A modified combination test for the analysis of clinical trials

Markus Neuhäuser

Department of Mathematics and Technique, RheinAhrCampus Remagen,

Koblenz University of Applied Sciences, Germany

In collaboration with Ann-Kristin Leuchs

Supported by DFG, Grant NE 1170/3-2

Introduction

In some **clinical trials**,

protocol amendments are required that change the inclusion criteria.

Possible reasons:

- too low recruitment rates
- regular violations of entry criteria

Introduction

In some **clinical trials**,

protocol amendments are required that change the inclusion criteria.

Possible reasons:
- too low recruitment rates
- regular violations of entry criteria

Example 1:

placebo-controlled trial in patients with asthma (Chow & Shao, 2005)

patient enrolment was slow $\rightarrow$ inclusion criteria were relaxed.

original protocol: baseline FEV_1 (l/sec) $\in [1.5; 2.0]$

1st amendment: baseline FEV_1 (l/sec) $\in [1.5; 2.5]$

2nd amendment: baseline FEV_1 (l/sec) $\in [1.5; 3.0]$

Introduction

Example 2:

long-term trial to investigate the time until relapse of cutaneous melanoma, amendment increased the inclusion limit for cholesterol

(Svolba & Bauer, 1999).

Introduction

Example 2:

long-term trial to investigate the time until relapse of cutaneous melanoma, amendment increased the inclusion limit for cholesterol

(Svolba & Bauer, 1999).

Here: <u>the patient populations before and after the amendment may differ.</u>

This difference in the populations is often ignored in the statistical analysis. The data are pooled:

→ bias, decreased power

Introduction

The “amendment should also cover any statistical consequences … and alterations to the planned statistical analysis” (Cleophas et al., 2006).

Introduction

> The "amendment should also cover any statistical consequences …
> and alterations to the planned statistical analysis" (Cleophas et al., 2006).

Division of the trial data according to the different phases,

a new phase is started after each amendment:

K amendments $\rightarrow$ $K + 1$ phases ($K \geq 1$).

- Weighted linear regression (Chow & Shao, 2005)
- Fisher's combination test (Lösch & Neuhäuser, 2008)

Fisher's combination test

seems to be a good strategy

(according to simulations, Lösch & Neuhäuser, 2008).

$K = 1$, 2 p-values, $\alpha = 0.05$:

Fisher's combination test significant if $p_1 \, p_2 \leq 0.0087 = c_\alpha$

Fisher's combination test

seems to be a good strategy

(according to simulations, Lösch & Neuhäuser, 2008).

$K = 1, 2$ p-values, $\alpha = 0.05$:

Fisher's combination test significant if $p_1 \, p_2 \leq 0.0087 = c_\alpha$

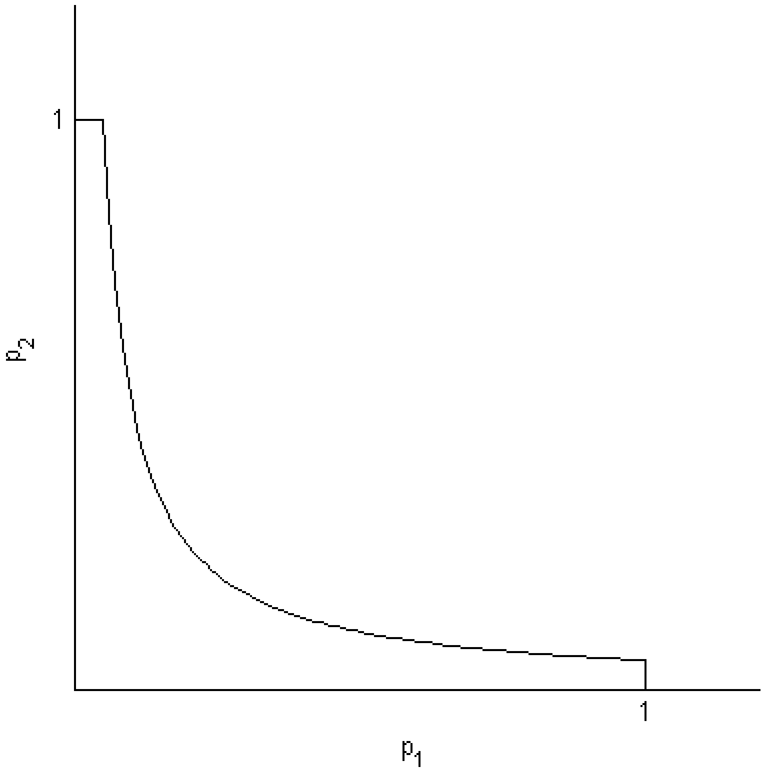
Bauer & Köhne's combination test

for clinical trials with an (adaptive) interim analysis (Bauer & Köhne, 1994)

