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

November 26, 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. The specific statistical
methodology was recently described in [Munk et. al. (2005)]. A two-arm design where a new experimen-
tal drug (E) is compared with the reference drug or active control (R) is common. However, trials with-
out 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 (E), the reference drug or active control (R) and a placebo control (P). 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 \( H_0: \frac{\mu_E - \mu_P}{\mu_R - \mu_P} \leq \theta \) 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 \( \theta \cdot 100\% \) 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 \( P_{a\text{O}_2} \) (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.

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 \( \theta \). Assuming heterogeneous - but "known" - variances, an optimal design can as well be calculated like [Hasler et al. (2007)], where the unbalancedness depends on the given threshold \( \theta \) 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.

<table>
<thead>
<tr>
<th>Treatment group</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placebo</td>
<td>16.5</td>
<td>7.5</td>
<td>14</td>
</tr>
<tr>
<td>ALM4+NO</td>
<td>26.5</td>
<td>10.4</td>
<td>14</td>
</tr>
<tr>
<td>ALM16+NO</td>
<td>36.7</td>
<td>13.2</td>
<td>14</td>
</tr>
</tbody>
</table>

**Table 1: Summary statistics for \( P_{a\text{O}_2} \) (kPa) 30 minutes after onset of one-lung ventilation of the clinical data set of Silva-Costa-Gomes et al. (2005)**
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 a clinical trial with three groups that correspond to the experimental, reference and placebo arms with means $\mu_1, \mu_2, \mu_3$, respectively. We focus on the previously introduced problem of finding a robust design for the non-inferiority hypothesis

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

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

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

is used where $\sigma^2_i$ 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 $\mu_i$ and variances $\sigma^2_i (i = 1, 2, 3)$. The formula (1) can be equivalently written as

$$T \sim N\left( \frac{\sqrt{n_1} (\mu_1 - \theta \mu_2 - (1 - \theta) \mu_3)}{\sqrt{\frac{1}{n_1} \sigma^2_1 + \frac{\theta^2}{n_2} \sigma^2_2 + \frac{(1 - \theta)^2}{n_3} \sigma^2_3}}, 1 \right)$$
with \( w_1 = \frac{z_1}{\bar{n}_1}, w_2 = \frac{n_2}{\bar{n}_1} \) being ratios of the sample sizes. For a given significance level \( \alpha \) and power level \( \beta \) we derive the formula

\[
\frac{\sqrt{\bar{n}_1} (\mu_1 - \theta \mu_2 - (1 - \theta) \mu_3)}{\sqrt{\sigma_1^2 + \frac{\theta^2}{w_1} \sigma_2^2 + \frac{(1-\theta)^2}{w_2} \sigma_3^2}} = z_{1-\alpha} + z_{\beta},
\]

where \( z_{\alpha} \) for \( \alpha \in [0, 1] \) denotes the \( \alpha \)-quantile of a standard normal distribution. This leads to

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

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 \( \sigma_2^2 \) and \( \sigma_3^2 \) with reference to \( \sigma_1^2 \).

With the specified nominal level \( \alpha \) and power \( \beta \), the minimum sample size \( n \) can be derived by minimizing the following function (2) with respect to \( w_1 \) and \( w_2 \)

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

This means that one has to determine the ratio of the variances \( b_1 \) and \( b_2 \) for an optimal design of the experiment. Since the function \( f(w_1, w_2|b_1, b_2) = \left( 1 + \frac{\theta^2}{w_1} b_1 + \frac{(1-\theta)^2}{w_2} b_2 \right) (1 + w_1 + w_2) \) is the only factor on the right side of equation (2) that involves \( w_1 \) and \( w_2 \), the minimum sample size can be derived by simply minimizing this function. 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 unique solution is given by

\[
w_1 = \theta \sqrt{b_1} \tag{3}
\]

\[
w_2 = (1-\theta) \sqrt{b_2}, \tag{4}
\]
which leads to the optimal sample sizes

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

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

\[
p_2 = \frac{w_1}{1+w_1+w_2} \quad p_3 = \frac{w_2}{1+w_1+w_2}
\]

which represent the proportion of observations allocated to group two and three with respect to the total sample size. Following \cite{Chernoff(1953)} the resulting design is called local optimal, because it depends on the (unknown) variance ratios \(b_1\) and \(b_2\). 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 \cite{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 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.

