Robust Designs in Non-Inferiority Three Arm Clinical Trials with Presence of Heteroscedasticity

Holger Dette, Ruhr University Bochum,
Universitätsstraße 150, D-44780 Bochum, Germany, E-mail: holger.dette@rub.de
Matthias Trampisch, Ruhr University Bochum,
Universitätsstraße 150, D-44780 Bochum, Germany, E-mail: matthias.trampisch@rub.de
Ludwig A. Hothorn, Leibniz University Hannover,
Herrenhauser Str. 2, D-30419 Hannover, Germany, E-mail: hothorn@biostat.uni-hannover.de

January 31, 2008

Abstract

In this paper, we describe an adjusted method to facilitate a non-inferiority trial by a three-arm robust design. Because local optimal designs derived in [Hasler et al. (2007)] require knowledge about the ratios of the population variances and are not necessarily robust with respect to possible misspecifications, a maximin approach is adopted. This method requires only the specification of an interval for the variance ratios and yields robust and efficient designs. We demonstrate that a maximin optimal design only depends on the boundary points specified for the intervals of the variance ratios and receive numerical and analytical solutions which are demonstrated in several examples. The derived designs are robust and very efficient for statistical analysis in non inferiority three arm trials.

Keywords: *maximin design, robust design, non-inferiority, three arm design, gold design trials, randomized clinical trial*

1 Introduction

Nowadays, randomized clinical trials claiming at least non-inferiority are performed. The specific statistical methodology was recently described in [Munk et. al. (2005)]. A two-arm design where a new experimental

Treatment group	Mean	Standard deviation	Sample size
Placebo	16.5	7.5	14
ALM4+NO	26.5	10.4	14
ALM16+NO	36.7	13.2	14

Table 1: *Summary statistics for Pa_{O_2} (kPa) 30 minutes after onset of one-lung ventilation of the clinical data set of Silva-Costa-Gomes et al. (2005)*

drug (with mean μ_1) is compared with the reference drug or active control (with mean μ_2) is common. However, trials without a placebo arm require an indirect inference, e.g. by meta-analysis and may be problematic (see e.g. [Hung et. al. (2007)]). Therefore, so-called "Gold design trials" are recommended as three-arm designs, which include the new experimental drug (with mean μ_1), the reference drug or active control (with mean μ_2) and an additional placebo control (with mean μ_3). For these trials, non-inferiority can be formulated as a fraction of the trial sensitivity, see e.g. [Pigeot et al. (2003)] or [Hung et. al. (2005)]. The null hypothesis is based on the ratio of the differences of the means $H_0 : \frac{\mu_1-\mu_3}{\mu_2-\mu_3} \leq \theta$ and is compared with the alternative $H_1 : \frac{\mu_1-\mu_3}{\mu_2-\mu_3} > \theta$ for a given non-inferiority threshold $\theta \in (0,1)$. The alternative hypothesis indicates that the relative efficacy of the experimental drug is more than $\theta \cdot 100\%$ of the efficacy of the reference compound compared to placebo. For this ratio hypothesis, a t-distributed test statistic was derived, assuming normal distribution and variance homogeneity. However, in real data it is more realistic that heterogeneous variances occur.

For example in [Silva-Costa-Gomes et al. (2005)], a randomized clinical trial was conducted comparing low and high-doses of almitrine combined with nitric oxide with a placebo group in the prevention of hydroxia during open-chest one-lung ventilation. Table 1 shows the related summary statistics of these three treatment arms for the primary respiratory endpoint Pa_{O_2} (kPa) for an administration 30 minutes after onset of one-lung ventilation. The experimental drug ALM4+NO was compared with the reference ALM16+NO and the placebo for at least non-inferiority of the low dose versus the high dose relative to the difference between the high dose and the placebo effect. Notice that the data shows a markedly lower variance in the placebo group, i.e. the assumption of homoscedasticity is hard to imagine.

One possibility to address the problem of heteroscedasticity is a logarithmic transformation to stabilize the variances, but there are many cases where this procedure does not yield homoscedastic data. Assuming homogeneous variances, an optimal design can be achieved as in [Pigeot et al. (2003)], where the unbalancedness now depends only on the given threshold θ. Assuming heterogeneous - but "known" - variances, an optimal design can as well be calculated like [Hasler et al. (2007)], but the unbalancedness now depends on the given threshold θ and the variances of the three treatments. However, the availability of the exact variances is rather unlikely in practice and a misspecification of these variances can lead to an experimental design with a low efficiency. In order to derive designs which are robust against such misspecification - but still efficient for a broad range of the parameters - we propose a maximin approach. In particular, we describe an adjusted method to facilitate a non-inferiority trial by a three-arm robust design in the case of heterogeneous variances. Only interval estimates of variance ratios have to be available for the construction of an experimental design of a randomized clinical trial. We consider this situation as more realistic from a practical point of view, because usually information from preliminary clinical trials do not yield precise information for the variance ratios, but often allows the experimenter to derive lower and upper bounds for such ratios. We prove that such robust optimal designs only depend on the boundary points of the specified region for the variance ratios and receive numerical and analytical solutions. Moreover, it is demonstrated that the derived designs are very efficient over a broad range of specified variance ratios. Thus, the new designs provide an interesting alternative to the commonly used designs, which may be inefficient if the ratios of the population variances have been misspecified. A MatLab program serving the purpose of calculating the robust designs can be downloaded at [Maximin-Program (2007)].

2 Local Optimal Design

We consider a clinical trial with three groups that correspond to the experimental, reference and placebo arms with means μ_1, μ_2, μ_3, respectively. We focus on the previously introduced problem of finding a robust design for the non-inferiority hypothesis

$$H_0 : \frac{\mu_1 - \mu_3}{\mu_2 - \mu_3} \leq \theta \quad vs. \quad H_1 : \frac{\mu_1 - \mu_3}{\mu_2 - \mu_3} > \theta$$

with a fixed retention fraction of $\theta \in (0,1)$.

For the motivation of a criterion for the comparison of competing designs the following statistic

$$T = \frac{\bar{x}_1 - \theta \bar{x}_2 - (1-\theta)\bar{x}_3}{\sqrt{\frac{1}{n_1}\sigma_1^2 + \frac{\theta^2}{n_2}\sigma_2^2 + \frac{(1-\theta)}{n_3}\sigma_3^2}} \quad \sim \quad N\left(\frac{\mu_1 - \theta\mu_2 - (1-\theta)\mu_3}{\sqrt{\frac{1}{n_1}\sigma_1^2 + \frac{\theta^2}{n_2}\sigma_2^2 + \frac{(1-\theta)}{n_3}\sigma_3^2}}, 1\right) \tag{1}$$

is used where σ_i^2 denotes the (unknown) variance, n_i the sample size and $\bar{x}_i$ the arithmetic mean of each group $i = \{1,2,3\}$. Furthermore, the observations in the different groups are assumed to be normally distributed with mean μ_i and variances σ_i^2 $(i = 1,2,3)$. The formula (1) can be equivalently written as

