Biostatistics
Two-Stage Sequential Designs

Helmut Schütz
BEBAC
Add-on / Two-Stage Designs

- Sometimes properly designed and executed studies fail due to
  - ‘true’ bioinequivalence,
  - poor study conduct (increasing variability),
  - pure chance (producer’s risk hit),
  - false (mainly over-optimistic) assumptions about the CV and/or T/R-ratio.

- The patient’s risk must be preserved
  - Already noticed at Bio-International Conferences (1989, 1992) and guidelines from the 1990s.
Sequential Designs

*Have a long and accepted tradition in clinical research (mainly phase III)*

  - First proposal by Gould (1995) in the field of BE did not get regulatory acceptance in Europe, but
  - new methods stated in recent guidelines.

AL Gould
*Group Sequential Extension of a Standard Bioequivalence Testing Procedure*
DOI: 10.1007/BF02353786
Sequential Designs

- Methods by Potvin et al. (2008) first validated framework in the context of BE
  - Supported by the ‘Product Quality Research Institute’
    (members: FDA/CDER, Health Canada, USP, AAPS, PhRMA…)
  - Three of BEBAC’s protocols accepted by German BfArM, first product approved in 06/2011.

Potvin D, DiLiberti CE, Hauck WW, Parr AF, Schuirmann DJ, and RA Smith
*Sequential design approaches for bioequivalence studies with crossover designs*
Review of Guidelines

- EMA (Jan 2010)
  Acceptable; Potvin et al. Method B preferred (?)

- Canada (May 2012)
  Potvin et al. Method C recommended

- FDA (Jun 2012)
  Potvin et al. Method C/D recommended
  API specific guidances: Loteprednol, (Dexamethasone / Tobramycin)

- Russia (2013)
  Acceptable; Potvin et al. Method B preferred (?)
Potvin et al. (Method B)

Evaluate BE at stage 1 ($\alpha 0.0294$)

- yes
  - BE met?
    - yes
      - Pass
    - no
      - no
        - yes
          - $\geq 80\%$?
            - yes
              - Estimate sample size based on $CV_{\text{intra}}$; T/R 0.95, $\alpha 0.0294$; continue to stage 2
            - no
              - Evaluate BE at stage 2 using pooled data from both stages ($\alpha 0.0294$)
      - Fail
    - no
      - Fail

- no
  - Evaluate power at stage 1 using $\alpha$-level of 0.0294
Potvin et al. (Method B)

1150 \cdot 10^6 \text{Sim’s (Method B)}

Two-Stage Sequential Designs
Two-Stage Sequential Designs

Potvin et al. (Method B)

1150 \cdot 10^6 \text{Sim's (Method B)}

% in stage 2

CV

n_1

Bioequivalence Studies in Russia: Pharmacokinetics, Statistics and Analytics
Moscow, 24 April 2014
Potvin et al. (Method B)

Sample size penalty (CV 14–40%, 80% power)

- \( n_{\text{total}} = 1.084n \) planned for 0.0500
- \( n_{\text{total}} = 1.023n \) planned for 0.0294

- \( n_{\text{total}} \): average sample size (two-stage)
- \( n \): sample size (fixed)
Potvin et al. (Method B)

**Technical Aspects**

- Only *one* Interim Analysis (after stage 1).
- Use software (wide step sizes in Diletti’s tables); preferably the exact method (avoid approximations).
- Should be termed ‘Interim Power Analysis’ *not* ‘Bioequivalence Assessment’ in the protocol.
- No *a posteriori* Power – only a validated method in the decision tree.
- No adjustment for T/R observed in stage 1 (not fully adaptive).
Potvin et al. (Method B)

Technical Aspects (cont’d)

- No futility rule preventing to go into stage 2 with a very high sample size!
  Must be clearly stated in the protocol (unfamiliar to the IEC because common in Phase III).