### 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 \(\frac{\sigma_2^2}{\sigma_1^2} \in V_1 := [V_1^L, V_1^U]\) and \(\frac{\sigma_3^2}{\sigma_1^2} \in V_2 := [V_2^L, V_2^U]\), where \(V_1^L, V_1^U, V_2^L, V_2^U\) are the boundary points of the postulated intervals for the variance ratios with respect to \(\sigma_1^2 \in \mathbb{R}^+\). 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. For this purpose we use the rate function (2). We mention 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

\[
eff(w_1, w_2, b_1, b_2) = \frac{f(v, \omega|b_1, b_2)}{f(w_1, w_2|b_1, b_2)} \in [0, 1], \tag{6}
\]

with \( f(v, \omega|b_1, b_2) := \min_{v, \omega} f(w_1, w_2|b_1, b_2) \). Equation (6) 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 \( w^* = (w_1^*, w_2^*) \) is called standardized maximin optimal or briefly maximin optimal design if it maximizes the minimum efficiency

\[
g(w_1, w_2) = \min_{b_1 \in V_1, b_2 \in V_2} \eff(w_1, w_2, b_1, b_2) \tag{7}
\]

over the rectangle \( V_1 \times V_2 \).

With our knowledge from the previous section it follows that for fixed variance ratios \( b_1 \) and \( b_2 \) the minimum of the function \( f(v, \omega|b_1, b_2) \) is attained at the point \( v = \theta \sqrt{b_1} \) and \( \omega = (1 - \theta) \sqrt{b_2} \) and thus formula (6) can be simplified to

\[
eff(w_1, w_2, b_1, b_2) = \frac{f(v, \omega|b_1, b_2)}{f(w_1, w_2|b_1, b_2)} = \frac{f(\theta \sqrt{b_1}, (1 - \theta) \sqrt{b_2}|b_1, b_2)}{f(w_1, w_2|b_1, b_2)}, \tag{8}
\]

where

\[
f(\theta \sqrt{b_1}, (1 - \theta) \sqrt{b_2}|b_1, b_2) = \left( 1 + \theta \sqrt{b_1} + (1 - \theta) \sqrt{b_2} \right)^2. \tag{9}
\]

This simplifies the analysis of formula (7) substantially since now

\[
g(w_1, w_2) = \min_{b_1 \in V_1, b_2 \in V_2} \frac{(1 + \theta \sqrt{b_1} + (1 - \theta) \sqrt{b_2})^2}{f(w_1, w_2|b_1, b_2)} \tag{10}
\]

\[
= \min_{b_1 \in V_1, b_2 \in V_2} \frac{(1 + \theta \sqrt{b_1} + (1 - \theta) \sqrt{b_2})^2}{\left( 1 + \frac{\theta^2}{w_1^2} b_1 + \frac{(1 - \theta)^2}{w_2^2} b_2 \right) (1 + w_1 + w_2)} \tag{11}
\]

The following Lemma states that the minimum on the right hand side of (10) with respect to \((b_1, b_2) \in V_1 \times V_2\)
may only be attained at the corners of the rectangle $V_1 \times V_2$. The proof can be found in the appendix.

**Lemma**

The minimum of the function $\mathit{eff}$ defined by (8) with respect to $(b_1, b_2) \in V_1 \times V_2$ may only be attained at the corners of the rectangle $V_1 \times V_2$, that is

$$g(w_1, w_2) = \min\{ \mathit{eff}(w_1, w_2, V_1^L, V_2^L), \mathit{eff}(w_1, w_2, V_1^L, V_2^U), \mathit{eff}(w_1, w_2, V_1^U, V_2^L), \mathit{eff}(w_1, w_2, V_1^U, V_2^U) \}$$

(12)

With this Lemma, one only has to numerically maximize the function (7) at the four corners of $V_1 \times V_2$. The resulting robust design is

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

(13)

