A BAYESIAN DESIGN TO TEST DOSE-RESPONSE SEQUENTIAL TRIALS USING ISOTONIC REGRESSION

Edcarlos Miranda de SOUZA[1]
Lucas Monteiro CHAVES[2]
Joel Augusto MUNIZ[2]
Devanil Jaques de SOUZA[2]

- ABSTRACT: The present work aims to achieve a dose with a pre-specified toxicity rate in a target population. It proposes, using Bayesian methods and isotonic regression, a sequential design to allow the researchers to update prior information. The theory behind the design is not easily assimilated to non-statisticians. What makes such approach attractive is its reduction to four steps of easy manipulation, without the need of computational effort. Simulation procedures confirm the effectiveness of the proposed methodology.

- KEYWORDS: Sequential design; up and down method; biased coin design; Narayana design.

1 Introduction

There are many experimental situations where one wants to measure the effectiveness of a given drug in a population. A usual procedure is the application of the drug in several levels, observing the response by counting the number of positive response. These are called dose-response trials. Several methods have been proposed in the literature to work in trials with these characteristics.

Sequential designs have received particular attention in dose-response trials where a reduced number of sampling units is desired, because of financial or ethical questions, as in the case of research with humans.

Among the situations that use dose-response trials, experiments with insecticides, clinical studies with cancer, acute toxicity or pharmaceutical research are, in general, highlighted. In these situations, it is reasonable to suppose that, with the increase of dose, the probability of positive response also increases, and an estimation procedure should incorporate this fact. Given an arbitrary curve, the mathematical procedure that obtains the growth curves as close as possible, in the sense of the least squares, is the isotonic regression. This theory is a natural tool in regards to dose-response trials.

[1] Federal University of Acre, Center for Technology and Exact Sciences, Postal Box: 500, Postal Code: 69915-900 , Rio Branco, AC, Brazil. E-mail: *profedcarlos@hotmail.com*
[2] Federal University of Lavras, Department of Exact Science, Postal Box: 3037, Postal Code: 37200-000, Lavras, MG, Brazil. E-mail: *lucas@dex.ufla.br*, *joamuniz@ufla.br*, *DevanilJaques@dex.b*r

In a dose-response trial, it is natural to suppose some prior information about the toxicity probabilities for the doses. In general, this knowledge is used previously to choose the doses to be tested. In this sense, it is reasonable to use Bayesian theory. However, the reported uses of Bayesian methods are overly complicated, what has hampered its practical use.

This work was carried out in order to propose a sequential design for dose-response trials that uses Bayesian methodology and the theory of isotonic regression, so that this design provides for the researcher, in a simple way, a route to be followed. The performance of the proposed procedures was evaluated using Mean Squared Error and Estimators Bias as criteria, through Monte Carlo simulation.

2 Dose-response trials

Dose-response trials are experimental designs used to study the effect, for instance, of increasing levels of a drug in a population. Bliss (1935) modeled the phenomenon defining a non-observable random variable, called Tolerance. Tolerance is the smallest amount of drug that produces an effect in an individual randomly picked in a population. The *probits* theory was developed assuming Tolerance as a normal random variable. Berkson (1944) supposed logistic distribution for the Tolerance and obtained the theory of *logits*.

One of the major problems studied is the determination of the 50% lethal dose ($LD_{50\%}$), which is the dose that produces, in average, effect in 50% of the population. The term lethal comes from the fact that Bliss used his theory to study the effect of insecticides. For reliable inferences bases on *probit* or *logit* analysis, large numbers of sampling units must receive either high or low doses. For some clinical trials, large samples and extreme doses may not be acceptable by ethical reasons, requiring the use of alternative methods.

3 Sequential designs

3.1 Up and down method

In 1948, working on experiments with explosives in the Explosives Research Laboratory in Bruceton, Pennsylvania, Dixon and Mood (1948) developed an alternative method to the *probit*, which was denominated by up and down method. The method is based on the theory of random walk and can be described in the language of dose-response trials as follows: Suppose the doses $d_1 < d_2 < \ldots < d_s$. A dose d_j is selected to start the process.

