - Appropriate steps to preserve the overall type I error (patients’ risk).
    - Stopping criteria should be defined *a priori*.
    - First stage data should be treated as an interim analysis.

‘Internal Pilot Study Design’
Two-Stage Design

- EMA GL on BE (2010)
  - Section 4.1.8 (cont’d)
    - Both analyses conducted at adjusted significance levels (with the confidence intervals accordingly using an adjusted coverage probability which will be higher than 90%). [...] 94.12% confidence intervals for both the analysis of stage 1 and the combined data from stage 1 and stage 2 would be acceptable, but there are many acceptable alternatives and the choice of how much alpha to spend at the interim analysis is at the company’s discretion.
Two-Stage Design

- EMA GL on BE (2010)
  - Section 4.1.8 (cont’d)
    - Plan to use a two-stage approach must be pre-specified in the protocol along with the adjusted significance levels to be used for each of the analyses.
    - When analysing the combined data from the two stages, a term for stage should be included in the ANOVA model.
Two-Stage Design

- Method by Potvin et al. (2008) promising
  - Supported by ‘The Product Quality Research Institute’ (members: FDA-CDER, Health Canada, USP, AAPS, PhRMA,…)
    - Acceptable by US-FDA
    - Acceptable as a Two-Stage Design in the EU
    - Three of BEBAC’s protocols already approved by German BfArM
Method ‘C’

Evaluate power at Stage 1 using $\alpha$-level of 0.050

If power $\geq 80\%$, evaluate BE at Stage 1 ($\alpha = 0.050$) and stop

IF BE met, stop

Pass or fail

Pass

If power $< 80\%$, evaluate BE at Stage 1 ($\alpha = 0.0294$)

If BE not met, calculate sample size based on Stage 1 and $\alpha = 0.0294$, continue to Stage 2

Evaluate BE at Stage 2 using data from both Stages ($\alpha = 0.0294$) and stop

Pass or fail
**Potvin et al. (2008)**

- **Technical Aspects**
  - Only *one* Interim Analysis (after Stage 1)
  - If possible, use software (too wide step sizes in Diletti’s tables)
  - Should be called ‘Interim Power Analysis’; *not* ‘Bioequivalence Assessment’ in the protocol
  - No *a-posteriori* Power – only a validated method in the decision tree
  - No adjustment for the PE observed in Stage 1
  - No stop criterion for Stage 2! Must be clearly stated in the protocol (may be unfamiliar to the IEC, because standard in Phase III).
Potvin et al. (2008)

- Technical Aspects (cont’d)
  - Adjusted $\alpha$ of 0.0294 (Pocock 1977)
  - If power is <80% in Stage 1 and in the pooled analysis (data from Stages 1 + 2), $\alpha$ 0.0294 is used (i.e., a $1-2\times\alpha = 94.12\%$ CI is calculated)
  - Overall patients’ risk is preserved at $\leq 0.0500$
Potvin et al. (2008)

- Technical Aspects (cont’d)
  - If the study is stopped after Stage 1, the (conventional) statistical model is:
    fixed: sequence + period + treatment
    random: subject(sequence)
  - If the study continues to Stage 2, the model for the combined analysis is:
    fixed: sequence + stage + period(stage) + treatment
    random: subject(sequence × stage)
  - No poolability criterion; combining is always allowed – even for significant differences between Stages.
Potvin et al. (2008)

Model Specification and User Settings
- Dependent variable: Cmax (ng/mL)
- Transform: LN
- Fixed terms: int+Sequence+Treatment+Period
- Random/repeated terms: Sequence*Subject

Final variance parameter estimates:
- Var(Sequence*Subject) = 0.0444152
- Var(Residual) = 0.071194
- Intrasubject CV = 0.271642

Bioequivalence Statistics
- User-Specified Confidence Level for CI's = 94.1200
- Percent of Reference to Detect for 2-1 Tests = 20.0%
- A.H.Lower = 0.800, A.H.Upper = 1.250
- Reference: Reference, LSMean = 1.593384, SE = 0.123689, GeoLSM = 4.920373

Test:
- Test, LSMean = 1.471058, SE = 0.123689, GeoLSM = 4.353839
- Difference = -0.1223, Diff_SE = 0.1958, df = 12.0
- Ratio(%Ref) = 88.4860

Classical
- CI 90% = (62.4145, 125.4478)
- CI User = (58.7888, 133.1845)

Failed to show average bioequivalence for confidence=94.12 and percent=20.0.