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. 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

$$c^T_\theta = (1, \theta, (1 - \theta)), \quad \theta \in (0, 1),$$

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

$$M(w, v) := \frac{1}{1 + w_1 + w_2} \text{diag} \left( \frac{\sigma_1^2}{\sigma_1^2}, \frac{w_1}{\sigma_1^2}, \frac{w_2}{\sigma_2^2} \right) = \frac{1}{\sigma_1^2} \cdot \left(1 + w_1 + w_2\right) \cdot \text{diag} \left(1, \frac{w_1}{b_1}, \frac{w_2}{b_2} \right).$$
The optimality criterion in (5) can be rewritten as

$$g(w) = \min_{b \in V} eff(w, b) = \min_{b \in V} \frac{e_{\theta}^T M^{-1}(w^*_b, b) c_{\theta}}{e_{\theta}^T M^{-1}(w, b) c_{\theta}},$$

(14)

where $w^*_b$ denotes the local optimal design assuming known ratios of the variances $b_1$ and $b_2$, that is $w^*_b = (\theta \sqrt{b_1}, (1 - \theta) \sqrt{b_2})$ (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

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

be the subset of $V$ consisting of those values of $b$, for which the efficiency (14) 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

$$\sum_{v \in N(w^*_M)} \pi^*(v) \cdot \left( \frac{e_{\theta}^T M^{-1}(w^*_M, v) x_i}{e_{\theta}^T M^{-1}(w^*_M, v) c_{\theta}} \right)^2 = 1, \quad i = 1, 2, 3,$$

(15)

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

$$v_1 = (V_L^L, V_L^U), v_2 = (V_U^L, V_L^U), v_3 = (V_L^L, V_U^U), v_4 = (V_U^L, V_U^U).$$

This means that all other points $v \in V$ have higher efficiencies.
3 ROBUST DESIGN WITH A MAXIMIN APPROACH

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

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

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 corner $v_i$ is not an element of the set $N(w_{M}^*)$.

We use the following notation to keep these equations more readable

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

and obtain the following system of nonlinear equations

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

These equations allow us to check whether a given design $(w_1, w_2)$ is standardized maximin optimal or not.

To find such an optimal design, one first solves the maximizing problem (13), 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 (16). If there exists a valid solution, one can be assured that the 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 $V^1 = [0.16, 0.64]$ 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 defined by (16)

$$
a_{11} = 0.5 \cdot \sqrt{0.16} = 0.2 \quad a_{21} = 0.5 \cdot \sqrt{0.49} = 0.35
$$

$$
a_{12} = 0.5 \cdot \sqrt{0.64} = 0.4 \quad a_{22} = 0.5 \cdot \sqrt{3.24} = 0.9
$$

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

$$
\arg \max_{w_1, w_2} \min \{ eff(w_1, w_2, a_{11}, a_{21}), eff(w_1, w_2, a_{12}, a_{21}),
\quad eff(w_1, w_2, a_{11}, a_{22}), eff(w_1, w_2, a_{12}, a_{22}) \}
$$

where the efficiency function in (6) is now defined for the new parameters $a_1 = \theta \sqrt{b_1}$, $a_2 = (1 - \theta) \sqrt{b_2}$, that is

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

In our considered case the numerical solution of the optimization problem (17) 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

$$
eff(w^*_1, w^*_2, a_{11}, a_{21}) = 0.9326 \quad eff(w^*_1, w^*_2, a_{12}, a_{21}) = 0.9326
$$

$$
eff(w^*_1, w^*_2, a_{11}, a_{22}) = 0.9326 \quad eff(w^*_1, w^*_2, a_{12}, a_{22}) = 0.9730
$$

in order to apply the Theorem of Section 3. 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 (16). Using
MatLab, Mathematica or any other adequate program, we calculate the weights to be $\pi(v_1) = 0.4603$ for the point $v_1 = (a_{11}, a_{21})$, $\pi(v_2) = 0.5072$ 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 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. Therefore, we assume that the variance ratios are located within the intervals $V^1 = [1.0, 2.0]$ and $V^2 = [0.40, 0.60]$. 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 $w^* = (0.9566, 0.1434)$. Using (5), 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 for a fixed sample size $n$. 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 $\theta$ (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 $V^1 = [V^1_L, V^1_U]$ is the specified interval for the variance ratio $b_1 = \sigma_2^2 / \sigma_1^2$, $V^2 = [V^2_L, V^2_U]$ is the interval of the variance ratio $b_2 = \sigma_3^2 / \sigma_1^2$, $p^*$ 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 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 $V^1 \times V^2$ 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
4 FURTHER DISCUSSION AND EXAMPLES