If the patient responds positively, the dose given to the next patient will be d_{j-1}. In the case of negative response the next dose will be d_{j+1}. The process is repeated until a predetermined number of patients is tested or the process is terminated by some sort of stopping rule. The idea is to see the trial as a random walk in the doses. If a patient receives dose d_i, let the transition probabilities p_i be the probability that the next patient receive the dose d_{i+1} and q_j the probability that the next patient receives a dose d_{i-1}. Assuming that the Tolerance is modeled by a $normal\ (\mu, \sigma^2)$ with cumulative distribution function Φ_{μ,σ^2}, the transition probabilities are given by:

$$p_i = \Phi_{\mu,\sigma^2}(d_i) \quad \textit{and} \quad q_i = 1 - p_i.$$

Now, with a result of random walk, the parameters μ and σ^2 can be estimated by the maximum likelihood method. The great advantage of such method is that the walk quickly converges to doses near LD_{50}. Thus, the sample size required to accurately estimate LD_{50} is about 35 percent less than the sample size necessary to obtain the same accuracy when using the *probit* method (DIXON; MOOD, 1948). Another advantage is, as the walk quickly stabilizes at doses near the LD_{50}, few individuals are subjected to very low or very high doses, a fundamental matter, by ethical reasons, when testing new drugs. A disadvantage of the method is that it is not efficient to estimate the lethal dose $LD_{\Gamma,}$ for Γ near 0% or 100%.

3.2 The biased coin up and down design

Several sequential designs has been proposed to overcome the problem of estimating proportions close to the tails, like LD_{10} or LD_{33}, for instance usually employed in phase I of oncology clinical trials.

Derman (1957) proposed a modification of the *up and down* method, in the sense that the random walk obtained quickly concentrates, for the specified Γ, on the dose θ_Γ:

$$\theta_\Gamma = F_T^{-1}(\Gamma) = inf\,\{\theta: F_T(\theta) \geq \Gamma\}, \tag{1}$$

with $0 < \Gamma < 1$. Such design, studied with modifications by Stylianou and Flournoy (2002), is called Biased Coin Design (BCD). There are two situations to be considered: $\Gamma \leq 0.5$ or $\Gamma > 0.5$.

In the situation that $\Gamma \leq 0.5$, the procedure begins by applying a dose d_i in a patient. If toxicity is observed, the dose to the next patient is d_{i-1}. If, otherwise, no toxicity is observed, the procedure is randomized as follows: with probability $b = \Gamma/(1 - \Gamma)$, the dose d_{i+1} is applied and with probability $1 - b$ the same dose d_i is applied.

When $\Gamma > 0.5$, the procedure starts applying a dose d_i in a patient. If no toxicity is observed, the dose to next patient is d_{i+1}. If, otherwise, toxicity is observed, the procedure is randomized: with probability $b = (1 - \Gamma)/\Gamma$, the dose d_{i-1} is applied and, with probability $1 - b$, the same dose d_i. Figure 1 shows the BCD summary.

In an interesting article, using the theory of Markov Chains, Giovagnoli and Pintacuda (1998) defined, in a general manner, designs based on random walks, generalizing the BCD. In that article, a rigorous proof based on invariant distributions for Markov chains is used and the dose θ_Γ is the most visited dose along the walk.

