VILNIUS UNIVERSITY

FACULTY OF MATHEMATICS AND INFORMATICS

Master thesis

Two-way Enriched Clinical Trial Design in Bayesian Framework

lšplėsto dvipakopio klinikinio randomizacijos plano analizė Bajeso metodu

Antanas Mainelis

VILNIUS, 2019

FACULTY OF MATHEMATICS AND INFORMATICS

DEPARTMENT OF STATISTICAL ANALYSIS

Supervisor: Assoc. Prof., Dr. Viktor Skorniakov

Darbo gynimo data: _____

Registracijos NR. _____

Contents

Abbreviations

- DF - degree of freedom;
- GEE - generalized estimating equations;
- KL - Kullback - Leibler;
- MCMC - Markov Chain Monte Carlo;
- MLE - maximum likelihood estimation;
- RMSE - root mean square error;
- SPCD - sequential parallel design;
- TED - two-way enriched design.

Abstract

A. Mainelis.Two-way Enriched Clinical Trial Design in Bayesian Framework: master thesis/ supervisor Assoc. Prof. Dr. V. Skorniakov; Vilnius university, faculty of Mathematics and Informatics, department of Statistical Analysis.

The objective of this master thesis is to apply Bayesian methodology to the two-way enriched design (TED) which helps to cope with the high rate of the placebo response. The main part is dedicated to the model parameters estimation using the maximum likelihood (ML) and Bayes approach. For the latter, the informative priors like constrained uniform, beta with appropriate parameters distributions and the objective priors like uniform distributions, Jeffreys, reference were tested. The parameters were estimated on the simulated data. The ML estimates had the lowest bias in almost all the cases, but Bayesian estimates had lower RMSE and Kullback-Leibler divergence in the majority of the configurations. Models with the informative priors showed the best results.

In the rest part of this thesis Bayes credible intervals for the hypothesis testing in TED are proposed. The type I error and the power of the new testing procedure were evaluated using the Markov Chain Monte Carlo simulations with the various parameters configurations. The power for testing was higher than in the original approach in almost all the cases. However, type I error rate was controlled well only in the models with the uniform distributions as priors.

Using the proposed method in TED, the sample size and the cost of the researches could be reduced. This work could be extend for the continuous outcome in Bayesian framework. Also, the alternative testing procedure might be developed.

Santrauka

A. Mainelis. lšplėsto dvipakopio klinikinio randomizacijos plano analizė Bajeso metodu: magistro baigiamasis darbas/ vadovas doc. dr. V. Skorniakov; Vilniaus universitetas, Matematikos ir informatikos fakultetas, Statistinės analizės katedra.

Šio magistro darbo tikslas - pritaikyti Bajeso metodologiją išplėsto dvipakopio klinikinio randomizacijos plano atvejui. Šis planas padeda suvaldyti aukštą placebo atsaką. Pagrindinė darbo dalis skirta parametrų vertinimui naudojant didžiausio tikėtinumo (DT) ir Bajeso metodus. Pastarajam buvo pasirinkti informatyvūs aprioriniai skirstiniai tokie kaip apribotas tolygusis arba beta skirstinys su atitinkamais parametrais ir neinformatyvūs - Jeffreys, ”reference” ir tolygūs intervale (0,1) skirstiniai. Parametrai buvo vertinami naudojant generuotus duomenis. DT įverčiai beveik visais atvejais turėjo mažiausią poslinkį, bet Bajeso įverčiai daugeliu situacijų turėjo mažesnes vidutines kvadratines paklaidas bei Kullback-Leibler divergenciją. Rezultatai, gauti su informatyviais aprioriniais skirstiniais, buvo geriausi.

Kita šio darbo dalis skirta hipotezių tikrinimui, kai naudojamas nagrinėjamas planas. Tam tikslui pasiūlyta naudoti Bajeso pasikliautinuosius intervalus. Naudojantis Markovo grandinių Monte Karlo metodu buvo įvertinta naujos hipotezių tikrinimo procedūros pirmos rūšies klaidos tikimybė ir galia keletui parametrų kombinacijų. Beveik visais atvejais gauta galia buvo didesnė negu pirminių pasiūlytų testų, tačiau pirmos rūšies klaidos tikimybė svyravo apie pasirinktą lygmenį tik naudojant tolygius apriorinius skirstinius.

Taigi, taikant Bajeso metodą imties dydis ir tyrimų kaina gali būti sumažinta kai naudojamas išplėstas dvipakopis klinikinis randomizacijos planas. Pratęsiant šio darbo tematiką būtų galima pasiūlyti alternatyvą hipotezių tikrinimui arba sukurti modelį tolydaus atsako atvejui.

Introduction

Randomization designs for the clinical trials are widely spread in many medical related fields. They are usually used to test the effect of a new drug or a treatment when the outcome is binary. As a gold standard, subjects participating in the trial are randomly divided into two groups: treatment and placebo. The treatment group gets the actual drug and another one gets only imitation - the placebo. This way, the actual effect could be estimated after the trial is finished. However, there are several disease areas, where the high rate of placebo response makes some difficulties to evidentiate the clinically and statistically significant benefit of medication. These often include psychiatric disorders like depression, anxiety, schizophrenia and other [2, 7, 13, 16]. In these cases, using traditional randomization design is quite complicated because the required sample size to find statistically significant effect is greatly increasing. This leads to several issues: the research becomes more expensive, more time consuming and needs much more effort.

To overcome previously mentioned challenges, new clinical trial designs were developed. The latest one is the two-way enriched design (from now on it will be abbreviated as TED) introduced by A. Ivanova and R. N. Tamura [5]. The authors of the later article shows that TED is superior over other proposed designs by presenting that the sample size needed to acquire 80% power is smaller in the most cases. The possible areas of application of TED, including epilepsy, periodontal and even medical device clinical studies, already mentioned in some works [4, 6, 11]. However, the sample size could be reduced more by incorporating Bayes theory. Therefore, the purpose of this work is to show that Bayesian framework works better than proposed maximum likelihood approach for TED and that the expenses of trials with the high rate of placebo response could be downsized. The Bayesian approach is applied for the two statistical problems in TED: the parameters estimation and the hypothesis testing.

This master thesis is organized as follows. The first chapter of this work is dedicated for a brief summary of different proposed clinical trial designs, the original approach for TED and other findings from literature review. In the second chapter, Bayesian approach for TED is introduced. The main focus of that chapter is on the parameters estimation and the usage of different prior distributions. Also, the comparison of Bayes and maximum likelihood estimates on the simulated data is produced. The third chapter is devoted to the hypothesis testing and power simulations. All three cases of tests from the original approach are tested. And finally, all the most important findings and results are presented in the conclusions section.

1. Literature review

1.1. Clinical trial designs

Clinical trials designs dedicated for coping with the high rate of placebo response are usually called designs with enrichment strategies. They often consist of two sequential stages: one for filtering out a part of subjects and another one for the actual randomization. Here three of the most popular, besides TED, will be introduced in brief.

First one is called placebo lead-in (run-in) design. It is the first one proposed and it is quite straightforward. The placebo is given for all subjects at so called stage 0 and then only placebo non-responders are included into the randomization process for stage 1. In this way, the rate of placebo response is expected to be lower after stage 1 is finished. However, a few drawbacks could be already distinguished. The bigger sample size is needed and only the part of it is used in the hypothesis testing. Moreover, the real benefit of using this design is arguable based on some works which show that the effect size and placebo response rate do not change [8, 15].

Another design is randomized withdrawal design (also noted as randomized discontinuation). The concept is similar to placebo lead-in design, except this time the drug (or treatment) is given to all patients at stage 0 and only those who respond to the medication are included into the further analysis. It is believed that trial subjects who switched to placebo after randomization are more likely to lose the effect and this way the effect size should increase in overall. However, there are resembling disadvantages using this procedure like the greater amount of sample in stage 0 and longer study time. This kind of design is especially suggested when population of patients is heterogeneous and the health of only of a subset of it could be improved.

One more proposal for improving traditional parallel design and not so well prepared placebo lead-in design is to use the sequential parallel comparison design (SPCD) [2]. Trials of this type are conducted by first randomizing all patients to two groups like in the standard approach. In the second stage all placebo non-responders are randomly reassigned again. It is similar to placebo lead-in, however SPCD uses randomization in the initial stage as well. For the efficacy analysis, all available data from first stage and results of the placebo non-responders from the second stage is used. However, usually all patients are kept in the trial during both stages (for others the initial treatment is kept) for blinding purpose. To make sure that sufficient amount of subjects participate in the second stage, often higher proportion is allocated to placebo group at initial stage. This design attempts to ease the placebo response rate and also to lessen the required sample size. That is why this approach is popular among researchers and various models are developed to test the treatment

effect. SPCD is the convenient way to reduce overall length of the research, however since all the patients are enrolled in the study from beginning to an end, the novel design, two-way enriched design (TED), was proposed. It will be discussed in more details in the next subsection because it is the base of this master thesis.

1.2. Original approach for TED

As it was mentioned previously, TED was introduced by A. Ivanova and R. N. Tamura, and this subsection will be based on their article [5]. This design combines two approaches from previous subsection: randomized withdrawal design and SPCD. In the first stage all patients are randomly assigned to the placebo or drug. In the stage 2 only drug responders and placebo non-responders are re-randomized. Other subjects keep their initial treatment in the second stage, again, for blinding purpose, but their data is not used in efficacy analysis. TED is depicted in Figure 1. This new design is based on assumption that a drug, which works significantly more effectively than the placebo, will also be more effective in the maintenance of efficacy. Therefore, using it the required sample size could be reduced more than with SPCD.

In the article it is stated that there are two approaches for randomization. One is by using definition of TED which was introduced previously. Another one is to randomly assign patients to one of the four sequences placebo-placebo, placebo-drug, drug-placebo and drug-drug at the beginning of the trial. If the randomization is via Bernoulli random variables these two approaches are equiv-

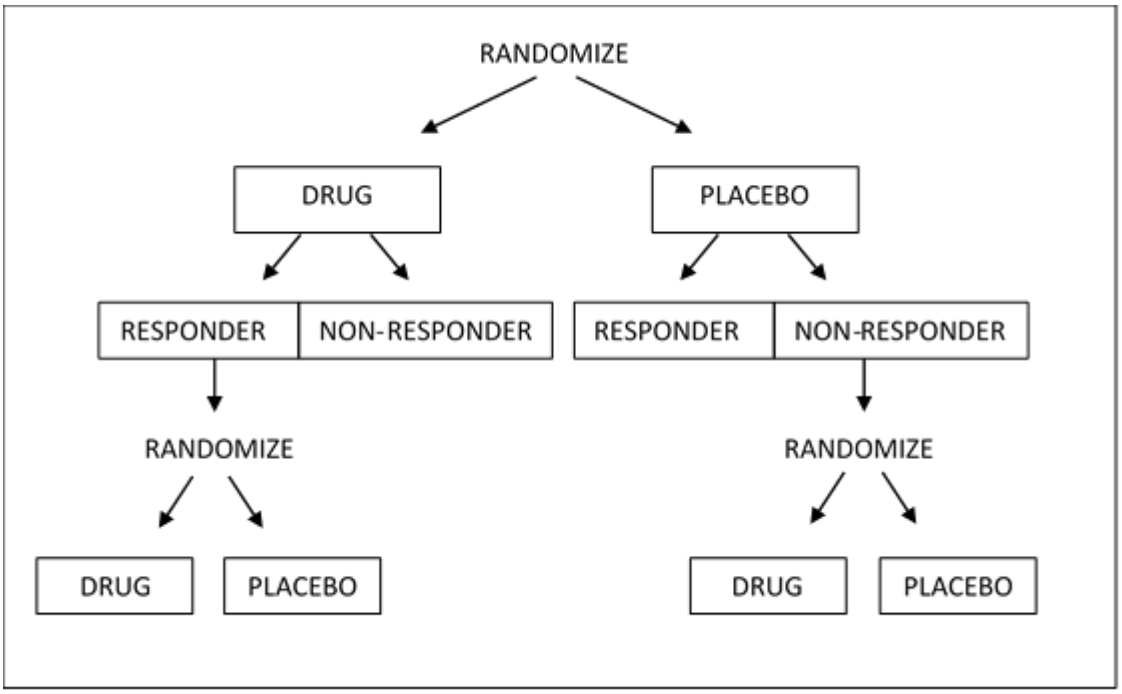

Figure 1. Two-way enriched design [5].

Table 1. Two-way enriched design.

Treatment		Response			
Stage 1	Stage 2	Stage 1	Stage 2	Count	Probability
Placebo	Placebo(n_1)	No	Yes	n_{11}	$s_2(1-q_1)q_2$
		No	No	n_{12}	$s_2(1-q_1)(1-q_2)$
		Yes	•	n_{13}	q_1
		No	missing	n_{14}	$(1-s_2)(1-q_1)$
Placebo	Drug (n_2)	No	Yes	n_{21}	$s_2(1-q_1)p_2$
		No	No	n_{22}	$s_2(1-q_1)(1-p_2)$
		Yes	•	n_{23}	q_1
		No	missing	n_{24}	$(1-s_2)(1-q_1)$
Drug	Placebo (n_3)	No	•	n_{31}	$(1-p_1)$
		Yes	Yes	n_{32}	$s_3p_1q_3$
		Yes	No	n_{33}	$s_3p_1(1-q_3)$
		Yes	missing	n_{34}	$(1-s_3)p_1$
Drug	Drug (n_4)	No	•	n_{41}	$(1-p_1)$
		Yes	Yes	n_{42}	$s_3p_1p_3$
		Yes	No	n_{43}	$s_3p_1(1-p_3)$
		Yes	missing	n_{44}	$(1-s_3)p_1$

Note. p_1: P(drug response in Stage 1), q_1: P(placebo response in Stage 1), p_2: P(drug response in Stage 2|placebo non-responder in Stage 1), q_2: P(placebo response in Stage 2|placebo non-responder in Stage 1), p_3: P(drug response in Stage 2|drug responder in Stage 1), q_3: P(placebo response in Stage 2|drug responder in Stage 1), s_2 is the proportion of placebo non-responders in Stage 1 who participate in Stage 2, s_3 is the proportion of drug responders in Stage 1 who participate in Stage 2. Responses denoted "•" are not included in the analysis by design, n_{14} and n_{24} are placebo non-responders and n_{34} and n_{44} are drug responders, who dropout and do not participate in Stage 2 [5].

alent. For simplification and consistency with the article, the latter one approach is used in this thesis. The overall sample size will be noted as n and corresponding sample sizes for sequences will be n_1, n_2, n_3 and n_4.

Some more definitions used in the article are necessary to understand the proposed approach. Define $b = (n_1 + n_2)/n$, as the proportion selected for placebo group in the first stage. To get the better view, further notations related to probabilities are introduced in Table 1. Based on these, the joint likelihood function for unknown parameters (p_1, q_1, p_2, q_2, p_3, q_3, s_2, s_3) with a little correction[1] is expressed like this:

$$\begin{aligned} L(p_1,q_1,p_2,q_2,p_3,q_3,s_2,s_3) \propto\ & p_1^{n_3+n_4-n_{31}-n_{41}}(1-p_1)^{n_{31}+n_{41}} q_1^{n_{23}+n_{13}}(1-q_1)^{n_1+n_2-n_{23}-n_{13}} \times \\ & \times p_2^{n_{21}}(1-p_2)^{n_{22}} q_2^{n_{11}}(1-q_2)^{n_{12}} p_3^{n_{42}}(1-p_3)^{n_{43}} q_3^{n_{32}}(1-q_3)^{n_{33}} \times \\ & \times s_2^{n_{11}+n_{12}+n_{21}+n_{22}}(1-s_2)^{n_{14}+n_{24}} s_3^{n_{32}+n_{33}+n_{42}+n_{43}}(1-s_3)^{n_{34}+n_{44}}. \end{aligned}$$

[1] in the article there are typing mistakes with the power of p_1 and $(1-p_1)$

However, for all the simulations and results, the authors consider parameters s_2 and s_3, called retention rates, fixed at 1. Therefore, in this thesis they also will be fixed and the simpler likelihood function will be exploited:

$$L(p_1, q_1, p_2, q_2, p_3, q_3) \propto p_1^{n_3+n_4-n_{31}-n_{41}}(1-p_1)^{n_{31}+n_{41}} q_1^{n_{23}+n_{13}}(1-q_1)^{n_1+n_2-n_{23}-n_{13}} \times \\ \times p_2^{n_{21}}(1-p_2)^{n_{22}} q_2^{n_{11}}(1-q_2)^{n_{12}} p_3^{n_{42}}(1-p_3)^{n_{43}} q_3^{n_{32}}(1-q_3)^{n_{33}}. \tag{1}$$

When the retention rates are not assumed to be equal to one, all the calculations could be quite easily modified. However, this thesis aims to compare obtained results with the presented in [5], so that part is omitted.

Ivanova and Tamura proposed three score tests for their newly developed design. They define the treatment effects $\Delta_1 = p_1 - q_1$, $\Delta_2 = p_2 - q_2$ and $\Delta_3 = p_3 - q_3$. For the first test - score test with one degree of freedom (DF), they assume that investigator might have some knowledge and define new parameters ρ_1 and ρ_2 such that $\Delta_1\rho_2 = \Delta_2$ and $\Delta_1\rho_3 = \Delta_3$. The test is derived under the assumption that $r_2 = \rho_2$ and $r_3 = \rho_3$, where r_2 and r_3 are known constants. Then new likelihood, which uses new parametrization of $p_1 = \Delta_1 + q_1$, $p_2 = r_2\Delta_1 + q_2$ and $p_3 = r_3\Delta_1 + q_3$, is presented. The hypothesis for testing is $H_0 : \Delta_1 = 0$, which implies that treatment effect equals 0 in both stages. For testing this hypothesis, score test is derived using maximum likelihood estimates under H_0. The presented test is asymptotic. More details could be found in the article. The test is dependent on three parameters b, r_2 and r_3. In fact, authors show that other clinical trial designs are the special cases of TED by changing these. For example when $b = 1$ it is placebo lead-in design, when $b = 0$ it is randomized withdrawal design, when $b = 0.5$ and $r_2 = r_3 = 0$ it is the traditional parallel design. Therefore, if the investigator is willing to make some assumptions, various combinations could be used, however, if making premise does not sound right, the recommended values are suggested $(b, r_2, r_3) = (0.5, 1, 1)$.

Other two tests are derived in a similar fashion. Score test with two degrees of freedom (DF) uses parametrization where only one of the treatment effect is proportional to the effect in stage 1: $p_2 = r_2\Delta_1 + q_2$ or $p_3 = r_3\Delta_1 + q_3$, and parameters to estimate are q_1, q_2, q_3, Δ_1 and one of Δ_2 or Δ_3. Score test with three DF uses parametrization with all three treatment effects Δ_1, Δ_2 and Δ_3.

For all three tests, the sample sizes, which are required to achieve desired power, are provided using optimal[2] and recommended test parameters b, r_1 and r_2. However, comparison with other designs are performed only for the first test. Authors show that the TED produces the lowest sample size in the most of the cases using recommended values and in all cases using the optimal values.

[2]the optimal $r_2 = (p_2 - q_2)/(q_1 - q_1), r_3 = (p_3 - q_3)/(q_1 - q_1)$ and the optimal b that maximizes power can be found via grid search

In this thesis, only the data related to recommended values $(b, r_2, r_2) = (0.5, 1, 1)$ will be analyzed.

1.3. Related works

During literature review, only one work which introduced a new approach for TED was found. It is a very recent article written by Liu et al [9] and published in June 2019. The authors briefly review the most important clinical trial designs and propose a novel method for the latest one - TED.

The new approach is based on the repeated measures model and it is justified because after the first and second stage the measurements are made for the same subjects. For binary outcome design, authors suggest to use logistic regression and, since there is a correlation between the outcomes in the two stages within the same patient, they apply generalized estimating equations (GEE) to estimate the treatment effect. Moreover, the approach with several different methods for continuous outcome is introduced in their work. These include repeated measures model, weighted repeated measures model with weights from propensity score and weights from K-means clustering. For more details refer to [9] because this master thesis is focused on the method for the clinical trials with binary outcome and no further analysis of that case is needed.

