Add-On and Sequential Designs

Helmut Schütz
BEBAC

Add-On Designs

- Were extensively discussed at the Bio-International ’92 Conference, Bad Homburg, Germany
  - Sometimes (rarely) accepted in the US and EU
  - Implemented in Canadian (1992, 1996) and Japanese guidelines (1997, 2006)
  - 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 (1st part: n=55, df 53)
    0.01718 (2nd part: n=14, df 12)

Add-On Designs

Sum of Squares

Hypothesis	DF	SSE	MSE	F_stat	P_value
Sequence	1	0.0271658	0.0271658	0.173639	0.6786
Sequence*Subject	53	8.29185	0.15645	11.844	<0.0001
Treatment	1	0.211196	0.211196	15.9885	0.0002
Period	1	0.0271536	0.0271536	2.05565	0.1575
Error	53	0.700088	0.0132092		

Sum of Squares

Hypothesis	DF	SSE	MSE	F_stat	P_value
Sequence	1	0.0489527	0.0489527	0.419204	0.5295
Sequence*Subject	12	1.4013	0.116775	6.79641	0.0012
Treatment	1	0.0349142	0.0349142	2.03203	0.1795
Period	1	0.0839476	0.0839476	4.88581	0.0472
Error	12	0.206183	0.0171819		

$$\hat{F} = \frac{MSE_{large}}{MSE_{small}} = \frac{0.0171819}{0.0132092} = 1.30075$$

$$F_{1-0.05,12,53} = 1.940$$

$$p\left(\hat{F}\right) = 0.24595 > 0.05$$ ✓

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 Designs

Tests of Model Effects

Hypothesis	Numer_DF	Denom_DF	F_stat	P_value
int	1	56.5	2144.16	<0.0001
Study	1	56.5	0.0007	0.9784
Treatment	1	64.6	9.9949	0.0024
Treatment*Study	1	64.6	0.1156	0.7349

$F_{1-0.05,1,64.6} = 3.989$

$p(0.1156) = 0.7349 > 0.05$

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 Designs

- Canada cont'd
  - Formulation by study interaction not significant (p 0.735), pooled analysis acceptable.
  - No α–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 ≥20 *or* sample size of both parts ≥30
  - Sample size of second part ≥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 α–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).
  - Based on work by Armitage *et al.* (1969), McPherson (1974), Pocock (1977), O'Brien and Fleming (1979) and others.
    - 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)

CV%	90% CI	94.12% CI	ratio
10	8	8	1.000
15	12	14	1.167
20	20	24	1.200
25	28	34	1.214
30	40	48	1.200

  - 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.

Sequential Designs

- LA Gould (1995)
  - Overall α–level preserved at ≤0.05. Individual CI level ~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 α–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 alter-natives and the choice of how much alpha to spend at the interim analysis is at the company's discretion.