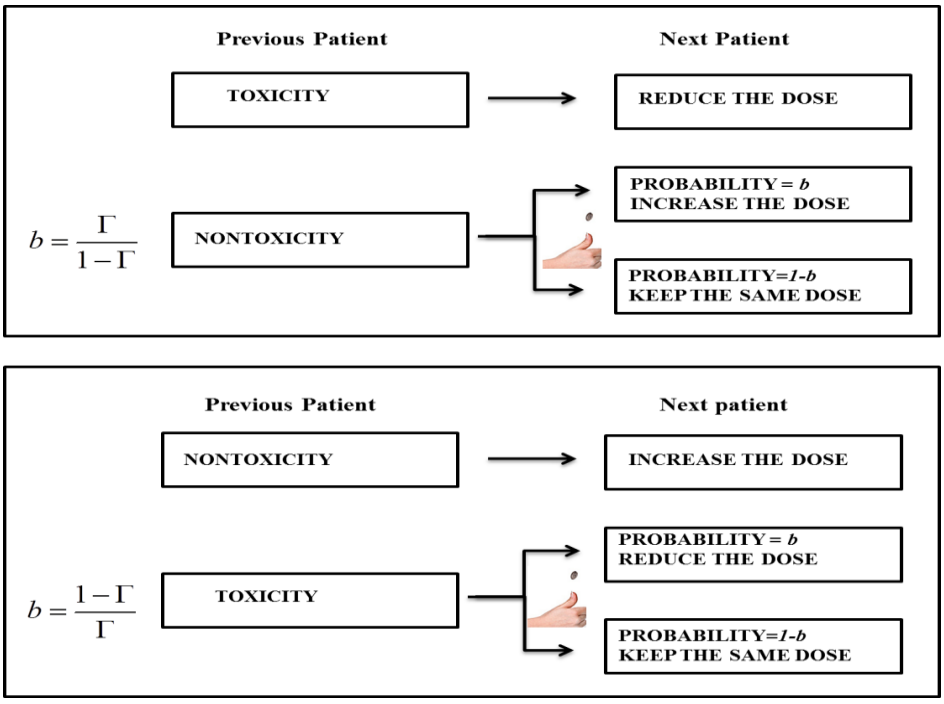

Figura 1 - Summary of the Biased Coin Design.

3.3 The Narayana *up and down* design

According to Ivanova et al. (2003), Narayana introduced, in his master's thesis, a rule to estimate $\theta_{0.50}$. The random walk is defined as follows: assume that the *n-th* patient received the dose d_j. Define $N_j(n)$ and $R_j(n)$ as the number of patients who underwent the dose d_j and the number of those patients who had positive response, respectively, including the *n-th* patient. The rule to provide the dose level for the *(n+1)-th* patient is:

1. Apply the dose d_{j-1}, if $R_j(n)/N_j(n) > 0.5$ and the *n-th* patient has positive response;
2. Apply the dose d_{j+1}, if $R_j(n)/N_j(n) < 0.5$ and the *n-th* patient has negative response;
3. d_j, otherwise.

3.4 The modified Narayana *up and down* design

Ivanova et al. (2003) changed the original Narayana design for $\Gamma = 1 - 0{,}5^{1/k}$, with k a fixed integer. This was done by adding a new condition, based on moving averages. The rule is defined as follows: suppose that the *n-th* patient was allocated at the level d_j. The next patient will have the dose level:

1. d_{j-1}, if $R_j(n)/N_j(n) > \Gamma$ and if there is at least one toxicity among the most recent *k* responses in the current dose level d_j;
2. d_{j+1}, if $R_j(n)/N_j(n) < \Gamma$ and if there is no toxicity among the most recent *k* responses in the current dose level d_j;
3. d_j, otherwise.

3.5 Bayesian designs

In many clinical trials, particularly in chemotherapy trials on patients with cancer, it is usual to have a good idea about the dose-response curve. Therefore, the Bayesian approach seems natural **(**O'Quigley, Pepe and Fisher (1990); Ji, Li and Yin (2007)). Although the Bayesian methodology has shown good results, it is not very used, in practice, due the need of a statistician carrying out the design and analysis of the experiment (TOURNEAU; LEE; SIU, 2009).

3.6 Estimation process

Several estimators for the dose θ_Γ (with specified Γ) have been proposed. Parametric estimators assume a class of parametric distribution for the Tolerance, such as normal or logistic. In the nonparametric approach, no assumption is made about the distribution of Tolerance.