Turning to direct relation to this thesis, the effectiveness of the novel method for binary outcome was tested through a broad simulation of type I error and power. For evaluation 1 000 000 and 10 000 data sets for each scenario, taken from Ivanova's and Tamura's manuscript, were generated, accordingly. Power comparisons were made only with the score test with one degree of freedom. Simulation showed that type I error is well controlled in almost every case. However, the notable increase in the power was captured only in the situations, where one of the actual Δ_1, Δ_2 and Δ_3 equals to 0 and only with certain model parameters, as can be seen from Table 2. So, for other situations, there is no benefit of using this novel approach. Nevertheless, Liu et al extended TED application for the clinical trials with the continuous outcome, and that is the main achievement of

Table 2. Power testing for $H_0 : \Delta_1 = 0 \cap \Delta_2 = 0 \cap \Delta_3 = 0$ in Liu's et al manuscript [9].

p_1	q_1	p_2	q_2	p_3	q_3	N	Score Test(1df)	GEE $\omega_1 = 0.5$	GEE $\omega_1 = 0.6$	GEE $\omega_1 = 0.7$	GEE $\omega_1 = 0.8$	GEE $\omega_1 = 0.9$
0.4	0.3	0.4	0.3	0.9	0.8	412	0.79	0.71	0.76	0.77	0.75	0.68
0.5	0.3	0.4	0.2	0.9	0.8	128	0.8	0.65	0.71	0.76	0.77	0.74
0.5	0.3	0.4	0.1	0.9	0.8	96	0.82	0.64	0.7	0.73	0.71	0.65
0.4	0.3	0.4	0.3	0.9	0.7	312	0.8	**0.81**	**0.82**	0.8	0.73	0.61
0.5	0.3	0.4	0.2	0.9	0.7	104	0.81	0.7	0.74	0.75	0.73	0.66
0.5	0.3	0.4	0.1	0.9	0.7	80	0.82	0.66	0.71	0.71	0.67	0.6
0.4	**0.4**	0.4	0.3	0.9	0.8	2612	0.79	**0.96**	**0.91**	0.74	0.43	0.14
0.4	0.3	0.4	0.3	**0.9**	**0.9**	728	0.8	0.48	0.62	0.72	**0.81**	**0.83**
0.4	0.3	**0.4**	**0.4**	0.9	0.8	644	0.8	0.72	0.8	**0.84**	**0.85**	**0.83**

their work because in some fields only observation of continuous quantities make sense.

Therefore, after literature review it was assured that TED could definitely be useful in clinical trials and reducing the sample size would make it even more attractive to investigators. Furthermore, no attempts to apply Bayes theory for TED were found and the latest approach works better for only a few cases. This means that Bayes approach still eligible for improving the design.

2. Parameters estimation in TED

Parameters estimation is fundamental task in many applied statistics fields. The more accurate estimates, the more adequate conclusions. That is why the main purpose of this thesis is to show that Bayesian estimation in TED is more precise for small samples, which is crucial in medicine, than a maximum likelihood estimation. For this task, both methods will be introduced and applied on the simulated data.

2.1. Maximum likelihood estimation

2.1.1. Maximum likelihood method

Maximum likelihood estimation (noted as MLE) is the common approach in statistics and, as mentioned previously, Ivanova and Tamura applied it for TED as the base of their proposed tests. That is why this method is used as reference in this master thesis and the core principles of it will be reminded based on [1]. Suppose, $\mathbf{X}$ is a random vector having density $f(x;\theta), \theta \in \Theta \subset \mathbf{R}^m$, with respect to some σ-finite measure on $\mathbf{R}^m$, and the estimator $\hat{\theta}$ is needed. First of all, the likelihood function $L(\theta) = L_{\mathbf{X}}(\theta) = f(\mathbf{X};\theta)$ is found. It could be used as approximation of the probability that the sample is close to x given the value of θ. Based on this, it seems plausible to look for estimator by maximizing $L_x(\theta)$ over the parameter's θ space. The maximum likelihood (ML) estimator $\hat{\theta} = \hat{\theta}(\mathbf{X})$ is then obtained by changing realization x to $\mathbf{X}$ in the maximum expression and it satisfies condition

$$L(\hat{\theta}) = \sup_{\theta \in \Theta} L(\theta).$$

Usually, ML estimator is found by maximizing the function $l(\theta) = \ln L_{\mathbf{X}}(\theta)$. Logarithmic transformation is monotonic and continuous, consequently, it achieves maximum at the same point while it's expression often is simpler and more convenient to handle analytically. If function $l(\theta)$ achieves maximum and is differentiable (derivative exists) with respect to θ, ML estimator satisfies likelihood equations

$$\dot{l}(\theta) = 0; \tag{2}$$

where

$$\dot{l}(\theta) = \left(\frac{\partial l(\theta)}{\partial \theta_1}, \ldots, \frac{\partial l(\theta)}{\partial \theta_m} \right)^T$$

is the appropriate vector of derivatives. ML estimators are found by solving the system of equations (2) with respect to θ, however, sometimes this system is not solvable by hand, therefore, numeric methods are used to find solutions.

2.1.2. MLE in TED

In case of TED, the vector of unknown parameters is $\theta = (p_1, q_1, p_2, q_2, p_3, q_3) = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$. Log-likelihood function is $l(\theta) = (n_3 + n_4 - n_{31} - n_{41})\ln p_1 + (n_{31} + n_{41})\ln(1 - p_1) + (n_{23} + n_{13})\ln q_1 + (n_1 + n_2 - n_{23} - n_{13})\ln(1 - q_1) + n_{21}\ln p_2 + n_{22}\ln(1 - p_2) + n_{11}\ln q_2 + n_{12}\ln(1 - q_2) + n_{42}\ln p_3 + n_{43}\ln(1 - p_3) + n_{32}\ln q_3 + n_{33}\ln(1 - q_3)$, which is derived from (1). The system of equations (2) can be easily found and solved in this case, so no numeric methods are needed:

$$\begin{aligned} &\frac{\partial l(\theta)}{\partial p_1} = \frac{n_3 + n_4 - n_{31} - n_{41}}{p_1} - \frac{n_{31} + n_{41}}{1 - p_1} = 0 \\ &\vdots \\ &\frac{\partial l(\theta)}{\partial q_3} = \frac{n_{32}}{q_3} - \frac{n_{33}}{1 - q_3} = 0 \end{aligned}$$

The solution for the above system is:

$$\begin{aligned} \hat{p_1} &= \frac{n_3 + n_4 - n_{31} - n_{41}}{n_3 + n_4}, \quad \hat{q_1} = \frac{n_{23} + n_{13}}{n_1 + n_2} \\ \hat{p_2} &= \frac{n_{21}}{n_{21} + n_{22}}, \quad \hat{q_2} = \frac{n_{11}}{n_{11} + n_{12}} \\ \hat{p_3} &= \frac{n_{42}}{n_{42} + n_{43}}, \quad \hat{q_3} = \frac{n_{32}}{n_{32} + n_{33}} \end{aligned} \tag{3}$$

2.2. Bayesian estimation

2.2.1. Bayes theory

Bayesian statistics is very important part of this thesis, so some theoretical background is provided for the further understanding. This subsection is based on [3]. The essential thing, which makes Bayesian approach different from the frequentist one, is that the parameter of interest θ is treated as random rather than fixed and the sample is collected with only one observed value of parameter. In what follows, the brief description of the model corresponding to the case of absolutely continuous parameter θ is provided. Denote a random sample (usually consisting of independent identically distributed observations) as $Y = (Y_1, \ldots, Y_n)$ and the observed data as $y = (y_1, \ldots, y_n)$. The joint density function for θ and Y can be expressed as a product of two densities:

$$p(\theta, y) = \pi(\theta)p(y \mid \theta),$$

where $\pi(\theta)$ is the density of θ, which is called prior (the distribution is denoted accordingly), and $p(y \mid \theta)$ is the density of the data. Then the conditional density of θ given $Y = y$ can be derived using Bayes' theorem:

$$p(\theta \mid y) = \frac{p(\theta, y)}{p(y)} = \frac{\pi(\theta)p(y \mid \theta)}{p(y)},$$

where $p(y) = \int \pi(\theta)p(y \mid \theta)\mathrm{d}\theta$. The multiplier $\frac{1}{p(y)}$ is independent of θ and, therefore, changes nothing when it comes to maximization with respect to θ, so it can be omitted, and equivalent expression can be used:

$$p(\theta \mid y) \propto \pi(\theta)p(y \mid \theta). \tag{4}$$

The left side of (4) is called the posterior density. Moreover, it can be seen that $p(y \mid \theta)$ is the likelihood function. The formula above is the core for making conclusions about unknown parameter θ in Bayesian interference. For example, the posterior expectation $\mathrm{E}(\theta \mid Y = y)$ is usually used as a point estimate.

Bayes method is highly dependent on the choice of the prior distributions. Wrong choice could lead to the devastating conclusions. That is why this method gets a bunch of criticism. However, the right choice of prior can significantly improve performance of the Bayesian based model and increase the accuracy of the estimator. Nevertheless, there is no unique choice which prior distribution should be used, but there are a few options. These will be described in the next subsection.

2.2.2. Prior distributions

Usually, the prior distribution belongs to some parametric family, e.g. $\theta \sim Beta(\alpha; \beta)$. The corresponding parameters (in the example above, α and β) are called hyperparameters. One of the options is to take the informative distribution (both family and hyperparameters are chosen using additional information). They are based on the previous knowledge about θ or the opinion of an expert. It is very valuable, when the information is reliable, because additional data could be incorporated into the theoretical model.

When there is no trustworthy insights about the parameter, another option is to take one of the noninformative (also called objective) priors. These have only the minor influence to the posterior distribution. For example, it could be the uniform distribution or the normal distribution with a large dispersion. One of the frequently met in the literature is the Jeffreys prior. It is proportional to the determinant of the Fisher information matrix $\pi(\theta) = \sqrt{det(I(\theta))}$, where

$$I(\theta) = -\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln p(Y \mid \theta)\right].$$

This one is popular because it is usually computationally simple. Moreover, the Jeffreys prior is left invariant, which is useful property in many statistical problems.

One more objective prior is the reference prior. This one is the right invariant and in plentiful problems it has more desirable properties than the previously described Jeffreys. More details could be found in the [3]. However, derivation of the reference prior is not simple and that is the reason for the rare applications. Nevertheless, it will be employed in this thesis and the brief description is therefore provided. In the deriving of the reference prior, the parameters are assumed as having different importance. Hence, the order of the parameters is important. Suppose, case of d parameters is of interest and $\theta = (\theta_1, \ldots, \theta_d)$ is in order of parameters importance, i.e. θ_1 is the most important. Let $S = (I(\theta))^{-1}$ where $I(\theta)$ is Fisher information matrix and S_j be the $j \times j$ principal submatrix of S. Let $H_j = S_j^{-1}$ and h_j be the (j,j)-th element of H_j. Moreover, an increasing sequence of compact (closed and bounded) hyperrectangles $K_{1i} \times \cdots \times K_{di}$ whose union is Θ should be fixed. Then $p_i(\theta_d \mid \theta_1, \ldots, \theta_{d-1}) = c'_{di}(\theta_1, \ldots, \theta_{d-1})\sqrt{h_d}$ and the other conditional distributions are determined iteratively:

$$p_i(\theta_j \mid \theta_1, \ldots, \theta_{j-1}) = \begin{cases} c'_{ji}(\theta_1, \ldots, \theta_{j-1})\psi_j(\theta_1, \ldots, \theta_j) & \text{on } K_{ji} \\ 0 & \text{elsewhere,} \end{cases}$$

where

$$\psi_j(\theta_1, \ldots, \theta_j) = \exp\left\{ \int \frac{1}{2} \log h_j(\theta) p_i(\theta_{j+1}, \ldots, \theta_d \mid \theta_1, \ldots, \theta_j) d\theta_{j+1} \ldots d\theta_d \right\}$$

and $c'_{ji}(\theta_1, \ldots, \theta_{j-1})$ is normalizing constant such that

$$\int_{K_{ji}} p_i(\theta_j \mid \theta_1, \ldots, \theta_{j-1}) d\theta_j = 1.$$

Using the above expressions, the reference prior is then taken proportional to

$$\pi(\theta) \propto p_i(\theta_d \mid \theta_1, \ldots, \theta_{d-1}) p_i(\theta_{d-1} \mid \theta_1, \ldots, \theta_{d-2}) \ldots p_i(\theta_2 \mid \theta_1) p_i(\theta_1). \tag{5}$$

There is also another procedure, but this one is more convenient algebraically.

To conclude, there are several options for the choice of the prior distribution. Some of them were not even introduced in this subsection since this thesis will be limited to the presented ones as these are considered to be the most appropriate for TED.

2.2.3. MCMC method

Quite often the posterior distributions are complex and numerical methods are required to find the estimates. One of the most popular methods in Bayesian interference is the Markov Chain Monte Carlo (MCMC) method [3]. In a nutshell, its description is as follows. First, one constructs a Marko chain having limiting stationary distribution which coincides with that of θ. After that, one simulates a sufficiently long trajectory of the chain and makes use of the tail of that trajectory to approximate distribution of θ as well as other related quantities (moments, quantiles, etc.). The dropped part of the chain at the beginning of the trajectory is referred as *burn in*, and the reason for dropping is due to the fact that initial members have distribution which is usually not close enough to the stationary one. Hence may bias inference.

There are already several algorithms introduced for the simulation of the MCMC, but in this thesis only one, the Metropolis-Hastings algorithm, was used. It is a modified version of a random walk with acceptance/rejection rules for the convergence to the target distribution. This algorithm works as follows.

1. The initial value θ^0 is drawn from initial distribution $p_0(\theta)$.

2. For the next steps $t = 1, 2, \cdots$:

   (a) The value θ^* is proposed from the transitional distribution $J_t(\theta^* \mid \theta^{t-1})$ at time t. The transitional distribution is required to be symmetric, i.e

   $$J_t(\theta_a \mid \theta_b) = J_t(\theta_b \mid \theta_a) \quad \forall \theta_a, \theta_b, t.$$

   (b) The ratio of the densities is calculated

   $$r = \frac{p(\theta^* \mid y)}{p(\theta^{t-1} \mid y)}.$$

   (c) Then

   $$\theta^t = \begin{cases} \theta^* & \text{with probability } \min(r,1) \\ \theta^{t-1} & \text{otherwise.} \end{cases}$$

Also, one more modification is required when parameters are constrained. E.g. if condition $0 < \theta < 1$ has to be satisfied, then $\theta^t = \theta^*$ with probability 0 in the case of $\theta^* \leq 0$ or $\theta^* \geq 1$. It is proved that the sequence of the iterations $\theta^1, \theta^2, \ldots$ converges to the target distribution (inclusion of conditions for bounded parameters needs more iterations) by showing that simulated sequence

is Markov chain with unique stationary distribution which is the target distribution. Hence, this method is really helpful in many cases and this work is not the exception.

2.2.4. Bayesian estimation in TED

This subsection is devoted to description of application of Bayes' theory for the parameters estimation in TED. Several previously described priors were used to find the best combination with the same likelihood function (1), and here more details are provided.

The first one and the most common approach is to use the uniform $U(0,1)$ prior distribution for all the parameters. Before proceeding to that case, the definition of the Beta distribution $Beta(\alpha, \beta)$ is reminded. It has two parameters $\alpha > 0$ and $\beta > 0$ and the density function:

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathrm{B}(\alpha,\beta)}, \quad x \in [0,1].$$

Here $\mathrm{B}(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ and Γ is the Gamma function, which, for any positive integer n, attains value $\Gamma(n) = (n-1)!$. It is quite obvious that $Beta(1,1)$ and $U(0,1)$ are equivalent distributions with the density equal to 1 when $x \in [0,1]$. Therefore, it suffices to analyze the case with Beta prior distribution.

In general, all the parameters are assumed to be independent and the joint prior $\pi(\theta) = \pi(\theta_1)\pi(\theta_2)\pi(\theta_3)\pi(\theta_4)\pi(\theta_5)\pi(\theta_6) = \pi(p_1)\pi(q_1)\pi(p_2)\pi(q_2)\pi(p_3)\pi(q_3)$. Assuming that all the parameters have Beta distribution with the corresponding α_i and β_i, $\quad i = 1,\ldots,6$, and taking into account (4), the posterior distribution is given by

$$\begin{aligned} p(\theta \mid y) \propto\ & p_1^{n_3+n_4-n_{31}-n_{41}}(1-p_1)^{n_{31}+n_{41}} q_1^{n_{23}+n_{13}}(1-q_1)^{n_1+n_2-n_{23}-n_{13}} p_2^{n_{21}}(1-p_2)^{n_{22}} \times \\ & \times q_2^{n_{11}}(1-q_2)^{n_{12}} p_3^{n_{42}}(1-p_3)^{n_{43}} q_3^{n_{32}}(1-q_3)^{n_{33}} p_1^{\alpha_1-1}(1-p_1)^{\beta_1-1} q_1^{\alpha_2-1}(1-q_1)^{\beta_2-1} \times \\ & \times p_2^{\alpha_3-1}(1-p_2)^{\beta_3-1} q_2^{\alpha_4-1}(1-q_2)^{\beta_4-1} p_3^{\alpha_5-1}(1-p_3)^{\beta_5-1} q_3^{\alpha_6-1}(1-q_3)^{\beta_6-1}. \end{aligned}$$

It can be seen, that the posterior distribution is actually proportional to the product of the Beta distributions with appropriate parameters. In this situation, the prior is called conjugate. These types of priors have special property that the posterior distribution is of the same type as the prior (in this case, Beta). Then the expectations of the corresponding Beta distributions can be taken as

the estimators for the unknown parameters. Therefore, Bayesian estimators in this case are:

$$\begin{aligned}
\tilde{p_1} &= \frac{n_3 + n_4 - n_{31} - n_{41} + \alpha_1}{n_3 + n_4 + \alpha_1 + \beta_1}, \quad \tilde{q_1} = \frac{n_{23} + n_{13} + \alpha_2}{n_1 + n_2 + \alpha_2 + \beta_2} \\
\tilde{p_2} &= \frac{n_{21} + \alpha_3}{n_{21} + n_{22} + \alpha_3 + \beta_3}, \quad \tilde{q_2} = \frac{n_{11} + \alpha_4}{n_{11} + n_{12} + \alpha_4 + \beta_4} \\
\tilde{p_3} &= \frac{n_{42} + \alpha_5}{n_{42} + n_{43} + \alpha_5 + \beta_5}, \quad \tilde{q_3} = \frac{n_{32} + \alpha_6}{n_{32} + n_{33} + \alpha_6 + \beta_6}.
\end{aligned} \tag{6}$$

Different estimators can be derived by varying values of α and β. For the special uniform case, $\alpha_1 = \cdots = \alpha_6 = 1$ and $\beta_1 = \cdots = \beta_6 = 1$. This is uninformative prior. To model situations where some prior information is available, several informative priors were investigated as well. The first group of priors was based on ranges of values of parameters considered by Ivanova and Tamura [5]. Parameters p_1, q_1, p_2, q_2 values that were used in Ivanova's and Tamura's manuscript were not greater than 0.5 (in most cases less than 0.5), and values for the parameters p_3, q_3 were not less than 0.7. That is why distributions $Beta(1,2)$, $Beta(1,3)$ were used in this thesis as the informative priors for the parameters p_1, q_1, p_2, q_2, and $Beta(2,1)$, $Beta(3,1)$ for the parameters p_3, q_3, which makes the corresponding cases more likely to occur (see Figure 2).

There was one more informative prior tested in this work. Since in the most analyzed cases probabilities p_1, p_2, p_3 were higher than their corresponding probabilities q_1, q_2, q_3, the modified

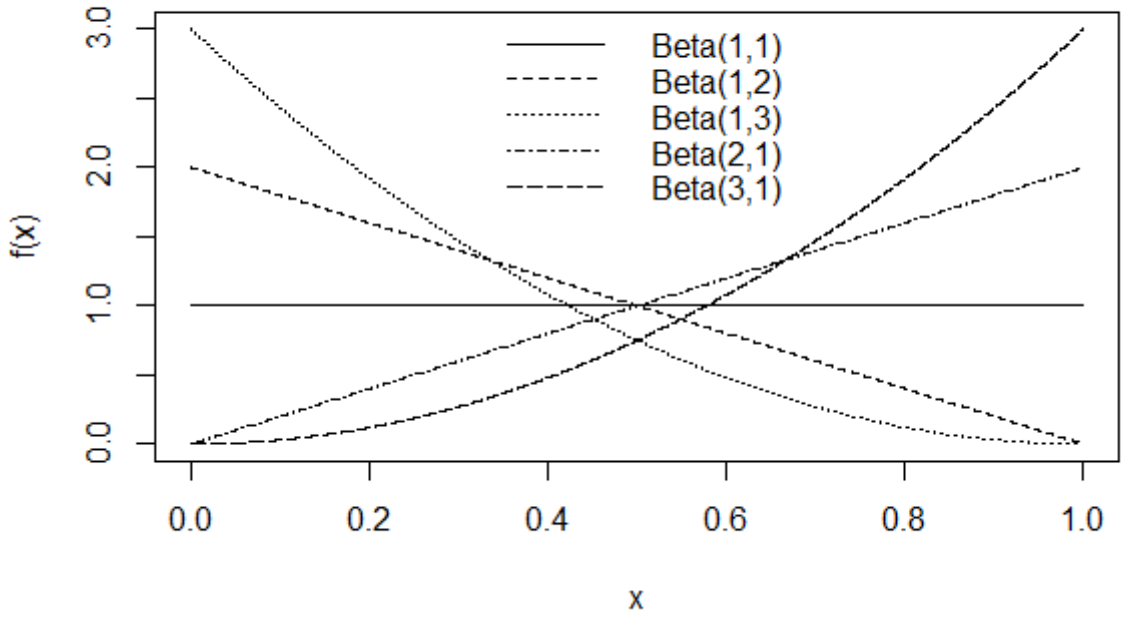

Figure 2. Density of the Beta priors.

uniform distribution was used: the same $U(0,1)$ prior was taken for the former parameters and $U(0,p_1), U(0,p_2), U(0,p_3)$ for the latter, accordingly. In this case, the posterior distribution was difficult to treat analytically, therefore the previously described MCMC method was applied for finding the estimates. Actually, the means of the MCMC simulations were taken as the estimates.

