Improved Testing of Multiple Endpoints in Group Sequential Trials

Susanne Urach, Martin Posch, Section for Medical Statistics, CEMSIIS
susanne.urach@meduniwien.ac.at, martin.posch@meduniwien.ac.at

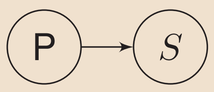
Group sequential test for two endpoints

Consider a two stage group sequential clinical trial with two normally distributed endpoints, one primary and one secondary endpoint. The means $\mu_i^T, i = 1, 2$ of the two outcomes in the treatment arm are compared to the corresponding values in the control arm $\mu_i^0, i = 1, 2$ testing the one sided hypotheses

$$H_P : \mu_1^T \leq \mu_1^0 \text{ vs. } H_1' : \mu_1^T > \mu_1^0 \quad \text{and} \quad H_S : \mu_2^T \leq \mu_2^0 \text{ vs. } H_2' : \mu_2^T > \mu_2^0.$$

Here $\delta_1 = \mu_1^T - \mu_1^0$, $\delta_2 = \mu_2^T - \mu_2^0$ denote the effect sizes and ρ the unknown true correlation between the endpoints.

Hierarchical Test: The secondary hypothesis is only tested after the primary hypothesis has been rejected:



Test each hypothesis with group sequential test at level $\alpha \rightarrow$ control of multiple type I error rate

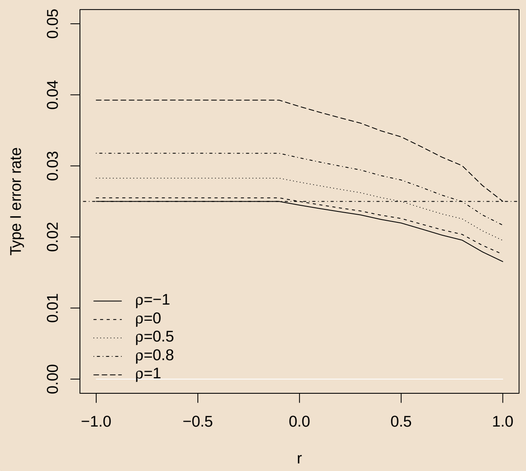
Stopping rule: Stop the trial in the interim analysis if H_P is rejected.

The stopping rule makes the test for the secondary endpoint strictly conservative [1, 3]. The actual type I error rate depends on the correlation. If the correlation were known we can improve the group sequential boundaries.

Problem of unknown correlation

Analyses of multiple endpoints in group sequential clinical trials use tests relying on the joint multivariate distribution which either presume known correlations or correlation estimates based on all data accumulated so far. As misspecification of true correlation ρ can have a large impact on the type I error rate, it is often proposed to use the most conservative adjustment across all possible correlations. But having observed data, we should not need to consider correlations completely unsupported by the data.

Type I error rate if correlation is misspecified



What is the type I error rate for H_S if we adjust group sequential boundaries assuming the correlation is r although the true correlation is ρ?

The maximum type I error rate maximized over all $\delta_1 > 0$ is shown in the left Figure.

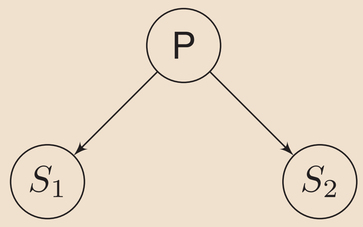
Methods to deal with unknown correlation ρ

- **Worst case approach:** Choose the correlation leading to the highest possible type I error rate to adjust the boundaries ($\rightarrow$ decrease in power). This is equivalent to not improving the boundaries at all.
- **Naive approach:** Use an interim estimate of the correlation to calculate the adjusted boundaries ($\rightarrow$ type I error rate inflation).
- **Berger Boos Method** = method to adjust for the worst case across all correlations in a confidence interval and still control the type I error rate of the trial [2]
- **Maximum approach**: lower the significance level such that when using an estimate of the correlation the type I error rate is still controlled.

