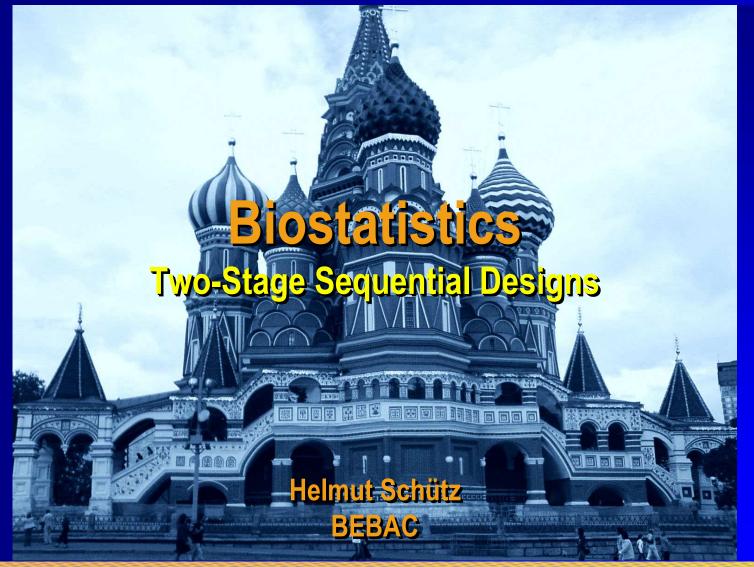


Wikimedia Commons • 2006 Schwallex • Creative Commons Attribution-ShareAlike 3.0 Unported





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)
 - Based on work by Armitage *et al.* (1969), McPherson (1974), Pocock (1977), O'Brien and Fleming (1979), Lan and DeMets (1983), ...
 - 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 J Pharmacokin Biopharm 23(1), 57–86 (1995) 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.

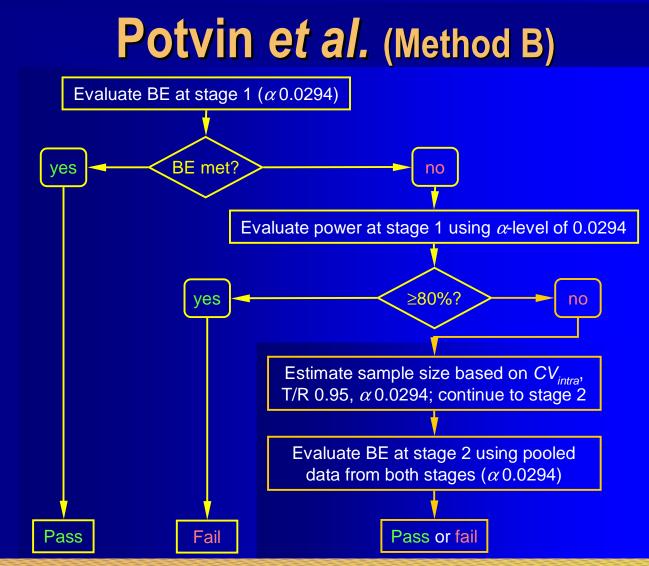
Potvin D, DiLiberti CE, Hauck WW, Parr AF, Schuirmann DJ, and RA Smith Sequential design approaches for bioequivalence studies with crossover designs Pharmaceut Statist 7(4), 245–62 (2008) DOI: 10.1002/pst.294