However, quite often assumptions about the data are not feasible and then the objective priors like $U(0,1)$ are used. One more previously mentioned non informative prior investigated in this thesis was Jeffreys. To derive its closed form, Fisher information matrix is needed. Let

$$\begin{aligned} &n_{p1} = n_3 + n_4 - n_{31} - n_{41}, \quad n_{p12} = n_{31} + n_{41}, \quad n_{q11} = n_{23} + n_{13}, \\ &n_{q12} = n_1 + n_2 - n_{23} + n_{13}, \quad n_{p21} = n_{21}, \quad n_{p22} = n_{22}, \quad n_{q21} = n_{11}, \quad n_{q22} = n_{12}, \\ &n_{p31} = n_{42}, \quad n_{p32} = n_{43}, \quad n_{q31} = n_{32}, \quad n_{q32} = n_{33} \end{aligned} \tag{7}$$

and $\theta = (p_1, q_1, p_2, q_2, p_3, q_3) = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$. Then $l(\theta) = n_{p11}\ln p_1 + n_{p12}\ln(1-p_1) + n_{q11}\ln q_1 + n_{q12}\ln(1-q_1) + n_{p21}\ln p_2 + n_{p22}\ln(1-p_2) + n_{q21}\ln q_2 + n_{q22}\ln(1-q_2) + n_{p31}\ln p_3 + n_{p32}\ln(1-p_3) + n_{q31}\ln q_3 + n_{q32}\ln(1-q_3)$ and

$$\begin{aligned} \frac{\partial l(\theta)}{\partial \theta_1} = \frac{\partial l(\theta)}{\partial p_1} &= \frac{n_{p11}}{p_1} - \frac{n_{p12}}{1-p_1}, \\ \frac{\partial l(\theta)}{\partial \theta_1^2} &= -\frac{n_{p11}}{p_1^2} - \frac{n_{p12}}{(1-p_1)^2} \\ -\mathrm{E}\left[\frac{\partial l(\theta)}{\partial \theta_1^2}\right] &= \frac{\mathrm{E}n_{p11}}{p_1^2} + \frac{\mathrm{E}n_{p12}}{(1-p_1)^2}, \end{aligned}$$

where $\mathrm{E}n_{p11} = n_3 + n_4 - n_3(1-p_1) - n_4(1-p_1) = (n_3+n_4)p_1$ and $\mathrm{E}n_{p12} = (n_3+n_4)(1-p_1)$ by using Table 1. It is pretty obvious that $-\mathrm{E}\left[\frac{\partial l(\theta)}{\partial \theta_i \theta_j}\right] = -\mathrm{E}[0], \quad i \neq j, \quad i,j = 1,\ldots,6$ and the other expectations $-\mathrm{E}\left[\frac{\partial l(\theta)}{\partial \theta_k^2}\right], \quad k = 2,\ldots,6$ could be derived similarly with

$$\begin{aligned} &\mathrm{E}n_{q11} = (n_1+n_2)q_1, \quad \mathrm{E}n_{q12} = (n_1+n_2)(1-q_1), \quad \mathrm{E}n_{p21} = n_2(1-q_1)p_2, \\ &\mathrm{E}n_{p22} = n_2(1-q_1)(1-p_2), \quad \mathrm{E}n_{q21} = n_1(1-q_1)q_2, \quad \mathrm{E}n_{q22} = n_1(1-q_1)(1-q_2), \\ &\mathrm{E}n_{p31} = n_4 p_1 p_3, \quad \mathrm{E}n_{p32} = n_4 p_1(1-p_3), \quad \mathrm{E}n_{q31} = n_3 p_1 q_3, \quad \mathrm{E}n_{q32} = n_3 p_1(1-q_3). \end{aligned}$$

Thus, the Fisher information matrix is

$$I(\theta) = \begin{pmatrix} \frac{n_3+n_4}{p_1(1-p_1)} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{n_1+n_2}{q_1(1-q_1)} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{n_2(1-q_1)}{p_2(1-p_2)} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{n_1(1-q_1)}{q_2(1-q_2)} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{n_4 p_1}{p_3(1-p_3)} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{n_3 p_1}{q_3(1-q_3)} \end{pmatrix},$$

and the prior $\pi(\theta) = \sqrt{\mid I(\theta) \mid}$, where the determinant equals to the product of the diagonal elements. Again, one can see that the posterior distribution is proportional to the product of the Beta distributions. Therefore, the estimators using Jeffreys prior are obtained from (6) with $\alpha_1 = \frac{3}{2}, \quad \beta_1 = \frac{1}{2}, \quad \alpha_2 = \frac{1}{2}, \quad \beta_2 = \frac{3}{2}, \quad \alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = \frac{1}{2}$ and $\beta_3 = \beta_4 = \beta_5 = \beta_6 = \frac{1}{2}$.

Finally, the last of the previously described reference prior distributions which was applied in TED was the reference prior. To derive its form, first of all the order of the parameters importance has to be declared. In this case, the results in the treatment group are considered to be more important than in the placebo group and the first stage is believed to be more meaningful rather than the second. That is why the vector of parameters $\theta = (p_1, p_2, p_3, q_1, q_2, q_3)$ is analyzed and the Fisher information matrix then has the different ordering on the diagonal:

$$I(\theta) = \begin{pmatrix} \frac{n_3+n_4}{p_1(1-p_1)} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{n_2(1-q_1)}{p_2(1-p_2)} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{n_4 p_1}{p_3(1-p_3)} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{n_1+n_2}{q_1(1-q_1)} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{n_1(1-q_1)}{q_2(1-q_2)} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{n_3 p_1}{q_3(1-q_3)} \end{pmatrix}.$$

Although, in this special case, the positioning brings no essential difference as it is seen later. Moreover, in this case, H_j coincides with (see Subsection 2.2.2) the $j \times j$ principal submatrix of $I(\theta)$ because only the elements on the diagonal are not equal 0. Consequently, $h_1 = \frac{n_3+n_4}{p_1(1-p_1)}, \ldots, h_6 = \frac{n_3 p_1}{q_3(1-q_3)}$ and, keeping the sequence of rectangles $K_{1i} \times \cdots \times K_{6i}$ fixed, the first conditional distribution is given by

$$p_i(\theta_6 \mid \theta_1, \ldots, \theta_5) = c'_{6i}(\theta_1, \ldots, \theta_5) \sqrt{\frac{n_3 \theta_1}{\theta_6(1-\theta_6)}} \quad \text{on } K_{6i},$$

where

$$\int_{K_{6i}} c'_{6i}(\theta_1,\dots,\theta_5)\sqrt{\frac{n_3\theta_1}{\theta_6(1-\theta_6)}}d\theta_6 = c'_{6i}(\theta_1,\dots,\theta_5)\sqrt{n_3\theta_1}arcsin(2\theta_6-1)\mid_{K_{6i}}=1 \implies$$

$$\implies c'_{6i}(\theta_1,\dots,\theta_5) = c_{6i}\frac{1}{\sqrt{\theta_1}}.$$

Hence, the expression of the first conditional distribution:

$$p_i(\theta_6 \mid \theta_1,\dots,\theta_5) = \frac{\tilde{c}_{6i}}{\sqrt{\theta_6(1-\theta_6)}}.$$

The next conditional distribution is derived like this:

$$\begin{aligned}
p_i(\theta_5 \mid \theta_1,\dots,\theta_4) &= c'_{5i}(\theta_1,\dots,\theta_4)\psi_5(\theta_1,\dots,\theta_5) \quad \text{on } K_{5i},\\
\psi_5(\theta_1,\dots,\theta_5) &= \exp\left\{\int_{K_{6i}} \frac{1}{2}\log h_5 p_i(\theta_6 \mid \theta_1,\dots,\theta_5)d\theta_6\right\}\\
&= \exp\left\{\int_{K_{6i}} \frac{1}{2}\log\frac{n_1(1-\theta_4)}{\theta_5(1-\theta_5)}\frac{\tilde{c}_{6i}}{\sqrt{\theta_6(1-\theta_6)}}d\theta_6\right\}\\
&= \sqrt{\frac{n_1(1-\theta_4)}{\theta_5(1-\theta_5)}}\exp\left\{\int_{K_{6i}} \frac{\tilde{c}_{6i}}{\sqrt{\theta_6(1-\theta_6)}}d\theta_6\right\}\\
&= \sqrt{\frac{n_1(1-\theta_4)}{\theta_5(1-\theta_5)}}c_6^*,\\
c'_{5i}(\theta_1,\dots,\theta_4) &= c_{5i}\frac{1}{\sqrt{1-\theta_4}} \implies\\
p_i(\theta_5 \mid \theta_1,\dots,\theta_4) &= \frac{\tilde{c}_{5i}}{\sqrt{\theta_5(1-\theta_5)}}.
\end{aligned}$$

The conditional distributions for parameters θ_4, θ_3 and the distribution $p_i(\theta_1)$ are obtained in the same way and are very similar. For θ_2, there is the difference in ψ function, because the element h_2 depends on the θ_4, but it does not make any significant impact:

$$\begin{aligned}
\psi_2(\theta_1,\theta_2) &= \exp\left\{\int_{K_{6i}\times\cdots\times K_{3i}} \frac{1}{2}\log h_2 p_i(\theta_3,\dots,\theta_6 \mid \theta_1,\theta_2)d\theta_6\dots d\theta_3\right\}\\
&= \exp\left\{\int_{K_{6i}\times\cdots\times K_{3i}} \frac{1}{2}\log\frac{n_2(1-\theta_4)}{\theta_2(1-\theta_2)}\frac{\tilde{c}_{6i}}{\sqrt{\theta_6(1-\theta_6)}}\dots\frac{\tilde{c}_{3i}}{\sqrt{\theta_3(1-\theta_3)}}d\theta_6\dots d\theta_3\right\}\\
&= \sqrt{\frac{n_2}{\theta_2(1-\theta_2)}}\exp\left\{\int_{K_{6i}\times\cdots\times K_{3i}} \frac{\log\sqrt{1-\theta_4}\tilde{c}_{6i}}{\sqrt{\theta_6(1-\theta_6)}}\dots\frac{\tilde{c}_{3i}}{\sqrt{\theta_3(1-\theta_3)}}d\theta_6\dots d\theta_3\right\}\\
&= \sqrt{\frac{n_2}{\theta_2(1-\theta_2)}}c_2^*.
\end{aligned}$$

Everything else is equivalent to the procedure that was showed before and the distribution has anal-

ogous expression. Note that all the distributions are proportional to the part of the corresponding h_j which is dependent on the θ_j regardless the order of the parameters. Furthermore, the expressions of h_j will also be the same, just with different indexes when the order is changed, because the Fisher information matrix has only the elements on the diagonal. Thus, in this case, the reference prior is proportional to the same combination of distributions regardless of the parameters placing and it is obtained by using (5):

$$\pi(\theta) \propto \frac{1}{\sqrt{q_3(1-q_3)}} \frac{1}{\sqrt{q_2(1-q_2)}} \frac{1}{\sqrt{q_1(1-q_1)}} \frac{1}{\sqrt{p_3(1-p_3)}} \frac{1}{\sqrt{p_2(1-p_2)}} \frac{1}{\sqrt{p_1(1-p_1)}}.$$

And once again, the posterior is proportional to the product of the Beta distributions and the estimators are obtained using (6) with $\alpha_1 = \cdots = \alpha_6 = \frac{1}{2}$ and $\beta_1 = \cdots = \beta_6 = \frac{1}{2}$.

At this point the theoretical part of the parameters estimation is finished. In many cases, the Bayesian estimators in TED are just special cases of Beta-Binomial model with different α and β values. In the next subsection, using careful simulations of the different situations, the question which one of them is the best in particular settings is answered.

2.3. Comparison of methods via data simulation

Data simulation was performed using R computing environment [12]. Nine different combinations of the parameters, taken from Ivanova's and Tamura's manuscript, were analysed. For each case, three sample sizes - 60, 100 and 160 - were tested. These sizes were selected because the main focus of this thesis was to show that the Bayesian estimates are more accurate for the small samples. For each scenario, 10 000 datasets were generated (Figure 3). Data was simulated by generating the $n_i \quad (i = 1,2,3,4)$ values from $U(0,1)$ distribution in every of the four groups: placebo-placebo, placebo-drug, drug-placebo and drug-drug. Then, for example in the placebo-placebo group, the number of generated values x which satisfy condition $x \leq (1-q_1)q_2$ is n_{11}, the number of x's

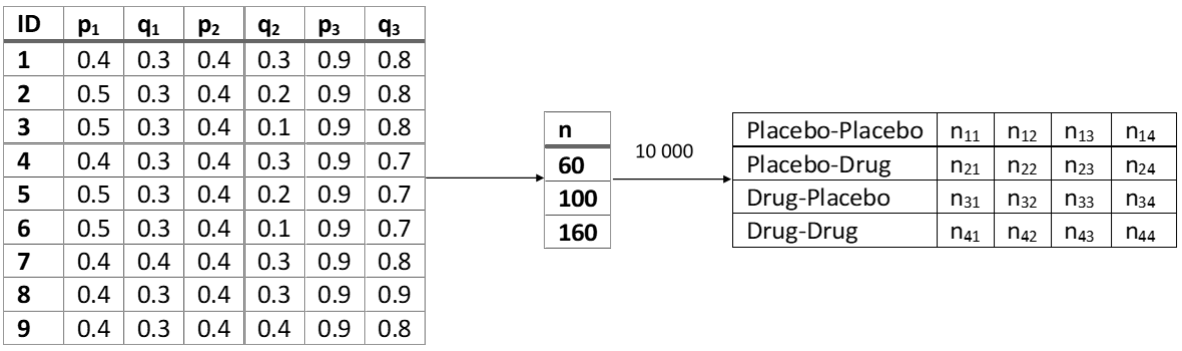

Figure 3. Combinations for the simulation of the data.

such that $(1-q_1)q_2 < x \leq 1-q_1$ is n_{12}, and the number of the other values is n_{13}. Overall, $9 \times 3 \times 6 \times 10000$ estimates were obtained for each method. The ML estimates were evaluated using (3) and the Bayesian estimates - using (6) with different parameters and MCMC method with 10 000 simulations and 1000 burn in as described in Subsection 2.2.4.

The main purpose of the simulations was to compare the different methods and the most important one — to compare the Bayes method versus MLE. Several metrics were calculated for that. The first one was the bias. Let m be the number of the simulations (in previously mentioned case $m = 10000$) and $\tilde{\theta}_i$ be the estimate of the parameter θ during the $i-th$ simulation. Then the estimator of the bias:

$$\widehat{bias}(\theta) = \theta - \frac{1}{m}\sum_{i=1}^{m}\tilde{\theta}_i.$$

The next one measure was the root of the mean squared errors (RMSE):

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(\theta - \tilde{\theta}_i)^2}.$$

These two metrics were calculated for every parameter in every configuration of the parameter set and n. The third one was quite different. It was the Kullback–Leibler (KL) divergence [10]. For discrete probability distributions P and Q defined on the same probability space:

$$D(P \mid\mid Q) = \sum_{x \in \Omega} P(x) \ln \frac{P(x)}{Q(x)}.$$

Thus, assuming that Q is the distribution with the true parameters and P - with estimated ones, the equivalence of these two distributions can be tested. E.g., using the first configuration with $q_1 = 0.3$ and $q_2 = 0.3$ in the placebo-placebo group, the discrete distribution Q can be defined this way: $\Omega = \{1, 2, 3\}$, $Q(1) = (1-0.3)0.3 = 0.21$, $\quad Q(2) = (1-0.3)(1-0.3) = 0.49$ and $Q(3) = 0.3$ (see Table 1). Let $\tilde{q}_{1i}$ and $\tilde{q}_{2i}$ be the estimates of the corresponding parameters during the $i-th$ iteration. Then the distribution P_i could be defined like this: $\Omega = \{1, 2, 3\}$, $P_i(1) = (1-\tilde{q}_{1i})\tilde{q}_{2i}$, $\quad P(2) = (1-\tilde{q}_{1i})(1-\tilde{q}_{2i})$ and $P(3) = \tilde{q}_{2i}$, and KL divergence $D_i(P_i \mid\mid Q)$ could be easily calculated. The KL divergence that was used to compare the different methods was the average of D_i over all the iterations: $D(P \mid\mid Q) = \frac{1}{m}\sum_{i=1}^{m} D_i(P_i \mid\mid Q)$. For other groups and configurations, KL divergence was estimated in an analogical manner. The KL divergence, which is close to zero, indicates that the two distributions are almost identical.

Therefore, in total, RMSE and bias were evaluated for $9 \times 3 \times 6$ configurations and KL divergence was calculated for $9 \times 3 \times 4$ cases for each of the seven approaches. Below, only the main insights

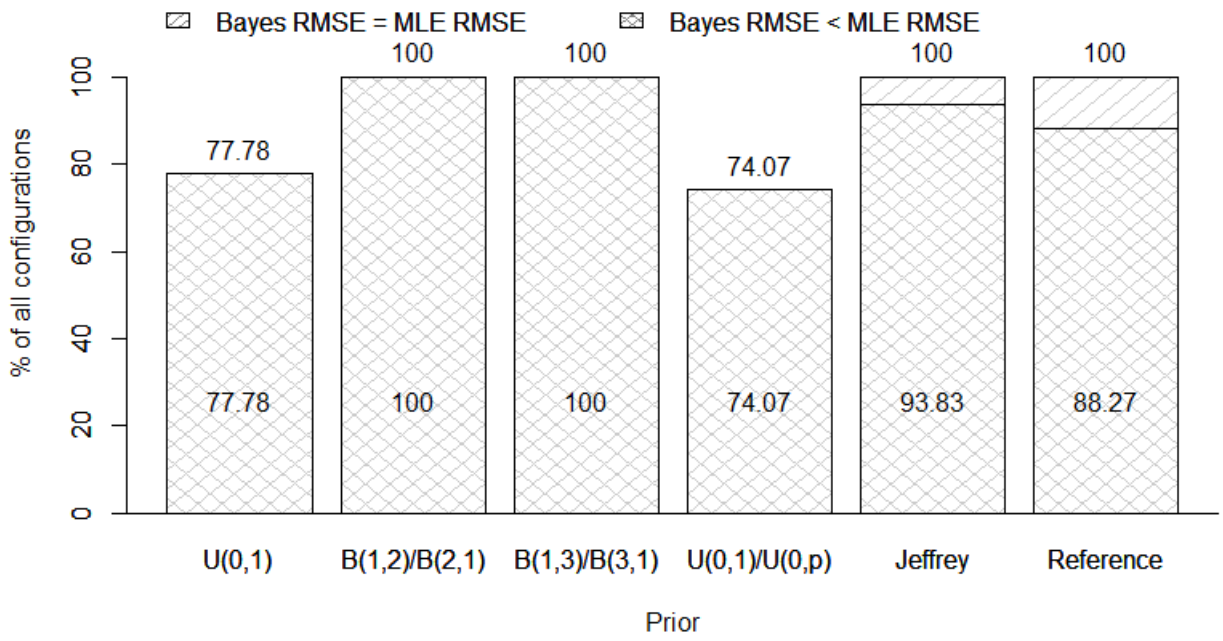

Figure 4. Comparison of the RMSE between Bayes method with the different priors and MLE.

regarding the results are discussed. All the evaluated measures from the simulation are provided in the appendix. First of all, after the examination of the results, it was found that none of the Bayesian approaches had lower bias in absolute value than ML approach in almost all the cases (MLE bias was the lowest in 99.38% cases). However, there was absolutely different situation when looking at the RMSE. Bayesian method with all the priors worked better when comparing RMSE (Figure 4), but some of them gave better results than the others. Case with the semi informative $Beta(1,2)$ or $Beta(2,1)$ [B(1,2)/B(2,1) in the Figure 4] and $Beta(1,3)$ or $Beta(3,1)$ [B(1,3)/B(3,1)] priors in all the cases returned lower RMSE than MLE. Meanwhile, using the other informative prior — combination of $U(0,1)$ and $U(0,p_1)$, $U(0,p_2)$, $U(0,p_3)$ [U(0,1)/U(0,p)] — gave inferior results and that is justified, because not all the cases satisfy assumption that was used in constructing the prior. Actually, in the 45.06% of the configurations, this approach has returned the minimum RMSE. Nevertheless, the combination of $Beta(1,3)$ and $Beta(3,1)$ as the prior surpassed it and, consequently, RMSE were the minimal in 56.79% of the cases. Objective priors did not excel other priors based on the minimum errors, however in the most cases they worked better than MLE. The best of them was Jeffreys (based on the percentage of the configurations with lower RMSE than obtained by using the MLE). Similar conclusions could be derived when comparing KL divergence (Figure 5), but the superiority of the Bayesian estimates was not so evident as it was with the RMSE. Again, the combination of the $Beta(1,3)$ and $Beta(3,1)$ as the prior worked better than MLE in

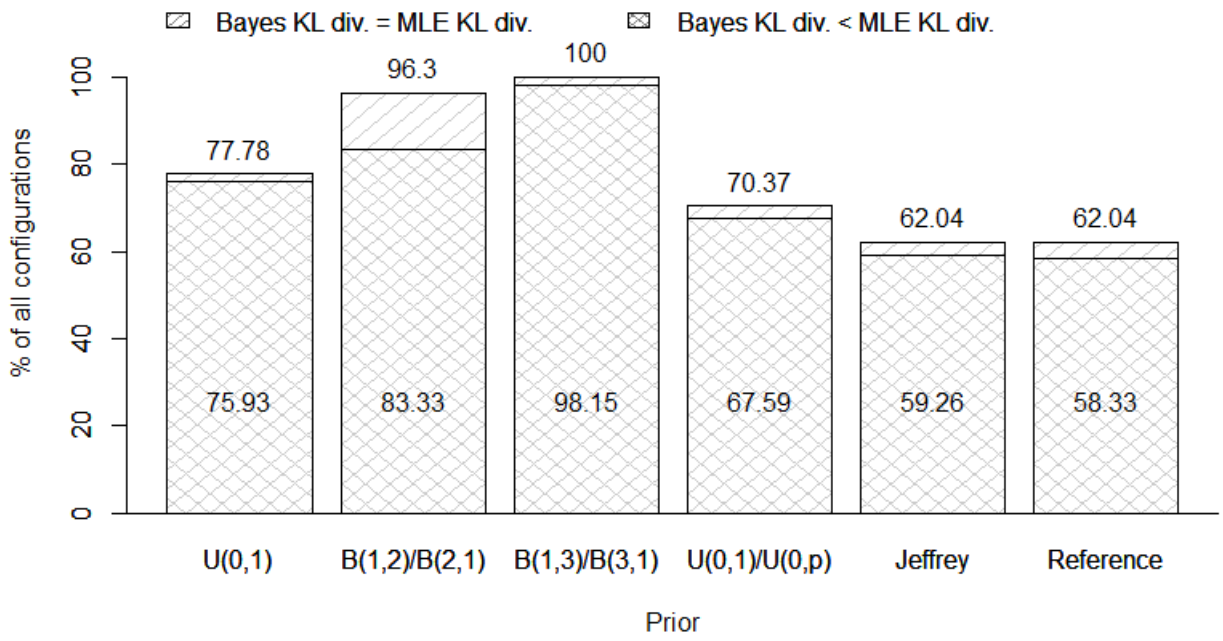

Figure 5. Comparison of the KL divergence between Bayes method with the different priors and MLE.

almost all the cases and the KL divergence was minimal in around 60% of them. However, the approach with $U(0,1)$ was the best among the not informative priors with 75.93% of cases having the divergence less than the divergence obtained by using the MLE.