Two-Stage Design

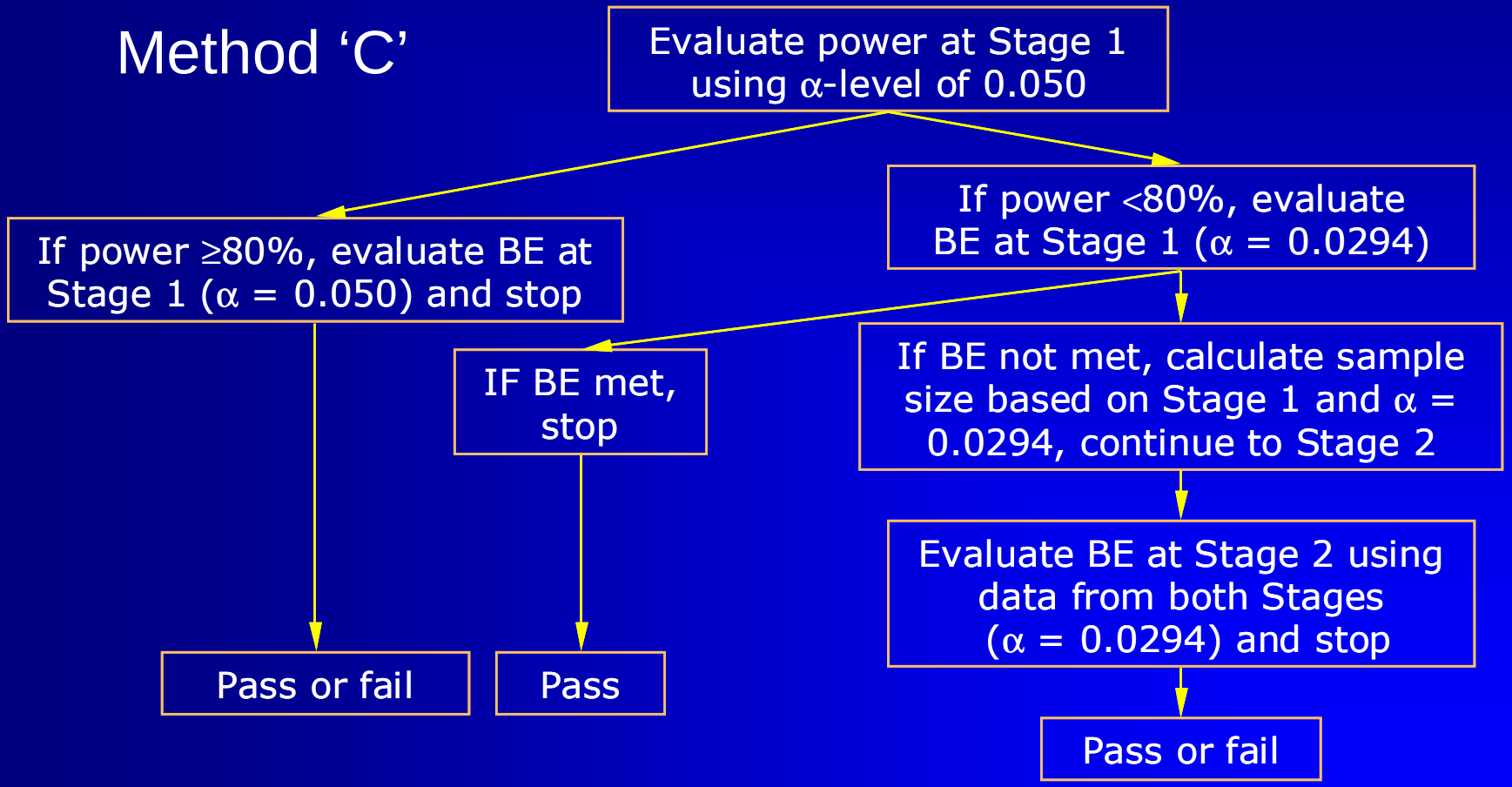
- EMA GL on BE (2010)
  - Section 4.1.8 (cont'd)
    - Plan to use a two-stage approach must be pre-specified in the protocol along with the adjusted significance levels to be used for each of the analyses.
    - When analysing the combined data from the two stages, a term for stage should be included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

Potvin *et al.* (2008)

Method 'C'



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 α of 0.0294 (Pocock 1977)
  - If power is <80% in Stage 1 and in the pooled analysis (data from Stages 1 + 2), α 0.0294 is used (*i.e.*, a 1–2×α = 94.12% CI is calculated)
  - Overall patients' risk is preserved at ≤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_{intra} 27.2%

α 0.0294 (if power <80%)

Failed 90% CI (if power ≥80%) and 94.12% CI (if power <80%)

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
```

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

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
```

Calculate total sample size: expected ratio 95%, CV_{intra} 27.2%, 80% power

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

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)
    CI User = (   91.2387,  110.6797)
    Average bioequivalence shown for confidence=94.12 and percent=20.0.
```

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

α 0.0294 in pooled analysis

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

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 Hepburne

References

- Collection of links to global documents
  http://bebac.at/Guidelines.htm
- ICH
  - E9: Statistical Principles for Clinical Trials (1998)
- EMA-CPMP/CHMP/EWP
  - Points to Consider on Multiplicity Issues in Clinical Trials (2002)
  - Guideline on the Investigation of BE (2010)
- H Melander
  Problems and Possibilities with the Add-On Subject Design
  In: KK Midha and HH Blume (eds.)
  Bio-International. Bioavailability, Bioequivalence and Pharmacokinetics
  medpharm Scientific Publishers, Stuttgart, pp 85–90 (1993)
  ISBN 388763-019-X
- LA Gould
  Group Sequential Extension of a Standard Bioequivalence Testing Procedure
  J Pharmacokin Biopharm 23/1, 57–86 (1995)
  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
  Pharmaceut Statist 7/4, 245–262 (2008)
  DOI: 10.1002/pst.294