- Pocock’s $\alpha 0.0294$ is used in stage 1 and in the pooled analysis (data from stages 1 + 2), 
  i.e., the $1 - 2 \times \alpha = 94.12\%$ CI is calculated.

- Overall patient’s risk preserved at $\leq 0.05$. 
Potvin et al. (Method B)

- Technical Aspects (cont’d) + EMA modification

- If the study is stopped after stage 1, the statistical model is:
  
  \[
  \text{fixed: sequence + period + treatment} \\
  + \text{subject(sequence)}
  \]

- If the study continues to stage 2, the model for the combined analysis is:
  
  \[
  \text{fixed: stage + sequence + sequence(stage)} \\
  + \text{subject(sequence × stage) + period(stage)} \\
  + \text{treatment}
  \]

- No poolability criterion! Combining is \textit{always allowed} – even if a significant difference between stages is observed. No need to test this effect.
Potvin et al. (Method B)

- Technical Aspects (cont’d) + EMA modification
  - Incomprehensible why this modification was introduced by EMA’s Biostatistical Working Party
    - Simulations performed or “gut feeling”?
      - Modification shown to be irrelevant.
    - Furthermore no difference whether subjects were treated as a fixed or random term (unless T/R > 1.20).

Karalis V and P Macheras
On the Statistical Model of the Two-Stage Designs in Bioequivalence Assessment
Technical Aspects (cont’d)

- Potvin et al. used a simple approximative power estimation based on the shifted central $t$-distribution.
- If possible use the exact method (Owen; $R$ package `PowerTOST` method = 'exact') or at least one based on the noncentral $t$-distribution (`PowerTOST` method = 'noncentral').
- Power obtained in stage 1 (example 2 from Potvin):

<table>
<thead>
<tr>
<th>method</th>
<th>% power</th>
</tr>
</thead>
<tbody>
<tr>
<td>approximative (shifted central $t$)</td>
<td>50.49</td>
</tr>
<tr>
<td>approximative (noncentral $t$)</td>
<td>52.16</td>
</tr>
<tr>
<td>exact (Owen’s $Q$)</td>
<td>52.51</td>
</tr>
</tbody>
</table>
Two-Stage Sequential Designs

Example (Potvin Method B)

Model Specification and User Settings
- Dependent variable: Response
- Transform: LN
- Fixed terms: int + Sequence + Period + Treatment
- Random/repeated terms: Sequence * Subject

Final variance parameter estimates:
- Var(Sequence * Subject) = 0.408682
- Var(Residual) = 0.0326336
- Intrasubject CV = 0.182132

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 LS Mean = 0.954668, SE = 0.191772, GeoLSM = 2.597808
- Test: Test LS Mean = 1.038626, SE = 0.191772, GeoLSM = 2.825331

- Difference = 0.0840, Diff_SE = 0.0737, df = 10.0
- Ratio(%Ref) = 108.7583

Classical
- CI User = (92.9330, 127.2838)

Failed with 94.12% Confidence Interval

- CV_intra = 18.2%
- \(\alpha = 0.0294\)
Example (Potvin Method B)

library(PowerTOST)
power.TOST(alpha=0.0294, theta0=0.95, 
          CV=0.182132, n=12, design='2x2', 
          method='exact')

[1] 0.5251476

Power 52.5% – initiate stage 2

sampleN.TOST(alpha=0.0294, targetpower=0.80,  
              theta0=0.95, CV=0.182132, design='2x2',  
              method='exact')

+++++++++++ 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 = 0.182132

Sample size
n power
20 0.829160

Total sample size 20: include another 8 in stage 2

α 0.0294, T/R 95% – not 108.76% 
observed in stage 1!