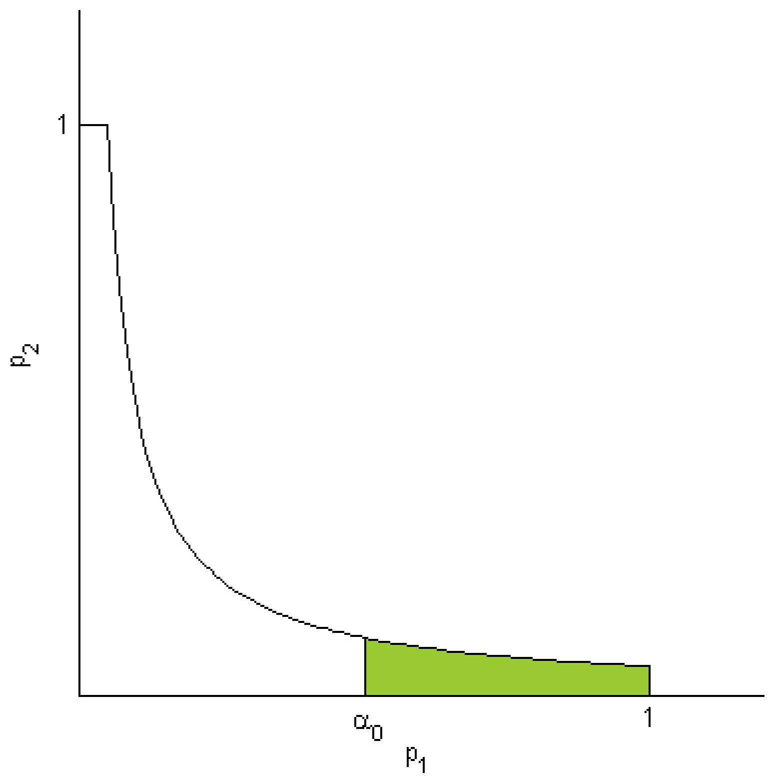
Trial is terminated due to insufficient effects if $p_1 \geq \alpha_0$ (e.g. $\alpha_0 = 0.5$).

Early stopping with the rejection of H_0 if $p_1 \leq \alpha_1$ (e.g. $\alpha_1 = 0.0233$ if $\alpha = 5\%$).

