MEDICAL UNIVERSITY OF VIENNA

LISBOA UNIVERSIDADE DE LISBOA — FACULDADE DE FARMÁCIA Universidade de Lisboa

BE·BAC

To pool, or not to pool, that is the question…

Wikimedia Commons • 2007 Sokoljan • Creative Commons SA 3.0 Unported

Helmut Schütz

Center for Medical Data Science
of the Medical University of Vienna
Faculty of Pharmacy
of the University of Lisbon
BEBAC, Vienna

ICH M13A Q&As: 4.1 Documentation

M13A recommends that all relevant BE studies conducted, regardless of the study outcome, should be provided. If […] multiple pivotal studies result in inconsistent BE conclusions, the totality of the evidence should be considered. The applicant should discuss the results and justify the BE claim. When relevant, a combined analysis of all studies may be considered as a sensitivity analysis in addition to the individual study analyses. It is not acceptable, however, to pool studies which fail to demonstrate BE without a study that passes.

ICH M13A Q&As: 4.1 Documentation

If there are differences in the study conditions, *e.g.*, sampling times, fasting or fed conditions, or method of administration, pooling is not justifiable. A different number of subjects is not considered a difference in study conditions.

An excursion into sample size estimation

ICH M13A, Section 2.1.3

The number of subjects to be included in the BE study should be based on an appropriate sample size determination to achieve a pre-specified power and pre-specified type 1 error. A sufficient number of subjects should be enrolled in the BE study to account for possible dropouts and/or withdrawals.

- Power is chosen by the applicant (≥80%)
- Type I error is fixed by the authority (≤5%)
- T/R-ratio is an assumption
- Variability is an assumption

} appropriate!

- Dropout- / withdrawal-rate is anticipated

Failed studies

Any study may fail because the *realized* values (*i.e.*, the observations) deviate from the assumptions

Impact

- Worse T/R-ratio (deviating more from 100% than assumed)
- Worse variability (higher than assumed)
- Worse dropout- / withdrawal-rate (higher than anticipated)

Types of failure

- Confidence interval *entirely* outside the acceptance range, *i.e.*,
  upper CL < lower acceptance limit *or*
  lower CL > upper acceptance limit
  → **bioinequivalence** proven (reformulate)
- One confidence limit within the acceptance range
  → **inconclusive** (underpowered)

Information from an underpowered study

The design might have been insufficient

- Too limited sampling in the area of t_{max} led to increased variability of C_{max}
- Unexpected "first time C_{max}"
- Pre-dose concentrations due to a too short washout phase
- If you have to change the study conditions, the new study will stand on its own
  → pooling is not justifiable

Even with a proper design, the sample size was too small

- The results (T/R-ratio, variability, dropout-rate) are more reliable than any pilot study or data from the literature
  → design the next study based on all new information

Appropriate cases

Insufficient design of first study (failed)

- Properly designed second study (passed)
- Pooling not acceptable due to a change of study conditions
- Discuss the results and justify the BE claim

Proper design of the first study (failed due to lack of power)

- *Larger* sample size of the second study (passed)
  - Discuss the results and justify the BE claim
  - Pooling is acceptable in principle, but why?

Questionable cases

Proper design of the first study (failed due to lack of power)

- *Same* sample size of the second study (passed)
  - Discuss the results and justify the BE claim
  - Pooling is acceptable in principle, but why?
  - Expect – rightly! – questions from the authority
    - Why does the applicant consider the second study to be more ‘reliable’ than the first?
    - Might have to provide ‘coverage probability’ or ‘tipping point analysis’
    - Repeating a study often enough, sooner or later one will pass simply by chance

Pooling and beyond

Serious poblems

- The entire *α* was already *spent* in the first study – it was not planned as a two-stage design, which would allow an *α*-adjustment to control the patient's risk
- The second study is assessed with the *same α*, allowing a confirmatory analysis
- Calculating *any* confidence interval of the pooled studies is *wrong*; only the point estimate is correct

If you are unsure about your assumptions

- Opt for a two-stage design instead
- Valid confidence interval; no problems with patient's risk
- Will be covered in ICH M13C