CV_{intra} 18.2%, 12 subjects in stage 1

Estimate total sample size:  
α 0.0294, T/R 95%, CV_{intra} 18.2%,  
80% power

Simulations (n, 12, CV 18.2%)  
• α_{emp} 0.042635  
• power 85.3%
Example (Potvin Method B / EMA)

Model Specification and User Settings

- Dependent variable: Cmax (ng/mL)
- Transform: LN
- Fixed terms: int + Stage + Sequence + Sequence*Stage + Sequence*Stage*Subject + Period(Stage) + Treatment

Final variance parameter estimates:
- Var(Sequence*Stage*Subject) = 0.549653
- Var(Residual) = 0.0458956
- Intrasubject CV = 0.216714

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  LS Mean = 1.133431  SE = 0.171385  GeoLSM = 3.106297

Test:  Test  LS Mean = 1.147870  SE = 0.171385  GeoLSM = 3.151473

Difference = 0.0144, Diff_SE = 0.0677, df = 17.0
Ratio(%Ref) = 101.4544

Classical CI  90% = (90.1729, 114.1472)

CI User = (88.4422, 116.3810)

Average bioequivalence shown for confidence=94.12 and percent=20.0.

α ≤ 0.05

BE shown with 94.12% CI; α ≤ 0.05

8 subjects in stage 2 (20 total), modified model in pooled analysis

Q&A Rev. 7 (March 2013)
Potvin et al. (Method C)

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

- yes
  - yes: Evaluate BE at stage 1 ($\alpha$ 0.050)
  - no: Evaluate BE at stage 1 ($\alpha$ 0.0294)

- no
  - yes: BE met?
    - yes: Estimate sample size based on $CV_{intra}^2$
      T/R 0.95, $\alpha$ 0.0294; continue to stage 2
    - no: Evaluate BE at stage 2 using pooled data from both stages ($\alpha$ 0.0294)
  - no: Pass or fail

Pass or fail

Pass

Pass or fail
Potvin et al. (Method C)

1150 - 10^6 Sim's (Method C)

α

CV

n

6% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

0.0324 0.0338 0.0352 0.0366 0.0380 0.0394 0.0408 0.0422 0.0436 0.0450 0.0464 0.0478 0.0492 0.0500 0.0506 0.0520

α

2 4 6 8 10 12

60 48 36 24 12

1150 - 10^6 Sim's (Method C)
Potvin et al. (Method B vs. C)

**Pros & cons**

- Method C (if power $\geq 80\%$) is a conventional BE study; no penalty in terms of $\alpha$ needs to be applied.
- Method C proceeds to stage 2 less often and has smaller average total sample sizes than Method B for cases where the initial sample size is reasonable for the CV.
- If the size of stage 1 is low for the actual CV both methods proceed to stage 2 almost all the time; total sample sizes are similar.
- Method B slightly more conservative than C.
Potvin et al. (Method B vs. C)

**Recommendations**

- Method C/D preferred due to slightly higher power than method B (FDA, HPFB). Method B for EMA & Russia (?)
- Plan the study *as if* the CV is known
  - If assumptions turn out to be true = no penalty
  - If lower power (CV higher than expected), BE still possible in first stage (penalty; 94.12% CI) or continue to stage 2 as a ‘safety net’.
- Don’t jeopardize! Smaller sample sizes in the first stage than in a fixed design don’t pay off. Total sample sizes are ~10–20% higher.
TSDs: Alternatives

- Methods by Potvin et al. (2008) limited to T/R of 0.95 and 80% power

  Follow-up publications (T/R 0.95...0.90, 80...90% power)

<table>
<thead>
<tr>
<th>reference</th>
<th>method</th>
<th>T/R</th>
<th>target power</th>
<th>CV</th>
<th>$\alpha_{adj.}$</th>
<th>max. $\alpha_{emp.}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potvin et al.</td>
<td>B</td>
<td>0.95</td>
<td>80%</td>
<td>10–100%</td>
<td>0.0294</td>
<td>0.0485</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0.95</td>
<td></td>
<td></td>
<td></td>
<td>0.0510</td>
</tr>
<tr>
<td>Montague et al.</td>
<td>D</td>
<td>0.90</td>
<td></td>
<td></td>
<td>0.0280</td>
<td>0.0518</td>
</tr>
<tr>
<td>Fuglsang</td>
<td>B</td>
<td>0.95</td>
<td>90%</td>
<td>10–80%</td>
<td>0.0284</td>
<td>0.0501</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>0.90</td>
<td></td>
<td></td>
<td>0.0274</td>
<td>0.0503</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td>0.0269</td>
<td>0.0501</td>
</tr>
</tbody>
</table>