In conclusion, the ML estimates are less biased than Bayesian. However, the more meaningful metric in this case is the KL divergence because it lets to look at the similarity of the distributions rather than the individual parameters. And based on that, the Bayes approach is better, because KL divergence is lower in more than 50% of the cases regardless of the prior. The RMSE are not contradictory to this conclusion either. Usage of the modest objective priors, i.e., incorporating just a little of additional information, gives the better estimates than the MLE. And if the researcher is willing to include more reliable information from the previous or pilot studies, then the estimates become even more accurate.

3. Hypothesis testing in TED using Bayes approach

There are two main fields in the applied statistics: the parameters estimation and the hypothesis testing. The first one was discussed previously and it was shown that Bayesian inference works better. In this section, the second one is analysed. However, in this thesis, only one quite simple approach will be presented because the core of this thesis was the fist part. That approach is to use the credible intervals [3, 14]. These are similar to the confidence intervals in the classical statistics, but interpretation is quite different.

Let the unknown parameter be $\theta \in \Theta \subset \mathbf{R}$ and $C \subset \Theta$ be the $(1-\alpha)100\%$ credible interval. Then it satisfies the condition:

$$P(\theta \in C \mid X = x) \geq 1 - \alpha,$$

where $\alpha \in (0,1)$ and x is the observed sample. For hypothesis testing, it is suggested to use the equal-tailed or one-sided $1-\alpha$ credible interval. With $q_\alpha(\theta \mid x)$ denoting the posterior quantile function of θ, such that $P(\theta < q_\alpha(\theta \mid x)) = \alpha$, these two intervals are $(q_{\alpha/2}(\theta \mid x), q_{1-\alpha/2}(\theta \mid x))$ and $(q_\alpha(\theta \mid x), +\infty)$ or $(-\infty, q_{1-\alpha}(\theta \mid x))$ accordingly. Moreover, only the objective priors are recommended to use when implementing this approach. Based on the previous section, the uniform and Jeffreys priors were tested because the reference prior is computationally difficult and no superiority over the Jeffreys was noticed in this case. The hypothesis testing using the credible intervals is analogous to testing with the confidence intervals in the classical statistics. Suppose $H_0 : \theta = \theta_0$ holds, then null hypothesis is accepted if $\theta_0 \in C$ and rejected otherwise.

Furthermore, there is another concept proposed from Bayesian perspective. Let the hypothesis for testing is $H_0 : \theta \in \Theta_0$ with alternative $H_1 : \theta \in \Theta_1$, where Θ_0 and Θ_1 are of the same dimension as for one-sided null and alternative. Then the posterior odds ratio $P(\Theta_0|x)/P(\Theta_1|x)$ is calculated and then using some threshold, e.g. $\frac{0.05}{0.95} = \frac{1}{19}$, conclusion about the hypothesis is made. However, by looking at this more closer, it can be seen that it is the same as using the credible interval.

3.1. New parametrization and priors for one DF test

The primary objective of this part is to present the hypothesis testing alternative for the one DF test (see Subsection 1.3). However, for the consistency purpose, performance of the same procedure is also examined for other two tests described in Subsection 1.3. Thus, the new parametrization for

likelihood function (1) is used:

$$L(\Delta_1,\Delta_2,\Delta_3,q_1,q_2,q_3) \propto (\Delta_1+q_1)^{n_3+n_4-n_{31}-n_{41}}(1-\Delta_1-q_1)^{n_{31}+n_{41}}q_1^{n_{23}+n_{13}}\times$$
$$\times(1-q_1)^{n_1+n_2-n_{23}-n_{13}}(\Delta_2+q_2)^{n_{21}}(1-\Delta_2-q_2)^{n_{22}}\times \quad (8)$$
$$\times q_2^{n_{11}}(1-q_2)^{n_{12}}(\Delta_3+q_3)^{n_{42}}(1-\Delta_3-q_3)^{n_{43}}q_3^{n_{32}}(1-q_3)^{n_{33}},$$

where $\Delta_i = p_i - q_i, \quad i = 1,2,3$ and the same assumptions, that $\Delta_2 = r_2\Delta_1$ and $\Delta_3 = r_3\Delta_1$ as for one DF test are valid. Only the case with $r_2 = r_3 = 1$ is analyzed. The general case could be studied similarly.

This means that the vector of parameters for the estimation $\theta = (\Delta_1, q_1, q_2, q_3) = (\theta_1, \theta_2, \theta_3, \theta_4)$ and the hypothesis for testing is $H_0 : \Delta_1 \leq 0$. In the original approach the equality was used, but, in general, the values less than zero are also in the favor that the treatment was not effective enough. This hypothesis was tested by using the 0.95 credible interval $(q_{0.05}(\Delta_1 \mid x), +\infty)$ of Δ_1. If the lower bound is more than zero, the null hypothesis is rejected, which means that the treatment is effective.

For Bayesian approach, the priors are needed, and, as mentioned before, two of them were tested. The first one was simple: $\Delta_1 \sim U(-1,1)$, $q_1,q_2,q_3 \sim U(0,1)$ and $\pi(\theta) \propto U(-1,1)U(0,1)U(0,1)U(0,1)$. The second one was Jeffreys and the determinant of the Fisher information matrix was needed, which is not as simple as it was in the second section. The following is the derivation. First, the log-likelihood function is obtained using the same definitions (7):

$$\begin{aligned} l(\theta) =& n_{p11}\ln(\theta_1+\theta_2) + n_{p12}\ln(1-\theta_1-\theta_2) + n_{q11}\ln\theta_2 + n_{q12}\ln(1-\theta_2) + n_{p21}\ln(\theta_1+\theta_3) + \\ &+ n_{p22}\ln(1-\theta_1-\theta_3) + n_{q21}\ln\theta_3 + n_{q22}\ln(1-\theta_3) + n_{p31}\ln(\theta_1+\theta_4) + \\ &+ n_{p32}\ln(1-\theta_1-\theta_4) + n_{q31}\ln\theta_4 + n_{q32}\ln(1-\theta_4). \end{aligned}$$

Then the first derivatives are found:

- $l_1 = \frac{\partial l(\theta)}{\partial\theta_1} = \frac{n_{p11}}{\theta_1+\theta_2} - \frac{n_{p12}}{1-\theta_1-\theta_2} + \frac{n_{p21}}{\theta_1+\theta_3} - \frac{n_{p22}}{1-\theta_1-\theta_3} + \frac{n_{p31}}{\theta_1+\theta_4} - \frac{n_{p32}}{1-\theta_1-\theta_4}$;
- $l_2 = \frac{\partial l(\theta)}{\partial\theta_2} = \frac{n_{p11}}{\theta_1+\theta_2} - \frac{n_{p12}}{1-\theta_1-\theta_2} + \frac{n_{q11}}{\theta_2} - \frac{n_{q12}}{1-\theta_2}$;
- $l_3 = \frac{\partial l(\theta)}{\partial\theta_3} = \frac{n_{p21}}{\theta_1+\theta_3} - \frac{n_{p22}}{1-\theta_1-\theta_3} + \frac{n_{q21}}{\theta_3} - \frac{n_{q22}}{1-\theta_3}$;
- $l_4 = \frac{\partial l(\theta)}{\partial\theta_4} = \frac{n_{p31}}{\theta_1+\theta_4} - \frac{n_{p32}}{1-\theta_1-\theta_4} + \frac{n_{q31}}{\theta_4} - \frac{n_{q32}}{1-\theta_4}$.

Let $l_{ij} = \frac{\partial l_i}{\partial\theta_j}, \quad i,j = 1,2,3,4$. Then the second derivatives are:

- $l_{11} = -\frac{n_{p11}}{(\theta_1+\theta_2)^2} - \frac{n_{p12}}{(1-\theta_1-\theta_2)^2} - \frac{n_{p21}}{(\theta_1+\theta_3)^2} - \frac{n_{p22}}{(1-\theta_1-\theta_3)^2} - \frac{n_{p31}}{(\theta_1+\theta_4)^2} - \frac{n_{p32}}{(1-\theta_1-\theta_4)^2}$;

- $l_{12} = l_{21} = -\frac{n_{p11}}{(\theta_1+\theta_2)^2} - \frac{n_{p12}}{(1-\theta_1-\theta_2)^2}$;
- $l_{13} = l_{31} = -\frac{n_{p21}}{(\theta_1+\theta_3)^2} - \frac{n_{p22}}{(1-\theta_1-\theta_3)^2}$;
- $l_{14} = l_{41} = -\frac{n_{p31}}{(\theta_1+\theta_4)^2} - \frac{n_{p32}}{(1-\theta_1-\theta_4)^2}$;
- $l_{22} = -\frac{n_{p11}}{(\theta_1+\theta_2)^2} - \frac{n_{p12}}{(1-\theta_1-\theta_2)^2} - \frac{n_{q11}}{\theta_2^2} - \frac{n_{q12}}{(1-\theta_2)^2}$;
- $l_{33} = -\frac{n_{p21}}{(\theta_1+\theta_3)^2} - \frac{n_{p22}}{(1-\theta_1-\theta_3)^2} - \frac{n_{q21}}{\theta_3^2} - \frac{n_{q22}}{(1-\theta_3)^2}$;
- $l_{44} = -\frac{n_{p31}}{(\theta_1+\theta_4)^2} - \frac{n_{p32}}{(1-\theta_1-\theta_4)^2} - \frac{n_{q31}}{\theta_4^2} - \frac{n_{q32}}{(1-\theta_4)^2}$;
- $l_{23} = l_{32} = l_{24} = l_{42} = l_{34} = l_{43} = 0$.

Now, the expectations can be easily obtained by again using Table 1 as in the Subsection 2.2.4, but with different parametrization. For short, put $s_{ij} = -\mathrm{E}l_{ij}$. The determinant of the Fisher information matrix is then:

$$|I(\theta)| = \begin{vmatrix} s_{11} & s_{12} & s_{13} & s_{14} \\ s_{12} & s_{22} & 0 & 0 \\ s_{13} & 0 & s_{33} & 0 \\ s_{14} & 0 & 0 & s_{44} \end{vmatrix} = -s_{14} \begin{vmatrix} s_{12} & s_{22} & 0 \\ s_{13} & 0 & s_{33} \\ s_{14} & 0 & 0 \end{vmatrix} + s_{44} \begin{vmatrix} s_{11} & s_{12} & s_{13} \\ s_{12} & s_{22} & 0 \\ s_{13} & 0 & s_{33} \end{vmatrix} =$$

$$= s_{14}s_{22} \begin{vmatrix} s_{13} & s_{33} \\ s_{14} & 0 \end{vmatrix} + s_{44}s_{13} \begin{vmatrix} s_{12} & s_{22} \\ s_{13} & 0 \end{vmatrix} + s_{44}s_{33} \begin{vmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{vmatrix} =$$

$$= -s_{14}^2 s_{22} s_{33} - s_{44} s_{13}^2 s_{22} + s_{44} s_{33} s_{11} s_{22} - s_{44} s_{33} s_{12}^2.$$

Thus, the posterior $p(\theta|x) \sim L(x|\theta)\sqrt{|I(\theta)|}$ when Jeffreys prior is used. In both situations, the posteriors were difficult to handle analitically and the MCMC method was used to get the credible intervals. All parameters were constrained: $\theta_1 \in (-1,1)$ and $0 \leq \theta_2, \theta_3, \theta_4 \leq 1$. That is why 100 000 simulations with 10 000 burn in were used to better evaluate the credible intervals. The rest 90 000 values were considered to be taken from the posterior distribution and the $5-th$ percentile of them was taken as the lower bound for the interval. If the value was greater than zero, null hypothesis was rejected. And that was the procedure for testing the hypothesis in the Bayesian framework.

3.2. New parametrization and priors for two DF test

For two DF test, the same likelihood function (8) was used, but, in this case, only one of the Δ_2 and Δ_3 was proportional to the Δ_1. Let $\Delta_1 = \Delta_2$. Then the hypothesis for testing the treatment

effect was $H_0 : \Delta_1 \leq 0 \cap \Delta_3 \leq 0$. That is why two credible intervals were used for testing and hypothesis was accepted only if both lower bounds were not greater than zero. However, the type I error increased by using this approach and the more appropriate concept would be better. One of possible remedies would be to apply closed testing with some type of correction of p-value as it is usually done in multiple testing procedures instead of applying closed testing without adjustments. That is why in this case, the one-sided 0.975 credible intervals were used instead of the 0.95. This solved the problem with the higher type I error.

Again, the first prior tested was based on the uniform priors: $\Delta_1, \Delta_3 \sim U(0,1)$ and q_1,q_2,$q_3 \sim U(0,1)$. In this case, the vector of the unknown parameters $\theta = (\Delta_1, \Delta_3, q_1, q_2, q_3) = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$. Then, the Jeffreys prior was obtained by using the same notations from the previous section of this work. Here comes the derivation. The first derivatives are:

- $l_1 = \frac{n_{p11}}{\theta_1+\theta_3} - \frac{n_{p12}}{1-\theta_1-\theta_3} + \frac{n_{p21}}{\theta_1+\theta_4} - \frac{n_{p22}}{1-\theta_1-\theta_4}$;
- $l_2 = \frac{n_{p31}}{\theta_2+\theta_5} - \frac{n_{p32}}{1-\theta_2-\theta_5}$;
- $l_3 = \frac{n_{p11}}{\theta_1+\theta_3} - \frac{n_{p12}}{1-\theta_1-\theta_3} + \frac{n_{q11}}{\theta_3} - \frac{n_{q12}}{1-\theta_3}$;
- $l_4 = \frac{n_{p21}}{\theta_1+\theta_4} - \frac{n_{p22}}{1-\theta_1-\theta_4} + \frac{n_{q21}}{\theta_4} - \frac{n_{q22}}{1-\theta_4}$;
- $l_5 = \frac{n_{p31}}{\theta_2+\theta_5} - \frac{n_{p32}}{1-\theta_2-\theta_5} + \frac{n_{q31}}{\theta_5} - \frac{n_{q32}}{1-\theta_5}$.

Using these, the second derivatives are calculated:

- $l_{11} = -\frac{n_{p11}}{(\theta_1+\theta_3)^2} - \frac{n_{p12}}{(1-\theta_1-\theta_3)^2} - \frac{n_{p21}}{(\theta_1+\theta_4)^2} - \frac{n_{p22}}{(1-\theta_1-\theta_4)^2}$;
- $l_{12} = l_{21} = l_{15} = l_{51} = 0$;
- $l_{13} = l_{31} = -\frac{n_{p11}}{(\theta_1+\theta_3)^2} - \frac{n_{p12}}{(1-\theta_1-\theta_3)^2}$;
- $l_{14} = l_{41} = -\frac{n_{p21}}{(\theta_1+\theta_4)^2} - \frac{n_{p22}}{(1-\theta_1-\theta_4)^2}$;
- $l_{22} = l_{25} = l_{52} = -\frac{n_{p31}}{(\theta_2+\theta_5)^2} - \frac{n_{p32}}{(1-\theta_2-\theta_5)^2}$;
- $l_{23} = l_{32} = l_{24} = l_{42} = 0$;
- $l_{33} = -\frac{n_{p11}}{(\theta_1+\theta_3)^2} - \frac{n_{p12}}{(1-\theta_1-\theta_3)^2} - \frac{n_{q11}}{\theta_3^2} - \frac{n_{q12}}{(1-\theta_3)^2}$;
- $l_{44} = -\frac{n_{p21}}{(\theta_1+\theta_4)^2} - \frac{n_{p22}}{(1-\theta_1-\theta_4)^2} - \frac{n_{q21}}{\theta_4^2} - \frac{n_{q22}}{(1-\theta_4)^2}$;
- $l_{55} = -\frac{n_{p31}}{(\theta_2+\theta_5)^2} - \frac{n_{p32}}{(1-\theta_2-\theta_5)^2} - \frac{n_{q31}}{\theta_5^2} - \frac{n_{q32}}{(1-\theta_5)^2}$;
- $l_{34} = l_{43} = l_{35} = l_{53} = l_{45} = l_{54} = 0$.

Then the Jeffreys prior is proportional to the square root of the

$$|I(\theta)| = \begin{vmatrix} s_{11} & 0 & s_{13} & s_{14} & 0 \\ 0 & s_{22} & 0 & 0 & s_{25} \\ s_{13} & 0 & s_{33} & 0 & 0 \\ s_{14} & 0 & 0 & s_{44} & 0 \\ 0 & s_{25} & 0 & 0 & s_{55} \end{vmatrix} = -s_{25} \begin{vmatrix} s_{11} & 0 & s_{13} & s_{14} \\ s_{13} & 0 & s_{33} & 0 \\ s_{14} & 0 & 0 & s_{44} \\ 0 & s_{25} & 0 & 0 \end{vmatrix} + s_{55} \begin{vmatrix} s_{11} & 0 & s_{13} & s_{14} \\ 0 & s_{22} & 0 & 0 \\ s_{13} & 0 & s_{33} & 0 \\ s_{14} & 0 & 0 & s_{44} \end{vmatrix} =$$

$$= -s_{25}^2 \begin{vmatrix} s_{11} & s_{13} & s_{14} \\ s_{13} & s_{33} & 0 \\ s_{14} & 0 & s_{44} \end{vmatrix} + s_{55}s_{22} \begin{vmatrix} s_{11} & s_{13} & s_{14} \\ s_{13} & s_{33} & 0 \\ s_{14} & 0 & s_{44} \end{vmatrix} =$$

$$= (s_{55}s_{22} - s_{25}^2) \left(s_{14} \begin{vmatrix} s_{13} & s_{33} \\ s_{14} & 0 \end{vmatrix} + s_{44} \begin{vmatrix} s_{11} & s_{13} \\ s_{13} & s_{33} \end{vmatrix} \right) =$$

$$= (s_{55}s_{22} - s_{25}^2)(-s_{14}^2 s_{33} + s_{44}s_{11}s_{33} - s_{44}s_{13}^2).$$

Finally, the same procedure with 100 000 MCMC simulations was applied to draw the conclusions about the treatment effect. Only with the previously mentioned adjustment, that the $2.5 - th$ percentiles of the generated posterior distributions values were taken as the lower bounds for the intervals.

3.3. New parametrization and priors for three DF test

For the third test, no assumptions were made about the treatment effects between the stages and all Δ_i were used in the parametrization, i.e. the likelihood function identical to (8) was incorporated into the testing procedure. The null hypothesis was $H_0 : \Delta_1 \leq 0 \cap \Delta_2 \leq 0 \cap \Delta_3 \leq 0$ and it was rejected if at least one lower bound of the three one-sided 0.984 credible intervals was more than zero. The adjustment was made to better control the type I error as for the 2 DF test.