Stylianou and Flournoy (2002) evaluated, via simulation, the performance of two parametric estimators based on the *logit* method, using weighted least squares and maximum likelihood. Maximum likelihood method is also evaluated in the works of Durham, Flournoy, and Rosenberger (1997), Ivanova et al. (2003), Pace and Stylianou (2007) and Gerke and Siedentop (2007).

In many dose-response trials, one do not have a pre-defined parametric model and, so, nonparametric estimators are required. One of these methods, *Empirical Estimator of the Media*, is the simple truncated average of some observations. Consider that a sequential dose-response trial is performed and n_i represents the number of times that the walk visited the dose d_i. If n_i is equal to 1 or 0, it means that the walk visited the dose only once or did not visit this dose, so that this dose will be discarded in the estimation process. A natural estimator for the target dose is:

$$\hat{\theta}_\Gamma = \frac{1}{M}\sum_{n_i \neq 1 \; ou \; 0} d_i n_i \qquad i = 1, 2, \ldots, k \;, \tag{2}$$

with $\mathrm{M} = \sum_{n_i \neq 1 \; ou \; 0} n_i \quad and \quad i = 1, 2, \ldots, k,$ where k represents the number of doses previously established in the experiment.

Variations of this estimator are found in Stylianou and Flournoy (2002) and Pace and Stylianou (2007). The estimator suggested in these works performs a truncation in some observations, which consists in discarding the first identical values, which generally appear to an initial value distant from the dose of interest. They argue that this estimator has been shown superior to the simple estimator of the media, especially when the initial dose is far from θ_Γ.

4 Isotonic regression

For a given set of pre-established values $X = \{x_1, x_2, \ldots, x_k\}$, with $x_1 < x_2 < \ldots < x_k$, let $y_j(x_i)$, with $j = 1,2, \ldots, m_j$ be measures of some variable of interest (response variable). For instance, X could be the drug doses being tested and $y_j(x_i)$ the response to the *j-th* repetition of dose x_i (1 if positive and 0 if negative).

In certain situations, it is reasonable to assume that $p(x) = E[Y|X = x]$, where $Y|X = x$ is a Bernoulli variable and $p(x)$ is a non-decreasing function of x. An estimation of $p(x)$ could be obtained by minimizing the sum of squares

$$\sum_{i,j}\left(y_j(x_i) - f(x_i)\right)^2$$

in the class of non-decreasing (***isotonic***) functions f on X, that is, in the class of functions f such that $x_i < x_j$ implies $f(x_i) \leq f(x_j)$.

Definiton. Given a function g, an isotonic function g^* is an isotonic regression of g with weights w, if it minimizes, in the class of isotonic functions f on X, the sum

$$\sum_{x \in X} [g(x) - f(x)]^2 w(x). \tag{3}$$

Isotonic regression calculation can be obtained by the algorithm PAVA (Pool Adjacent Violators) (BARLOW et al., 1972). A version of this algorithm, implemented in statistical software R, may be found in Leeuw, Hornik and Mair (2009). In that version, more general results of the isotonic regression was implemented, such as isotonic regression on a partially ordered sets.

The problem that arises in dose-response trials, with the use of the observed success rate $p_i = \sum_{j=1}^{m_i} y_j(x_i)/m_i$ to estimate $p_i = p(x_i)$, is that the estimates may not be increasing with the dose. In that case isotonic estimation can be used. Stylianou and Flournoy (2002), using *up and down* designs, presents two isotonic estimators. These estimators are also discussed later in Stylianou, Proschan and Flournoy (2003) and Pace and Stylianou (2007). Paul, Rosenberger and Flournoy (2004) developed another one for the case of non-binary response.

