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

July 5, 2007

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 of 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. 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. A two-arm design where a new experimental drug (E) is compared with the reference drug or active control (R) is common. "Gold design trials" are performed as three-arm designs, including the new experimental drug (E), the reference drug or active control (R) and a placebo control (P). For these trials, Pigeot et al. [2003] formulate non-inferiority as a fraction of the trial sensitivity. The null hypothesis is based on the ratio of the differences of the means $H_0 : \frac{\mu_E-\mu_P}{\mu_R-\mu_P} \leq \theta$ and is compared with the alternative $H_1 : \frac{\mu_E-\mu_P}{\mu_R-\mu_P} > \theta$ for a given threshold $\theta \in (0,1)$. The alternative hypothesis indicates that the relative efficacy of the experimental drug is more than $\theta * 100\%$ 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, Table 1 shows the summary statistics for the primary respiratory endpoint Pa_{O_2} (kPa) of a clinical trial data set of Silva-Costa-Gomes et al. [2005]. The experimental drug ALM4+NO was compared

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)*

with the reference ALM16+NO and the placebo which all were administered 30 minutes after onset of one-lung ventilation. The data shows a markedly lower variance in the placebo group.

Assuming homogeneous variances, an optimal design can be achieved (see 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 (see Hasler et al. [2007]), where the unbalancedness 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 substantial errors in the experimental design. 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 does 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 three groups which correspond to an experimental, a reference and a placebo group with means μ_1, μ_2, μ_3 in medical trials and concentrate on the previously introduced problem of finding a robust design for the 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 effectiveness threshold $\theta \in (0,1)$ in a so called non-inferiority three arm design.
We scrutinize the following statistic

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

where σ_i^2 denotes the (unknown) variance, n_i the sample size and $\bar{x}_i$ the arithmetic mean of each group $i = \{1,2,3\}$, and 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 changed to

$$T \sim N\left(\frac{\sqrt{n_1}\,(\mu_1 - \theta\mu_2 - (1-\theta)\mu_3)}{\sqrt{\sigma_1^2 + \frac{\theta^2}{w_1}\sigma_2^2 + \frac{(1-\theta)^2}{w_2}\sigma_3^2}}, 1\right)$$

with $w_1 = \frac{n_2}{n_1}$, $w_2 = \frac{n_3}{n_1}$ being ratios of the sample sizes. For a given significance level α and power level $1-\beta$ we derive the formula

$$\frac{\sqrt{n_1}\,(\mu_1 - \theta\mu_2 - (1-\theta)\mu_3)}{\sqrt{\sigma_1^2 + \frac{\theta^2}{w_1}\sigma_2^2 + \frac{(1-\theta)^2}{w_2}\sigma_3^2}} \quad = \quad z_{1-\alpha} - z_\beta\,,$$

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

$$\begin{aligned} n_1 &= \left(z_{1-\alpha} - z_\beta\right)^2 (\mu_1 - \theta\mu_2 - (1-\theta)\mu_3)^{-2} \left(\sigma_1^2 + \frac{\theta^2}{w_1}\sigma_2^2 + \frac{(1-\theta)^2}{w_2}\sigma_3^2\right) \\ &= \left(\frac{z_{1-\alpha} - z_\beta}{\mu_1 - \theta\mu_2 - (1-\theta)\mu_3}\right)^2 \cdot \sigma_1^2 \left(1 + \frac{\theta^2}{w_1}b_1 + \frac{(1-\theta)^2}{w_2}b_2\right) \end{aligned}$$

as sample size for group one, where $b_1 = \sigma_2^2/\sigma_1^2$ and $b_2 = \sigma_3^2/\sigma_1^2$ denote the (fixed) ratios of the variances σ_2^2 and σ_3^2 with reference to σ_1^2. This means that one has to determine the ratios of the variances for the design of the experiment.

The minimum sample size n to achieve the required power is now determined via

$$n = n_1(1 + w_1 + w_2) = \left(\frac{z_{1-\alpha} - z_\beta}{\mu_1 - \theta\mu_2 - (1-\theta)\mu_3}\right)^2 \cdot \sigma_1^2 \left(1 + \frac{\theta^2}{w_1}b_1 + \frac{(1-\theta)^2}{w_2}b_2\right)(1 + w_1 + w_2)\,, \tag{2}$$

where $\left(1 + \frac{\theta^2}{w_1}b_1 + \frac{(1-\theta)^2}{w_2}b_2\right)(1 + w_1 + w_2)$ is the only part which can be minimized with respect to the sample sizes (or more precisely: with respect to the ratios of sample sizes w_1 and w_2).