Montague TH, Potvin D, DiLiberti CE, Hauck WW, Parr AF, and DJ Schuirmann

*Additional results for ‘Sequential design approaches for bioequivalence studies with crossover designs’*
Pharmaceut Statist 11(1), 8–13 (2011) DOI: 10.1002/pst.483

A Fuglsang

*Sequential Bioequivalence Trial Designs with Increased Power and Controlled Type I Error Rates*
Montague et al. (Method D)

1150 \cdot 10^6 \text{ Sim's (Method D)}

Two-Stage Sequential Designs

Montague et al. (Method D)
TSDs: Alternatives

  - Based on Method C ($\alpha_{adj.} 0.0294$) or D ($\alpha_{adj.} 0.0280$)
  - Sample size re-estimation based on observed T/R-ratio in stage 1 (fully adaptive)
  - Upper sample size limit (UL)
  - Frameworks:
    - $n_1$ 12–96, CV 10–60%, $n_1+n_2 \leq UL 150$
    - $n_1$ 18–96, CV 20–40%, $n_1+n_2 \leq UL 100$

Karalis V and P Macheras
An Insight into the Properties of a Two-Stage Design in Bioequivalence Studies

V Karalis
The role of the upper sample size limit in two-stage bioequivalence designs
Two-Stage Sequential Designs

Karalis & Macheras

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

- yes
  - $\geq 80\%$?
    - no
      - Evaluate BE at stage 1 ($\alpha$ 0.050)
    - yes
      - Evaluate BE at stage 1 ($\alpha$ 0.0294)

Evaluate BE at stage 1 ($\alpha$ 0.050)

- yes
  - $n_1 + n_2 > UL$?
    - no
      - BE met?
        - yes
          - Pass or fail
        - no
          - Fail
    - yes
      - Estimate sample size based on $CV_{intra}$ & $T/R_{stage_1}$, $\alpha$ 0.0294

Evaluate BE at stage 1 ($\alpha$ 0.0294)

- yes
  - BE met?
    - yes
      - Pass
    - no
      - Fail

- no
  - $T/R_{stage_1}$ {0.8, 1.25}
    - yes
      - Pass or fail
    - no
      - Fail

Evaluate BE at stage 2 using pooled data from both stages ($\alpha$ 0.0294)
Karalis & Macheras (n ≤150)

578·10^6 Sim’s (Karalis/Macheras)
Karalis & Macheras (n ≤150)

library(PowerTOST)
power.TOST(alpha=0.05, theta0=1.0876, 
CV=0.182132, n=12, design='2x2',
method='exact')

[1] 0.531698

sampleN.TOST(alpha=0.0294, targetpower=0.80,
theta0=1.0876, CV=0.182132, design='2x2',
method='exact')

+++++++++++ 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 = 1.0876, CV = 0.182132

Sample size
n  power
28  0.813921

α 0.05, observed T/R 108.76%, CV_{intra} 18.2%, 12 subjects in stage 1

Power 53.2% – initiate stage 2

Estimate total sample size:
α 0.0294, T/R 108.76%,
CV_{intra} 18.2%, 80% power

Simulations (n, 12, CV 18.2%, UL 150)
• α_{emp} 0.049681
• power 89.1%

Total sample size 28 (≤150): include another 16 in stage 2
Karalis & Macheras (Expl. a)