Simulation

2×2×2 crossover

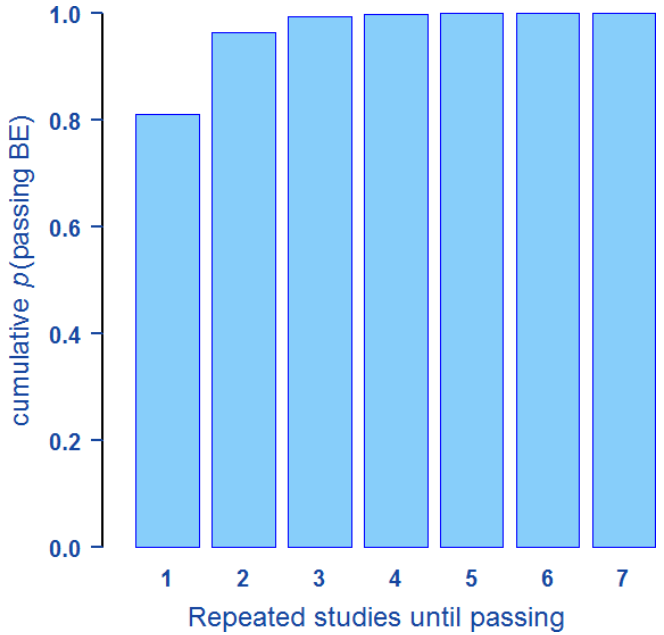
- Study planned assuming a T/R-ratio of 0.95 and CV of 20%, powered for ≥80% → n=20
  - n=18: PE 90.25% (90% CI: 79.52 – 102.42%), CV 22.0%
    'coverage probability' 94.22%
- Repeated study planned with the outcomes of the first → n=42
  - n=36: PE 92.56% (90% CI: 85.61 – 100.09%), CV 19.8%
    'coverage probability' 99.83%
- Pooled analysis: N=54, PE 91.79%, CV 20.5%
  - 90.00% CI: 87.81 – 95.95%
  - 94.12% CI: 87.23 – 96.60%
    'coverage probability' 99.99994%

Multiple repeats

2×2×2 crossover

- Study planned under the same assumptions as before
- If it fails, repeated studies with the *same* sample size as the first
  - Since the studies are planned for ≥80% power, each of them has still a fair chance of passing
  - If the *next* study fails as well, yet another repeat
  - Continue until a study passes

Repeat #	Σ *p*(BE)
1	0.8000
2	0.9600
3	0.9920
4	0.9984
5	0.9997
6	0.9999
7	≈1.0000



Conclusions

- If a study failed by design, change conditions of the second; provide both studies and justify the BE claim
  → no pooling
- If a study is underpowered, repeat it in a larger sample size; provide both studies and justify the BE claim
  → pooling possible but of doubtful value
- Never – ever! – repeat a study in the same sample size

Medical researchers can be divided into two sorts: those who think that meta is better and those who believe that pooling is fooling.

Stephen Senn

Meta-analyst: One who thinks that if manure is piled high enough it will smell like roses.

Stephen Senn

To pool or not to pool – that is not a question

Thank You!

Helmut Schütz
Center for Medical Data Science of the Medical University of Vienna
1090 Vienna, Austria
helmut.schuetz@muv.ac.at
Faculty of Pharmacy of the University of Lisbon
1649-003 Lisbon, Portugal
helmut@edu.ulisboa.pt
BEBAC
1070 Vienna, Austria
helmut.schuetz@bebac.at

Backup

It is not acceptable to pool studies which fail to demonstrate BE without a study that passes. (M13A Q&As: 4.1)

- Two *failed* studies, both with 18 subjects, where the PE of the second study is the reciprocal of the PE of the first study
- PE 90.25% (90% CI: 79.52 – 102.42%)
  'coverage probability' 94.22%
- PE 110.80% (90% CI: 97.63 – 125.75%)
  'coverage probability' 94.22%
- Pooled analysis: PE 100.00% (!!)
  - 90.00% CI: 92.00 – 109.05%
  - 94.12% CI: 90.47 – 110.54%
    'coverage probability' 99.9942%