The optimal values for w_1 and w_2 are determined by solving the system of equations

$$0 = \frac{\delta}{\delta w_1}\left(1 + \frac{\theta^2}{w_1}b_1 + \frac{(1-\theta)^2}{w_2}b_2\right)(1 + w_1 + w_2)$$

$$0 = \frac{\delta}{\delta w_2}\left(1 + \frac{\theta^2}{w_1}b_1 + \frac{(1-\theta)^2}{w_2}b_2\right)(1 + w_1 + w_2)\,.$$

For $0 < \theta < 1$ the only solution is

$$w_1 = \theta\sqrt{b_1} \tag{3}$$

$$w_2 = (1-\theta)\sqrt{b_2}\,, \tag{4}$$

which leads to the optimal sample sizes

$$\begin{aligned} n_1 &= \left(\frac{z_{1-\alpha}-z_\beta}{\mu_1-\theta\mu_2-(1-\theta)\mu_3}\right)^2 \cdot \sigma_1^2\left(1+\theta\sqrt{b_1}+(1-\theta)\sqrt{b_2}\right) \\ n_2 &= \theta\cdot\sqrt{b_1}\cdot n_1 \\ n_3 &= (1-\theta)\cdot\sqrt{b_2}\cdot n_1 \\ N &= n_1\cdot\left(1+\theta\sqrt{b_1}+(1-\theta)\sqrt{b_2}\right). \end{aligned}$$

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

$$p_2 = \frac{w_1}{1+w_1+w_2} \qquad p_3 = \frac{w_2}{1+w_1+w_2} \tag{5}$$

which indicate 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. 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 b_1 and b_2, which are usually not available before the experiment. In particular, a misspecification of these ratios may result in serious errors of the allocation of the treatments thus making that specific trial rather inefficient. In the following section, we will propose a robust design, which is less sensitive with respect to misspecified variance ratios and which is additionally excessively efficient for the three-arm clinical trial.

3 Robust Design With A Maximin Approach

A more realistic approach to the problem considered in Section 1 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 $\frac{\sigma_1^2}{\sigma_2^2}\in V^1 := \left[V_L^1, V_U^1\right]$ and $\frac{\sigma_1^2}{\sigma_3^2}\in V^2 := \left[V_L^2, V_U^2\right]$, where $V_L^1, V_U^1, V_L^2, V_U^2$ are the boundary points of the postulated intervals for the variance ratios with respect to $\sigma_1^2\in\mathbb{R}^+$. We want to minimize the required total population sample size N to achieve a given power. For this purpose we use the rate function

$$f(w_1,w_2,b_1,b_2) = \left(1+\frac{\theta^2}{w_1}b_1+\frac{(1-\theta)^2}{w_2}b_2\right)(1+w_1+w_2) \tag{6}$$

and mention that - if the ratios of the variances are fixed and known - this function has exactly one minimum (see the previous chapter or Hasler et al. [2007]). Nevertheless, this local optimal design might not be a good choice if the ratios of the variances are misspecified. In order to derive designs which are less sensitive with respect to such misspecifications, we consider the efficiency

$$eff(w_1,w_2,b_1,b_2) = \frac{\min_{\nu,\omega} f(\nu,\omega,b_1,b_2)}{f(w_1,w_2,b_1,b_2)} \in [0,1], \tag{7}$$

which measures the performance of an arbitrary design $w = (w_1, w_2)$ (in the denominator) with respect to the best design (in the numerator) calculated under the assumption that b_1 and b_2 are the "true" ratios of the population variances. Following Dette [1997] a design $w_M^* = \left(w_1^*, w_2^*\right)$ is called standardized maximin optimal if it maximizes the minimum efficiency

$$g(w_1, w_2) \quad = \quad \min_{b_1 \in V^1, b_2 \in V^2} eff(w_1, w_2, b_1, b_2) \tag{8}$$

over the rectangle $V^1 \times V^2$.