- CV assumed as 20%, T/R 95%
  - In a fixed sample design for 80% power sample sizes would be 20 ($\alpha 0.05$) or 24 ($\alpha 0.0294$).
  - The sponsor chooses $n_1$ 24 and UL 100.
  - $10^6$ simulations (Potvin C), $10^5$ (K & M)

<table>
<thead>
<tr>
<th>method</th>
<th>(overall) power</th>
<th>power (stage 1)</th>
<th>% studies to stage 2</th>
<th>$n_{95%}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potvin et al.</td>
<td>90.1</td>
<td>88.1</td>
<td>4.2</td>
<td>24</td>
</tr>
<tr>
<td>Karalis &amp; Macheras</td>
<td>94.8</td>
<td>83.5</td>
<td>11.4</td>
<td>66</td>
</tr>
</tbody>
</table>

- ~Three times as many studies forced to stage 2 with a high probability of large sample sizes.
Karalis & Macheras (Expl. b)

- CV assumed as 40%, T/R 95%
  - Fixed sample design n 66 (α 0.05) or 80 (α 0.0294).
  - The sponsor chooses n₁ 60 and UL 150.
  - $10^6$ simulations (Potvin C), $10^5$ (K & M)

<table>
<thead>
<tr>
<th>method</th>
<th>(overall) power</th>
<th>power (stage 1)</th>
<th>% studies to stage 2</th>
<th>n₉₅%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potvin et al.</td>
<td>83.6</td>
<td>69.7</td>
<td>23.8</td>
<td>98</td>
</tr>
<tr>
<td>Karalis &amp; Macheras</td>
<td>74.2</td>
<td>67.2</td>
<td>7.2</td>
<td>130</td>
</tr>
</tbody>
</table>

- Power <80%; only ~$\frac{1}{3}$ of studies proceed to stage 2, although with considerably larger sample sizes.

Labes D and H Schütz
An Insight into the Properties of a Two-Stage Design in Bioequivalence Studies: A Rejoinder
Pharm Res (submitted 2013)
Futility Rules revised

- EMA GL Section 4.1.8 ‘Two-stage design’
  “[...] the stopping criteria should be clearly defined prior to the study.”

  - What does that mean?
    - Failing in stage 1 or the pooled analysis according to the chosen method.
      → Part of the validated frameworks.
    - Early stopping for futility (e.g., ‘bad’ ratio, extreme stage 2 sample size caused by high CV – better to opt for reference-scaling...).
      → Not validated. A misunderstanding by regulators (stopping criterion ≠ futility rule).
Futility Rules revised

- Introduction of a futility rule does not inflate the patient’s risk, but power may drop substantially!
  - State stopping criteria unambiguously in the protocol.
  - If you want to introduce a futility rule, simulations are mandatory in order to maintain sufficient power.

“Introduction of [...] futility rules may severely impact power in trials with sequential designs and under some circumstances such trials might be unethical.”

A Fuglsang

*Futility Rules in Bioequivalence Trials with Sequential Designs*

APPS J 16(19), 79–82 (2014) DOI: 10.1208/s12248-013-9540-0
Advanced Example

- ‘Must pass’ BE in stage 1 (first to file)
  - Fixed T/R 90% (pessimistic; very likely better)
  - Expected CV 20% (pilot study with two references)
  - ~30% expected drop-out rate; start with 88 to have \( n_1 \geq 60 \)
- Targets
  - >90% power for \( n_1 = 60 \) – even for extreme CV of 45%
  - 90% power for \( n_1 \geq 60 \) (CV 20%) in stage 1
  - Not <80% power for CV \( \geq 25\% \) in stage 1
  - Low probability to proceed to stage 2
Advanced Example

- ‘Must pass’ BE in stage 1 (first to file)
  - Sponsor prefered Method B (EU submission…)
  - Fuglsang published $\alpha_{adj.} = 0.0269$ for T/R 0.90 and 90% power – but only for Method C…
  - Same $\alpha_{adj.}$ applicable?
  - Likely…
    - Potvin et al. showed less inflation with Method B.
    - Fuglsang needed less adjustment in Method B.
    - But we have to justify that!
  - $10^6$ sim’s for $\alpha$ and $10^5$ for power.
    - Thanks to Detlew Labes for R package *Power2Stage*!
Advanced Example