Therefore, six parameters were estimated $\theta = (\Delta_1, \Delta_2, \Delta_3, q_1, q_2, q_3) = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$ and using the first approach:

$$\pi(\theta) \propto U(-1,1)U(-1,1)U(-1,1)U(0,1)U(0,1)U(0,1).$$

The Jeffreys prior was found equivalently to the both cases provided before. Here comes a short reasoning. The first derivatives are:

- $l_1 = \frac{n_{p11}}{\theta_1+\theta_4} - \frac{n_{p12}}{1-\theta_1-\theta_4}$;

- $l_2 = \frac{n_{p21}}{\theta_2+\theta_5} - \frac{n_{p22}}{1-\theta_2-\theta_5}$;
- $l_3 = \frac{n_{p31}}{\theta_3+\theta_6} - \frac{n_{p32}}{1-\theta_3-\theta_6}$
- $l_4 = \frac{n_{p11}}{\theta_1+\theta_4} - \frac{n_{p12}}{1-\theta_1-\theta_4} + \frac{n_{q11}}{\theta_4} - \frac{n_{q12}}{1-\theta_4}$;
- $l_5 = \frac{n_{p21}}{\theta_2+\theta_5} - \frac{n_{p22}}{1-\theta_2-\theta_5} + \frac{n_{q21}}{\theta_5} - \frac{n_{q22}}{1-\theta_5}$;
- $l_6 = \frac{n_{p31}}{\theta_3+\theta_6} - \frac{n_{p32}}{1-\theta_3-\theta_6} + \frac{n_{q31}}{\theta_3} - \frac{n_{q32}}{1-\theta_6}$.

Then the second derivative are:

- $l_{11} = l_{14} = l_{41} = -\frac{n_{p11}}{(\theta_1+\theta_4)^2} - \frac{n_{p12}}{(1-\theta_1-\theta_4)^2}$;
- $l_{22} = l_{25} = l_{52} = -\frac{n_{p21}}{(\theta_2+\theta_5)^2} - \frac{n_{p22}}{(1-\theta_2-\theta_5)^2}$;
- $l_{33} = l_{36} = l_{63} = -\frac{n_{p31}}{(\theta_3+\theta_6)^2} - \frac{n_{p32}}{(1-\theta_3-\theta_6)^2}$
- $l_{44} = -\frac{n_{p11}}{(\theta_1+\theta_4)^2} - \frac{n_{p12}}{(1-\theta_1-\theta_4)^2} - \frac{n_{q11}}{\theta_4^2} - \frac{n_{q12}}{(1-\theta_4)^2}$;
- $l_{55} = -\frac{n_{p21}}{(\theta_2+\theta_5)^2} - \frac{n_{p22}}{(1-\theta_2-\theta_5)^2} - \frac{n_{q21}}{\theta_5^2} - \frac{n_{q22}}{(1-\theta_5)^2}$;
- $l_{66} = -\frac{n_{p31}}{(\theta_3+\theta_6)^2} - \frac{n_{p32}}{(1-\theta_3-\theta_6)^2} - \frac{n_{q31}}{\theta_3^2} - \frac{n_{q32}}{(1-\theta_6)^2}$;
- all other $l_{ij} = 0$.

Lastly, the Fisher information is found and its determinant

$$|I(\theta)| = \begin{vmatrix} s_{11} & 0 & 0 & s_{14} & 0 & 0 \\ 0 & s_{22} & 0 & 0 & s_{25} & 0 \\ 0 & 0 & s_{33} & 0 & 0 & s_{36} \\ s_{14} & 0 & 0 & s_{44} & 0 & 0 \\ 0 & s_{25} & 0 & 0 & s_{55} & 0 \\ 0 & 0 & s_{36} & 0 & 0 & s_{66} \end{vmatrix} =$$

$$= s_{33} \begin{vmatrix} s_{11} & 0 & s_{14} & 0 & 0 \\ 0 & s_{22} & 0 & s_{25} & 0 \\ s_{14} & 0 & s_{44} & 0 & 0 \\ 0 & s_{25} & 0 & s_{55} & 0 \\ 0 & 0 & 0 & 0 & s_{66} \end{vmatrix} - s_{36} \begin{vmatrix} s_{11} & 0 & s_{14} & 0 & 0 \\ 0 & s_{22} & 0 & s_{25} & 0 \\ 0 & 0 & 0 & 0 & s_{36} \\ s_{14} & 0 & s_{44} & 0 & 0 \\ 0 & s_{25} & 0 & s_{55} & 0 \end{vmatrix} =$$

$$= (s_{66}s_{33} - s_{36}^2)\begin{vmatrix} s_{11} & 0 & s_{14} & 0 \\ 0 & s_{22} & 0 & s_{25} \\ s_{14} & 0 & s_{44} & 0 \\ 0 & s_{25} & 0 & s_{55} \end{vmatrix} = (s_{66}s_{33} - s_{36}^2)(s_{22}s_{55} - s_{25}^2)\begin{vmatrix} s_{11} & s_{14} \\ s_{14} & s_{44} \end{vmatrix} =$$

$$= (s_{66}s_{33} - s_{36}^2)(s_{22}s_{55} - s_{25}^2)(s_{11}s_{44} - s_{14}^2).$$

Therefore, in this situation like in those examined before, the expression of the posterior density was too complex in case of the Jeffreys prior and the MCMC simulations were used to get the credible intervals.

3.4. Power simulation results

The main objective of these proposed tests in the Bayes framework, is to reduce the sample size required to achieve 80% power. For that purpose, the broad simulation was performed using R software. Recall that the power is the probability to reject the null hypothesis when it is really not true. This means that the data for the power estimation has to hold the alternative hypothesis, which in TED case meant that the treatment was effective. Nine different configurations of the parameters were taken from the Ivanova's and Tamura's manuscript. The sample sizes for each case and test were also taken from there to be able to compare the results.

Moreover, the type I error, probability to reject the null hypothesis when it is actually true, was simulated, because all the results in Ivanova's and Tamura's manuscript are provided for type I error rate of 0.05. For that reason, the data has to hold the null hypothesis, which in TED case meant that the treatment was not effective enough, i.e. $p_i = q_i, \quad i = 1,2,3$. Nine configurations of the parameters were taken from [9] and the samples sizes were the same as for the power simulation.

The data was generated using the procedure described in Section 2.3 with the corresponding set of the parameters and sample size n. In total, 10 000 simulations were executed for each combination of the parameters with n and test. For each combination two estimates were found: for test with uniform priors and Jeffreys.

First of all, the type I error α was estimated by taking the part of the cases where null hypothesis was rejected when it was true. The hypothesis testing procedures from section 3.1, 3.2 and 3.3 of this thesis were used as appropriate. In all cases, the size of MCMC simulations was 100 000 with 10 000 burn in. Results are provided in Table 3. One can see that type I error rate is quite well controlled for the alternative of 1 DF test when using the uniform distributions as the prior. Meanwhile, it exceeds 0.05 in almost all the cases when using the Jeffreys prior. Similar situation is observed with the alternatives of the other two tests: uniform gives better results than the Jeffreys.

Table 3. Type I error for testing $H_0 : \Delta_1 \leq 0 \cap \Delta_2 \leq 0 \cap \Delta_3 \leq 0$ in Bayesian framework.

						1 DF		2 DF		3 DF	
p_1	q_1	p_2	q_2	p_3	q_3	Uniform	Jeffreys	Uniform	Jeffreys	Uniform	Jeffreys
0.4	0.4	0.4	0.4	0.9	0.9	**0.048**	**0.057**	**0.045**	**0.055**	**0.047**	**0.055**
						(n = 412)		(n = 504)		(n = 572)	
0.5	0.5	0.4	0.4	0.9	0.9	**0.041**	**0.064**	**0.037**	**0.060**	**0.043**	**0.062**
						(n = 128)		(n = 152)		(n = 170)	
0.3	0.3	0.4	0.4	0.9	0.9	**0.043**	**0.060**	**0.031**	**0.051**	**0.041**	**0.057**
						(n = 96)		(n = 108)		(n = 114)	
0.4	0.4	0.4	0.4	0.7	0.7	**0.049**	**0.056**	**0.053**	**0.060**	**0.047**	**0.055**
						(n = 312)		(n = 328)		(n = 372)	
0.5	0.5	0.4	0.4	0.7	0.7	**0.047**	**0.059**	**0.049**	**0.062**	**0.048**	**0.063**
						(n = 104)		(n = 128)		(n = 142)	
0.3	0.3	0.4	0.4	0.7	0.7	**0.048**	**0.065**	**0.042**	**0.066**	**0.039**	**0.058**
						(n = 80)		(n = 96)		(n = 100)	
0.4	0.4	0.4	0.4	0.8	0.8	**0.049**	**0.050**	**0.050**	**0.052**	**0.052**	**0.055**
						(n= 2612)		(n = 1916)		(n = 1446)	
0.4	0.4	0.5	0.5	0.8	0.8	**0.049**	**0.054**	**0.049**	**0.055**	**0.049**	**0.054**
						(n = 728)		(n= 648)		(n = 734)	
0.4	0.4	0.3	0.3	0.8	0.8	**0.052**	**0.057**	**0.050**	**0.055**	**0.044**	**0.048**
						(n = 644)		(n = 772)		(n = 716)	

Moreover, it could be seen that the adjustments helped to control the type I error around 0.05 for other two tests.

Thereafter, the power was evaluated by summing the cases where the null hypothesis was re-

Table 4. Power for testing $H_0 : \Delta_1 \leq 0 \cap \Delta_2 \leq 0 \cap \Delta_3 \leq 0$ in Bayesian framework.

						1 DF		2 DF		3 DF	
p_1	q_1	p_2	q_2	p_3	q_3	Uniform	Jeffreys	Uniform	Jeffreys	Uniform	Jeffreys
0.4	0.3	0.4	0.3	0.9	0.8	**0.86**	**0.87**	**0.84**	**0.86**	**0.79**	**0.82**
						(n = 412)		(n = 504)		(n = 572)	
0.5	0.3	0.4	0.2	0.9	0.8	**0.85**	**0.87**	**0.85**	**0.88**	**0.83**	**0.87**
						(n = 128)		(n = 152)		(n = 170)	
0.5	0.3	0.4	0.1	0.9	0.8	**0.85**	**0.88**	**0.84**	**0.88**	**0.79**	**0.84**
						(n = 96)		(n = 108)		(n = 114)	
0.4	0.3	0.4	0.3	0.9	0.7	**0.87**	**0.88**	**0.80**	**0.82**	**0.77**	**0.80**
						(n = 312)		(n = 328)		(n = 372)	
0.5	0.3	0.4	0.2	0.9	0.7	**0.85**	**0.88**	**0.83**	**0.87**	**0.80**	**0.85**
						(n = 104)		(n = 128)		(n = 142)	
0.5	0.3	0.4	0.1	0.9	0.7	**0.84**	**0.88**	**0.82**	**0.88**	**0.74**	**0.80**
						(n = 80)		(n = 96)		(n = 100)	
0.4	0.4	0.4	0.3	0.9	0.8	**0.86**	**0.87**	**0.83**	**0.84**	**0.80**	**0.81**
						(n = 2612)		(n = 1916)		(n = 1446)	
0.4	0.3	0.4	0.3	0.9	0.9	**0.87**	**0.88**	**0.87**	**0.89**	**0.85**	**0.86**
						(n = 728)		(n= 648)		(n = 734)	
0.4	0.3	0.4	0.4	0.9	0.8	**0.87**	**0.88**	**0.83**	**0.85**	**0.83**	**0.85**
						(n = 644)		(n = 772)		(n = 716)	

jected when it was false. All the simulations were made same as for the type I error just with the different configurations of parameters. Power was estimated for all three alternative tests. The results of simulations are presented in Table 4. The highest power was obtained when the Jeffreys prior was used. However, it is worth to remember that α was not controlled very well for Jeffreys prior. Despite that, the approach with the uniform ditribution also has shown higher than 80% power in the majority of the configurations (for 1 DF test in all the configurations) and the type I error rate was controlled quite well. The power for the 2 DF and 3 DF tests is lower compared to 1 DF test, however, as it was mentioned previously, those two tests were not the main concern in this thesis.

To conclude, the alternative testing to the one DF test in Bayesian framework has higher power, which means that smaller sample sizes are needed. Although the usage of the uniform priors does not give the highest power, it is the best choice of the two tested because the type I error was controlled around 0.05. The actual sample size for achieving the 80% power could be estimated by selecting the configuration of parameters which is expected and then running the simulations.

Conclusions

In this thesis, the Bayesian models with six different prior distributions were created for the first statistical problem in TED - parameters estimation. The comparison with the MLE was conducted on the simulated data using various combinations of the parameters and sample sizes. The less biased estimates were obtained using MLE approach, but Bayesian estimates had lower RMSE and KL divergence in the majority of the configurations. The best results were achieved for the cases with the informative priors like Beta distributions. When there is not much additional information, the objective priors could be used and two of the best were Jeffreys and the combination of the uniform distributions. The reference prior is more complex and the superiority over these two was not detected. In conclusion, the Bayesian estimates are better than ML because the estimated distribution of the data is closer to the actual one using the former approach based on the KL divergence.

Hypothesis testing in the Bayesian approach was performed using the credible intervals. These intervals were evaluated using the MCMC simulations and the parametrization of the models for each of the three tests from the original approach. For the alternative of the 1 DF test, the simulated power was higher than 80% for both priors: combination of the uniform distributions and the Jeffreys. For other two tests, there were a few cases with the power less than 80% because of the adjustment. However, it helped to control the type I error rate for those two tests and in the majority of the cases the power was still higher. Thus, the smaller sample sizes are needed to achieve the same power as in the original approach. Yet, the concept with Jeffreys prior did not controlled the type I error rate around 0.05 for all the tests and the actual benefit of it could be questioned. Nevertheless, the simulations from the Bayesian model with the uniform priors showed that the error rate is controlled well, and the power is also higher than using the original approach.

To conclude, the results of this thesis could be applied directly. Using the previously mentioned Bayesian methodology, the sample size could be reduced and the cost of the researches could be cut without loss of the precision of the results. Furthermore, this work could be extended by proposing the alternative Bayesian testing for the presented hypothesis using the other methodology. Besides that, the model with continuous outcome could be developed in Bayesian framework as well.

References

[1] V. Bagdonavičius and J. J. Kruopis. *Matematinė statistika I*. Vilniaus universiteto leidykla, Vilnius, 2015. 91-95 p.

[2] M. Fava, A. Evans, D. Dorer, and D. Schoenfeld. The problem of the placebo response in clinical trials for psychiatric disorders: culprits, possible remedies, and a novel study design approach. *Psychotherapy and Psychosomatics*, 72:115–127 p. 2003.

[3] J. Ghosh, M. Delampady, and T. Samanta. *An Introduction to Bayesian Analysis: Theory and Methods*. Springer, 2006. 29-50, 132-148, 215-231 p.

[4] D. M. Goldenholz and S. R. Goldenholz. Response to placebo in clinical epilepsy trials—old ideas and new insights. *Epilepsy research*, 122:15–25, 2016.

[5] A. Ivanova and R. N. Tamura. A two-way enriched clinical trial design: combining advantages of placebo lead-in and randomized withdrawal. *Statistical Methods in Medical Research*, 24(6):871–890 p. 2015.

[6] A. Ivanova, Z. Zhang, L. Thompson, Y. Yang, R. M. Kotz, and X. Fang. Can sequential parallel comparison design and two-way enriched design be useful in medical device clinical trials? *Journal of biopharmaceutical statistics*, 26:(1):167–177, 2016.

[7] B. Kinon, A. Pots, and S. Watson. Placebo response in clinical trials with schizophrenia patients. *Current opinion in psychiatry*, 24:107–113 p. 2011.

[8] S. Lee, J. Walker, L. Jakul, and K. Sexton. Does elimination of placebo responders in a placebo run-in increase the treatment effect in randomized clinical trials? a meta-analytic evaluation. *Depression and anxiety*, 19:10–19 p. 2004.

[9] Y. Liu, D. Rybin, T. C. Heeren, and G. Doros. Comparison of novel methods in two-way enriched clinical trial design. *Statistics in medicine*, 38:4112–4130 p. 2019.

[10] MacKay and J. C. David. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003. 34 p.

[11] S. Nagarajan. Limiting the placebo and hawthorne effect in periodontal clinical trials: current concepts and future directions. *Journal of Research and Advancement in Dentistry*, 7:137–145, 2018.

[12] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2018. URL: `https://www.R-project.org/`.

[13] E. Schweizer and K. Rickels. Placebo response in generalized anxiety: its effect on the outcome of clinical trials. *The Journal of clinical psychiatry*, 58:30–38 p. 1997.

[14] M. Thulin. Decision-theoretic justifications for bayesian hypothesis testing using credible sets. *Journal of Statistical Planning and Inference*, 146:133–138 p. 2014.

[15] M. Trivedi and H. Rush. Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? *Neuropsychopharmacology*, 11:33 p. 1994.

[16] K. Weimer, L. Colloca, and P. Enck. Placebo effects in psychiatry: mediators and moderators. *The Lancet Psychiatry*, 2:246–257 p. 2015.

Appendix Nr. 1. Estimated bias for all cases

par	true.val	n	m	par.set	MLE	U(0,1)	B12/B21	B13/B31	U(0,p)	Jeffreys	Reference
p_1	0.4	60	10000	1	-0.001	-0.007	0.005	0.016	-0.019	-0.023	-0.005
q_1	0.3	60	10000	1	0.001	-0.011	-0.002	0.007	-0.008	0.004	-0.005
p_2	0.4	60	10000	1	0.001	-0.015	0.016	0.043	-0.026	-0.008	-0.008
q_2	0.3	60	10000	1	0.000	-0.033	-0.008	0.014	0.015	-0.018	-0.018
p_3	0.9	60	10000	1	0.000	0.106	0.081	0.062	0.144	0.062	0.062
q_3	0.8	60	10000	1	0.001	0.080	0.047	0.021	0.143	0.047	0.047
p_1	0.4	100	10000	1	0.002	-0.002	0.006	0.013	-0.014	-0.012	0.000
q_1	0.3	100	10000	1	0.000	-0.008	-0.002	0.004	-0.016	0.002	-0.004
p_2	0.4	100	10000	1	0.001	-0.010	0.011	0.029	-0.021	-0.005	-0.005
q_2	0.3	100	10000	1	-0.001	-0.022	-0.006	0.008	0.003	-0.012	-0.012
p_3	0.9	100	10000	1	-0.001	0.069	0.055	0.044	0.111	0.038	0.038
q_3	0.8	100	10000	1	0.001	0.053	0.033	0.015	0.110	0.030	0.030
p_1	0.4	160	10000	1	0.000	-0.002	0.003	0.007	-0.015	-0.008	-0.001
q_1	0.3	160	10000	1	0.000	-0.005	-0.002	0.002	-0.022	0.001	-0.003
p_2	0.4	160	10000	1	0.000	-0.007	0.007	0.019	-0.018	-0.003	-0.003
q_2	0.3	160	10000	1	0.000	-0.013	-0.003	0.007	-0.006	-0.007	-0.007
p_3	0.9	160	10000	1	0.000	0.046	0.038	0.031	0.091	0.025	0.025
q_3	0.8	160	10000	1	-0.001	0.034	0.021	0.010	0.086	0.018	0.018
p_1	0.5	60	10000	2	-0.001	-0.001	0.015	0.029	-0.001	-0.016	-0.001
q_1	0.3	60	10000	2	0.000	-0.012	-0.003	0.006	-0.003	0.004	-0.006
p_2	0.4	60	10000	2	0.001	-0.015	0.016	0.043	-0.015	-0.008	-0.008
q_2	0.2	60	10000	2	0.000	-0.049	-0.030	-0.014	-0.008	-0.027	-0.027
p_3	0.9	60	10000	2	0.000	0.088	0.069	0.054	0.088	0.050	0.050
q_3	0.8	60	10000	2	0.002	0.067	0.041	0.019	0.101	0.039	0.039
p_1	0.5	100	10000	2	0.000	0.000	0.009	0.018	0.000	-0.010	0.000
q_1	0.3	100	10000	2	0.000	-0.008	-0.002	0.004	-0.005	0.002	-0.004
p_2	0.4	100	10000	2	0.002	-0.009	0.011	0.030	-0.009	-0.004	-0.004
q_2	0.2	100	10000	2	-0.001	-0.032	-0.020	-0.010	-0.011	-0.017	-0.017
p_3	0.9	100	10000	2	0.000	0.057	0.047	0.038	0.057	0.031	0.031
q_3	0.8	100	10000	2	-0.001	0.042	0.026	0.012	0.067	0.022	0.022
p_1	0.5	160	10000	2	0.000	0.000	0.006	0.012	0.000	-0.006	0.000
q_1	0.3	160	10000	2	0.000	-0.005	-0.001	0.002	-0.004	0.001	-0.003
p_2	0.4	160	10000	2	-0.001	-0.007	0.006	0.018	-0.008	-0.004	-0.004
q_2	0.2	160	10000	2	0.001	-0.019	-0.012	-0.006	-0.011	-0.010	-0.010
p_3	0.9	160	10000	2	0.000	0.037	0.031	0.025	0.037	0.019	0.019
q_3	0.8	160	10000	2	-0.001	0.027	0.017	0.008	0.044	0.014	0.014
p_1	0.5	60	10000	3	-0.001	-0.001	0.014	0.029	-0.001	-0.017	-0.001
q_1	0.3	60	10000	3	0.000	-0.013	-0.003	0.006	-0.004	0.003	-0.007
p_2	0.4	60	10000	3	0.000	-0.016	0.015	0.042	-0.016	-0.009	-0.009
q_2	0.1	60	10000	3	0.000	-0.066	-0.053	-0.042	-0.051	-0.036	-0.036
p_3	0.9	60	10000	3	0.002	0.090	0.070	0.055	0.089	0.052	0.052
q_3	0.8	60	10000	3	0.000	0.066	0.040	0.018	0.100	0.037	0.037
p_1	0.5	100	10000	3	0.000	0.000	0.009	0.018	0.000	-0.010	0.000
q_1	0.3	100	10000	3	0.001	-0.007	-0.001	0.004	-0.004	0.002	-0.003
p_2	0.4	100	10000	3	0.000	-0.011	0.009	0.028	-0.011	-0.006	-0.006
q_2	0.1	100	10000	3	0.000	-0.042	-0.035	-0.028	-0.037	-0.022	-0.022
p_3	0.9	100	10000	3	-0.001	0.056	0.045	0.036	0.056	0.030	0.030
q_3	0.8	100	10000	3	0.002	0.044	0.028	0.014	0.069	0.025	0.025
p_1	0.5	160	10000	3	0.000	0.000	0.006	0.012	0.000	-0.006	0.000
q_1	0.3	160	10000	3	0.000	-0.005	-0.001	0.002	-0.005	0.001	-0.003
p_2	0.4	160	10000	3	0.002	-0.005	0.008	0.021	-0.005	-0.002	-0.002
q_2	0.1	160	10000	3	0.000	-0.027	-0.023	-0.019	-0.026	-0.014	-0.014
p_3	0.9	160	10000	3	0.000	0.037	0.031	0.026	0.037	0.020	0.020
q_3	0.8	160	10000	3	0.001	0.029	0.019	0.009	0.045	0.016	0.016