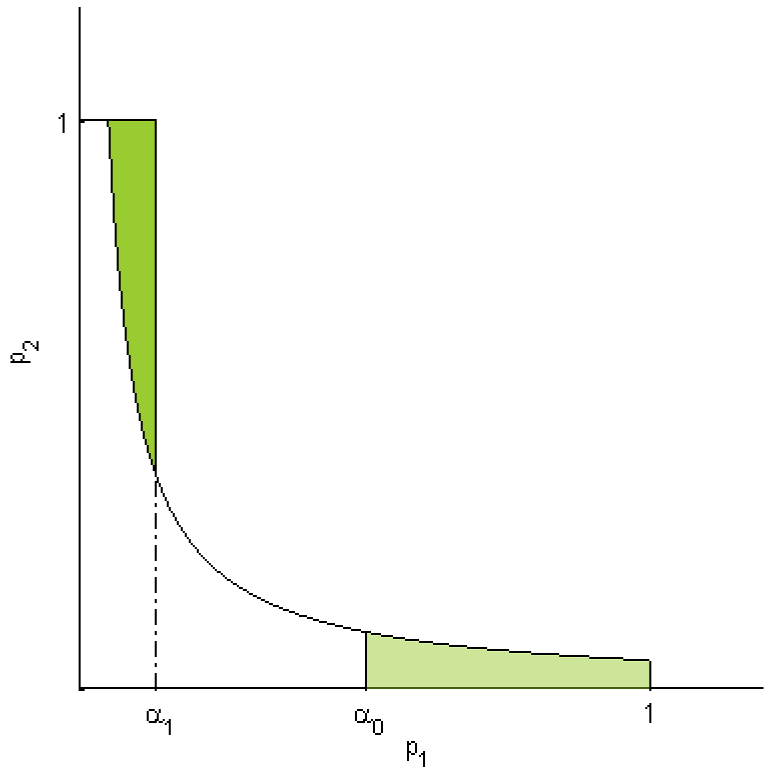
Bauer & Köhne's combination test



Bauer & Köhne's combination test



Bauer & Köhne's combination test



Bauer & Köhne's combination test

Diagram for a two-stage (adaptive) procedure

P-value of the 1st stage: p_1

I. $p_1 \leq \alpha_1$ → Stop and reject H_0

II. $\alpha_1 < p_1 < \alpha_0$ → go to 2nd stage

III. $p_1 \geq \alpha_0$ → stop without rejection of H_0

↓

P-value of the 2nd stage: p_2

II.1 $p_1\, p_2 \leq c_\alpha$ → reject H_0

II.2 $p_1\, p_2 > c_\alpha$ → no rejection of H_0

The (new) modified combination test

In contrast to a clinical study with an interim analysis, blinding can be maintained during the study: both phases are analysed at the end of the study.

→ no asymmetric decision rules:

α_0 and α_1 should be applied to both p_1 and p_2

The (new) modified combination test

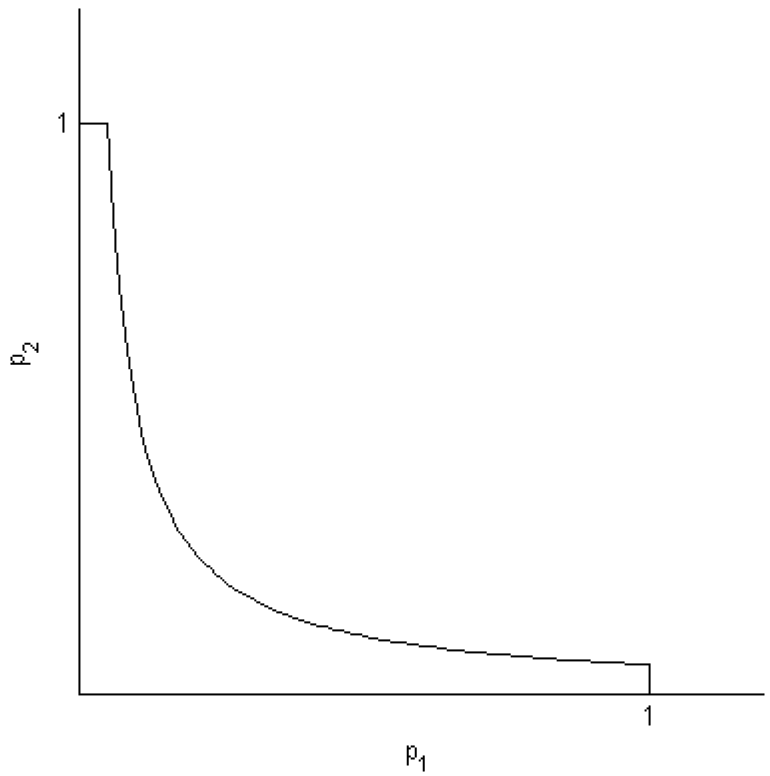
In contrast to a clinical study with an interim analysis, blinding can be maintained during the study: both phases are analysed at the end of the study.

→ no asymmetric decision rules:
α_0 and α_1 should be applied to both p_1 and p_2

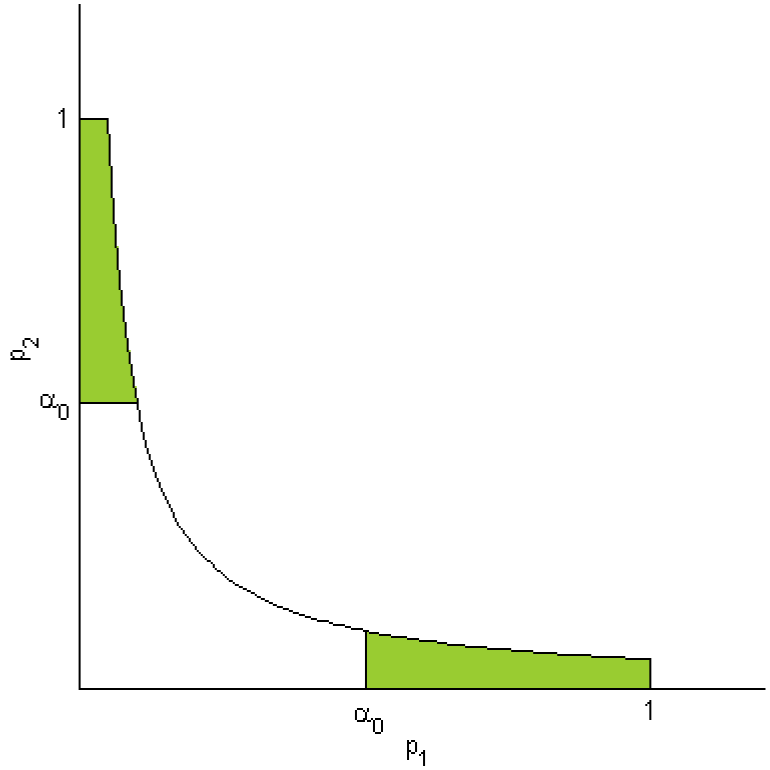
→ Modified combination test significant if
$\max(p_1, p_2) \leq \alpha_1$, or if
$\max(p_1, p_2) \leq \alpha_0$ and $p_1\, p_2 \leq c_\alpha$.