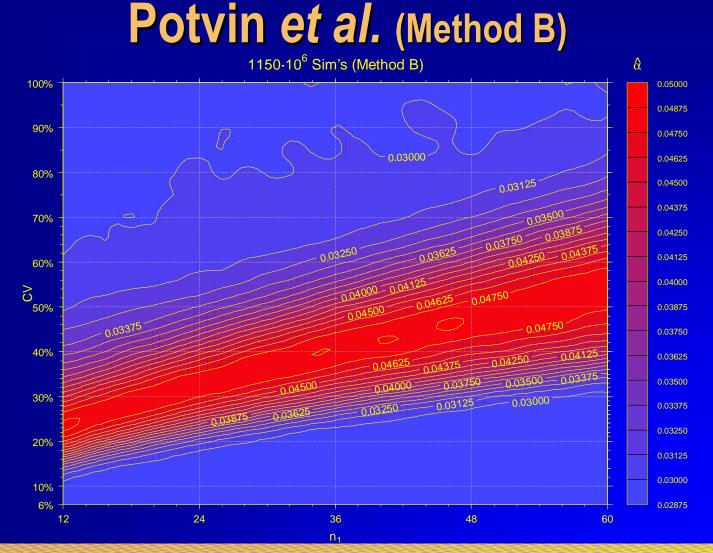
Review of Guidelines

•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 (?)