With our knowledge from the previous chapter we know that for fixed variance ratios b_1 and b_2 the minimum of $f(\nu, \omega, b_1, b_2)$ is attained at the point $\nu = \theta\sqrt{b_1}$ and $\omega = (1-\theta)\sqrt{b_2}$ and thus formula (7) has the form

$$eff(w_1, w_2, b_1, b_2) = \frac{\min_{\nu,\omega} f(\nu, \omega, b_1, b_2)}{f(w_1, w_2, b_1, b_2)} = \frac{f\left(\theta\sqrt{b_1}, (1-\theta)\sqrt{b_2}, b_1, b_2\right)}{f(w_1, w_2, b_1, b_2)}, \tag{9}$$

where

$$f\left(\theta\sqrt{b_1}, (1-\theta)\sqrt{b_2}, b_1, b_2\right) = \left(1 + \theta\sqrt{b_1} + (1-\theta)\sqrt{b_2}\right)^2. \tag{10}$$

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

$$g(w_1, w_2) \quad = \quad \min_{b_1 \in V^1, b_2 \in V^2} \frac{\left(1 + \theta\sqrt{b_1} + (1-\theta)\sqrt{b_2}\right)^2}{f(w_1, w_2, b_1, b_2)} \tag{11}$$

$$= \quad \min_{b_1 \in V^1, b_2 \in V^2} \frac{\left(1 + \theta\sqrt{b_1} + (1-\theta)\sqrt{b_2}\right)^2}{\left(1 + \frac{\theta^2}{w_1} b_1 + \frac{(1-\theta)^2}{w_2} b_2\right)(1 + w_1 + w_2)} \tag{12}$$

The following Lemma states that the minimum on the right hand side of (11) with respect to (b_1, b_2) may only be attained at the edges of the rectangle $V^1 \times V^2$. The proof can be found in the appendix.

Lemma

The minimum of the function eff defined by (9) with respect to $(b_1, b_2) \in V^1 \times V^2$ may only be attained at the edges of the rectangle $V_1 \times V_2$, that is

$$\begin{aligned} g(w_1, w_2) \quad = \quad & \min\{eff\left(w_1, w_2, V_L^1, V_L^2\right), eff\left(w_1, w_2, V_U^1, V_L^2\right), \\ & eff\left(w_1, w_2, V_L^1, V_U^2\right), eff\left(w_1, w_2, V_U^1, V_U^2\right)\} \end{aligned} \tag{13}$$

With this Lemma, one only has to numerically maximize the function (8) at the four edges of $V_1 \times V_2$. The resulting robust design is

$$\arg\max_{w_1, w_2} g(w_1, w_2) \quad = \quad (w_1^*, w_2^*) = w^*, \tag{14}$$

which 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. This condition can then be used to check the optimality of the numerically calculated design. For a detailed discussion the reader is referred to e.g. Pukelsheim [1993] or Müller [1995]. For this purpose we introduce the following notation

$$c_\theta^T = \left(\begin{array}{ccc} 1 & \theta & (1-\theta) \end{array} \right), \qquad \theta \in (0,1)\,,$$

and the set

$$V = V^1 \times V^2\,.$$

For fixed variance ratios $v = (b_1, b_2) \in V$ and arbitrary group ratios $w = (w_1, w_2)$ we define

$$M(w,v) := \frac{1}{1+w_1+w_2} diag\left(\sigma_1^2, \frac{w_1}{\sigma_2^2}, \frac{w_2}{\sigma_3^2}\right) = \frac{1}{\sigma_1^2 \cdot (1+w_1+w_2)} \cdot diag\left(1, \frac{w_1}{b_1}, \frac{w_2}{b_2}\right)$$

The optimality criterion in (7) can be rewritten as

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

where w_b^* denotes the locally optimal design assuming known ratios for the variances b_1 and b_2, that is $w_b^* = \left(\theta\sqrt{b_1}, (1-\theta)\sqrt{b_2}\right)$ (see the discussion in the previous chapter). The following characterization of the standardized maximin optimal design is a consequence of Theorem 2 in Biedermann et al. [2006].

Theorem

Let

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

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 positive weight $\pi^*(v)$ such that the following equality is valid for $i = \{1,2,3\}$

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

where

$$x_1 = \left(\begin{array}{c} 1 \\ 0 \\ 0 \end{array} \right), x_2 = \left(\begin{array}{c} 0 \\ 1 \\ 0 \end{array} \right), x_3 = \left(\begin{array}{c} 0 \\ 0 \\ 1 \end{array} \right)$$

and

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

By our Lemma derived in this Section, the set $N(w)$ consists for any design at most of the 4 edges of the

rectangle V, namely

$$v_1 = (V_1^L, V_2^L),\ v_2 = (V_1^U, V_2^L),\ v_3 = (V_1^L, V_2^U),\ v_4 = (V_1^U, V_2^U)\,.$$

Numerical evaluations show that for the standardized maximin optimal design w_M^* the set $N(w_M^*)$ normally degenerates to just two or three points which means that the remaining edges of the set V have higher efficiencies.