$$T \sim N\left(\sqrt{n_1}\frac{(\mu_1 - \theta\mu_2 - (1-\theta)\mu_3)}{\sqrt{\sigma_1^2 + \frac{\theta^2}{w_2}\sigma_2^2 + \frac{(1-\theta)^2}{w_3}\sigma_3^2}}, 1\right)$$

with $w_2 = \frac{n_2}{n_1}$, $w_3 = \frac{n_3}{n_1}$ being ratios of the sample sizes. For a given significance level α and power level $1-\beta$ we derive the formula

$$\sqrt{n_1}\frac{(\mu_1 - \theta\mu_2 - (1-\theta)\mu_3)}{\sqrt{\sigma_1^2 + \frac{\theta^2}{w_2}\sigma_2^2 + \frac{(1-\theta)^2}{w_3}\sigma_3^2}} \quad = \quad z_{1-\alpha} + z_{1-\beta}\,,$$

where z_u for $u \in [0,1]$ denotes the u-quantile of a standard normal distribution. This leads to

$$\begin{aligned} n_1 &= \left(z_{1-\alpha} + z_{1-\beta}\right)^2 (\mu_1 - \theta\mu_2 - (1-\theta)\mu_3)^{-2} \left(\sigma_2^2 + \frac{\theta^2}{w_2}\sigma_3^2 + \frac{(1-\theta)^2}{w_3}\sigma_3^2\right) \\ &= \left(\frac{z_{1-\alpha} + z_{1-\beta}}{\mu_1 - \theta\mu_2 - (1-\theta)\mu_3}\right)^2 \cdot \sigma_2^2 \left(1 + \frac{\theta^2}{w_2} r_2 + \frac{(1-\theta)^2}{w_3} r_3\right) \end{aligned}$$

as sample size n_1 for group one, where $r_2 = \sigma_2^2/\sigma_1^2$ and $r_3 = \sigma_3^2/\sigma_1^2$ denote the (fixed) ratios of the variances σ_2^2 and σ_3^2 with reference to σ_1^2.

With the specified nominal level α and power $1-\beta$, the minimum total sample size n can be derived by minimizing the following function (2) with respect to w_2 and w_3

$$n = n_1(1 + w_2 + w_3) = \left(\frac{z_{1-\alpha} + z_{1-\beta}}{\mu_1 - \theta\mu_2 - (1-\theta)\mu_3}\right)^2 \cdot \sigma_2^2 \left(1 + \frac{\theta^2}{w_2} r_2 + \frac{(1-\theta)^2}{w_3} r_3\right)(1 + w_2 + w_3)\,. \qquad (2)$$

This means that one has to determine the ratio of the variances r_2 and r_3 for an optimal design of the experiment. Since the function

$$f(w_2, w_3 | r_2, r_3) = \left(1 + \frac{\theta^2}{w_2} r_2 + \frac{(1-\theta)^2}{w_3} r_3\right)(1 + w_2 + w_3) \qquad (3)$$

is the only factor on the right side of equation (2) that involves w_2 and w_3, the minimum sample size can be derived by simply minimizing this function. The optimal values for w_2 and w_3 are determined by solving the system of equations

$$0 = \frac{\delta}{\delta w_2}\left(1 + \frac{\theta^2}{w_2} r_2 + \frac{(1-\theta)^2}{w_3} r_3\right)(1 + w_2 + w_3)$$

$$0 = \frac{\delta}{\delta w_3}\left(1 + \frac{\theta^2}{w_2} r_2 + \frac{(1-\theta)^2}{w_3} r_3\right)(1 + w_2 + w_3)\,.$$

For $0 < \theta < 1$, the unique - and surprisingly simple - solution is given by

$$w_2 = \theta\sqrt{r_2} \quad (4)$$

$$w_3 = (1-\theta)\sqrt{r_3}, \quad (5)$$

which leads to the optimal sample sizes

$$\begin{aligned} n_1 &= \left(\frac{z_{1-\alpha}+z_{1-\beta}}{\mu_1-\theta\mu_2-(1-\theta)\mu_3}\right)^2 \cdot \sigma_2^2\,(1+\theta\sqrt{r_2}+(1-\theta)\sqrt{r_3}) \\ n_2 &= \theta\cdot\sqrt{r_2}\cdot n_1 \\ n_3 &= (1-\theta)\cdot\sqrt{r_3}\cdot n_1 \\ n &= n_1\cdot(1+\theta\sqrt{r_2}+(1-\theta)\sqrt{r_3})\,. \end{aligned}$$

For the calculation of the optimal group size allocation for a fixed sample size n, we introduce the following two parameters

$$p_2 = \frac{w_2}{1+w_2+w_3} \qquad p_3 = \frac{w_3}{1+w_2+w_3} \quad (6)$$

which represent the proportion of observations allocated to group two and three with respect to the total sample size. Following [Chernoff (1953)] the resulting design is called local optimal, because it depends on the (unknown) variance ratios r_2 and r_3. Thus, the local optimal design advises the experimenter to take $n_1 = (1-p_2-p_3)\cdot n$, $n_2 = p_2\cdot n$ and $n_3 = p_3\cdot n$ observations at group one, two and three, respectively. These results coincide with the recent findings in the article of [Hasler et al. (2007)], if one substitutes $i \in \{1,2,3\}$ with $i \in \{E,R,P\}$.

Note that the optimal sample sizes depend on the unknown variance ratios r_2 and r_3, which are usually not available before the experiment. In particular, a misspecification of these ratios may result in errors of the optimal allocation of the treatments thus making that specific trial less efficient. In the following section, we will propose a robust design, which is less sensitive with respect to misspecified variance ratios and very efficient for the three-arm clinical trial. This and additional properties are illustrated in section 4.

3 Robust Design With A Maximin Approach

A more realistic approach to the problem considered in section 2 is that the ratios of the variances are not exactly known, but interval estimates are available based on previous (similar) trials. This means that we have access to information of the form $\frac{\sigma_2^2}{\sigma_1^2} \in V^2 := \left[V_L^2, V_U^2\right]$ and $\frac{\sigma_3^2}{\sigma_1^2} \in V^3 := \left[V_L^3, V_U^3\right]$, where $V_L^2, V_U^2, V_L^3, V_U^3$ are the boundary points of the postulated intervals for the variance ratios r_2 and r_3 with respect to $\sigma_2^2 \in \mathbb{R}^+$. As an alternative to a rectangular region for the variance ratios elliptical or circular regions could be considered as well, but for the sake of brevity we restrict ourselves to the rectangle. We want to minimize the required total population sample size n to achieve a given power $1-\beta$. For this purpose we use the rate function (3). We mentioned that - if the ratios of the variances are fixed and known - this function has exactly one minimum (see the previous section or [Hasler et al. (2007)]). Nevertheless, this local optimal design might not be a good choice if the ratios of the variances have been misspecified. In order to derive designs which are less sensitive with respect to such misspecifications, we consider the efficiency

$$eff\,(w_2, w_3, r_2, r_3) \quad = \quad \frac{f\,(\nu, \omega | r_2, r_3)}{f\,(w_2, w_3 | r_2, r_3)} \in [0,1]\,, \tag{7}$$

with $f\,(\nu, \omega | r_2, r_3) := \min_{w_2, w_3} f\,(w_2, w_3 | r_2, r_3)$. Equation (7) measures the performance of an arbitrary design $w = (w_2, w_3)$ (in the denominator) with respect to the best design (in the numerator) calculated under the assumption that r_2 and r_3 are the "true" ratios of the population variances. Following [Dette (1997)] a design $w^* = (w_2^*, w_3^*)$ is called standardized maximin optimal or briefly maximin optimal design if it maximizes the minimum efficiency

$$g\,(w_2, w_3) \quad = \quad \min_{r_2 \in V^2, r_3 \in V^3} eff\,(w_2, w_3, r_2, r_3) \tag{8}$$

over the rectangle $V^2 \times V^3$.