par	true.val	n	m	par.set	MLE	U(0,1)	B12/B21	B13/B31	U(0,p)	Jeffreys	Reference
p_1	0.4	60	10000	4	0.001	-0.005	0.007	0.019	-0.017	-0.021	-0.002
q_1	0.3	60	10000	4	0.001	-0.012	-0.002	0.006	-0.008	0.004	-0.006
p_2	0.4	60	10000	4	0.002	-0.015	0.017	0.044	-0.026	-0.007	-0.007
q_2	0.3	60	10000	4	0.002	-0.031	-0.006	0.015	0.016	-0.016	-0.016
p_3	0.9	60	10000	4	-0.001	0.106	0.081	0.062	0.144	0.061	0.061
q_3	0.7	60	10000	4	-0.001	0.052	0.011	-0.022	0.093	0.030	0.030
p_1	0.4	100	10000	4	0.001	-0.003	0.005	0.012	-0.016	-0.013	-0.001
q_1	0.3	100	10000	4	0.000	-0.008	-0.002	0.004	-0.017	0.002	-0.004
p_2	0.4	100	10000	4	0.003	-0.008	0.012	0.031	-0.020	-0.003	-0.003
q_2	0.3	100	10000	4	0.002	-0.019	-0.003	0.011	0.004	-0.009	-0.009
p_3	0.9	100	10000	4	-0.001	0.069	0.056	0.044	0.112	0.038	0.038
q_3	0.7	100	10000	4	0.000	0.035	0.008	-0.015	0.069	0.019	0.019
p_1	0.4	160	10000	4	0.000	-0.002	0.003	0.007	-0.014	-0.008	-0.001
q_1	0.3	160	10000	4	0.001	-0.004	-0.001	0.003	-0.021	0.002	-0.002
p_2	0.4	160	10000	4	-0.003	-0.009	0.004	0.017	-0.021	-0.006	-0.006
q_2	0.3	160	10000	4	0.000	-0.013	-0.003	0.006	-0.006	-0.007	-0.007
p_3	0.9	160	10000	4	0.001	0.047	0.039	0.031	0.091	0.025	0.025
q_3	0.7	160	10000	4	0.000	0.023	0.005	-0.010	0.052	0.012	0.012
p_1	0.5	60	10000	5	0.000	0.000	0.015	0.029	0.000	-0.016	0.000
q_1	0.3	60	10000	5	-0.002	-0.014	-0.004	0.005	-0.005	0.002	-0.008
p_2	0.4	60	10000	5	0.001	-0.016	0.016	0.043	-0.016	-0.008	-0.008
q_2	0.2	60	10000	5	0.001	-0.049	-0.030	-0.014	-0.008	-0.026	-0.026
p_3	0.9	60	10000	5	-0.001	0.087	0.069	0.053	0.087	0.049	0.049
q_3	0.7	60	10000	5	0.003	0.046	0.012	-0.016	0.066	0.027	0.027
p_1	0.5	100	10000	5	0.001	0.001	0.010	0.019	0.001	-0.009	0.001
q_1	0.3	100	10000	5	0.000	-0.008	-0.002	0.004	-0.005	0.002	-0.004
p_2	0.4	100	10000	5	0.001	-0.009	0.011	0.029	-0.009	-0.004	-0.004
q_2	0.2	100	10000	5	0.000	-0.031	-0.020	-0.009	-0.011	-0.016	-0.016
p_3	0.9	100	10000	5	0.000	0.057	0.046	0.037	0.057	0.031	0.031
q_3	0.7	100	10000	5	0.000	0.028	0.007	-0.012	0.039	0.015	0.015
p_1	0.5	160	10000	5	0.000	0.000	0.006	0.012	0.000	-0.006	0.000
q_1	0.3	160	10000	5	0.001	-0.004	-0.001	0.003	-0.004	0.002	-0.002
p_2	0.4	160	10000	5	0.000	-0.007	0.006	0.019	-0.007	-0.004	-0.004
q_2	0.2	160	10000	5	0.001	-0.020	-0.013	-0.006	-0.011	-0.010	-0.010
p_3	0.9	160	10000	5	0.000	0.037	0.031	0.026	0.037	0.020	0.020
q_3	0.7	160	10000	5	0.000	0.019	0.004	-0.009	0.024	0.010	0.010
p_1	0.5	60	10000	6	0.001	0.001	0.016	0.030	0.001	-0.015	0.001
q_1	0.3	60	10000	6	0.000	-0.013	-0.003	0.006	-0.004	0.003	-0.007
p_2	0.4	60	10000	6	0.004	-0.013	0.018	0.045	-0.013	-0.005	-0.005
q_2	0.1	60	10000	6	-0.001	-0.067	-0.054	-0.043	-0.051	-0.037	-0.037
p_3	0.9	60	10000	6	0.002	0.090	0.071	0.055	0.090	0.052	0.052
q_3	0.7	60	10000	6	-0.001	0.043	0.008	-0.019	0.062	0.023	0.023
p_1	0.5	100	10000	6	0.000	0.000	0.009	0.018	0.000	-0.010	0.000
q_1	0.3	100	10000	6	-0.001	-0.009	-0.003	0.003	-0.006	0.001	-0.005
p_2	0.4	100	10000	6	-0.001	-0.012	0.009	0.027	-0.012	-0.007	-0.007
q_2	0.1	100	10000	6	0.001	-0.041	-0.034	-0.027	-0.036	-0.021	-0.021
p_3	0.9	100	10000	6	0.000	0.057	0.046	0.037	0.057	0.031	0.031
q_3	0.7	100	10000	6	-0.002	0.027	0.005	-0.014	0.038	0.014	0.014
p_1	0.5	160	10000	6	0.001	0.001	0.007	0.013	0.001	-0.005	0.001
q_1	0.3	160	10000	6	0.000	-0.004	-0.001	0.003	-0.004	0.002	-0.002
p_2	0.4	160	10000	6	0.002	-0.005	0.008	0.020	-0.005	-0.002	-0.002
q_2	0.1	160	10000	6	0.000	-0.027	-0.023	-0.019	-0.026	-0.014	-0.014
p_3	0.9	160	10000	6	0.000	0.037	0.031	0.025	0.037	0.019	0.019
q_3	0.7	160	10000	6	0.001	0.020	0.006	-0.007	0.025	0.011	0.011

par	true.val	n	m	par.set	MLE	U(0,1)	B12/B21	B13/B31	U(0,p)	Jeffreys	Reference
p_1	0.4	60	10000	7	0.001	-0.006	0.007	0.018	-0.018	-0.021	-0.003
q_1	0.4	60	10000	7	-0.001	-0.007	0.005	0.017	0.051	0.009	-0.004
p_2	0.4	60	10000	7	0.000	-0.019	0.017	0.047	-0.029	-0.010	-0.010
q_2	0.3	60	10000	7	0.002	-0.036	-0.007	0.017	0.020	-0.019	-0.019
p_3	0.9	60	10000	7	0.000	0.107	0.082	0.063	0.145	0.063	0.063
q_3	0.8	60	10000	7	0.003	0.082	0.048	0.022	0.143	0.049	0.049
p_1	0.4	100	10000	7	0.000	-0.004	0.004	0.011	-0.016	-0.014	-0.002
q_1	0.4	100	10000	7	-0.001	-0.005	0.003	0.010	0.038	0.005	-0.003
p_2	0.4	100	10000	7	0.001	-0.012	0.012	0.033	-0.023	-0.006	-0.006
q_2	0.3	100	10000	7	0.000	-0.024	-0.006	0.010	0.007	-0.013	-0.013
p_3	0.9	100	10000	7	-0.001	0.069	0.056	0.044	0.112	0.038	0.038
q_3	0.8	100	10000	7	-0.001	0.052	0.031	0.014	0.109	0.028	0.028
p_1	0.4	160	10000	7	0.000	-0.002	0.003	0.007	-0.015	-0.008	-0.001
q_1	0.4	160	10000	7	0.000	-0.002	0.002	0.007	0.028	0.004	-0.001
p_2	0.4	160	10000	7	-0.001	-0.009	0.007	0.021	-0.021	-0.005	-0.005
q_2	0.3	160	10000	7	0.000	-0.016	-0.004	0.007	-0.004	-0.008	-0.008
p_3	0.9	160	10000	7	-0.001	0.045	0.037	0.030	0.092	0.023	0.023
q_3	0.8	160	10000	7	0.000	0.034	0.022	0.010	0.087	0.018	0.018
p_1	0.4	60	10000	8	0.000	-0.006	0.006	0.018	-0.018	-0.022	-0.003
q_1	0.3	60	10000	8	0.001	-0.012	-0.002	0.007	-0.008	0.004	-0.006
p_2	0.4	60	10000	8	-0.001	-0.017	0.014	0.041	-0.028	-0.010	-0.010
q_2	0.3	60	10000	8	0.000	-0.033	-0.008	0.014	0.015	-0.018	-0.018
p_3	0.9	60	10000	8	0.001	0.107	0.082	0.063	0.145	0.063	0.063
q_3	0.9	60	10000	8	0.000	0.106	0.081	0.062	0.196	0.062	0.062
p_1	0.4	100	10000	8	0.000	-0.004	0.004	0.011	-0.016	-0.013	-0.002
q_1	0.3	100	10000	8	0.000	-0.007	-0.001	0.004	-0.016	0.002	-0.003
p_2	0.4	100	10000	8	0.000	-0.010	0.010	0.028	-0.022	-0.005	-0.005
q_2	0.3	100	10000	8	0.001	-0.020	-0.004	0.011	0.004	-0.010	-0.010
p_3	0.9	100	10000	8	0.000	0.070	0.056	0.044	0.112	0.039	0.039
q_3	0.9	100	10000	8	0.001	0.070	0.057	0.045	0.159	0.039	0.039
p_1	0.4	160	10000	8	0.000	-0.002	0.002	0.007	-0.015	-0.008	-0.001
q_1	0.3	160	10000	8	0.000	-0.005	-0.001	0.003	-0.021	0.002	-0.002
p_2	0.4	160	10000	8	-0.001	-0.007	0.006	0.018	-0.019	-0.004	-0.004
q_2	0.3	160	10000	8	-0.001	-0.014	-0.004	0.006	-0.007	-0.008	-0.008
p_3	0.9	160	10000	8	0.000	0.046	0.038	0.031	0.092	0.025	0.025
q_3	0.9	160	10000	8	-0.001	0.045	0.037	0.030	0.132	0.023	0.023
p_1	0.4	60	10000	9	-0.001	-0.007	0.005	0.017	-0.019	-0.022	-0.004
q_1	0.3	60	10000	9	0.000	-0.012	-0.003	0.006	-0.008	0.003	-0.006
p_2	0.4	60	10000	9	0.003	-0.014	0.017	0.044	-0.025	-0.006	-0.006
q_2	0.4	60	10000	9	0.000	-0.017	0.015	0.042	0.085	-0.009	-0.009
p_3	0.9	60	10000	9	-0.001	0.105	0.081	0.062	0.144	0.061	0.061
q_3	0.8	60	10000	9	0.001	0.080	0.047	0.021	0.142	0.047	0.047
p_1	0.4	100	10000	9	0.000	-0.004	0.003	0.011	-0.016	-0.014	-0.002
q_1	0.3	100	10000	9	0.000	-0.008	-0.002	0.004	-0.016	0.002	-0.004
p_2	0.4	100	10000	9	0.003	-0.008	0.012	0.030	-0.020	-0.003	-0.003
q_2	0.4	100	10000	9	-0.001	-0.011	0.009	0.027	0.068	-0.006	-0.006
p_3	0.9	100	10000	9	-0.001	0.069	0.055	0.044	0.112	0.038	0.038
q_3	0.8	100	10000	9	0.004	0.055	0.035	0.017	0.111	0.032	0.032
p_1	0.4	160	10000	9	0.000	-0.003	0.002	0.007	-0.015	-0.009	-0.001
q_1	0.3	160	10000	9	0.000	-0.005	-0.002	0.002	-0.022	0.001	-0.003
p_2	0.4	160	10000	9	0.001	-0.006	0.007	0.019	-0.018	-0.003	-0.003
q_2	0.4	160	10000	9	-0.001	-0.008	0.006	0.018	0.053	-0.004	-0.004
p_3	0.9	160	10000	9	0.000	0.046	0.038	0.031	0.092	0.024	0.024
q_3	0.8	160	10000	9	0.001	0.035	0.023	0.011	0.087	0.019	0.019

Appendix Nr. 2. Estimated RMSE for all cases

par	true.val	n	m	par.set	MLE	U(0,1)	B12/B21	B13/B31	U(0,p)	Jeffreys	Reference
p_1	0.4	60	10000	1	0.090	0.085	0.082	0.081	0.080	0.088	0.088
q_1	0.3	60	10000	1	0.084	0.080	0.076	0.074	0.062	0.079	0.081
p_2	0.4	60	10000	1	0.153	0.128	0.119	0.117	0.118	0.139	0.139
q_2	0.3	60	10000	1	0.144	0.125	0.111	0.104	0.072	0.132	0.132
p_3	0.9	60	10000	1	0.128	0.143	0.116	0.096	0.177	0.125	0.125
q_3	0.8	60	10000	1	0.177	0.150	0.120	0.101	0.173	0.154	0.154
p_1	0.4	100	10000	1	0.070	0.067	0.066	0.066	0.065	0.068	0.068
q_1	0.3	100	10000	1	0.065	0.063	0.061	0.060	0.058	0.062	0.064
p_2	0.4	100	10000	1	0.119	0.107	0.102	0.101	0.099	0.113	0.113
q_2	0.3	100	10000	1	0.110	0.101	0.094	0.089	0.065	0.105	0.105
p_3	0.9	100	10000	1	0.097	0.106	0.092	0.080	0.146	0.095	0.095
q_3	0.8	100	10000	1	0.130	0.120	0.103	0.091	0.139	0.121	0.121
p_1	0.4	160	10000	1	0.055	0.054	0.053	0.053	0.053	0.054	0.054
q_1	0.3	160	10000	1	0.051	0.050	0.049	0.049	0.056	0.050	0.051
p_2	0.4	160	10000	1	0.093	0.087	0.084	0.084	0.082	0.090	0.090
q_2	0.3	160	10000	1	0.087	0.082	0.078	0.076	0.061	0.084	0.084
p_3	0.9	160	10000	1	0.077	0.083	0.075	0.069	0.129	0.077	0.077
q_3	0.8	160	10000	1	0.101	0.096	0.087	0.081	0.117	0.097	0.097
p_1	0.5	60	10000	2	0.091	0.085	0.084	0.085	0.085	0.087	0.088
q_1	0.3	60	10000	2	0.084	0.079	0.076	0.074	0.067	0.079	0.081
p_2	0.4	60	10000	2	0.156	0.131	0.121	0.119	0.131	0.142	0.142
q_2	0.2	60	10000	2	0.126	0.116	0.101	0.091	0.066	0.117	0.117
p_3	0.9	60	10000	2	0.113	0.125	0.105	0.089	0.125	0.111	0.111
q_3	0.8	60	10000	2	0.153	0.136	0.113	0.097	0.138	0.138	0.138
p_1	0.5	100	10000	2	0.071	0.068	0.068	0.068	0.068	0.069	0.070
q_1	0.3	100	10000	2	0.065	0.063	0.062	0.061	0.059	0.063	0.064
p_2	0.4	100	10000	2	0.118	0.106	0.101	0.100	0.107	0.112	0.112
q_2	0.2	100	10000	2	0.096	0.092	0.084	0.078	0.065	0.092	0.092
p_3	0.9	100	10000	2	0.087	0.094	0.084	0.075	0.095	0.086	0.086
q_3	0.8	100	10000	2	0.116	0.108	0.096	0.087	0.103	0.109	0.109
p_1	0.5	160	10000	2	0.056	0.055	0.054	0.055	0.055	0.055	0.055
q_1	0.3	160	10000	2	0.051	0.050	0.049	0.049	0.049	0.050	0.051
p_2	0.4	160	10000	2	0.094	0.088	0.085	0.084	0.088	0.091	0.091
q_2	0.2	160	10000	2	0.076	0.073	0.069	0.066	0.061	0.074	0.074
p_3	0.9	160	10000	2	0.068	0.072	0.066	0.061	0.072	0.067	0.067
q_3	0.8	160	10000	2	0.090	0.086	0.080	0.075	0.080	0.087	0.087
p_1	0.5	60	10000	3	0.090	0.085	0.083	0.085	0.085	0.086	0.087
q_1	0.3	60	10000	3	0.085	0.080	0.077	0.075	0.068	0.079	0.082
p_2	0.4	60	10000	3	0.154	0.129	0.120	0.118	0.130	0.140	0.140
q_2	0.1	60	10000	3	0.095	0.103	0.091	0.080	0.079	0.094	0.094
p_3	0.9	60	10000	3	0.115	0.128	0.107	0.091	0.128	0.113	0.113
q_3	0.8	60	10000	3	0.151	0.134	0.112	0.096	0.136	0.136	0.136
p_1	0.5	100	10000	3	0.070	0.068	0.067	0.068	0.068	0.068	0.069
q_1	0.3	100	10000	3	0.065	0.063	0.061	0.060	0.059	0.063	0.064
p_2	0.4	100	10000	3	0.118	0.106	0.101	0.099	0.106	0.112	0.112
q_2	0.1	100	10000	3	0.072	0.077	0.070	0.065	0.069	0.072	0.072
p_3	0.9	100	10000	3	0.086	0.093	0.082	0.074	0.093	0.084	0.084
q_3	0.8	100	10000	3	0.116	0.109	0.096	0.088	0.105	0.110	0.110
p_1	0.5	160	10000	3	0.056	0.055	0.054	0.055	0.055	0.055	0.055
q_1	0.3	160	10000	3	0.051	0.050	0.049	0.049	0.049	0.050	0.051
p_2	0.4	160	10000	3	0.094	0.088	0.085	0.085	0.088	0.091	0.091
q_2	0.1	160	10000	3	0.057	0.060	0.056	0.053	0.058	0.057	0.057
p_3	0.9	160	10000	3	0.068	0.072	0.066	0.062	0.072	0.067	0.067
q_3	0.8	160	10000	3	0.091	0.087	0.081	0.076	0.081	0.088	0.088