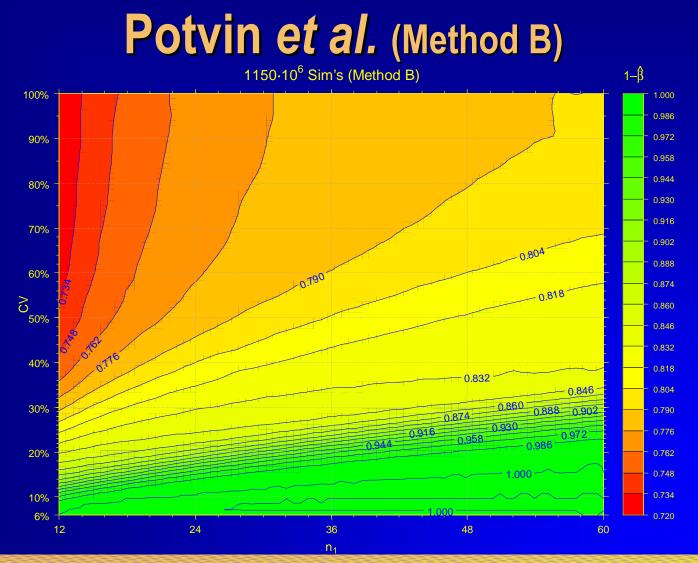




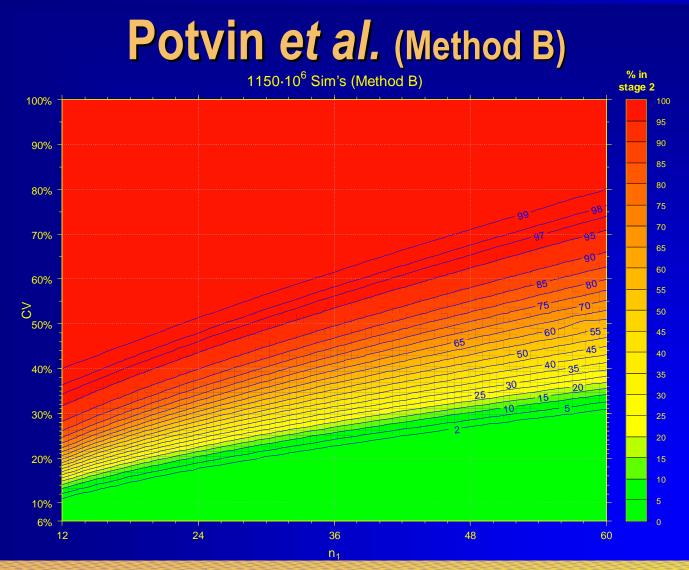




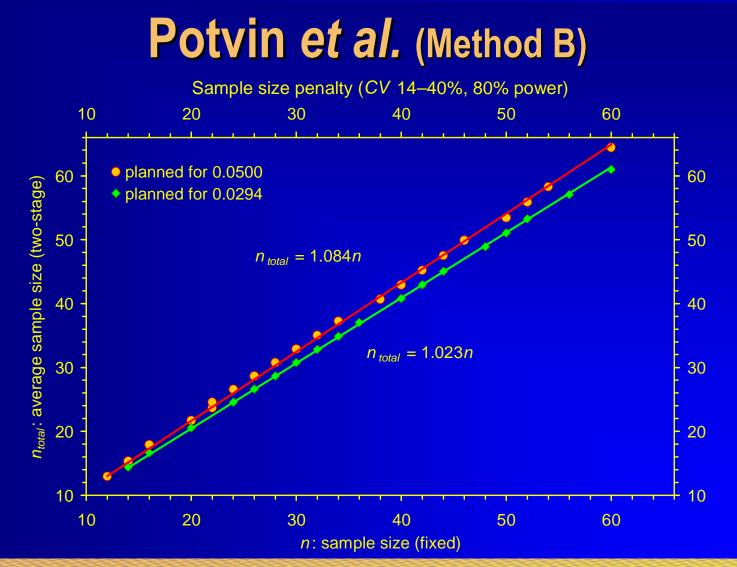














Technical Aspects

- Only one Interim Analysis (after stage 1).
- Use software (wide step sizes in Diletti's tables); preferrably 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).



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 α 0.0294 is used in stage 1 and in the pooled analysis (data from stages 1 + 2), *i.e.*, the 1 – 2× α = 94.12% CI is calculated.
- Overall patient's risk preserved at ≤0.05.



Technical Aspects (cont'd) + EMA modification

If the study is stopped after stage 1, the statistical model is:

fixed: sequence + period + treatment

+ subject(sequence)

If the study continues to stage 2, the model for the combined analysis is:

fixed: stage + sequence + sequence(stage)

- + subject(sequence × stage) + period(stage)
- + treatment

No poolability criterion! Combining is always allowed – even if a significant difference between stages is observed. No need to test this effect.



Technical Aspects (cont'd) + EMA modification Incomprehensible why this modification was introducted 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 J Pharm Pharmacol 66(1), 48–52 (2014) DOI: 10.1111/jphp.12164



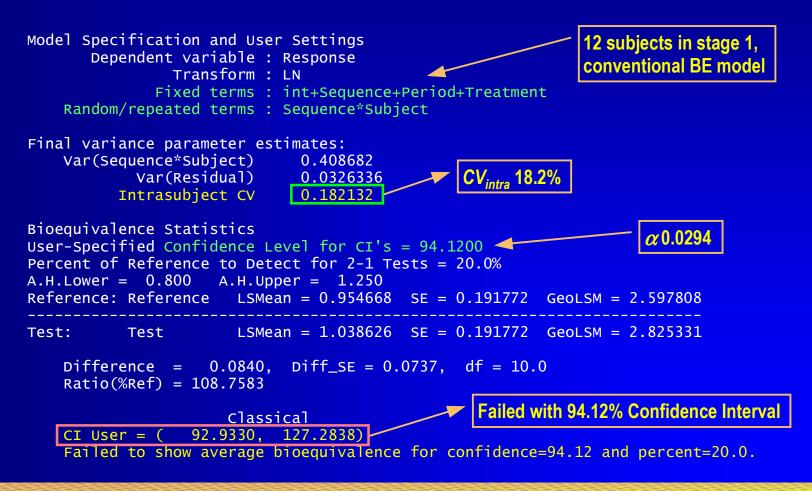
Technical Aspects (cont'd)

- 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):

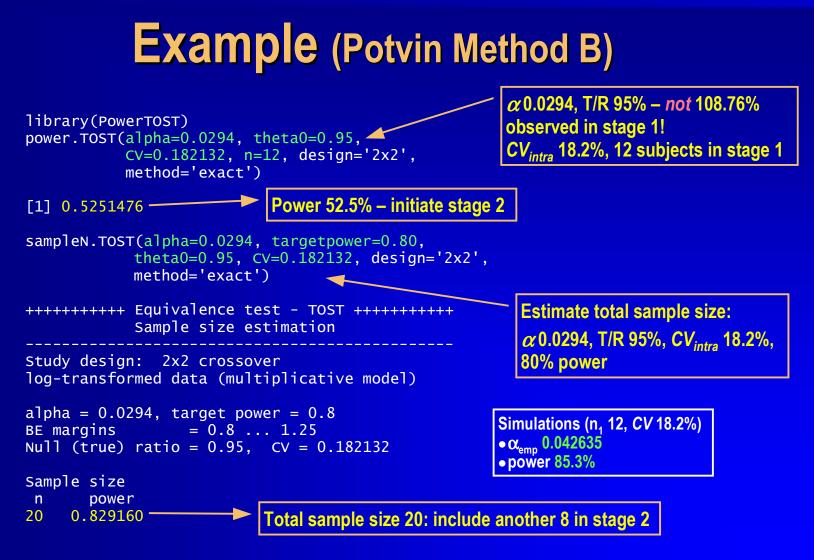
method	% power	
approximative (shifted central <i>t</i>)	50.49	
approximative (noncentral t)	52.16	
exact (Owen's <i>Q</i>)	52.51	