Equality of (16) is always attained at the points x_1, x_2 and x_3 which 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 $\pi(v_1), \pi(v_2), \pi(v_3), w_1$ and w_2 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 edge v_i is not an element of the set $N(w_M^*)$.
We use the following notation to keep the equations more readable

$$a_{11} = \theta\sqrt{V_L^1},\ a_{12} = \theta\sqrt{V_U^1},\ a_{21} = (1-\theta)\sqrt{V_L^2},\ a_{22} = (1-\theta)\sqrt{V_U^2}\,,$$

and obtain the following system of nonlinear equations

$$\pi(v_1) \cdot \frac{(1+w_1+w_2)}{1+\frac{a_{11}^2}{w_1}+\frac{a_{21}^2}{w_2}} + \pi(v_2) \cdot \frac{(1+w_1+w_2)}{1+\frac{a_{12}^2}{w_1}+\frac{a_{21}^2}{w_2}} + \pi(v_3) \cdot \frac{(1+w_1+w_2)}{1+\frac{a_{11}^2}{w_1}+\frac{a_{22}^2}{w_2}} + \pi(v_4) \cdot \frac{(1+w_1+w_2)}{1+\frac{a_{12}^2}{w_1}+\frac{a_{22}^2}{w_2}} = 1 \tag{17}$$

$$\pi(v_1) \cdot \frac{(1+w_1+w_2)\left(\frac{a_{11}}{w_1}\right)^2}{1+\frac{a_{11}^2}{w_1}+\frac{a_{21}^2}{w_2}} + \pi(v_2) \cdot \frac{(1+w_1+w_2)\left(\frac{a_{12}}{w_1}\right)^2}{1+\frac{a_{12}^2}{w_1}+\frac{a_{21}^2}{w_2}} + \pi(v_3) \cdot \frac{(1+w_1+w_2)\left(\frac{a_{11}}{w_1}\right)^2}{1+\frac{a_{11}^2}{w_1}+\frac{a_{22}^2}{w_2}} + \pi(v_4) \cdot \frac{(1+w_1+w_2)\left(\frac{a_{12}}{w_1}\right)^2}{1+\frac{a_{12}^2}{w_1}+\frac{a_{22}^2}{w_2}} = 1$$

$$\pi(v_1) \cdot \frac{(1+w_1+w_2)\left(\frac{a_{21}}{w_2}\right)^2}{1+\frac{a_{11}^2}{w_1}+\frac{a_{21}^2}{w_2}} + \pi(v_2) \cdot \frac{(1+w_1+w_2)\left(\frac{a_{21}}{w_2}\right)^2}{1+\frac{a_{12}^2}{w_1}+\frac{a_{21}^2}{w_2}} + \pi(v_3) \cdot \frac{(1+w_1+w_2)\left(\frac{a_{22}}{w_2}\right)^2}{1+\frac{a_{11}^2}{w_1}+\frac{a_{22}^2}{w_2}} + \pi(v_4) \cdot \frac{(1+w_1+w_2)\left(\frac{a_{22}}{w_2}\right)^2}{1+\frac{a_{12}^2}{w_1}+\frac{a_{22}^2}{w_2}} = 1$$

These equations allow us to check whether a given design (w_1, w_2) is optimal or not. To find such an optimal design, one first solves the maximizing problem (14), evaluates the efficiencies at the edges 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 the 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]). Several examples of the described procedure can be found in the following section.

4 Examples

4.1 Example 1

In a first example we illustrate the use of the checking condition. For this purpose let us assume that the variance ratios are located in the intervals $V^1 = [0.16, 0.64]$ and $V^2 = [0.49, 3.24]$, 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

(17)

$$a_{11} = 0.5 \cdot \sqrt{0.16} = 0.2 \qquad a_{21} = 0.5 \cdot \sqrt{0.49} = 0.35$$
$$a_{12} = 0.5 \cdot \sqrt{0.64} = 0.4 \qquad a_{22} = 0.5 \cdot \sqrt{3.24} = 0.9$$