For $\alpha = 0.05$ and $\alpha_0 = 0.5$: $c_\alpha = 0.0087$ and $\alpha_1 = 0.1793$.

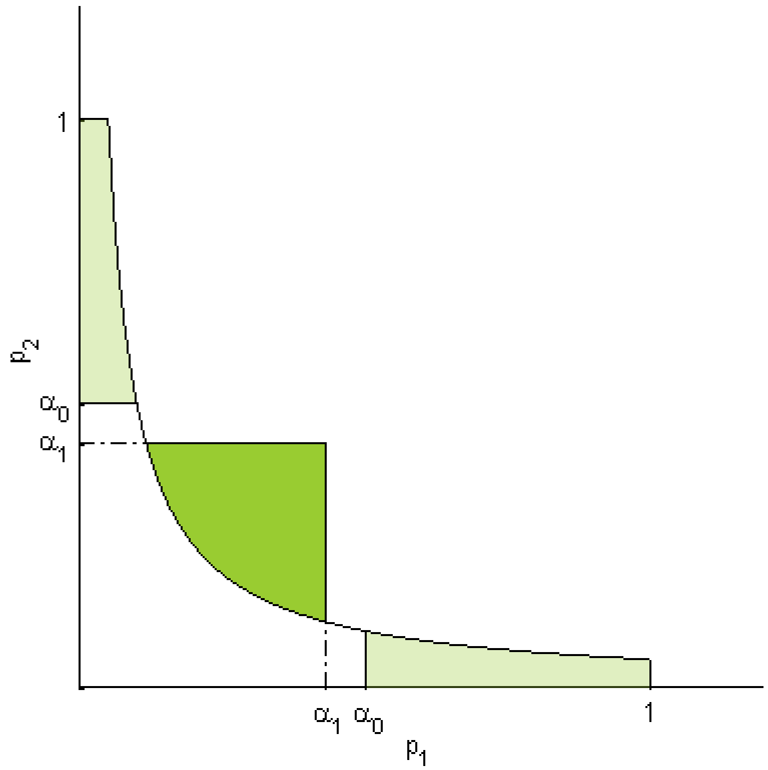
The (new) modified combination test



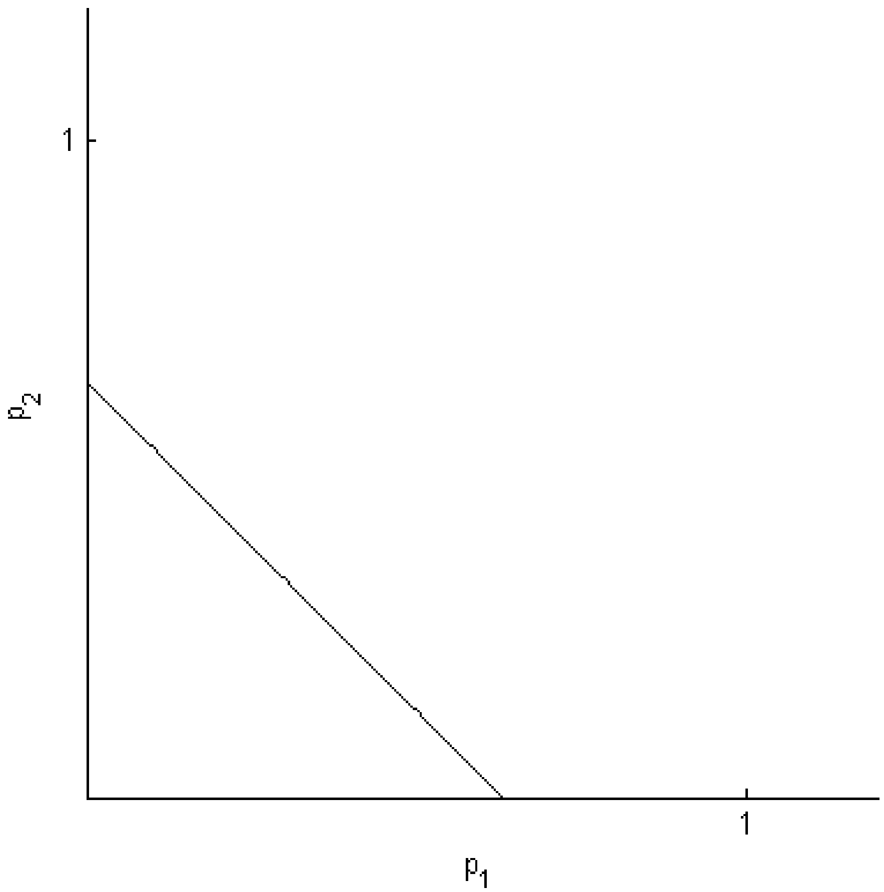
The (new) modified combination test



The (new) modified combination test