With our knowledge from the previous section it follows that for fixed variance ratios r_2 and r_3 the minimum of the function $f\,(\nu, \omega | r_2, r_3)$ is attained at the point $\nu = \theta\sqrt{r_2}$ and $\omega = (1-\theta)\sqrt{r_3}$ and thus formula (7) can be simplified to

$$eff(w_2, w_3, r_2, r_3) = \frac{f\,(\nu, \omega | r_2, r_3)}{f\,(w_2, w_3 | r_2, r_3)} = \frac{f\left(\theta\sqrt{r_2},\, (1-\theta)\sqrt{r_3} | r_2, r_3\right)}{f\,(w_2, w_3 | r_2, r_3)}\,, \tag{9}$$

where

$$f\,(\theta\sqrt{r_2},\, (1-\theta)\sqrt{r_3} | r_2, r_3) = \left(1 + \theta\sqrt{r_2} + (1-\theta)\sqrt{r_3}\right)^2. \tag{10}$$

This simplifies the analysis of formula (8) substantially since now

$$g(w_2,w_3) = \min_{r_2\in V^2, r_3\in V^3} \frac{\left(1+\theta\sqrt{r_2}+(1-\theta)\sqrt{r_3}\right)^2}{f(w_2,w_3|r_2,r_3)} \quad (11)$$

$$= \min_{r_2\in V^2, r_3\in V^3} \frac{\left(1+\theta\sqrt{r_2}+(1-\theta)\sqrt{r_3}\right)^2}{\left(1+\frac{\theta^2}{w_2}r_2+\frac{(1-\theta)^2}{w_3}r_3\right)(1+w_2+w_3)} \quad (12)$$

The following Lemma states that the minimum on the right hand side of (11) with respect to $(r_2,r_3)\in V^2\times V^3$ may only be attained at the corners of the rectangle $V^2\times V^3$. The proof can be found in the appendix.

Lemma

The minimum of the function eff defined by (9) with respect to $(r_2,r_3)\in V^2\times V^3$ may only be attained at the corners of the rectangle $V_2\times V_3$, that is

$$g(w_2,w_3) = \min\{eff\left(w_2,w_3,V_L^2,V_L^3\right), eff\left(w_2,w_3,V_U^2,V_L^3\right), eff\left(w_2,w_3,V_L^2,V_U^3\right), eff\left(w_2,w_3,V_U^2,V_U^3\right)\} \quad (13)$$

With this Lemma, one only has to numerically maximize the function (8) at the four corners of $V_2\times V_3$ rather than the whole rectangle area. The resulting robust design is

$$\arg\max_{w_2,w_3} g(w_2,w_3) = (w_2^*, w_3^*) = w^*. \quad (14)$$

The actual value has to be calculated numerically using e.g. [Maximin-Program (2007)].

Note that such numerical optimization may yield local maxima and it is not clear that a numerically found maximum corresponds to the global maximum, i.e. the standardized maximin optimal design. In the following, we state a necessary and sufficient checking condition for the standardized maximin optimal design. For a more detailed discussion the reader is referred to e.g. [Pukelsheim (1993)] or [Müller (1995)]. The following Theorem can be used to check the optimality of the numerically calculated design. For this purpose we introduce the following notation

$$c_\theta^T = (1,\theta,(1-\theta)), \qquad \theta\in(0,1),$$

and the set

$$V = V^2\times V^3.$$

For fixed variance ratios $v=(r_2,\, r_3)\in V$ and arbitrary group ratios $w=(w_2,w_3)$ we define

$$M(w,v) := \frac{1}{1+w_2+w_3} diag\left(\sigma_1^2, \frac{w_2}{\sigma_2^2}, \frac{w_3}{\sigma_3^2}\right) = \frac{1}{\sigma_1^2\cdot(1+w_2+w_3)}\cdot diag\left(1, \frac{w_2}{r_2}, \frac{w_3}{r_3}\right)$$

The optimality criterion in (8) can be rewritten as

$$g(w) = \min_{r\in V} eff(w,r) = \min_{r\in V} \frac{c_\theta^T M^{-1}(w_r^*, r)\, c_\theta}{c_\theta^T M^{-1}(w,r)\, c_\theta}, \tag{15}$$

where w_r^* denotes the local optimal design assuming known ratios of the variances r_2 and r_3, that is $w_r^* = \left(\theta\sqrt{r_2}, (1-\theta)\sqrt{r_3}\right)$ (see the discussion in the previous section). The following characterization of the standardized maximin optimal design is a consequence of Theorem 2 in [Biedermann et al. (2006)].

Theorem

Let

$$N(w) = \left\{\tilde{r}\in V \,|\, eff(w,\tilde{r}) = \min_{r\in V} eff(w,\, r)\right\}$$

be the subset of V consisting of those values of b, for which the efficiency (15) of a design w takes its minimal value over V. A design w_M^* is standardized maximin optimal if and only if for each $v\in N(w_M^*)$ there exists a nonnegative weight $\pi^*(v)$ such that the following equations are valid

$$\sum_{v\in N\left(w_M^*\right)} \pi^*(v)\cdot\frac{\left(c_\theta^T M^{-1}(w_M^*, v)\, x_i\right)^2}{c_\theta^T M^{-1}(w_M^*, v)\, c_\theta} = 1, \qquad i = 1,2,3, \tag{16}$$

where

$$x_1 = \begin{pmatrix} 1\\ 0\\ 0\end{pmatrix}, x_2 = \begin{pmatrix} 0\\ 1\\ 0\end{pmatrix}, x_3 = \begin{pmatrix} 0\\ 0\\ 1\end{pmatrix}$$

and

$$\sum_{v\in N(w_M^*)} \pi^*(v) = 1$$

By our Lemma derived in this section, the set $N(w)$ for any design w consists of at most the four corners of the rectangle V, namely

$$v_1 = (V_2^L, V_3^L),\, v_2 = (V_2^U, V_3^L),\, v_3 = (V_2^L, V_3^U),\, v_4 = (V_2^U, V_3^U)\,.$$

This means that all other points $v \in V$ have higher efficiencies.