| $\theta$ | $V^1$ | $V^2$ | $p^* = (p^*_2, p^*_3)$ | 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 

| $\theta$ | $V^1$ | $V^2$ | $p^* = (p^*_2, p^*_3)$ | 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 

| $\theta$ | $V^1$ | $V^2$ | $p^* = (p^*_2, p^*_3)$ | 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 

| $\theta$ | $V^1$ | $V^2$ | $p^* = (p^*_2, p^*_3)$ | 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 size and minimal efficiency for different non-inferiority parameters $\theta$ and variance ratios $V^1$ and $V^2$ 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 a non-inferiority parameter $\theta = 0.8$ or $\theta = 0.6$. Note that the observed variance ratios in the data examples are given by $b_1 = \sigma_2^2 / \sigma_1^2 = 1.61$ and $b_2 = \sigma_3^2 / \sigma_1^2 = 0.52$.

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 derived by formula (3) for the point $b = (1.61, 0.52)$ 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 (left part) and the local optimal design (right part) if the parameters have been misspecified. For example, if the “true” ratios of the variances would be $b_1 = 4$ and $b_2 = 3$ 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 $b_1 = 1.61$ and $b_2 = 0.52$ 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 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 $V^1 = [0.81, 3.22]$ 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 (8) for varying parameters $b_1$ and $b_2$, 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 $b = (1.61, 0.52)$.

![Figure 1: Level curves of the efficiencies (8) for the standardized maximin optimal design for the intervals $V^1 = [0.64, 4.03]$ and $V^2 = [0.21, 1.3]$ (marked by the rectangle) and the local optimal design for $b = (1.61, 0.52)$.](image_url)

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, 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 design. On the other
4 FURTHER DISCUSSION AND EXAMPLES

Figure 2: Level curves of the efficiencies (8) for the standardized maximin optimal design for the intervals $V^1 = [0.81, 3.22]$ and $V^1 = [0.26, 1.04]$ (marked by the rectangle) and the design corresponding to the distribution $w = (1.61, 0.52)$.

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 preassign level $\alpha$ and power $\beta$ can be derived from Formula (15) in [Hasler et al. (2007)], and is in our notation given by

$$n_1 \geq \left(t_{1-\alpha}(\bar{\theta}) + t_\beta(\bar{\theta})\right)^2 \frac{\sigma_1^2 + \frac{\theta^2}{w_1} \sigma_2^2 + \frac{(1-\theta)^2}{w_2} \sigma_3^2}{(\mu_1 - \theta \mu_2 - (1-\theta) \mu_3)^2},$$
where $t_u(\vartheta)$ for $u \in [0, 1]$ denotes the $u$-quantile of a $t_\theta$-distribution with

$$
\vartheta = \frac{1}{\sigma_1^2 + \frac{\vartheta^2}{\sigma_2^2} + \frac{\vartheta^2}{\sigma_3^2}} \left( \frac{1}{\sigma_1^2} + \frac{\vartheta^2}{\sigma_2^2} + \frac{\vartheta^2}{\sigma_3^2} \right).
$$

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.

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 $\beta = 0.8$ with level $\alpha = 0.025$ and non-inferiority parameter $\theta = 0.6$ are $n_1 = 862$, $n_2 = 656$ and $n_3 = 249$ for the misspecified local optimal design and $n_1 = 766$, $n_2 = 644$ and $n_3 = 275$ for the standardized maximin optimal design for $V = [0.64, 4.03] \times [0.21, 1.3]$, which yields to a reduction of 4.9 % in the total sample size. Note that in this case the observed ratio of the means is $(M_1 - M_3)/(M_2 - M_3) = 0.49$.