270·10^6 Sim’s (Fuglsang mod. B: T/R 90%, power 90%)

α

CV

n_1

Bioequivalence Studies in Russia: Pharmacokinetics, Statistics and Analytics
Moscow, 24 April 2014
**Advanced Example**

- **‘Must pass’ BE in stage 1 (first to file)**
  - Targets met
    - 93% power for $n_1 \geq 60$ (CV 20%) and 90% for extreme CV of 45%
    - 90% power for $n_1 \geq 60$ (CV 20%) in stage 1
    - Low chances to proceed to stage 2 with CV 20%:
      - $n_1 = 60$: 6%, $n_1 = 72$: 1%
    - $\geq 80\%$ power for $CV \geq 20\%$, even for a more extreme drop-out rate
    - $\alpha_{adj.} = 0.0271$ would work as well (with 0.0278 < 0.052)
  - Study passed in the first stage (February 2014)
Two-Stage Sequential Designs

TSDs: Parallel Design

- A Fuglsang (2014)
  - Based on Potvin’s Methods B/C ($\alpha_{adj.}$ 0.0294, 80% power)
  - Framework: $n_1$ 48–120, CV 10–100%
  - Explored
    - equal and unequal variances of groups
    - conventional $t$-test and Welch-Satterthwaite approximation
- Results
  - No significant $\alpha$-inflation
  - Power $\geq 78.4$

A Fuglsang
Sequential Bioequivalence Approaches for Parallel Designs
AAPS J, Epub ahead of print (Feb 2014), DOI: 10.1208/s12248-014-9571-1
Case Study 1 (EMA)

- Method C: Study passed BE in stage 1 (49 subjects, CV 30.65%, 90% CI)
  - UK/Ireland: Unadjusted $\alpha$ in stage 1 not acceptable.
    - Study passed BE with 94.12% CI as well (post hoc switch to Method B).
  - Austria: The Applicant should demonstrate that the type I error inflation, which can be expected from the chosen approach, did not impact on the decision of bioequivalence.
    - One million simulations based on the study’s sample size and CV.
      - $\alpha_{emp}$ 0.0494 (95% CI: 0.0490 – 0.0498)
Case Study 2 (EMA)

- Method C: Study stopped in stage 1
  AUC power >80%: passed BE with 90% CI
  $C_{\text{max}}$ power <80%: passed BE with 94.12% CI

  - The Netherlands: Adapting the confidence intervals based upon power is not acceptable and also not in accordance with the EMA guideline. Confidence intervals should be selected *a priori*, without evaluation of the power. Therefore, the applicant should submit the 94.12% confidence intervals for AUC.

    - AUC fails BE with 94.12% CI
    - Sponsor repeated the study with a very (!) large sample size and failed on $C_{\text{max}}$. Project cancelled…
Case Study 3 (EMA)

- Method C: Two studies passed in stage 1
  \( (n=15 \text{ SD, } n=16 \text{ MD, } C_{\text{max}} \text{ CV } 17.93\%, 8.54\%, 90\% \text{ CIs}) \)
- Would have passed with Method B as well; however, 94.12\% CIs were not reported.
  - RMS Germany. Accepted by CMSs Austria, Denmark, Sweden, and The Netherlands.
  - Spain: Statistical analysis should be GLM. Please justify.
    - Evaluated with all-fixed effects model.
      Both studies passed.
      Issue resolved (September 2013)
Outlook

- Feasibility / futility rules.
- Arbitrary expected T/R and/or power.
- Methods without interim power.
- Dropping a candidate formulation from a higher-order cross-over; continue with 2×2.
- Full adaptive methods.
- Exact method (not depending on simulations).
- Application to replicate designs / scaling.
Don’t panic!

conventional 2×2 cross-over (fixed sample design)
Thank You!