par	true.val	n	m	par.set	MLE	U(0,1)	B12/B21	B13/B31	U(0,p)	Jeffreys	Reference
p_1	0.4	60	10000	4	0.089	0.083	0.081	0.080	0.078	0.086	0.086
q_1	0.3	60	10000	4	0.083	0.079	0.076	0.074	0.062	0.078	0.081
p_2	0.4	60	10000	4	0.152	0.127	0.118	0.117	0.118	0.138	0.138
q_2	0.3	60	10000	4	0.144	0.124	0.111	0.104	0.073	0.132	0.132
p_3	0.9	60	10000	4	0.130	0.143	0.116	0.095	0.177	0.125	0.125
q_3	0.7	60	10000	4	0.200	0.151	0.125	0.114	0.145	0.168	0.168
p_1	0.4	100	10000	4	0.069	0.067	0.065	0.065	0.064	0.068	0.068
q_1	0.3	100	10000	4	0.065	0.063	0.061	0.060	0.059	0.062	0.063
p_2	0.4	100	10000	4	0.118	0.106	0.101	0.100	0.098	0.112	0.112
q_2	0.3	100	10000	4	0.109	0.100	0.093	0.089	0.066	0.104	0.104
p_3	0.9	100	10000	4	0.098	0.107	0.093	0.081	0.148	0.096	0.096
q_3	0.7	100	10000	4	0.150	0.127	0.112	0.105	0.121	0.136	0.136
p_1	0.4	160	10000	4	0.054	0.053	0.052	0.052	0.052	0.053	0.053
q_1	0.3	160	10000	4	0.051	0.050	0.049	0.049	0.056	0.050	0.051
p_2	0.4	160	10000	4	0.094	0.088	0.085	0.083	0.083	0.091	0.091
q_2	0.3	160	10000	4	0.087	0.082	0.078	0.076	0.061	0.084	0.084
p_3	0.9	160	10000	4	0.076	0.082	0.074	0.068	0.128	0.076	0.076
q_3	0.7	160	10000	4	0.116	0.105	0.097	0.092	0.103	0.109	0.109
p_1	0.5	60	10000	5	0.091	0.086	0.084	0.086	0.086	0.087	0.089
q_1	0.3	60	10000	5	0.084	0.080	0.076	0.074	0.067	0.078	0.081
p_2	0.4	60	10000	5	0.154	0.129	0.119	0.118	0.129	0.140	0.140
q_2	0.2	60	10000	5	0.125	0.115	0.100	0.090	0.066	0.116	0.116
p_3	0.9	60	10000	5	0.115	0.126	0.106	0.090	0.126	0.112	0.112
q_3	0.7	60	10000	5	0.175	0.142	0.121	0.111	0.133	0.154	0.154
p_1	0.5	100	10000	5	0.071	0.068	0.067	0.068	0.068	0.068	0.069
q_1	0.3	100	10000	5	0.065	0.063	0.061	0.060	0.058	0.063	0.064
p_2	0.4	100	10000	5	0.118	0.106	0.100	0.100	0.106	0.111	0.111
q_2	0.2	100	10000	5	0.096	0.091	0.084	0.078	0.064	0.092	0.092
p_3	0.9	100	10000	5	0.086	0.094	0.083	0.074	0.094	0.085	0.085
q_3	0.7	100	10000	5	0.133	0.117	0.106	0.100	0.109	0.123	0.123
p_1	0.5	160	10000	5	0.056	0.055	0.055	0.055	0.055	0.055	0.056
q_1	0.3	160	10000	5	0.051	0.050	0.049	0.049	0.049	0.050	0.051
p_2	0.4	160	10000	5	0.093	0.087	0.084	0.083	0.087	0.089	0.089
q_2	0.2	160	10000	5	0.076	0.074	0.070	0.067	0.061	0.074	0.074
p_3	0.9	160	10000	5	0.068	0.072	0.067	0.062	0.072	0.068	0.068
q_3	0.7	160	10000	5	0.103	0.095	0.089	0.086	0.090	0.099	0.099
p_1	0.5	60	10000	6	0.091	0.085	0.084	0.086	0.085	0.087	0.088
q_1	0.3	60	10000	6	0.083	0.079	0.076	0.074	0.067	0.078	0.081
p_2	0.4	60	10000	6	0.154	0.129	0.120	0.119	0.129	0.140	0.140
q_2	0.1	60	10000	6	0.094	0.103	0.091	0.080	0.079	0.093	0.093
p_3	0.9	60	10000	6	0.116	0.128	0.107	0.091	0.128	0.113	0.113
q_3	0.7	60	10000	6	0.174	0.140	0.119	0.109	0.131	0.152	0.152
p_1	0.5	100	10000	6	0.071	0.068	0.068	0.068	0.068	0.069	0.070
q_1	0.3	100	10000	6	0.065	0.063	0.061	0.060	0.058	0.062	0.064
p_2	0.4	100	10000	6	0.118	0.106	0.100	0.099	0.106	0.111	0.111
q_2	0.1	100	10000	6	0.072	0.077	0.070	0.065	0.069	0.072	0.072
p_3	0.9	100	10000	6	0.087	0.094	0.083	0.074	0.094	0.085	0.085
q_3	0.7	100	10000	6	0.132	0.115	0.105	0.099	0.107	0.122	0.122
p_1	0.5	160	10000	6	0.056	0.054	0.054	0.055	0.055	0.055	0.055
q_1	0.3	160	10000	6	0.051	0.050	0.050	0.049	0.049	0.050	0.051
p_2	0.4	160	10000	6	0.093	0.087	0.084	0.084	0.087	0.090	0.090
q_2	0.1	160	10000	6	0.057	0.060	0.056	0.053	0.058	0.057	0.057
p_3	0.9	160	10000	6	0.068	0.072	0.067	0.062	0.073	0.068	0.068
q_3	0.7	160	10000	6	0.106	0.098	0.091	0.088	0.092	0.101	0.101

par	true.val	n	m	par.set	MLE	U(0,1)	B12/B21	B13/B31	U(0,p)	Jeffreys	Reference
p_1	0.4	60	10000	7	0.089	0.083	0.081	0.080	0.078	0.086	0.086
q_1	0.4	60	10000	7	0.090	0.084	0.082	0.081	0.068	0.085	0.087
p_2	0.4	60	10000	7	0.169	0.137	0.125	0.124	0.126	0.151	0.151
q_2	0.3	60	10000	7	0.157	0.131	0.115	0.107	0.074	0.141	0.141
p_3	0.9	60	10000	7	0.130	0.144	0.117	0.097	0.179	0.126	0.126
q_3	0.8	60	10000	7	0.175	0.150	0.120	0.100	0.173	0.153	0.153
p_1	0.4	100	10000	7	0.069	0.067	0.065	0.065	0.064	0.068	0.068
q_1	0.4	100	10000	7	0.069	0.067	0.065	0.065	0.054	0.067	0.068
p_2	0.4	100	10000	7	0.130	0.115	0.108	0.107	0.106	0.122	0.122
q_2	0.3	100	10000	7	0.121	0.109	0.100	0.095	0.068	0.113	0.113
p_3	0.9	100	10000	7	0.098	0.107	0.093	0.081	0.148	0.096	0.096
q_3	0.8	100	10000	7	0.131	0.120	0.104	0.092	0.139	0.122	0.122
p_1	0.4	160	10000	7	0.055	0.053	0.053	0.053	0.053	0.054	0.054
q_1	0.4	160	10000	7	0.055	0.054	0.053	0.053	0.045	0.054	0.055
p_2	0.4	160	10000	7	0.100	0.093	0.089	0.088	0.087	0.096	0.096
q_2	0.3	160	10000	7	0.094	0.088	0.083	0.081	0.063	0.090	0.090
p_3	0.9	160	10000	7	0.076	0.081	0.074	0.067	0.129	0.075	0.075
q_3	0.8	160	10000	7	0.102	0.097	0.088	0.081	0.118	0.098	0.098
p_1	0.4	60	10000	8	0.088	0.083	0.080	0.080	0.078	0.085	0.085
q_1	0.3	60	10000	8	0.084	0.079	0.076	0.074	0.061	0.079	0.081
p_2	0.4	60	10000	8	0.154	0.130	0.120	0.118	0.120	0.141	0.141
q_2	0.3	60	10000	8	0.143	0.123	0.110	0.103	0.072	0.131	0.131
p_3	0.9	60	10000	8	0.132	0.145	0.117	0.097	0.179	0.127	0.127
q_3	0.9	60	10000	8	0.131	0.144	0.117	0.097	0.211	0.127	0.127
p_1	0.4	100	10000	8	0.069	0.066	0.065	0.065	0.064	0.068	0.068
q_1	0.3	100	10000	8	0.065	0.063	0.061	0.060	0.059	0.062	0.063
p_2	0.4	100	10000	8	0.118	0.106	0.101	0.100	0.099	0.112	0.112
q_2	0.3	100	10000	8	0.110	0.100	0.093	0.089	0.066	0.104	0.104
p_3	0.9	100	10000	8	0.099	0.108	0.093	0.082	0.148	0.097	0.097
q_3	0.9	100	10000	8	0.099	0.108	0.094	0.082	0.175	0.097	0.097
p_1	0.4	160	10000	8	0.055	0.054	0.053	0.053	0.054	0.054	0.054
q_1	0.3	160	10000	8	0.052	0.051	0.050	0.049	0.057	0.050	0.051
p_2	0.4	160	10000	8	0.094	0.088	0.085	0.084	0.082	0.091	0.091
q_2	0.3	160	10000	8	0.088	0.083	0.079	0.077	0.062	0.085	0.085
p_3	0.9	160	10000	8	0.076	0.082	0.074	0.068	0.129	0.076	0.076
q_3	0.9	160	10000	8	0.076	0.081	0.074	0.067	0.149	0.075	0.075
p_1	0.4	60	10000	9	0.087	0.082	0.080	0.079	0.077	0.085	0.085
q_1	0.3	60	10000	9	0.084	0.080	0.077	0.075	0.062	0.079	0.082
p_2	0.4	60	10000	9	0.153	0.128	0.119	0.118	0.118	0.139	0.139
q_2	0.4	60	10000	9	0.153	0.129	0.119	0.117	0.104	0.140	0.140
p_3	0.9	60	10000	9	0.130	0.143	0.116	0.096	0.178	0.126	0.126
q_3	0.8	60	10000	9	0.174	0.148	0.118	0.099	0.171	0.151	0.151
p_1	0.4	100	10000	9	0.070	0.068	0.066	0.066	0.065	0.069	0.069
q_1	0.3	100	10000	9	0.065	0.063	0.061	0.060	0.059	0.063	0.064
p_2	0.4	100	10000	9	0.120	0.107	0.102	0.101	0.099	0.113	0.113
q_2	0.4	100	10000	9	0.119	0.107	0.102	0.100	0.085	0.113	0.113
p_3	0.9	100	10000	9	0.099	0.107	0.093	0.081	0.147	0.097	0.097
q_3	0.8	100	10000	9	0.132	0.121	0.104	0.092	0.141	0.122	0.122
p_1	0.4	160	10000	9	0.054	0.053	0.053	0.052	0.053	0.054	0.054
q_1	0.3	160	10000	9	0.051	0.050	0.049	0.049	0.057	0.050	0.051
p_2	0.4	160	10000	9	0.093	0.087	0.084	0.083	0.082	0.090	0.090
q_2	0.4	160	10000	9	0.093	0.087	0.084	0.083	0.070	0.090	0.090
p_3	0.9	160	10000	9	0.077	0.082	0.075	0.068	0.129	0.076	0.076
q_3	0.8	160	10000	9	0.104	0.098	0.089	0.082	0.119	0.099	0.099

Appendix Nr. 3. Estimated KL divergence for all cases

sample	n	par	p_1	q_1	p_2	q_2	p_3	q_3	MLE	U(0,1)	B12/B21	B13/B31	U(0,p)	Jeffreys	Ref
plac.-plac.	60	1	0.4	0.3	0.4	0.3	0.9	0.8	0.053	0.042	0.037	0.036	0.020	0.049	0.050
plac.-drug	60	1	0.4	0.3	0.4	0.3	0.9	0.8	0.057	0.042	0.038	0.039	0.031	0.049	0.050
drug-plac.	60	1	0.4	0.3	0.4	0.3	0.9	0.8	0.040	0.037	0.030	0.026	0.041	0.042	0.043
drug-drug	60	1	0.4	0.3	0.4	0.3	0.9	0.8	0.043	0.039	0.031	0.027	0.048	0.035	0.035
plac.-plac.	100	1	0.4	0.3	0.4	0.3	0.9	0.8	0.033	0.027	0.025	0.024	0.016	0.029	0.030
plac.-drug	100	1	0.4	0.3	0.4	0.3	0.9	0.8	0.034	0.027	0.026	0.026	0.023	0.030	0.030
drug-plac.	100	1	0.4	0.3	0.4	0.3	0.9	0.8	0.025	0.024	0.021	0.020	0.027	0.028	0.028
drug-drug	100	1	0.4	0.3	0.4	0.3	0.9	0.8	0.021	0.024	0.021	0.018	0.034	0.023	0.023
plac.-plac.	160	1	0.4	0.3	0.4	0.3	0.9	0.8	0.020	0.017	0.017	0.016	0.014	0.019	0.019
plac.-drug	160	1	0.4	0.3	0.4	0.3	0.9	0.8	0.020	0.018	0.017	0.017	0.017	0.019	0.019
drug-plac.	160	1	0.4	0.3	0.4	0.3	0.9	0.8	0.019	0.016	0.015	0.014	0.019	0.018	0.019
drug-drug	160	1	0.4	0.3	0.4	0.3	0.9	0.8	0.014	0.016	0.014	0.013	0.026	0.017	0.017
plac.-plac.	60	2	0.5	0.3	0.4	0.2	0.9	0.8	0.044	0.040	0.035	0.032	0.022	0.046	0.047
plac.-drug	60	2	0.5	0.3	0.4	0.2	0.9	0.8	0.058	0.043	0.039	0.039	0.039	0.050	0.051
drug-plac.	60	2	0.5	0.3	0.4	0.2	0.9	0.8	0.040	0.037	0.032	0.029	0.037	0.043	0.044
drug-drug	60	2	0.5	0.3	0.4	0.2	0.9	0.8	0.039	0.038	0.032	0.028	0.038	0.035	0.036
plac.-plac.	100	2	0.5	0.3	0.4	0.2	0.9	0.8	0.031	0.026	0.024	0.023	0.019	0.030	0.030
plac.-drug	100	2	0.5	0.3	0.4	0.2	0.9	0.8	0.034	0.027	0.026	0.026	0.026	0.030	0.030
drug-plac.	100	2	0.5	0.3	0.4	0.2	0.9	0.8	0.028	0.025	0.023	0.022	0.022	0.029	0.029
drug-drug	100	2	0.5	0.3	0.4	0.2	0.9	0.8	0.021	0.024	0.022	0.020	0.024	0.025	0.025
plac.-plac.	160	2	0.5	0.3	0.4	0.2	0.9	0.8	0.021	0.017	0.017	0.016	0.014	0.019	0.019
plac.-drug	160	2	0.5	0.3	0.4	0.2	0.9	0.8	0.020	0.018	0.017	0.017	0.018	0.019	0.019
drug-plac.	160	2	0.5	0.3	0.4	0.2	0.9	0.8	0.020	0.017	0.016	0.015	0.014	0.019	0.019
drug-drug	160	2	0.5	0.3	0.4	0.2	0.9	0.8	0.015	0.016	0.015	0.014	0.016	0.017	0.017
plac.-plac.	60	3	0.5	0.3	0.4	0.1	0.9	0.8	0.037	0.039	0.034	0.030	0.027	0.039	0.040
plac.-drug	60	3	0.5	0.3	0.4	0.1	0.9	0.8	0.058	0.042	0.039	0.039	0.039	0.049	0.050
drug-plac.	60	3	0.5	0.3	0.4	0.1	0.9	0.8	0.039	0.037	0.031	0.029	0.036	0.043	0.043
drug-drug	60	3	0.5	0.3	0.4	0.1	0.9	0.8	0.040	0.039	0.032	0.028	0.039	0.036	0.036
plac.-plac.	100	3	0.5	0.3	0.4	0.1	0.9	0.8	0.023	0.025	0.023	0.021	0.022	0.027	0.027
plac.-drug	100	3	0.5	0.3	0.4	0.1	0.9	0.8	0.033	0.027	0.026	0.026	0.026	0.030	0.030
drug-plac.	100	3	0.5	0.3	0.4	0.1	0.9	0.8	0.027	0.025	0.023	0.021	0.022	0.029	0.029
drug-drug	100	3	0.5	0.3	0.4	0.1	0.9	0.8	0.021	0.024	0.021	0.019	0.024	0.024	0.025
plac.-plac.	160	3	0.5	0.3	0.4	0.1	0.9	0.8	0.017	0.017	0.016	0.015	0.016	0.018	0.018
plac.-drug	160	3	0.5	0.3	0.4	0.1	0.9	0.8	0.020	0.018	0.017	0.017	0.018	0.019	0.019
drug-plac.	160	3	0.5	0.3	0.4	0.1	0.9	0.8	0.020	0.017	0.016	0.015	0.014	0.019	0.019
drug-drug	160	3	0.5	0.3	0.4	0.1	0.9	0.8	0.015	0.016	0.015	0.014	0.016	0.017	0.017
plac.-plac.	60	4	0.4	0.3	0.4	0.3	0.9	0.7	0.052	0.041	0.037	0.036	0.020	0.049	0.049
plac.-drug	60	4	0.4	0.3	0.4	0.3	0.9	0.7	0.056	0.041	0.038	0.038	0.031	0.048	0.049
drug-plac.	60	4	0.4	0.3	0.4	0.3	0.9	0.7	0.042	0.036	0.030	0.029	0.031	0.046	0.046
drug-drug	60	4	0.4	0.3	0.4	0.3	0.9	0.7	0.043	0.038	0.030	0.026	0.048	0.034	0.034
plac.-plac.	100	4	0.4	0.3	0.4	0.3	0.9	0.7	0.033	0.026	0.025	0.024	0.016	0.029	0.029
plac.-drug	100	4	0.4	0.3	0.4	0.3	0.9	0.7	0.033	0.027	0.026	0.026	0.023	0.029	0.030
drug-plac.	100	4	0.4	0.3	0.4	0.3	0.9	0.7	0.030	0.025	0.022	0.022	0.021	0.030	0.030
drug-drug	100	4	0.4	0.3	0.4	0.3	0.9	0.7	0.021	0.024	0.021	0.018	0.034	0.023	0.023
plac.-plac.	160	4	0.4	0.3	0.4	0.3	0.9	0.7	0.020	0.017	0.017	0.016	0.014	0.019	0.019
plac.-drug	160	4	0.4	0.3	0.4	0.3	0.9	0.7	0.020	0.018	0.017	0.017	0.017	0.019	0.019
drug-plac.	160	4	0.4	0.3	0.4	0.3	0.9	0.7	0.020	0.017	0.015	0.015	0.015	0.018	0.019
drug-drug	160	4	0.4	0.3	0.4	0.3	0.9	0.7	0.013	0.015	0.014	0.013	0.025	0.016	0.016
plac.-plac.	60	5	0.5	0.3	0.4	0.2	0.9	0.7	0.043	0.039	0.034	0.032	0.022	0.046	0.046
plac.-drug	60	5	0.5	0.3	0.4	0.2	0.9	0.7	0.056	0.042	0.038	0.039	0.038	0.049	0.050
drug-plac.	60	5	0.5	0.3	0.4	0.2	0.9	0.7	0.047	0.039	0.034	0.033	0.035	0.048	0.048
drug-drug	60	5	0.5	0.3	0.4	0.2	0.9	0.7	0.040	0.038	0.032	0.028	0.039	0.036	0.036