Edgington's (1972) combination test



Simulation Results

t-test (one-sided) for normally distributed data,

Standard deviation in phase 2 larger (group 1: 1, group 2: $\sqrt{1.5}$),

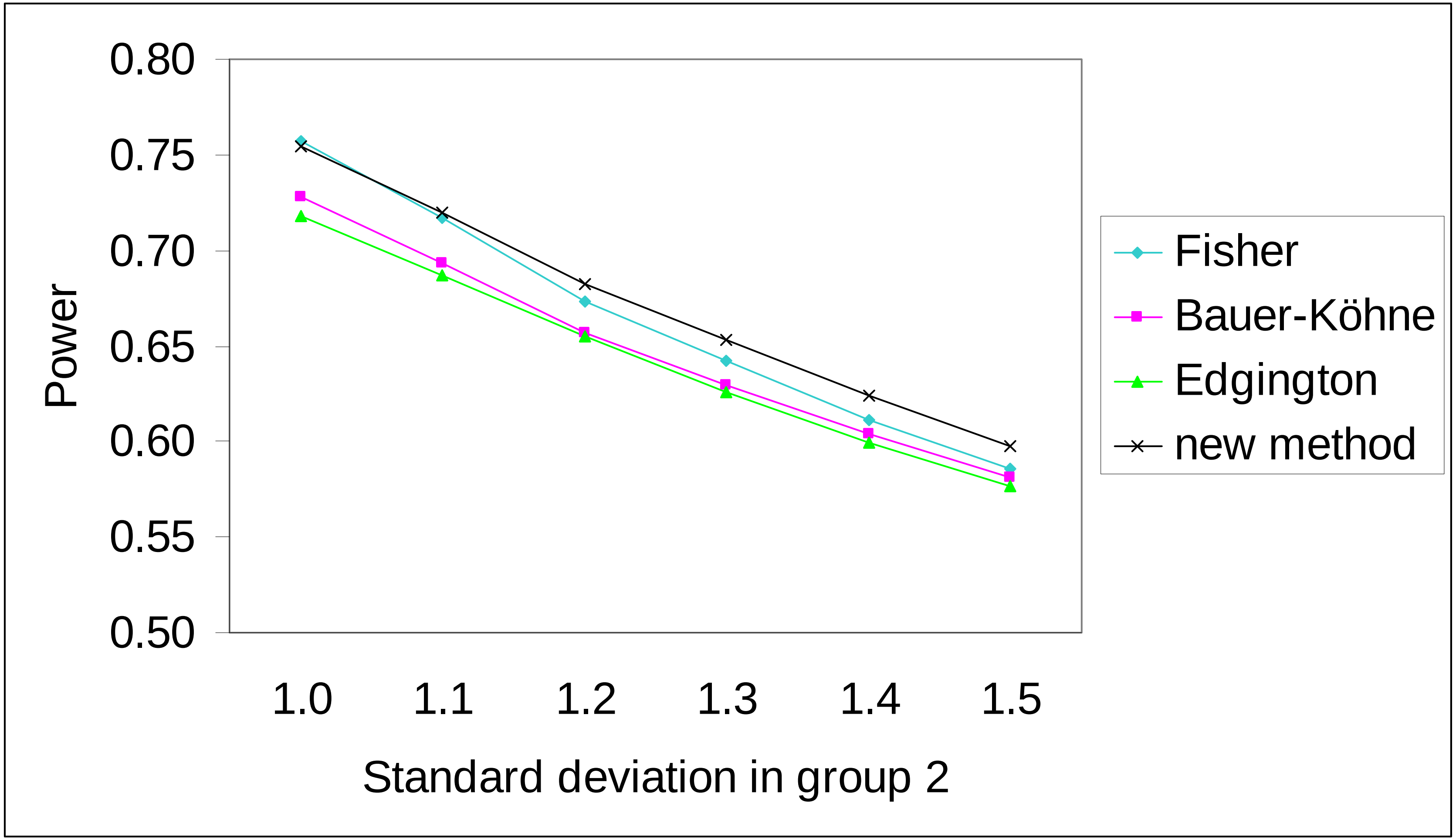
Sample size per group and phase: 50, $\alpha = 0.05$