14 subjects in Stage 1, conventional BE model

\[ CV_{\text{intra}} = 27.2\% \]

\[ \alpha = 0.0294 \]

(if power <80%)

Failed 90% CI (if power \geq 80%) and 94.12% CI (if power <80%).
Add-On and Sequential Designs

Potvin et al. (2008)

```
require(PowerTOST)
power.TOST(alpha=0.05, logscale=TRUE, theta1=0.8, theta2=1.25, theta0=0.95, CV=0.271642, n=14, design = "2x2", exact = TRUE)

[1] 0.3189318

Power 31.9% – initiate Stage 2
```

```
sampleN.TOST(alpha=0.0294, targetpower=0.8, logscale=TRUE, theta1=0.8, theta2=1.25, theta0=0.95, CV=0.271642, design = "2x2", exact = TRUE, print = TRUE)

+++++++++++ Equivalence test - TOST ++++++++++++ Sample size estimation

Study design: 2x2 crossover log-transformed data (multiplicative model)
alpha = 0.0294, target power = 0.8
BE margins = 0.8 ... 1.25
Null (true) ratio = 0.95, CV = 0.271642

Sample size
n  power
40  0.817146

Total sample size 40: 26 in Stage 2 (28 recruited)
```

Expected ratio 95% – not 88.5% observed in stage 1! CV$_{intra}$ 27.2%, 14 subjects in Stage 1

Calculate total sample size: expected ratio 95%, CV$_{intra}$ 27.2%, 80% power
Potvin et al. (2008)

Model Specification and User Settings
Dependent variable : Cmax (ng/mL)
  Transform : LN
Fixed terms : int+Sequence+Stage+Period(Stage)+Treatment
Random/repeated terms : Sequence*Stage*Subject

Final variance parameter estimates:
Var(Sequence*Stage*Subject) 0.0430110
Var(Residual) 0.0376772
Intrasubject CV 0.1959489

Bioequivalence Statistics
User-Specified Confidence Level for CI's = 94.1200
Percent of Reference to Detect for 2-1 Tests = 20.0%
A.H.Lower = 0.800  A.H.Upper = 1.250
Formulation variable: Treatment
Reference: Reference  LSMean= 1.520255  SE= 0.047872  GeoLSM= 4.573390

Test:  Test  LSMean= 1.525145  SE= 0.047872  GeoLSM= 4.595809

Difference = 0.0049, Diff_SE= 0.0496, df= 38.0
Ratio(%Ref) = 100.4902

Classical
CI 90% = ( 92.4329, 109.2499)

BE shown with 94.12% CI; overall \( \alpha \leq 0.05! \)

\( \alpha \) 0.0294 in pooled analysis

27 subjects in Stage 2 (41 total), modified model for pooled analysis

Average bioequivalence shown for confidence=94.12 and percent=20.0.
Add-On and Sequential Designs

Helmut Schütz

BEBAC

Consultancy Services for Bioequivalence and Bioavailability Studies

1070 Vienna, Austria

helmut.schuetz@bebac.at
To bear in Remembrance...

The sample size calculation is an excuse for a sample size and not a reason. *Stephen Senn*

Science is a way of trying not to fool yourself. The first principle is that you must not fool yourself, and you are the easiest person to fool. *Richard Feynman*

If you obey all the rules, you will miss all the fun. *Katherine Hepburn*
References

• Collection of links to global documents
  http://bebac.at/Guidelines.htm
• ICH
• EMA-CPMP/CHMP/EWP
  ▪ Points to Consider on Multiplicity Issues in Clinical Trials (2002)
• H Melander
  Problems and Possibilities with the Add-On Subject Design
  In: KK Midha and HH Blume (eds.)
  Bio-International. Bioavailability, Bioequivalence and Pharmacokinetics
  ISBN 388763-019-X
• LA Gould
  Group Sequential Extension of a Standard Bioequivalence Testing Procedure
  DOI: 10.1007/BF02353786
• Potvin D, Diliberti CE, Hauck WW, Parr AF, Schuirmann DJ, and RA Smith
  Sequential design approaches for bioequivalence studies with crossover designs
  DOI: 10.1002/pst.294