In the next step we numerically maximize the minimal efficiency at the edges of the rectangle $V = V^1 \times V^2$ in terms of w_1 and w_2:

$$\begin{aligned} argmax_{w_1,w_2} \quad min \quad \{&eff(w_1,w_2,a_{11},a_{21}), eff(w_1,w_2,a_{12},a_{21}), \\ &eff(w_1,w_2,a_{11},a_{22}), eff(w_1,w_2,a_{12},a_{22})\} \end{aligned} \tag{18}$$

using a slightly modified version efficiency function (7), that is

$$eff(w_1,w_2,a_1,a_2) \quad = \quad \frac{(1+a_1+a_2)^2}{\left(1+\frac{a_1^2}{w_1}+\frac{a_2^2}{w_2}\right)(1+w_1+w_2)}$$

The numerical solution of the optimization problem (18) is in our considered case $w^* = (0.3818, 0.6249)$ yielding a minimal efficiency of at least 93.26%. To check whether the calculated design is optimal or not, we calculate the efficiencies

$$eff(w_1^*,w_2^*,a_{11},a_{21}) = 0.9326 \qquad eff(w_1^*,w_2^*,a_{12},a_{21}) = 0.9326$$

$$eff(w_1^*,w_2^*,a_{11},a_{22}) = 0.9326 \qquad eff(w_1^*,w_2^*,a_{12},a_{22}) = 0.9730$$

in order to apply Theorem. Because the efficiency at the point $v_4 = (a_{12},a_{22})$ 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 and/or Mathematica, we calculate the weights to be $\pi(v_1) = 0.4603$ for the point $v_1 = (a_{11},a_{21})$, $\pi(v_2) = 0.5072$ for the point $v_2 = (a_{12},a_{21})$ and the remaining mass to be $\pi(v_3) = 0.0325$ at the point $v_3 = (a_{12},a_{21})$.

With this weight distribution we validated that the solution w^* is indeed the optimal solution. Using the conversion (5), the optimal allocation w^* means that we have to take about $p_2^* = 17\%$ of our observations at group two, about $p_3^* = 32\%$ of our observations at group three, and the remaining 51% of our observations at group one. To achieve a predefined power, one can use the formula (2) to compute the approximately total sample size N.

4.2 Example 2

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. Consider the case that we will design a new randomized clinical trial with a further experimental drug and that we expect similar results as presented in Table 1. Therefore, we assume that the variance ratios are located within the intervals $V^1 = [1.0, 2.0]$ and $V^2 = [0.40, 0.60]$. The non-inferiority parameter is given by $\theta = 0.8$. Numerical calculations similar to Example 4.1 yield the optimal weight distribution to be $w^* = (0.9566, 0.1434)$. Using (5), we obtain that this means that

θ	V^1	V^2	$p^* = (p_2^*, p_3^*)$	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^1	V^2	$p^* = (p_2^*, p_3^*)$	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^1	V^2	$p^* = (p_2^*, p_3^*)$	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^1	V^2	$p^* = (p_2^*, p_3^*)$	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 allocation and minimal efficiency for different non-inferiority parameters θ and variance ratios V_1 and V_2*

the standardized maximin designs allocates approximately $n_1 = 0.4762 \cdot N$, $n_2 = 0.4555 \cdot N$ and $n_3 = 0.0683 \cdot N$ at the three trials for a fixed sample size N. The efficiency of this design is at least 0.9910.

If we consider a possible total sample size of 100 patients, this design advises the experimenter to prescribe about 46 persons the experimental drug, 7 persons a placebo treatment, and the remaining 57 persons the standard treatment.

4.3 Example 3

In our final example we list some optimal robust designs for certain non-inferiority parameters and variance ratios and compare their efficiencies. For a fixed variance $\sigma_1^2 \in \mathbb{R}$ of the first group, numerical evaluations yield Table 2 where $V^1 = [V_1^L, V_1^U]$ is the specified interval for the variance ratio of our second group, $V^2 = [V_2^L, V_2^U]$ is the interval of the variance ratio of our third group, p^* is the optimal allocation of group two and three, 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 group 2, $p_3^* \cdot N$ observations at group 3, and the remaining observations at group one. 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 region $V^1 \times V^2$ and are always very high. This and further results, which are not shown for the sake of simplicity, 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.

5 Concluding Remarks

Most 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 very inefficient 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 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 is available. These data might not grant 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 efficiency over the range of the provided variance ratios. The numerical results indicate that the new results 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 my be inefficient, if the variance ratios have been misspecified. A MatLab program for the numerical construction of the standardized designs can be downloaded at Maximin-Program [2007].

6 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.