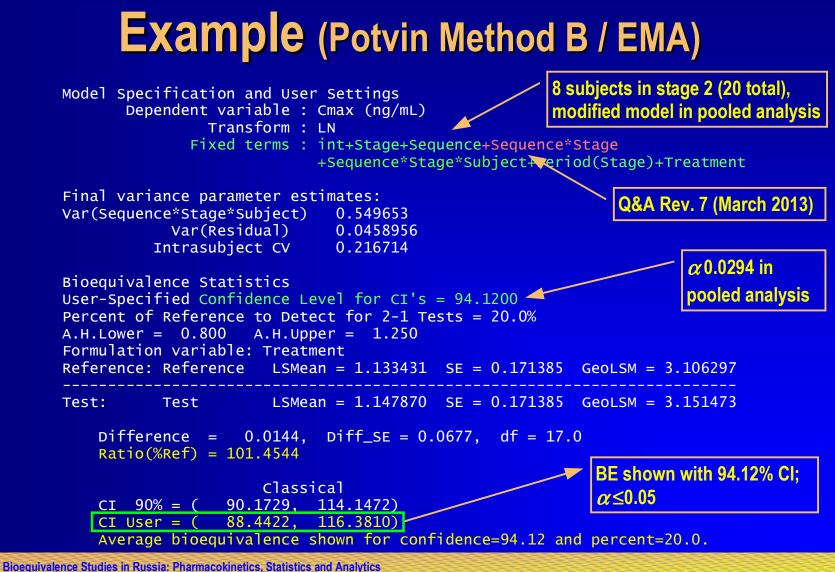






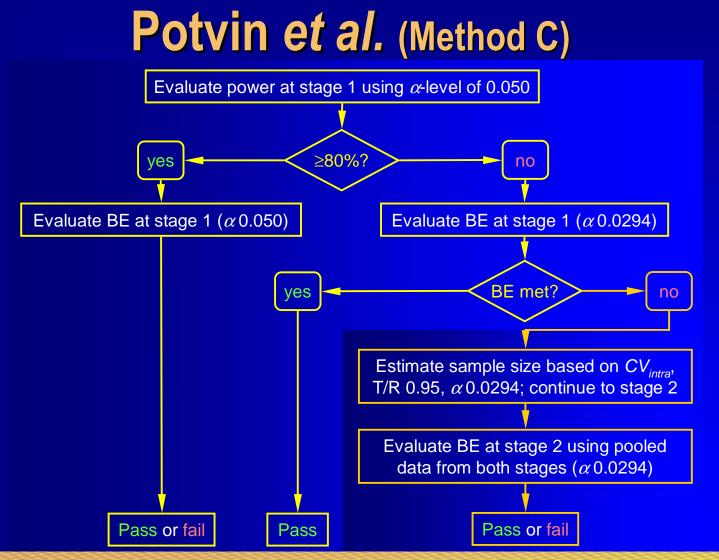




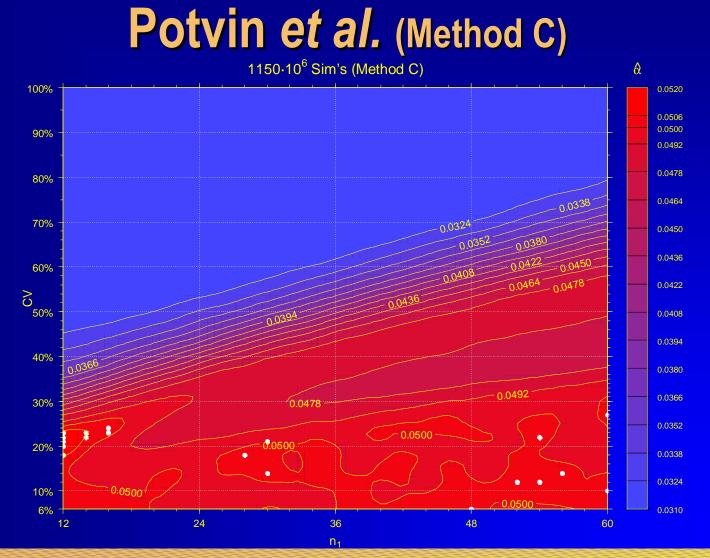


Moscow, 24 April 2014











Potvin et al. (Method B vs. C)

•Pros & cons

- Method C (*if power* ≥80%) is a conventional BE study; no penalty in terms of α 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 reason-able 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)

reference	method	T/R	target power	CV	$lpha_{adj.}$	max. $\alpha_{emp.}$
Potvin <i>et al.</i>	В	0.95		10–100%	0.0294	0.0485
	С	0.95				0.0510
Montague et al.	D	0.90			0.0280	0.0518
Fuglsang	В	0.95	90%	10–80%	0.0284	0.0501
	D				0.0274	0.0503
	D	0.90			0.0269	0.0501

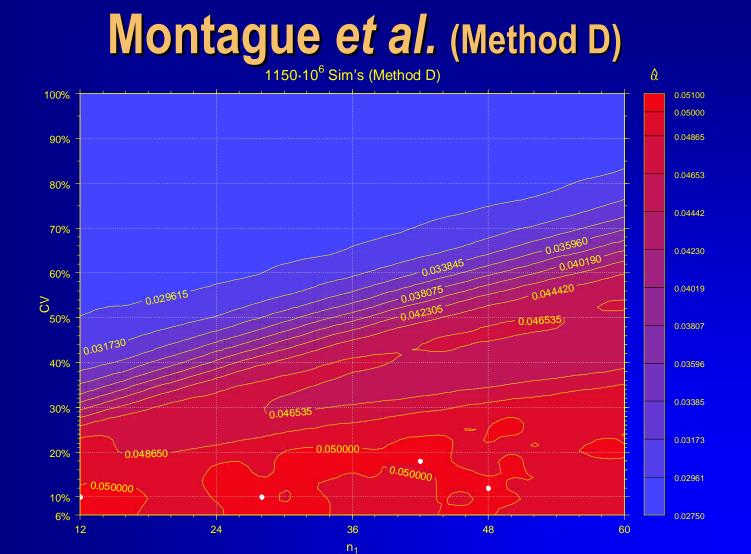
Montague TH, Potvin D, DiLiberti CE, Hauck WW, Parr AF, and DJ Schuirmann

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 AAPS J 15(3), 659–61 (2013) DOI: 10.1208/s12248-013-9475-5







TSDs: Alternatives

•Karalis & Macheras (2013), Karalis (2013)