The Theorem leads to the following three equations for $i = 1,2,3$:

$$\sum_{j=1}^{4} \pi(v_j) \cdot \frac{\left(c_\theta^T M^{-1}(w_M^*, v_j)\, x_i\right)^2}{c_\theta^T M^{-1}(w_M^*, v_j)\, c_\theta} = 1$$

These equations contain the unknown parameters $\pi(v_1), \pi(v_2), \pi(v_3), w_2$ and w_3 since $\pi(v_4) = 1 - \pi(v_1) - \pi(v_2) - \pi(v_3)$. Note that some of the probabilities $\pi(v_i)$ may be zero because the corresponding corner v_i is not an element of the set $N(w_M^*)$.

We use the following notation to keep these equations more readable

$$a_{2L} = \theta\sqrt{V_L^2},\, a_{2U} = \theta\sqrt{V_U^2},\, a_{3L} = (1-\theta)\sqrt{V_L^3},\, a_{3U} = (1-\theta)\sqrt{V_U^3}\,,$$

and obtain the following system of nonlinear equations

$$\pi(v_1)\cdot\frac{(1+w_2+w_3)}{1+\frac{a_{2L}^2}{w_2}+\frac{a_{3L}^2}{w_3}} + \pi(v_2)\cdot\frac{(1+w_2+w_3)}{1+\frac{a_{2U}^2}{w_2}+\frac{a_{3L}^2}{w_3}} + \pi(v_3)\cdot\frac{(1+w_2+w_3)}{1+\frac{a_{2L}^2}{w_2}+\frac{a_{3U}^2}{w_3}} + \pi(v_4)\cdot\frac{(1+w_2+w_3)}{1+\frac{a_{2U}^2}{w_2}+\frac{a_{3U}^2}{w_3}} = 1 \tag{17}$$

$$\pi(v_1)\cdot\frac{(1+w_2+w_3)\left(\frac{a_{2L}}{w_2}\right)^2}{1+\frac{a_{2L}^2}{w_2}+\frac{a_{3L}^2}{w_3}} + \pi(v_2)\cdot\frac{(1+w_2+w_3)\left(\frac{a_{2U}}{w_2}\right)^2}{1+\frac{a_{2U}^2}{w_2}+\frac{a_{3L}^2}{w_3}} + \pi(v_3)\cdot\frac{(1+w_2+w_3)\left(\frac{a_{2L}}{w_2}\right)^2}{1+\frac{a_{2L}^2}{w_2}+\frac{a_{3U}^2}{w_3}} + \pi(v_4)\cdot\frac{(1+w_2+w_3)\left(\frac{a_{2U}}{w_2}\right)^2}{1+\frac{a_{2U}^2}{w_2}+\frac{a_{3U}^2}{w_3}} = 1$$

$$\pi(v_1)\cdot\frac{(1+w_2+w_3)\left(\frac{a_{3L}}{w_3}\right)^2}{1+\frac{a_{2L}^2}{w_2}+\frac{a_{3L}^2}{w_3}} + \pi(v_2)\cdot\frac{(1+w_2+w_3)\left(\frac{a_{3L}}{w_3}\right)^2}{1+\frac{a_{2U}^2}{w_2}+\frac{a_{3L}^2}{w_3}} + \pi(v_3)\cdot\frac{(1+w_2+w_3)\left(\frac{a_{3U}}{w_3}\right)^2}{1+\frac{a_{2L}^2}{w_2}+\frac{a_{3U}^2}{w_3}} + \pi(v_4)\cdot\frac{(1+w_2+w_3)\left(\frac{a_{3U}}{w_3}\right)^2}{1+\frac{a_{2U}^2}{w_2}+\frac{a_{3U}^2}{w_3}} = 1$$

These equations allow us to check whether a given design (w_2, w_3) is standardized maximin optimal or not. To find such an optimal design, one first solves the maximizing problem (14), evaluates the efficiencies at the corners of the rectangle V and then picks the point(s) where the minimum efficiency is attained (the weights of the remaining points are set to zero). Now one numerically evaluates the remaining weights $\pi(v_j)$ using the system of equations (17). If there exists a valid solution, one can be assured that a standardized maximin optimal design has been found. All of these calculations can easily be done using e.g. MatLab [The-MathWorks (1984)] and/or Mathematica [Wolfram-Research (1988)]. Numerical evaluations show that for the standardized maximin optimal design w_M^* the set $N(w_M^*)$ usually contains only two or three points. Several examples of the described procedure can be found in the following section.

4 Further discussion and examples

4.1 Verifying the optimality of a given design

We begin with an example illustrating the use of the checking condition. For this purpose let us assume that the variance ratios are located in the intervals $V^2 = [0.16, 0.64]$ and $V^3 = [0.49, 3.24]$, and that the non-inferiority parameter is given by $\theta = 0.5$. We first convert these parameters to the previously used terms in the system of equations defined by (17)

$$a_{2L} = 0.5 \cdot \sqrt{0.16} = 0.2 \qquad a_{3L} = 0.5 \cdot \sqrt{0.49} = 0.35$$
$$a_{2U} = 0.5 \cdot \sqrt{0.64} = 0.4 \qquad a_{3U} = 0.5 \cdot \sqrt{3.24} = 0.9$$

In the next step we numerically maximize the minimal efficiency at the corners of the rectangle $V = V^2 \times V^3$ in terms of w_2 and w_3:

$$\begin{aligned} argmax_{w_2,w_3} \quad min \quad \{ & eff\,(w_2, w_3, a_{2L}, a_{3L})\,, eff\,(w_2, w_3, a_{2U}, a_{3L})\,, \\ & eff\,(w_2, w_3, a_{2L}, a_{3U})\,, eff\,(w_2, w_3, a_{2U}, a_{3U})\} \end{aligned} \tag{18}$$

where the efficiency function in (7) is now defined for the new parameters $a_2 = \theta\sqrt{r_2},\ a_3 = (1-\theta)\sqrt{r_3}$, that is

$$eff\,(w_2, w_3, a_2, a_3) \quad = \quad \frac{(1 + a_2 + a_3)^2}{\left(1 + \frac{a_2^2}{w_2} + \frac{a_3^2}{w_3}\right)(1 + w_2 + w_3)}$$

In our considered case the numerical solution of the optimization problem (18) is $w^* = (0.3818, 0.6249)$ yielding a minimal efficiency of at least 93.26% over the rectangle $V = [0.16, 0.64] \times [0.49, 3.24]$. To check whether the numerically calculated design is optimal or not, we calculate the efficiencies

$$eff\,(w_2^*, w_3^*, a_{2L}, a_{3L}) = 0.9326 \qquad eff\,(w_2^*, w_3^*, a_{2U}, a_{3L}) = 0.9326$$

$$eff\,(w_2^*, w_3^*, a_{2L}, a_{3U}) = 0.9326 \qquad eff\,(w_2^*, w_3^*, a_{2U}, a_{3U}) = 0.9730$$

in order to apply the Theorem of Section 3. Because the efficiency at the point $v_4 = (a_{2U}, a_{3U})$ is greater than the efficiencies at the other three points, we set the weight $\pi(v_4)$ equal zero. Thus, we have to numerically find the weights $\pi(v_1)$ and $\pi(v_2)$ (since $\pi(v_3) = 1 - \pi(v_1) - \pi(v_2)$) to fulfill the three equations in (17). Using MatLab, Mathematica or any other adequate program, we calculate the weights to be $\pi(v_1) = 0.4603$ for the point $v_1 = (a_{2L}, a_{3L})$, $\pi(v_2) = 0.5072$ for the point $v_2 = (a_{2U}, a_{3L})$ and the remaining mass to be $\pi(v_3) = 0.0325$ at the point $v_3 = (a_{2U}, a_{3L})$.