Shift	Fisher	Bauer-Köhne	Edgington	New method
0	0.05	0.05	0.05	0.05
0.1	0.15	0.15	0.15	0.15
0.3	0.58	0.58	0.56	**0.59**
0.5	0.92	0.92	0.90	**0.93**
0.7	1.00	1.00	0.99	1.00

Simulation Results

t-test (one-sided) for normally distributed data, $\alpha = 0.05$, shift = 0.4

Sample size per group in phase 1: 25, in phase 2: 50



The (new) modified combination test

The proposed modified combination test can also be useful

- in trials with an (adaptive) interim analysis when a stop after the first phase with rejection of the null hypothesis is not desired. A stop for futility is still possible in case of $p_1 > \alpha_0$.

The (new) modified combination test

The proposed modified combination test can also be useful

- in trials with an (adaptive) interim analysis when a stop after the first phase with rejection of the null hypothesis is not desired. A stop for futility is still possible in case of $p_1 > \alpha_0$.
- for the analysis of multicentre studies (Neuhäuser & Senske, 2009).

The (new) modified combination test

Two-treatment multicenter trial, nonparametric analysis:

→ Rank-sum test for grouped data (**van Elteren test**).

If there are no ties and no differences between centers with regard to the groups' sample sizes, the van Elteren test is equivalent to the inverse-normal combination test (using center-specific rank-sum tests).

The (new) modified combination test

Two-treatment multicenter trial, nonparametric analysis:

→ Rank-sum test for grouped data (**van Elteren test**).

If there are no ties and no differences between centers with regard to the groups' sample sizes, the van Elteren test is equivalent to the inverse-normal combination test (using center-specific rank-sum tests).

→ Other combination tests may be applied,
e.g. Fisher's combination test or the modification proposed here.

The (new) modified combination test

Neuhäuser & Senske (2009):

Fisher’s combination test is more powerful than van Elteren’s test when

- there are large differences between the centers’ *p*-values,
- some quantitative interaction between treatment and center, and/or
- heterogeneity in variability.

The (new) modified combination test

Neuhäuser & Senske (2009):

Fisher's combination test is more powerful than van Elteren's test when

- there are large differences between the centers' *p*-values,
- some quantitative interaction between treatment and center, and/or
- heterogeneity in variability.

→ The (new) modified combination test might be a good alternative.
Maybe, in a further modification, a few p-values larger than α_0
may be acceptable in case of a large number of centers.

References

Bauer P & Köhne K (1994): Evaluation of experiments with adaptive interim analyses. *Biometrics* 50, 1029-1041.

Cleophas TJ et al. (2006): *Statistics applied to clinical trials*. Springer, 3rd edition.

Chow SC & Shao J (2005): Inference for clinical trials with some protocol amendments. *Journal of Biopharmaceutical Statistics* 15, 659-666.

Edgington ES (1972): A normal curve method for combining probability values from independent experiments. *Journal of Psychology* 82, 85–89.

Lösch C & Neuhäuser M (2008): The statistical analysis of a clinical trial when a protocol amendment changed the inclusion criteria. *BMC Medical Research Methodology* 8, 16.

Neuhäuser M & Senske R (2009): The analysis of multicentre clinical trials when there is heterogeneity between centres. *Journal of Statistical Computation and Simulation* 79, 1381-1387.

Svolba G & Bauer P (1999): Statistical quality control in clinical trials. *Controlled Clinical Trials* 20, 519-530.