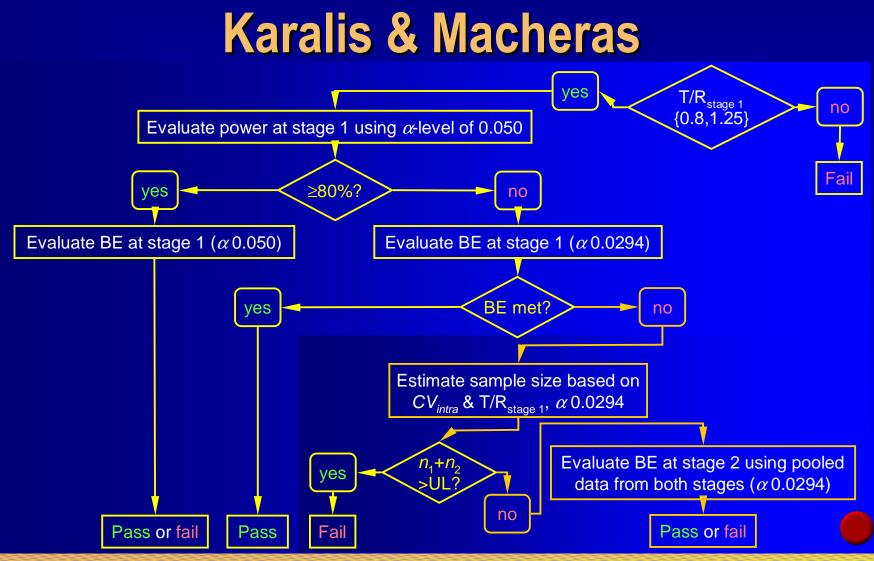
- Based on Method C ($\alpha_{adj.}$ 0.0294) or D ($\alpha_{adj.}$ 0.0280)
- Sample size re-estimation based on observed T/R-ratio in stage 1 (fully adaptive)
- Upper sample size limit (UL)
- Frameworks:

■ n_1 12–96, CV 10–60%, $n_1+n_2 \le$ UL 150 ■ n_1 18–96, CV 20–40%, $n_1+n_2 \le$ UL 100

Karalis V and P Macheras

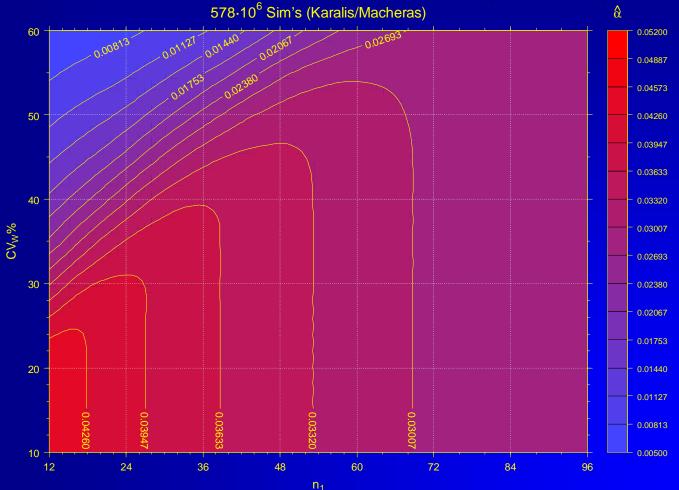
An Insight into the Properties of a Two-Stage Design in Bioequivalence Studies Pharm Res 30(7), 1824–35 (2013), <u>DOI: 10.1007/s11095-013-1026-3</u> V Karalis The role of the upper sample size limit in two-stage bioequivalence designs Int J Pharm 456(1), 87–84 (2013), DOI: 10.1016/j.jipharm.2013.08.013





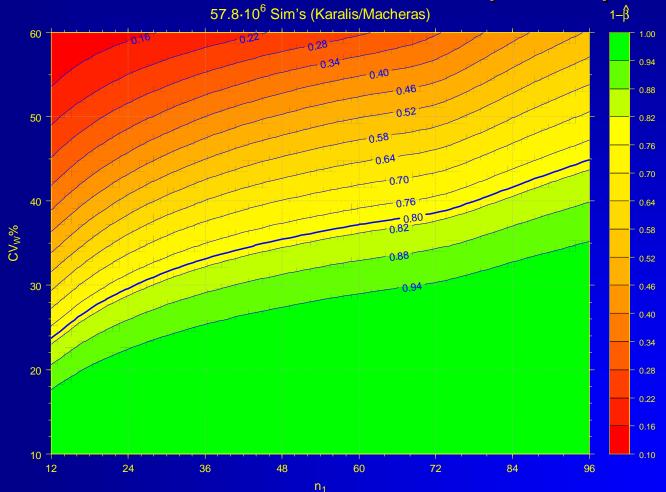


Karalis & Macheras (n ≤150)