With this weight distribution we validated that the solution w^* is indeed the optimal solution. Using the conversion (6), the optimal allocation w^* means that we have to take about $p_2^* = 17\%$ of our observations at the reference arm, about $p_3^* = 32\%$ of our observations at the placebo arm, and the remaining 51% of our observations at the experimental arm.

4.2 Some optimal designs for the Pa$_{O_2}$ example

In the second example we illustrate how the new methodology can be used to derive a robust and efficient design for a similar clinical trial as considered in the introduction. Assume that we have to design a new randomized clinical trial with a new experimental drug and that we expect similar results as presented in Table 1. In the context of this paper, the experimental drug is ALM4+NO, the reference drug is ALM16+NO and, of course, the placebo takes the part of the placebo. Since the variance ratios in this example are $r_2 = 1.61$ and $r_3 = 0.52$, we assume that the real variance ratios are located within the intervals $V^2 = [1.0, 2.0]$ and $V^3 = [0.40, 0.60]$. If the non-inferiority parameter is given by $\theta = 0.8$, numerical calculations similar to Example 4.1 yield the optimal weight distribution to be $w^* = (0.9566, 0.1434)$. For fixed sample size n and using conversion (6), the standardized maximin design allocates approximately $n_1 = 0.4762 \cdot n$, $n_2 = 0.4555 \cdot n$ and $n_3 = 0.0683 \cdot n$ to the three groups. The efficiency of this design over the rectangle $[1.0, 2.0] \times [0.4, 0.6]$ is at least 0.9910. The reason for the surprisingly small sample size of the placebo group originates from its variance ratio and the nature of how a three arm clinical trial depends on the parameter θ (compare (1)). If the total sample size is 100, then this design advises the experimenter to prescribe about 46 persons the experimental drug, 47 persons the standard treatment, and the remaining 7 persons to placebo treatment.

Further maximin optimal designs are shown in Table 2. Here $V^2 = [V_2^L, V_2^U]$ is the specified interval for the variance ratio $r_2 = \sigma_2^2/\sigma_1^2$, $V^3 = [V_3^L, V_3^U]$ is the interval of the variance ratio $r_3 = \sigma_3^2/\sigma_1^2$, p^* is the optimal allocation of the reference (p_2^*) and placebo arm (p_3^*), and the column labeled with eff shows the minimal (worst case) efficiency. Rather than listing the values of w^* we list the values of p^* because they are easier to interpret: for a sample of size n this means to take $p_2^* \cdot n$ observations at the reference, $p_3^* \cdot n$ observations

θ	V^2	V^3	$p^* = (p_2^*, p_3^*)$	eff
0.6	[0.4, 0.5]	[3, 4]	(0.1875, 0.3474)	0.9978
0.6	[3, 4]	[0.4, 0.5]	(0.4685, 0.1127)	0.9980
0.6	[0.8, 1.2]	[0.4, 0.5]	(0.3197, 0.1443)	0.9969
0.6	[0.8, 1.2]	[0.4, 1.7]	(0.3057, 0.1938)	0.9753

θ	V^2	V^3	$p^* = (p_2^*, p_3^*)$	eff
0.7	[0.4, 0.5]	[3, 4]	(0.2315, 0.2760)	0.9979
0.7	[3, 4]	[0.4, 0.5]	(0.5205, 0.0805)	0.9982
0.7	[0.8, 1.2]	[0.4, 0.5]	(0.3664, 0.1065)	0.9969
0.7	[0.8, 1.2]	[0.4, 1.7]	(0.3544, 0.1464)	0.9795

θ	V^2	V^3	$p^* = (p_2^*, p_3^*)$	eff
0.8	[0.4, 0.5]	[3, 4]	(0.2809, 0.1957)	0.9981
0.8	[3, 4]	[0.4, 0.5]	(0.5677, 0.0513)	0.9984
0.8	[0.8, 1.2]	[0.4, 0.5]	(0.4116, 0.0699)	0.9970
0.8	[0.8, 1.2]	[0.4, 1.7]	(0.4031, 0.0981)	0.9846

θ	V^2	V^3	$p^* = (p_2^*, p_3^*)$	eff
0.9	[0.4, 0.5]	[3, 4]	(0.3369, 0.1046)	0.9985
0.9	[3, 4]	[0.4, 0.5]	(0.6108, 0.0246)	0.9986
0.9	[0.8, 1.2]	[0.4, 0.5]	(0.4552, 0.0344)	0.9972
0.9	[0.8, 1.2]	[0.4, 1.7]	(0.4517, 0.0501)	0.9906

Table 2: *Optimal group size and minimal efficiency for different non-inferiority parameters θ and variance ratios V^2 and V^3*

at the placebo arm, and the remaining observations at the experimental arm. The MatLab program used to derive the optimal designs may be attained at [Maximin-Program (2007)]. It is worthwhile to mention that the efficiency values in Table 2 represent the minimal efficiency value over the rectangle $V^2 \times V^3$ and are always very high. These results indicate that the derived results are rather robust and efficient. If one chooses the optimal allocation p^* of the standardized maximin optimal design, one can be assured that the design is close to being "perfect" for the considered range of variance ratios.

4.3 Robustness of optimal designs

In this section we investigate the efficiencies of various designs if the parameters have been misspecified. Again, we will use the values of Table 1 and non-inferiority parameters $\theta = 0.8$ and $\theta = 0.6$. Note that the observed variance ratios in Table 1 are given by $r_2 = \frac{\sigma_3^2}{\sigma_2^2} = 1.61$ and $r_3 = \frac{\sigma_3^2}{\sigma_2^2} = 0.52$.

In our first example we study the efficiency of the local optimal design and the minimax design if the initial parameters have been misspecified. If the non-inferiority parameter is given by $\theta = 0.6$, the local optimal design for the point $b = (1.61, 0.52)$ is derived by formula (4) and has sample size distribution $w = (0.76, 0.29)$. As a typical example for a robust design we consider the standardized maximin optimal design for the rectangle $[0.64, 4.03] \times [0.21, 1.3]$, which yields the optimal weight $w^* = (0.84, 0.36)$. In Figure 1 we display the level curves of the standardized maximin optimal design (left part) and the local optimal design (right part) for varying variance parameters. A very important property is how the designs efficiencies are competing if the initial variance parameters have been misspecified. For example, if the "true" ratios of the variances would be $r_2 = 4$ and $r_3 = 3$ the efficiency of the local optimal design would be 89% while the standardized maximin optimal design yields an efficiency of 94 %. On the other hand, if the "true" variance ratios would be exactly $r_2 = 1.61$ and $r_3 = 0.52$ the local optimal design had efficiency 100 %, while the minimax optimal design yields 99.5 % efficiency. However, whenever the variance ratios are incorrectly specified, a large efficiency is obtained by the standardized maximin optimal design.