Outlook: Generalisation to three endpoints

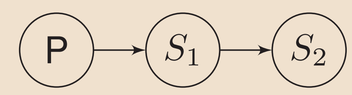
Gate-keeping Test

Testing two secondary endpoints as soon as the primary endpoint can be rejected



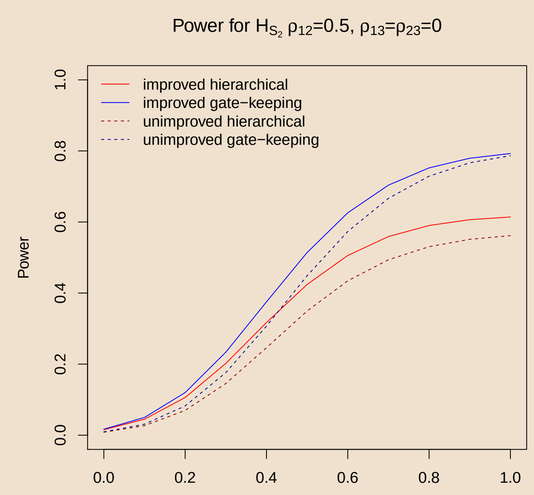
Hierarchical Test

Testing the secondary endpoint only if the primary endpoint can be rejected and the tertiary only if the secondary endpoint can be rejected

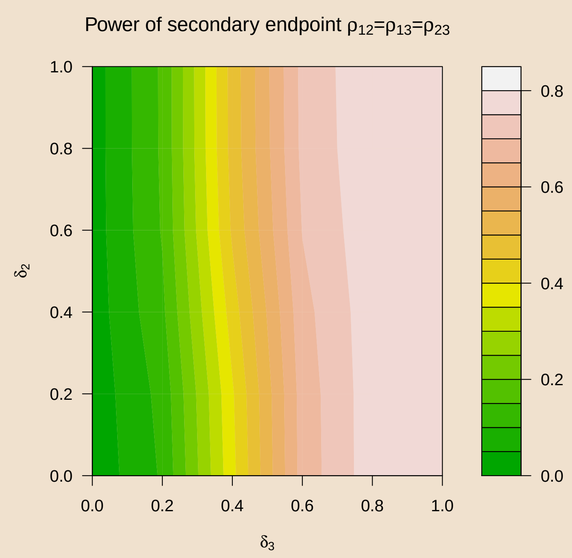


Using the correlation between the endpoints leads to improved boundaries and an increase in power. The secondary power of the gate-keeping test is bounded by the power of the primary endpoint while the tertiary power of the hierarchical procedure approximates the power of the power of the secondary endpoint.

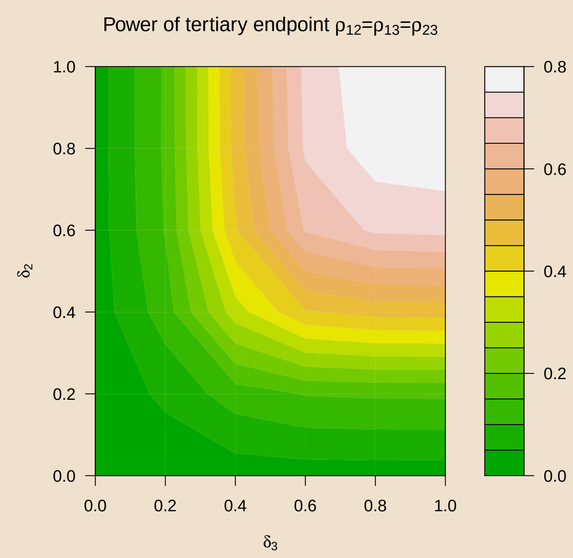
For $\delta_1 = \delta_2 = 0.5$ we calculated the power to reject the third hypothesis:



Improved Gate-keeping Test



Improved Hierarchical Test



Future work:
Improved tests for three endpoints based on estimated correlations.

Conclusions

- It is possible to use multiple testing methods based on the joint multivariate distribution together with an estimate for the correlation and still control the type I error rate.
- The power of testing procedures for multiple endpoints in group sequential trials can be improved by estimating the correlation. For the case of three endpoints the optimal testing strategy depends on the alternative considered.

References

[1] Tamhane et al. *Testing a primary and a secondary endpoint in a group sequential design*. Biometrics 2010.

[2] Tamhane et al. *Adaptive extensions of a two-stage group sequential procedure for testing primary and secondary endpoints (I): unknown correlation between the endpoints.* Statistics in Medicine 2012.

[3] Glimm et al. *Hierarchical testing of multiple endpoints in group-sequential trials.* Statistics in Medicine 2010.

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement number FP HEALTH 2013-603160. ASTERIX Project - http://www.asterix-fp7.eu/