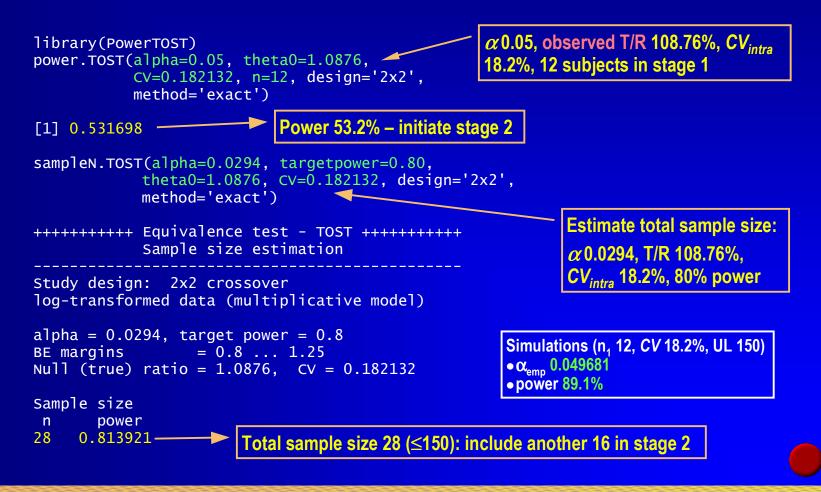


Karalis & Macheras (n ≤150)





Karalis & Macheras (n ≤150)





Karalis & Macheras (Expl. a)

• CV assumed as 20%, T/R 95%

- In a fixed sample design for 80% power sample sizes would be 20 (α 0.05) or 24 (α 0.0294).
- The sponsor chooses n₁ 24 and UL 100.
- 10⁶ simulations (Potvin C), 10⁵ (K & M)

method	(overall) power	power (stage 1)	% studies to stage 2	n _{95%}
Potvin <i>et al.</i>	90.1	88.1	4.2	24
Karalis & Macheras	94.8	83.5	11.4	66

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.
- 10⁶ simulations (Potvin C), 10⁵ (K & M)

method	(overall) power	power (stage 1)	% studies to stage 2	n _{95%}
Potvin e <i>t al.</i>	83.6	69.7	23.8	98
Karalis & Macheras	74.2	67.2	7.2	130

Power <80%; only ~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.

 \rightarrow 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 referencescaling...).

 \rightarrow Not validated. A misunderstanding by regulators (stopping criterion \neq 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 circum-

stances 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



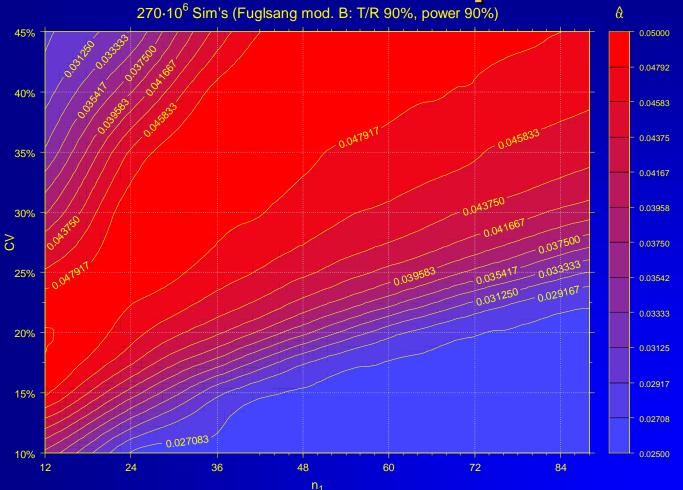
•'Must pass' BE in stage 1 (first to file)

- 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 $n_1 \ge 60$
- Targets
 - **Solution Solution Solution**
 - **90%** power for $n_1 \ge 60$ (*CV* 20%) in stage 1
 - Not <80% power for CV ≥25% in stage 1
 - Low probability to proceed to stage 2



- •'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 α and 10⁵ for power.
 - Thanks to Detlew Labes for *R* package *Power2Stage*!