In our second example we consider the the case $\theta = 0.8$ and compare the standardized maximin optimal design with the more heuristic allocation rule $w = (1.61, 0.52)$, which yields a relative sample size apportionment corresponding directly to the variance ratios. For the maximin approach we will use the intervals $V^2 = [0.81, 3.22]$ and $V^2 = [0.26, 1.04]$ resulting in $w^* = (1.03, 0.15)$. The corresponding level curves in Figure 2 show the values of formula (9) for varying parameters r_2 and r_3, where the left part of Figure 2 corresponds to minimax design and the right part to the heuristic allocation rule. The standardized maximin optimal design clearly outperforms the heuristic design in a very broad area around the point $b = (1.61, 0.52)$.

Summarizing these and similar numerical studies, which are not shown for the sake of brevity, we obtain the following picture: there is no evident loss in efficiency in the application of the standardized maximin optimal design, even if precise knowledge of the variance ratios is available. The standardized maximin optimal design offers about the same high efficiencies in the considered variance intervals as the local optimal designs. On the other hand – whenever the ratios of the variances have been moderately misspecified – the standardized maximin optimal design is more efficient than the local optimal design.

4.4 Sample size calculation

In this section we compare the effects of the new standardized maximin and the local optimal designs with respect to the sample size required to achieve a specific power. [Hasler et al. (2007)] recommended to use the Welch type test [Welch (1938)] to address for possible heteroscedasticity in the data. For this test the necessary sample sizes to keep a preassigned level α and power $1 - \beta$ can be derived from formula (15) in

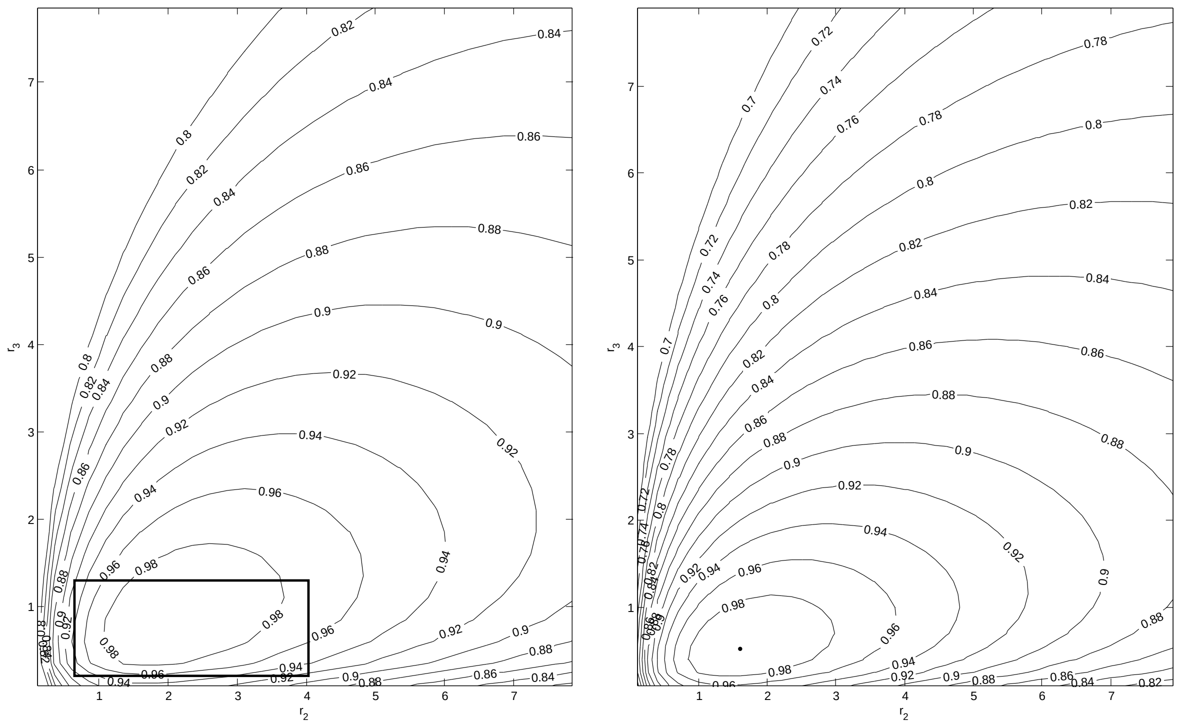

Figure 1: *Level curves of the efficiencies (9) for the standardized maximin optimal design for the intervals* $V^2 = [0.64, 4.03]$ *and* $V^3 = [0.21, 1.3]$ *(marked by the rectangle) and the local optimal design for* $b = (1.61, 0.52)$.

[Hasler et al. (2007)], and is (in our notation) given by

$$n_1 \geq \left(t_{1-\alpha}(\vartheta) + t_{1-\beta}(\vartheta)\right)^2 \frac{\sigma_1^2 + \frac{\theta^2}{w_2}\sigma_2^2 + \frac{(1-\theta)^2}{w_3}\sigma_3^2}{(\mu_1 - \theta\mu_2 - (1-\theta)\mu_3)^2},$$

where $t_u(\vartheta)$ for $u \in [0,1]$ denotes the u-quantile of a t_ϑ-distribution with

$$\vartheta = \frac{\left(\frac{1}{n_1}\sigma_1^2 + \frac{\theta^2}{w_2 n_1}\sigma_2^2 + \frac{(1-\theta)^2}{w_3 n_1}\sigma_3^2\right)^2}{\frac{1}{n_1^2(n_1-1)}\sigma_1^4 + \frac{\theta^4}{(w_2 n_1)(w_2 n_1 - 1)}\sigma_2^4 + \frac{(1-\theta)^4}{(w_3 n_1)(w_3 n_1 - 1)}\sigma_3^4}$$

degrees of freedom.

It was pointed out in the previous subsection that the standardized maximin optimal design yields better efficiencies than the local optimal design if the variance ratios have been misspecified. On the other hand, the standardized maximin optimal design offers about the same high efficiencies in the considered variance intervals as the local optimal design, if the variance ratios have been correctly specified. In the following discussion we demonstrate that these differences are also reflected in the necessary sample size to achieve a given power.