7 Appendix

7.1 Proof of Lemma

We will investigate the previously used efficiency function $g(w_1, w_2)$ from (8) for fixed w_1 and w_2 and vary the variance ratios b_1 and b_2 to see where a minimum is attained. For this purpose we consider the function

$$h(b_1,b_2) = \frac{\left(1+\theta\sqrt{b_1}+(1-\theta)\sqrt{b_2}\right)^2}{\left(1+\frac{\theta^2}{w_1}b_1+\frac{(1-\theta)^2}{w_2}b_2\right)(1+w_1+w_2)} \tag{19}$$

and simplify it to

$$f(a_1,a_2) = \frac{(1+a_1+a_2)^2}{\left(1+\frac{a_1^2}{w_1}+\frac{a_2^2}{w_2}\right)(1+w_1+w_2)},$$

where $a_1 = \theta\sqrt{b_1}$ and $a_2 = (1-\theta)\sqrt{b_2}$. The gradient of $grad\, f(a_1,a_2)$ is given by

$$grad\, f(a_1,a_2) \quad = 2\frac{1+a_1+a_2}{(1+\frac{a_1^2}{w_1}+\frac{a_2^2}{w_2})(1+w_1+w_2)} \quad \begin{pmatrix} 1-\dfrac{a_1(1+a_1+a_2)}{w_1(1+\frac{a_1^2}{w_1}+\frac{a_2^2}{w_2})} \\ 1-\dfrac{a_2(1+a_1+a_2)}{w_2(1+\frac{a_1^2}{w_1}+\frac{a_2^2}{w_2})} \end{pmatrix},$$

which equals zero only at the point

$$\begin{aligned} a_1^* &= w_1 \\ a_2^* &= w_2 \end{aligned} \tag{20}$$

The Hessian Matrix at this point is obtained as

$$H\left(f(a_1^*,a_2^*)\right) \quad = \frac{2}{(1+w_1+w_2)^2} \quad \begin{pmatrix} -\frac{1+w_2}{w_1} & 1 \\ 1 & -\frac{1+w_1}{w_2} \end{pmatrix}$$

This matrix is indefinite: the signs of the two minors alternate starting with a negative value. With this information it follows that the minimum of (8) must be attained at the boundary of the set $V = V^1 \times V^2$. But looking at the one-directional derivatives with respect to a_1 and a_2 yield even more: the minimum value must be attained at one of the four edges of the rectangle. This follows because the function $\frac{\delta f}{\delta a_1}$ has only one possible extrema at the point

$$\tilde{a_1} \quad = \quad \frac{w_1(a_2^2+w_2)}{w_2(1+a_2)}$$

where the second derivative is always negative. Thus this point always corresponds to a maximum. The same argument is valid for the function $\frac{\delta f}{\delta a_2}$ and leads to the conclusion that the minimal value of $f(a_1,a_2)$ (and of $h(b_1,b_2)$ for fixed w_1,w_2 and θ, of course) is taken at one of the four edges of the rectangle.

Thus, (19) has only a single, global extrema which is a maximum, and the directional derivatives in direction

of b_1 and b_2 (a_1 and a_2, respectively) have only one critical point corresponding to a local maximum, too. Since the set $V = V^1 \times V^2$ is compact, we conclude that the minimal value of h (with respect to (b_1, b_2) is attained at one of the four edges of the rectangle V.

References

Stefanie Biedermann, Holger Dette, and Andrey Pepelyshev. Some robust design strategies for percentile estimation in binary response models. *The Canadian Journal of Statistics*, 34:603–622, 2006.

Herman Chernoff. Locally optimal designs for estimating parameters. *Annals of Mathematical Statistics*, 24: 586–602, 1953.

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

Mario Hasler, Richardus Vonk, and Ludwig A. Hothorn. Assessing non-inferiority of a new treatment in a three-arm trial in the presence of heteroscedasticity. *Statistics in Medicine*, 26:xx–yy, 2007.

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

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

Iris Pigeot, Juliane Schäfer, Joachim Röhmel, and Dieter Hauschke. Assessing non-inferiority of a new treatment in a three-arm clinical trial including a placebo. *Statistics in Medicine*, 22:883–899, 2003.

Friedrich Pukelsheim. *Optimal Design of Experiments*. J. Wiley, 1993.

Teresa Silva-Costa-Gomes, Lluis Gallart, Jordi Valles, Lurdes Trillo, Juan Minguella, and Margarita 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. Matlab. *www.matlab.com*, 1984.

Wolfram-Research. Mathematica. *www.wolfram.com*, 1988.