•'Must pass' BE in stage 1 (first to file)

- Targets met
 - 93% power for n₁ 60 (CV 20%) and 90% for extreme CV of 45%
 - **90%** power for $n_1 \ge 60$ (*CV* 20%) in stage 1
 - Low chances to proceed to stage 2 with CV 20%: n₁ 60: 6%, n₁ 72: 1%
 - ■≥80% power for CV ≥20%, even for a more extreme drop-out rate
- α_{adj.} 0.0271 would work as well (with 0.0278 < 0.052)
 Study passed in the first stage (February 2014)



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 *a*-inflation
 - Power ≥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





 Method C: Study passed BE in stage 1 (49 subjects, CV 30.65%, 90% CI)

- **UK**/Ireland: Unadjusted α 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.

 α_{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_{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_{max}. Project cancelled...



Case Study 3 (EMA)

 Method C: Two studies passed in stage 1 (n=15 SD, n=16 MD, C_{max} CV 17.93%, 8.54%, 90% CIs)

- Would have passed with Method B as well; however, 94.12% Cls 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 higherorder cross-over; continue with 2×2.
- •Full adaptive methods.
- Exact method (not depending on simulations).

Application to replicate designs / scaling.







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 α 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

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)
- 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)
 - Bioequivalence Recommendations for Specific Products (2007–2013):

Draft Guidance on Lotepredenol (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

Int J Clin Pharm Ther Toxicol 29(1), 1-8 (1991)

AL Gould

Group Sequential Extension of a Standard Bioequivalence Testing Procedure

J Pharmacokin Biopharm 23(1), 57–86 (1995) DOI: 10.1007/BF02353786

- Hauck WW, Preston PE, and FY Bois A Group Sequential Approach to Crossover Trials for Average Bioequivalence J Biopharm Stat 71(1), 87–96 (1997)
- DOI: 10.1080/10543409708835171
- Patterson S and B Jones Determining Sample Size, in: Bioequivalence and Statistics in Clinical Pharmacology Chapman & Hall/CRC, Boca Raton (2006)
- SA Julious

Sample Sizes for Clinical Trials Chapman & Hall/CRC, Boca Raton (2010)

D Labes

Package 'PowerTOST', Version 1.1-10 (2014-01-31) http://cran.r-project.org/web/packages/PowerTOST/PowerTOST.pdf



References

- D Labes Package 'Power2Stage', Version 0.0-8 (2014-04-11) http://cran.rproject.org/web/packages/Power2Stage/Power2Stage.pdf Potvin D et al. Sequential design approaches for bioequivalence studies with crossover designs Pharmaceut Statist 7(4), 245–62 (2008) 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 García-Arieta A and J Gordon **Bioequivalence Requirements in the European Union: Critical** Discussion AAPS J 14(4), 738-48 (2012) DOI: 10.1208/s12248-012-9382-1 BM Davit Sequential Designs and Interim Analyses in Bioequivalence: FDA's Experience AAPS Annual Meeting, Chicago, IL, October 13–18, 2012
- Bioequivalence Studies in Russia: Pharmacokinetics, Statistics and Analytics Moscow, 24 April 2014

A Fuglsang

Sequential Bioequivalence Trial Designs with Increased Power and Controlled Type I Error Rates AAPS J 15(3) 659–61 (2013) 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 Pharm Res 30(7), 1824–35 (2013) DOI: 10.1007/s11095-013-1026-3
- Karalis V and P Macheras On the Statistical Model of the Two-Stage Designs in Bioequivalence Assessment
- J Pharm Pharmacol 66(1), 48–52 (2014)
- DOI: 10.1111/jphp.12164
- A Fuglsang Futility Rules in Bioequivalence Trials with Sequential Designs APPS J 16(19), 79–82 (2014) DOI: 10.1208/s12248-013-9540-0
- A Fuglsang Sequential Bioequivalence Approaches for Parallel Designs AAPS J, Epub ahead of print (Feb 2014) DOI: 10.1208/s12248-014-9571-1