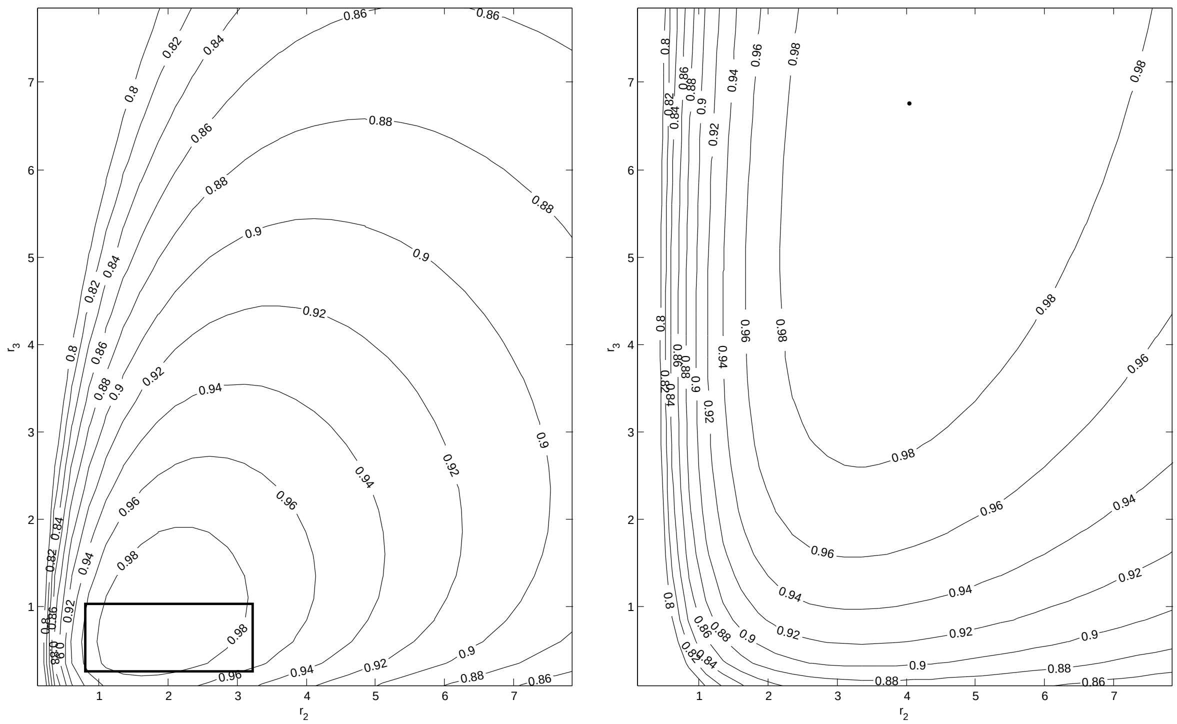

Figure 2: *Level curves of the efficiencies (9) for the standardized maximin optimal design for the intervals* $V^2 = [0.81, 3.22]$ *and* $V^2 = [0.26, 1.04]$ *(marked by the rectangle) and the design corresponding to the distribution* $w = (1.61, 0.52)$.

For example, consider the situation where the clinical team uses the preliminary information from Table 1 for the construction of the local optimal design, but the "true" variances are given by $\sigma_1^2 = 10.4^2$, $\sigma_2^2 = 4\sigma_1^2$ and $\sigma_3^2 = 3\sigma_1^2$. In this case the minimal sample sizes to achieve a power of $1-\beta = 0.8$ with level $\alpha = 0.025$ and non-inferiority parameter $\theta = 0.6$ are $n = 1767$ ($n_1 = 862$, $n_2 = 656$ and $n_3 = 249$) for the misspecified local optimal design and $n = 1685$ ($n_1 = 766$, $n_2 = 644$ and $n_3 = 275$) for the standardized maximin optimal design for $V = [0.64, 4.03] \times [0.21, 1.3]$. This equals reduction of 4.9 % in the total sample size. Note that in this case the observed ratio of the means is $(\bar{x}_1 - \bar{x}_3)/(\bar{x}_2 - \bar{x}_3) = 0.49$.

As a further example we consider the above case but with $\mu_1 = 33.67$, $\mu_2 = 36.7$ and $\mu_3 = 16.5$, which corresponds to a ratio $(\mu_1 - \mu_3)/(\mu_2 - \mu_3) = 0.85$. In this case the misspecified local optimal design requires $n = 314$ ($n_1 = 154$, $n_2 = 116$ and $n_3 = 44$) observations to achieve a power of 0.8, while the standardized maximin optimal design yields $n = 299$ ($n_1 = 136$, $n_2 = 114$ and $n_3 = 49$) for the required sample sizes of the non-inferiority trial. This corresponds to a reduction of 5.4 % in the total sample size.

If the variances ratios are correctly specified the efficiencies of the local and the standardized maximin optimal designs are very similar and this similarity is also reflected in the sample sizes required for the two

$V^2 \times V^3$		$[0.64, 4.03] \times [0.21, 1.30]$			$[0.81, 3.22] \times [0.26, 1.04]$			$[1.61, 1.61] \times [0.52, 0.52]$		
θ	$\frac{\mu_1-\mu_3}{\mu_2-\mu_3}$	n_1	n_2	n	n_1	n_2	n	n_1	n_2	n
0.6	0.85	65	55	143	67	55	144	70	53	143
0.6	0.90	46	39	101	47	38	101	49	37	100
0.6	0.95	34	29	75	35	29	75	36	27	74
0.6	1	26	22	57	27	22	58	28	21	57
0.8	0.85	1733	1844	3899	1749	1583	3891	1799	1826	3885
0.8	0.90	434	462	977	438	397	975	451	458	974
0.8	0.95	194	206	437	196	177	436	201	204	434
0.8	1	110	117	248	111	100	247	114	116	246

Table 3: *Sample sizes n needed for $\alpha = 0.025$ and $1-\beta = 0.8$ using the minimax approach with the indicated variance ratio intervals V^2 and V^3. The designs in the last row are local optimal.*

different designs to achieve a given power. To illustrate this fact we again investigate the situation considered in Table 1 and compare the required minimum sample size to achieve a power of 80 percent. The following parameters are chosen: the non-inferiority threshold $\theta = 0.6$ and $\theta = 0.8$, the expected values of the reference $\mu_2 = 36.7$ and the placebo $\mu_3 = 16.5$, and the three standard deviations $\sigma_1 = 10.4$, $\sigma_2 = 13.2$, $\sigma_3 = 7.5$ and thus variance ratios of $r_2 = \frac{\sigma_2^2}{\sigma_1^2} = 1.61$, $r_3 = \frac{\sigma_3^2}{\sigma_1^2} = 0.52$. The expected value of the new experimental drug μ_1 will be varied as the only parameter. In the following we compare the minimal sample sizes for two standardized maximin optimal designs and the local optimal design which uses the "true" variance ratios. For the sake of comparison we display the results for the same ratios $(\mu_1-\mu_3)/(\mu_2-\mu_3)$ as considered by Hasler et. al. (2007). It is clearly visible that the maximin approach specifies very efficient designs which are about as good as the local optimal choice. In particular, the total sample size to achieve the required power is at most 0.3 % larger for the standardized maximin optimal design as for the local optimal design, which requires the exact and correct specification of the variance ratios.

5 Concluding Remarks

Most optimal experimental designs for three-arm clinical trials depend on the ratios of the population variances, which are not available before the trial. An erroneous specification of these ratios can lead to a loss of the efficiency of the local optimal experimental designs, and notable care is necessary in choosing these variance ratios. In this paper we have proposed a new method for robust designs in three-arm non-inferiority trials which is less sensitive to such misspecifications. In particular, only intervals of variance ratios have to be specified for the design of the clinical trial in advance. These estimates may even be very conservative and the resulting standardized maximin design still allows to conduct economic and highly efficient studies. We feel that this situation is more realistic in practical applications, because in many cases preliminary information from previous similar trials are available. These data might not provide a precise classification of the variance ratios, but might allow to specify - sometimes very large - intervals of the required ratios of the population variances.