sample	n	par	p_1	q_1	p_2	q_2	p_3	q_3	MLE	U(0,1)	B12/B21	B13/B31	U(0,p)	Jeffreys	Ref
plac.-plac.	100	5	0.5	0.3	0.4	0.2	0.9	0.7	0.031	0.026	0.024	0.023	0.018	0.029	0.030
plac.-drug	100	5	0.5	0.3	0.4	0.2	0.9	0.7	0.033	0.027	0.025	0.026	0.026	0.029	0.030
drug-plac.	100	5	0.5	0.3	0.4	0.2	0.9	0.7	0.032	0.026	0.024	0.023	0.022	0.030	0.030
drug-drug	100	5	0.5	0.3	0.4	0.2	0.9	0.7	0.021	0.024	0.021	0.020	0.024	0.024	0.025
plac.-plac.	160	5	0.5	0.3	0.4	0.2	0.9	0.7	0.021	0.017	0.017	0.016	0.014	0.019	0.019
plac.-drug	160	5	0.5	0.3	0.4	0.2	0.9	0.7	0.020	0.017	0.017	0.017	0.017	0.018	0.019
drug-plac.	160	5	0.5	0.3	0.4	0.2	0.9	0.7	0.020	0.017	0.016	0.016	0.015	0.019	0.019
drug-drug	160	5	0.5	0.3	0.4	0.2	0.9	0.7	0.015	0.016	0.015	0.014	0.016	0.017	0.018
plac.-plac.	60	6	0.5	0.3	0.4	0.1	0.9	0.7	0.036	0.039	0.033	0.030	0.027	0.039	0.039
plac.-drug	60	6	0.5	0.3	0.4	0.1	0.9	0.7	0.056	0.042	0.039	0.039	0.038	0.049	0.050
drug-plac.	60	6	0.5	0.3	0.4	0.1	0.9	0.7	0.046	0.038	0.033	0.033	0.034	0.047	0.048
drug-drug	60	6	0.5	0.3	0.4	0.1	0.9	0.7	0.040	0.039	0.032	0.029	0.039	0.036	0.037
plac.-plac.	100	6	0.5	0.3	0.4	0.1	0.9	0.7	0.022	0.025	0.023	0.021	0.021	0.027	0.027
plac.-drug	100	6	0.5	0.3	0.4	0.1	0.9	0.7	0.033	0.027	0.025	0.025	0.026	0.029	0.029
drug-plac.	100	6	0.5	0.3	0.4	0.1	0.9	0.7	0.032	0.026	0.024	0.023	0.022	0.030	0.030
drug-drug	100	6	0.5	0.3	0.4	0.1	0.9	0.7	0.021	0.024	0.021	0.020	0.024	0.025	0.025
plac.-plac.	160	6	0.5	0.3	0.4	0.1	0.9	0.7	0.017	0.017	0.016	0.015	0.016	0.019	0.019
plac.-drug	160	6	0.5	0.3	0.4	0.1	0.9	0.7	0.020	0.018	0.017	0.017	0.017	0.019	0.019
drug-plac.	160	6	0.5	0.3	0.4	0.1	0.9	0.7	0.021	0.018	0.017	0.016	0.016	0.019	0.019
drug-drug	160	6	0.5	0.3	0.4	0.1	0.9	0.7	0.015	0.016	0.015	0.014	0.016	0.018	0.018
plac.-plac.	60	7	0.4	0.4	0.4	0.3	0.9	0.8	0.050	0.040	0.036	0.034	0.020	0.049	0.049
plac.-drug	60	7	0.4	0.4	0.4	0.3	0.9	0.8	0.056	0.042	0.038	0.038	0.032	0.050	0.050
drug-plac.	60	7	0.4	0.4	0.4	0.3	0.9	0.8	0.039	0.036	0.029	0.026	0.040	0.041	0.042
drug-drug	60	7	0.4	0.4	0.4	0.3	0.9	0.8	0.043	0.038	0.031	0.026	0.048	0.034	0.035
plac.-plac.	100	7	0.4	0.4	0.4	0.3	0.9	0.8	0.033	0.027	0.025	0.024	0.014	0.030	0.030
plac.-drug	100	7	0.4	0.4	0.4	0.3	0.9	0.8	0.034	0.027	0.026	0.026	0.021	0.030	0.031
drug-plac.	100	7	0.4	0.4	0.4	0.3	0.9	0.8	0.025	0.024	0.021	0.020	0.027	0.028	0.028
drug-drug	100	7	0.4	0.4	0.4	0.3	0.9	0.8	0.022	0.024	0.020	0.018	0.034	0.023	0.023
plac.-plac.	160	7	0.4	0.4	0.4	0.3	0.9	0.8	0.020	0.017	0.017	0.016	0.010	0.019	0.019
plac.-drug	160	7	0.4	0.4	0.4	0.3	0.9	0.8	0.020	0.017	0.017	0.017	0.014	0.018	0.019
drug-plac.	160	7	0.4	0.4	0.4	0.3	0.9	0.8	0.019	0.016	0.015	0.014	0.019	0.019	0.019
drug-drug	160	7	0.4	0.4	0.4	0.3	0.9	0.8	0.013	0.015	0.014	0.013	0.026	0.017	0.017
plac.-plac.	60	8	0.4	0.3	0.4	0.3	0.9	0.9	0.052	0.041	0.037	0.035	0.020	0.049	0.049
plac.-drug	60	8	0.4	0.3	0.4	0.3	0.9	0.9	0.057	0.042	0.039	0.039	0.032	0.050	0.050
drug-plac.	60	8	0.4	0.3	0.4	0.3	0.9	0.9	0.044	0.038	0.031	0.026	0.062	0.034	0.035
drug-drug	60	8	0.4	0.3	0.4	0.3	0.9	0.9	0.043	0.038	0.031	0.026	0.048	0.034	0.035
plac.-plac.	100	8	0.4	0.3	0.4	0.3	0.9	0.9	0.033	0.027	0.025	0.024	0.016	0.029	0.030
plac.-drug	100	8	0.4	0.3	0.4	0.3	0.9	0.9	0.033	0.027	0.025	0.025	0.023	0.030	0.030
drug-plac.	100	8	0.4	0.3	0.4	0.3	0.9	0.9	0.022	0.024	0.021	0.018	0.044	0.023	0.023
drug-drug	100	8	0.4	0.3	0.4	0.3	0.9	0.9	0.022	0.024	0.021	0.018	0.034	0.023	0.023
plac.-plac.	160	8	0.4	0.3	0.4	0.3	0.9	0.9	0.021	0.018	0.017	0.017	0.014	0.019	0.019
plac.-drug	160	8	0.4	0.3	0.4	0.3	0.9	0.9	0.020	0.018	0.017	0.017	0.017	0.019	0.019
drug-plac.	160	8	0.4	0.3	0.4	0.3	0.9	0.9	0.014	0.015	0.014	0.013	0.033	0.016	0.017
drug-drug	160	8	0.4	0.3	0.4	0.3	0.9	0.9	0.014	0.016	0.014	0.013	0.026	0.017	0.017
plac.-plac.	60	9	0.4	0.3	0.4	0.4	0.9	0.8	0.057	0.042	0.038	0.039	0.028	0.049	0.050
plac.-drug	60	9	0.4	0.3	0.4	0.4	0.9	0.8	0.057	0.042	0.039	0.039	0.031	0.049	0.050
drug-plac.	60	9	0.4	0.3	0.4	0.4	0.9	0.8	0.038	0.035	0.028	0.025	0.039	0.040	0.041
drug-drug	60	9	0.4	0.3	0.4	0.4	0.9	0.8	0.043	0.038	0.030	0.026	0.047	0.034	0.034
plac.-plac.	100	9	0.4	0.3	0.4	0.4	0.9	0.8	0.034	0.027	0.026	0.026	0.020	0.030	0.030
plac.-drug	100	9	0.4	0.3	0.4	0.4	0.9	0.8	0.034	0.028	0.026	0.026	0.023	0.030	0.030
drug-plac.	100	9	0.4	0.3	0.4	0.4	0.9	0.8	0.026	0.025	0.022	0.020	0.027	0.028	0.029
drug-drug	100	9	0.4	0.3	0.4	0.4	0.9	0.8	0.022	0.024	0.021	0.019	0.034	0.024	0.024
plac.-plac.	160	9	0.4	0.3	0.4	0.4	0.9	0.8	0.020	0.018	0.017	0.017	0.015	0.019	0.019
plac.-drug	160	9	0.4	0.3	0.4	0.4	0.9	0.8	0.020	0.017	0.017	0.017	0.017	0.018	0.018
drug-plac.	160	9	0.4	0.3	0.4	0.4	0.9	0.8	0.019	0.017	0.015	0.015	0.019	0.019	0.019
drug-drug	160	9	0.4	0.3	0.4	0.4	0.9	0.8	0.014	0.016	0.014	0.013	0.026	0.017	0.017

Appendix Nr. 4. R code for parameters estimation of one configuration

```
library(FME)
library(tidyverse)
library(doParallel)
rezults <- list()
p.all <- list(c(0.4, 0.4, 0.9), c(0.5, 0.4, 0.9), c(0.5, 0.4, 0.9),
              c(0.4, 0.4, 0.9), c(0.5, 0.4, 0.9), c(0.5, 0.4, 0.9),
              c(0.4, 0.4, 0.9), c(0.4, 0.4, 0.9), c(0.4, 0.4, 0.9))
q.all <- list(c(0.3, 0.3, 0.8), c(0.3, 0.2, 0.8), c(0.3, 0.1, 0.8),
              c(0.3, 0.3, 0.7), c(0.3, 0.2, 0.7), c(0.3, 0.1, 0.7),
              c(0.4, 0.3, 0.8), c(0.3, 0.3, 0.9), c(0.3, 0.4, 0.8))
priors <- 7
k <- 1
n <- 60
# sample sizes
b <- 1/2
n <- c(b/2*n, b/2*n, (1-b)/2*n,(1-b)/2*n)
# true parameters values
p <- p.all[[k]]
q <- q.all[[k]]
# both s2 and s3 are fixed at ratio 1
s2 <- 1
s3 <- 1
# generating data
pp <- c(s2*(1-q[1])*q[2], s2*(1-q[1])*(1-q[2]), q[1], (1-s2)*(1-q[1]))
pd <- c(s2*(1-q[1])*p[2], s2*(1-q[1])*(1-p[2]), q[1], (1-s2)*(1-q[1]))
dp <- c(1-p[1], s3*p[1]*q[3], s3*p[1]*(1-q[3]), (1-s3)*p[1])
dd <- c(1-p[1], s3*p[1]*p[3], s3*p[1]*(1-p[3]), (1-s3)*p[1])
prob <- list(pp, pd, dp, dd)

m <- 10000
cl <- makeCluster(4)
```

```
registerDoParallel(cl)
M <- foreach(i = 1:m, .combine = rbind, .packages = c("FME")) %dopar% {
  count <- c()
  M <- numeric(priors*6)
  for (j in 1:4) {
    pr <- prob[[j]]
    u <- runif(n = n[j])
    u <- ifelse(u <= pr[1], 1, ifelse(u <= pr[1]+pr[2], 2, ifelse(u <= pr
        [1]+pr[2]+pr[3], 3, 4)))
    count <- c(count, sum(u==1), sum(u==2), sum(u==3), sum(u==4))
  }

  # ML estimates
  M[1] <- (n[3]+n[4]-count[9]-count[13])/(n[3]+n[4])
  M[2] <- (count[3]+count[7])/(n[1]+n[2])
  M[3] <- (count[5])/(count[5]+count[6])
  M[4] <- (count[1])/(count[1]+count[2])
  M[5] <- (count[14])/(count[14]+count[15])
  M[6] <- (count[10])/(count[10]+count[11])

  # Bayes uniform
  alpha <- 1
  beta <- 1
  M[7] <- (n[3]+n[4]-count[9]-count[13]+alpha)/(n[3]+n[4]+alpha+beta)
  M[8] <- (count[3]+count[7]+alpha)/(alpha+n[1]+n[2]+beta)
  M[9] <- (count[5]+alpha)/ (count[5]+alpha+count[6]+beta)
  M[10] <- (count[1]+alpha)/(count[1]+alpha+count[2]+beta)
  M[11] <- (count[14]+alpha)/(count[14]+alpha+count[15]+beta)
  M[12] <- (count[10]+alpha)/(count[10]+alpha+count[11]+beta)

  # Bayes beta(1,2) or beta (2,1)
  M[13] <- (n[3]+n[4]-count[9]-count[13]+1)/(n[3]+n[4]+1+2)
  M[14] <- (count[3]+count[7]+1)/(1+n[1]+n[2]+2)
```

```
M[15] <- (count[5]+1)/ (count[5]+1+count[6]+2)
M[16] <- (count[1]+1)/(count[1]+1+count[2]+2)
M[17] <- (count[14]+2)/(count[14]+2+count[15]+1)
M[18] <- (count[10]+2)/(count[10]+2+count[11]+1)

# Bayes beta(1,3) or beta (3,1)
M[19] <- (n[3]+n[4]-count[9]-count[13]+1)/(n[3]+n[4]+1+3)
M[20] <- (count[3]+count[7]+1)/(1+n[1]+n[2]+3)
M[21] <- (count[5]+1)/ (count[5]+1+count[6]+3)
M[22] <- (count[1]+1)/(count[1]+1+count[2]+3)
M[23] <- (count[14]+3)/(count[14]+3+count[15]+1)
M[24] <- (count[10]+3)/(count[10]+3+count[11]+1)
#
ll2 <- function(par) {
  p1 <- par[1]
  f1=(n[3]+n[4]-count[9]-count[13])*log(par[1]) + (count[9]+count[13])*
     log(1-par[1]) +
    (count[3]+count[7])*log(par[2]) + (n[1]+n[2]-count[3]-count[7])*log(1-
       par[2]) +
    count[5]*log(par[3]) + count[6]*log(1-par[3]) + count[1]*log(par[4]) +
        count[2]*log(1-par[4]) +
    count[14]*log(par[5]) + count[15]*log(1-par[5]) + count[10]*log(par
       [6]) + count[11]*log(1-par[6])
  f = -2*f1
  return(f)
}

prior <- function(par) {
  f2=dunif(par[1], 0, 1, log=TRUE) + dunif(par[2], 0, p[1], log=TRUE) +
    dunif(par[3], 0, 1, log=TRUE) + dunif(par[4], 0, p[2], log=TRUE) +
    dunif(par[5], 0, 1, log=TRUE) + dunif(par[6], 0, p[3], log=TRUE)
  f = -2*f2
  return(f)
```

```
  }

  # Initial values
  ini <- c(1/2, 1/2, 1/2, 1/2, 1/2, 1/2)
  low <- rep(1e-10, 6)
  up <- rep(0.99999999, 6)

  # Bayes bounded u(0,p)
  m1 <- modMCMC(ll2, ini, lower = low, upper = up, prior = prior,
              niter = 10000, burninlength = 1000, verbose = FALSE)
  M[25:30] <- apply(m1$pars, 2, mean)

  # Jeffreys
  alpha <- 1/2
  beta <- 1/2
  M[31] <- (n[3]+n[4]-count[9]-count[13]+3/2)/(n[3]+n[4]+3/2+1/2)
  M[32] <- (count[3]+count[7]+1/2)/(1/2+n[1]+n[2]+3/2)
  M[33] <- (count[5]+alpha)/(count[5]+alpha+count[6]+beta)
  M[34] <- (count[1]+alpha)/(count[1]+alpha+count[2]+beta)
  M[35] <- (count[14]+alpha)/(count[14]+alpha+count[15]+beta)
  M[36] <- (count[10]+alpha)/(count[10]+alpha+count[11]+beta)

  # Reference
  M[37] <- (n[3]+n[4]-count[9]-count[13]+alpha)/(n[3]+n[4]+alpha+beta)
  M[38] <- (count[3]+count[7]+alpha)/(alpha+n[1]+n[2]+beta)
  M[39] <- (count[5]+alpha)/(count[5]+alpha+count[6]+beta)
  M[40] <- (count[1]+alpha)/(count[1]+alpha+count[2]+beta)
  M[41] <- (count[14]+alpha)/(count[14]+alpha+count[15]+beta)
  M[42] <- (count[10]+alpha)/(count[10]+alpha+count[11]+beta)
  M
}
stopCluster(cl)
colnames(M) <- rep(c("p1", "q1", "p2", "q2", "p3", "q3"), priors)
```

Appendix Nr. 5. R code for power simulations of one test

```
library(FME)
library(doParallel)
modMCMC1 <- modMCMC
# modify MCMC
body(modMCMC1)[[59]][[4]][[4]][[2]] <-
  substitute(any(parnew < lower) | any(parnew > upper) |
              any(parnew[1] + parnew[2:4] > 1) | any(parnew[1] + parnew[2:4]
                  < 0))

body(modMCMC1)[[59]][[4]][[6]][[3]][[11]][[3]][[5]][[2]] <-
  substitute(any(parnew < lower) | any(parnew > upper) |
              any(parnew[1] + parnew[2:4] > 1) | any(parnew[1] + parnew[2:4]
                  < 0))

### Values from article
p.all <- list(c(0.4, 0.4, 0.9), c(0.5, 0.4, 0.9), c(0.5, 0.4, 0.9),
              c(0.4, 0.4, 0.9), c(0.5, 0.4, 0.9), c(0.5, 0.4, 0.9),
              c(0.4, 0.4, 0.9), c(0.4, 0.4, 0.9), c(0.4, 0.4, 0.9))
q.all <- list(c(0.3, 0.3, 0.8), c(0.3, 0.2, 0.8), c(0.3, 0.1, 0.8),
              c(0.3, 0.3, 0.7), c(0.3, 0.2, 0.7), c(0.3, 0.1, 0.7),
              c(0.4, 0.3, 0.8), c(0.3, 0.3, 0.9), c(0.3, 0.4, 0.8))
n.all1 <- c(412, 128, 96, 312, 104, 80, 2612, 728, 644)
power11 <- numeric(9)
power13 <- numeric(9)

for (k in 1:9) {
  # sample sizes
  b <- 1/2
  n <- n.all1[k]
  n1 <- c(b/2*n, b/2*n, (1-b)/2*n,(1-b)/2*n)
  # true parameters values
```

```
p <- p.all[[k]]
q <- q.all[[k]]
# both s2 and s3 are fixed at ratio 1
s2 <- 1
s3 <- 1
# generating data
pp <- c(s2*(1-q[1])*q[2], s2*(1-q[1])*(1-q[2]), q[1], (1-s2)*(1-q[1]))
pd <- c(s2*(1-q[1])*p[2], s2*(1-q[1])*(1-p[2]), q[1], (1-s2)*(1-q[1]))
dp <- c(1-p[1], s3*p[1]*q[3], s3*p[1]*(1-q[3]), (1-s3)*p[1])
dd <- c(1-p[1], s3*p[1]*p[3], s3*p[1]*(1-p[3]), (1-s3)*p[1])
prob <- list(pp, pd, dp, dd)
# Likelihood functions
ll1 <- function(par) {
  f1=(n1[3]+n1[4]-count1[9]-count1[13])*log(par[1]+par[2]) + (count1[9]+
      count1[13])*log(1-par[1]-par[2]) +
    (count1[3]+count1[7])*log(par[2]) + (n1[1]+n1[2]-count1[3]-count1[7])*
       log(1-par[2]) +
    count1[5]*log(par[1]+par[3]) + count1[6]*log(1-par[1]-par[3]) + count1
       [1]*log(par[3]) + count1[2]*log(1-par[3]) +
    count1[14]*log(par[1]+par[4]) + count1[15]*log(1-par[1]-par[4]) +
       count1[10]*log(par[4]) + count1[11]*log(1-par[4])
  f = -2*f1
  return(f)
}

# Priors
prior11 <- function(par) {
  f2=dunif(par[1], -1, 1, log=TRUE) + dunif(par[2], 0, 1, log=TRUE) +
    dunif(par[3], 0, 1, log=TRUE) + dunif(par[4], 0, 1, log=TRUE)
  f = -2*f2
  return(f)
}
prior13 <- function(par) {
```

```
  d1 <- par[1]
  q1 <- par[2]
  q2 <- par[3]
  q3 <- par[4]
  f11 = -(n1[4]*(d1+q1)/(d1+q3)+n1[4]*(d1+q1)/(1-d1-q3))^2*
    ((n1[3]+n1[4])/(d1+q1)+(n1[3]+n1[4])/(1-d1-q1)+(n1[1]+n1[2])/q1+(n1
        [1]+n1[2])/(1-q1))*
    (n1[2]*(1-q1)/(d1+q2)+n1[2]*(1-q1)/(1-d1-q2)+n1[1]*(1-q1)/q2+n1[1]*(1-
        q1)/(1-q2))
  f12 = -(n1[4]*(d1+q1)/(d1+q3)+n1[4]*(d1+q1)/(1-d1-q3)+n1[3]*(d1+q1)/q3+
      n1[3]*(d1+q1)/(1-q3))*
    (n1[2]*(1-q1)/(d1+q2)+n1[2]*(1-q1)/(1-d1-q2))^2*
    ((n1[3]+n1[4])/(d1+q1)+(n1[3]+n1[4])/(1-d1-q1)+(n1[1]+n1[2])/q1+(n1
        [1]+n1[2])/(1-q1))
  f13 = (n1[4]*(d1+q1)/(d1+q3)+n1[4]*(d1+q1)/(1-d1-q3)+n1[3]*(d1+q1)/q3+n1
      [3]*(d1+q1)/(1-q3))*
    (n1[2]*(1-q1)/(d1+q2)+n1[2]*(1-q1)/(1-d1-q2)+n1[1]*(1-q1)/q2+n1[1]*(1-
        q1)/(1-q2))*
    (((n1[3]+n1[4])/(d1+q1)+(n1[3]+n1[4])/(1-d1-q1)+n1[2]*(1-q1)/(d1+q2)+
        n1[2]*(1-q1)/(1-d1-q2)+
        n1[4]*(d1+q1)/(d1+q3)+n1[4]*(d1+q1)/(1-d1-q3))*
       ((n1[3]+n1[4])/(d1+q1)+(n1[3]+n1[4])/(1-d1-q1)+(n1[1]+n1[2])/q1+(n1
           [1]+n1[2])/(1-q1))-
       ((n1[3]+n1[4])/(d1+q1)+(n1[3]+n1[4])/(1-d1-q1))^2
    )
  f2 = sqrt(f11+f12+f13)
  f = -2*log(f2)
  return(f)
}

# Initial values
ini1 <- c(0.2, 0.5, 0.5, 0.5)
low1 <- c(-1, rep(1e-10, 3))
```

```
  up1 <- rep(0.99999999, 4)

  m <- 10000
  cl <- makeCluster(4)
  registerDoParallel(cl)
  M <- foreach(i = 1:m, .combine = rbind, .packages = c("FME")) %dopar% {
    # Data simulations
    count1 <- c()

    for (j in 1:4) {
      pr <- prob[[j]]
      u <- runif(n = n1[j])
      u <- ifelse(u <= pr[1], 1, ifelse(u <= pr[1]+pr[2], 2, ifelse(u <= pr
          [1]+pr[2]+pr[3], 3, 4)))
      count1 <- c(count1, sum(u==1), sum(u==2), sum(u==3), sum(u==4))
    }

     # Models
     m11 <- modMCMC1(ll1, ini1, lower = low1, upper = up1, prior = prior11,
                 niter = 100000, burninlength = 10000, verbose = FALSE)
     m13 <- modMCMC1(ll1, ini1, lower = low1, upper = up1, prior = prior13,
                  niter = 100000, burninlength = 10000, verbose = FALSE)
     # Results
     data.frame(hip11 = (quantile(m11$pars[,1], 0.05) > 0),
              hip13 = quantile(m13$pars[,1], 0.05) > 0)
  }
  stopCluster(cl)
  power11[k] <- sum(M$hip11)/m
  power13[k] <- sum(M$hip13)/m
}
```