Two-Stage Sequential Designs

Open Questions?

Helmut Schütz
BEBAC
Consultancy Services for Bioequivalence and Bioavailability Studies
1070 Vienna, Austria
helmut.schuetz@bebac.at
To bear in Remembrance...

The fundamental cause of trouble in the world today is that the stupid are cocksure while the intelligent are full of doubt.  

_**Bertrand Russell**_

In bioequivalence we must not forget the only important – _the patient_! He/she is living person, not just $\alpha 0.05$. 

_**Dirk Marteen Barends**_

It is a good morning exercise for a research scientist to discard a pet hypothesis every day before breakfast. It keeps him young. 

_**Konrad Lorenz**_
References

- ICH
- EMA-CPMP/CHMP/EWP
  - Points to Consider on Multiplicity Issues in Clinical Trials (2002)
  - Questions & Answers: Positions on specific questions addressed to the EWP therapeutic subgroup on Pharmacokinetics (2013)
- US-FDA
  - Center for Drug Evaluation and Research (CDER)
    - Statistical Approaches Establishing Bioequivalence (2001)
      - Draft Guidance on Loteprednol (Jun 2012)
      - Draft Guidance on Dexamethasone/Tobramycin (Jun 2012)
  - DB Owen
    - A special case of a bivariate non-central t-distribution
      - Biometrika 52(3/4), 437–46 (1965)
- Diletti E, Hauschke D, and VW Steinijans
  - Sample size determination for bioequivalence assessment by means of confidence intervals
- AL Gould
  - Group Sequential Extension of a Standard Bioequivalence Testing Procedure
    - DOI: 10.1007/BF02353786
- Hauck WW, Preston PE, and FY Bois
  - A Group Sequential Approach to Crossover Trials for Average Bioequivalence
    - DOI: 10.1080/10543409708835171
- Patterson S and B Jones
  - Determining Sample Size, in: Bioequivalence and Statistics in Clinical Pharmacology
- SA Julious
  - Sample Sizes for Clinical Trials
- D Labes
    - http://cran.r-project.org/web/packages/PowerTOST/PowerTOST.pdf
Two-Stage Sequential Designs

References

- D Labes
  Package ‘Power2Stage’, Version 0.0-8 (2014-04-11)
  http://cran.r-project.org/web/packages/Power2Stage/Power2Stage.pdf

- Potvin D et al.
  Sequential design approaches for bioequivalence studies with crossover designs
  DOI: 10.1002/pst.294

- Montague TH et al.
  Additional results for ‘Sequential design approaches for bioequivalence studies with crossover designs’
  Pharmaceut Statist 11(1), 8–13 (2011) DOI: 10.1002/pst.483

- Garcia-Arieta A and J Gordon
  Bioequivalence Requirements in the European Union: Critical Discussion

- BM Davit
  Sequential Designs and Interim Analyses in Bioequivalence: FDA’s Experience
  AAPS Annual Meeting, Chicago, IL, October 13–18, 2012

- A Fuglsang
  Sequential Bioequivalence Trial Designs with Increased Power and Controlled Type I Error Rates
  DOI: 10.1208/s12248-013-9475-5

- Karalis V and P Macheras
  An Insight into the Properties of a Two-Stage Design in Bioequivalence Studies
  DOI: 10.1007/s11095-013-1026-3

- Karalis V and P Macheras
  On the Statistical Model of the Two-Stage Designs in Bioequivalence Assessment
  DOI: 10.1111/jphp.12164

- A Fuglsang
  Futility Rules in Bioequivalence Trials with Sequential Designs

- A Fuglsang
  Sequential Bioequivalence Approaches for Parallel Designs
  AAPS J, Epub ahead of print (Feb 2014)
  DOI: 10.1208/s12248-014-9571-1