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 $n_1 = 154$, $n_2 = 116$ and $n_3 = 44$ observations to achieve a power of 0.8, while the standardized maximin optimal design yields $n_1 = 136$, $n_2 = 114$ 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 different designs to achieve a given power. To illustrate this fact we investigate again 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 $b_1 = \frac{\sigma_2^2}{\sigma_1^2} = 1.61$, $b_2 = \frac{\sigma_3^2}{\sigma_1^2} = 0.52$. The expected value of the new experimental drug $\mu_1$ 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
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 is 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 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

\[
V_1 \times V_2 = [0.64, 4.03] \times [0.21, 1.30] \\
[0.81, 3.22] \times [0.26, 1.04] \\
[1.61, 1.61] \times [0.52, 0.52]
\]

<table>
<thead>
<tr>
<th>(\theta)</th>
<th>(\frac{\mu_1 - \mu_3}{\mu_2 - \mu_3})</th>
<th>(n_1)</th>
<th>(n_2)</th>
<th>(n)</th>
<th>(n_1)</th>
<th>(n_2)</th>
<th>(n)</th>
<th>(n_1)</th>
<th>(n_2)</th>
<th>(n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6</td>
<td>0.85</td>
<td>65</td>
<td>55</td>
<td>143</td>
<td>67</td>
<td>55</td>
<td>144</td>
<td>70</td>
<td>53</td>
<td>143</td>
</tr>
<tr>
<td>0.6</td>
<td>0.90</td>
<td>46</td>
<td>39</td>
<td>101</td>
<td>47</td>
<td>38</td>
<td>101</td>
<td>49</td>
<td>37</td>
<td>100</td>
</tr>
<tr>
<td>0.6</td>
<td>0.95</td>
<td>34</td>
<td>29</td>
<td>75</td>
<td>35</td>
<td>29</td>
<td>75</td>
<td>36</td>
<td>27</td>
<td>74</td>
</tr>
<tr>
<td>0.6</td>
<td>1</td>
<td>26</td>
<td>22</td>
<td>57</td>
<td>27</td>
<td>22</td>
<td>58</td>
<td>28</td>
<td>21</td>
<td>57</td>
</tr>
<tr>
<td>0.8</td>
<td>0.85</td>
<td>1733</td>
<td>1844</td>
<td>3899</td>
<td>1749</td>
<td>1583</td>
<td>3891</td>
<td>1799</td>
<td>1826</td>
<td>3885</td>
</tr>
<tr>
<td>0.8</td>
<td>0.90</td>
<td>434</td>
<td>462</td>
<td>977</td>
<td>438</td>
<td>397</td>
<td>975</td>
<td>451</td>
<td>458</td>
<td>974</td>
</tr>
<tr>
<td>0.8</td>
<td>0.95</td>
<td>194</td>
<td>206</td>
<td>437</td>
<td>196</td>
<td>177</td>
<td>436</td>
<td>201</td>
<td>204</td>
<td>434</td>
</tr>
<tr>
<td>0.8</td>
<td>1</td>
<td>110</td>
<td>117</td>
<td>248</td>
<td>111</td>
<td>100</td>
<td>247</td>
<td>114</td>
<td>116</td>
<td>246</td>
</tr>
</tbody>
</table>

Table 3: Sample size \(n\) needed for \(\alpha = 0.025\) and \(\beta = 0.8\) using the minimax approach with the indicated variance ratio intervals \(V_1\) and \(V_2\). The last design is local optimal.

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.
6 Appendix

6.1 Proof of Lemma

We will investigate the previously used efficiency function $g(w_1, w_2)$ from (7) 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

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

and simplify it to

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

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

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

1 - \frac{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

$$\begin{align*}
a_1^* &= w_1 \\
a_2^* &= w_2
\end{align*}$$

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 an earlier version of this paper.
The Hessian Matrix at this point is obtained as

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

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 (7) 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 corners of the rectangle. This follows because the function \( \frac{\delta f}{\delta a_1} \) has only one possible extrema at the point

\[
\tilde{a}_1 = \frac{w_1(a_2^2 + w_2)}{w_2(1+a_2^*)}
\]

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 \( \theta \), of course) is taken at one of the four corners of the rectangle.

Thus, (18) 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 corners of the rectangle \( V \).

References


REFERENCES