Our approach is based on the standardized maximin principle, and determines the design which maximizes the worst case efficiency over the range of the specified variance ratios. The numerical results indicate that standardized maximin optimal designs are very efficient for all values of specified variances. Therefore, standardized maximin optimal designs provide an interesting alternative to the commonly used local optimal designs, which may be inefficient, if the variance ratios have been misspecified. A MatLab program for the numerical construction of the standardized designs may be downloaded at [Maximin-Program (2007)].

Acknowledgments The work of the authors was supported in part by the Sonderforschungsbereich 475, Komplexitätsreduktion in multivariaten Datenstrukturen (Teilprojekt A2) and by a NIH grant award IR01GM072876:01A1. The authors would also like to thank the associate editor and two unknown referees for their constructive comments on earlier versions of this paper.

6 Appendix

6.1 Proof of Lemma

We will investigate the previously used efficiency function $g(w_2, w_3)$ from (8) for fixed w_2 and w_3 and vary the variance ratios r_2 and r_3 to see where possible minima are attained. For this purpose we consider the function

$$h(r_2,r_3) = \frac{\left(1+\theta\sqrt{r_2}+(1-\theta)\sqrt{r_3}\right)^2}{\left(1+\frac{\theta^2}{w_2}r_2+\frac{(1-\theta)^2}{w_3}r_3\right)(1+w_2+w_3)} \tag{19}$$

and simplify it to

$$f(a_2,a_2) = \frac{(1+a_2+a_3)^2}{\left(1+\frac{a_2^2}{w_2}+\frac{a_3^2}{w_3}\right)(1+w_2+w_3)},$$

with $a_2 = \theta\sqrt{r_2}$ and $a_3 = (1-\theta)\sqrt{r_3}$. The gradient of $grad\, f(a_2,a_3)$ is given by

$$grad\, f(a_2,a_3) \quad = 2\frac{1+a_2+a_3}{(1+\frac{a_2^2}{w_2}+\frac{a_3^2}{w_3})(1+w_2+w_3)} \quad \begin{pmatrix} 1-\dfrac{a_2(1+a_2+a_3)}{w_2(1+\frac{a_2^2}{w_2}+\frac{a_3^2}{w_3})} \\ 1-\dfrac{a_3(1+a_2+a_3)}{w_3(1+\frac{a_2^2}{w_2}+\frac{a_3^2}{w_3})} \end{pmatrix},$$

which equals zero only at the point

$$
\begin{aligned}
a_2^* &= w_2 \\
a_3^* &= w_3
\end{aligned} \tag{20}
$$

The Hessian Matrix at this point is obtained as

$$
H\left(f(a_2^*, a_3^*)\right) = \tfrac{2}{(1+w_2+w_3)^2} \begin{pmatrix} -\frac{1+w_3}{w_2} & 1 \\ 1 & -\frac{1+w_2}{w_3} \end{pmatrix}.
$$

This matrix is negative definite: the signs of the two minors alternate starting with a negative value. Thus, the point (a_2^*, a_3^*) is a global maximum. With this information it follows that the minimum of (8) must be attained at the boundary of the set $V = V^2 \times V^3$. But looking at the one-directional derivatives with respect to a_2 and a_3 yield even more: the minimum value must be attained at one of the four corners of the rectangle. This follows because the function $\frac{\delta f}{\delta a_2}$ has only one possible extrema at the point

$$
\tilde{a_2} = \frac{w_2(a_3^2 + w_3)}{w_3(1+a_3)}
$$

where the second derivative is always negative. Thus this point always corresponds to a local maximum. The same argument is valid for the function $\frac{\delta f}{\delta a_3}$ and leads to the conclusion that the minimal value of $f(a_2, a_3)$ (and of $h(r_2, r_3)$ for fixed w_2, w_3 and θ, of course) is taken at one of the four corners of the rectangle.

Thus, (19) has only a single, global extrema which is a maximum, and the directional derivatives in direction of r_2 and r_3 (a_2 and a_3, respectively) have only one critical point corresponding to a local maximum, too. Since the set $V = V^2 \times V^3$ is compact, we conclude that the minimal value of h (with respect to (r_2, r_3) is attained at one of the four corners of the rectangle V.

□

References

[Biedermann et al. (2006)] S. Biedermann, H. Dette, and A. Pepelyshev. Some robust design strategies for percentile estimation in binary response models. *The Canadian Journal of Statistics*, 34:603–622, 2006.

[Chernoff (1953)] H. Chernoff. Locally optimal designs for estimating parameters. *Annals of Mathematical Statistics*, 24:586–602, 1953.

[Dette (1997)] H. Dette. Designing experiments with respect to "standardized" optimality criteria. *Journal of the Royal Statistical Society. Series B Methodological*, 59:97–110, 1997.

[Hasler et al. (2007)] M. Hasler, R. Vonk, and L. A. Hothorn. Assessing non-inferiority of a new treatment in a three-arm trial in the presence of heteroscedasticity. *Statistics in Medicine*, 26:, 2007.

[Hung et. al. (2005)] H.M.J. Hung , S.J. Wang, R. O'Neill. A regulatory perspective on choice of margin and statistical inference issue in non-inferiority trials. *Biometrical Journal*, 47 :28-36, 2007.

[Hung et. al. (2007)] H.M.J. Hung , S.J. Wang, R. O'Neill. Issues with statistical risks for testing methods in noninferiority trial without a placebo arm. *Journal of Biopharmaceutical Statistics*, 17 :201-213, 2007.

[Maximin-Program (2007)] Maximin-Program. Matlab program to calculate optimal robust segmentation of groups in three-arm non inferiority trials. *www.rub.de/mathematik3*, 2007.

[Müller (1995)] C. Müller. Maximin efficienct designs for estimating nonlinear aspects in linear models. *Journal of Statistical Planning and Inference*, 44:117–132, 1995.

[Munk et. al. (2005)] A. Munk, H.J. Trampisch. Therapeutic equivalence clinical issues and statistical methodology in noninferiority trials. *Biometrical Journal*, 47:7-9, 2005.

[Pigeot et al. (2003)] I. Pigeot, J. Schäfer, J. Röhmel, and D. Hauschke. Assessing non-inferiority of a new treatment in a three-arm clinical trial including a placebo. *Statistics in Medicine*, 22:883–899, 2003.

[Pukelsheim (1993)] F. Pukelsheim. *Optimal Design of Experiments*. J. Wiley, 1993.

[Silva-Costa-Gomes et al. (2005)] T. Silva-Costa-Gomes, L. Gallart, J. Valles, L. Trillo, J. Minguella, and M. Puig. Low- vs. high-dose almitrine combined with nitric oxide to prevent hydroxia during open-chest one-lung ventilation. *British Journal of Anaesthesia*, 95:410–416, 2005.

[The-MathWorks (1984)] The-MathWorks. Matlab. *www.matlab.com*, 1984.

[Welch (1938)] B.L. Welch The significance of the difference between two means when the population variances are unequal. *Biometrika*, 29:350–362, 1938.

[Wolfram-Research (1988)] Wolfram-Research. Mathematica. *www.wolfram.com*, 1988.