4.1 Maximum likelihood estimation of ordered parameters in exponential family

Consider a family of densities with two parameters, θ and h,

$$f(y; \theta, h) = exp\{[\Phi(\theta) + (y - \theta)\varphi(\theta)]h\}, \tag{4}$$

with respect to some measure $v_h(dy)$ (not necessarily specified), which may depend only on h. The function $\Phi(.)$ is strictly convex, $\varphi(.)$ is a determination of its derivative, say its derivative on the right. The argument θ takes values in a range of real numbers and h is positive. Barlow et al. (1972) showed that θ is the mean of these distributions and the variance is:

$$\frac{1}{h\,\varphi'(\theta)}.$$

The Bernoulli distribution with parameter p belongs to this family. Just take $\theta = p$, $\Phi(\theta) = \theta log\theta + (1-\theta)\log(1-\theta)$, $h = 1$ and $v_h(dy) = dy$. With convenient parameterization it can be shown that usual distributions, such as Normal, Poisson, Gamma, Exponential and so on, also belongs to this family (SOUZA, 2011; SOUZA, CHAVES and MUNIZ, 2012).

4.2 Bayesian estimation and isotonic regression

Suppose, for each $x \in X$, $m(x)$ observations of the distribution (4), with $\theta = \mu(x)$, $h = v(x)$, where $\mu(.)$ is unknown and $v(.)$ is known. Namely:

$$f(y; \mu(x), v(x)) = exp\{(\Phi(\mu(x)) + [y - \mu(x)]\varphi(\mu(x))v(x)\}. \quad (5)$$

Suppose it is also known that $\mu(.)$ is isotonic with respect to a partial order on X. The researcher can previously select a distribution for the parameters $\mu(x)$, for each $x \in X$, according to some previous knowledge, as it is usual in Bayesian approach.

Consider a priori density $\pi(\theta)$, with respect to some measure $d_\rho(\theta)$, with parameters θ_0 and $0 < h_0$:

$$\pi(\theta) \propto exp\{[\Phi(\theta) + (\theta_0 - \theta)\varphi(\theta)]h\} \quad (6)$$

Statement 4.2.1 The function $\Psi(\theta) = \Phi(\theta) + (\theta_0 - \theta)\varphi(\theta)$ assumes its maximum at $\theta = \theta_0$.

It follows from Statement 4.2.1 that θ_0 is the mode of the density (6). The larger the parameter h_0, the more leptokurtic the density is; if $h_0 = 0$ the density is uniform (non informative). The parameter h_0 is called *precision parameter*.

Suppose that, for each $x \in X$, a priori distribution for $\mu(x)$ is considered, with a precision parameter $h_0(x)$ and:

$$exp\{(\Phi(\theta) + [\mu_0(x) - \theta]\varphi(\theta))h_0(x)\}.$$

To obtain a priori for the vector $\boldsymbol{\mu} = (\mu(x_1), \mu(x_2), \ldots, \mu(x_k))$, the components are considered independent. Thus, the priori distribution for the vector $\boldsymbol{\mu}$ is:

$$\pi[\boldsymbol{\mu}] \propto \prod_{x \in X} exp\{(\Phi(\mu(x)) + [\mu_0(x) - \mu(x)]\varphi(\mu(x)))h_0(x)\}$$

$$\propto exp\left\{\sum_{x \in X}\left(\left(\Phi(\mu(x)) + [\mu_0(x) - \mu(x)]\varphi(\mu(x))\right)h_0(x)\right)\right\}.$$

If $\mu(x)$ is assumed to be isotonic, the vector $\boldsymbol{\mu}$ must have increasing coordinates and, therefore, the researcher must select an isotonic modal function $\mu_0(x)$. One way to incorporate the order restriction to the priori for the vector $\boldsymbol{\mu}$ is to define this priori as follows:

If $\mu(.)$ is isotonic on X:

$$\pi[\mu] \propto \prod_{x \in X} exp\left\{\left(\Phi\big(\mu(x)\big) + [\mu_0(x) - \mu(x)]\varphi\big(\mu(x)\big)\right) h_0(x)\right\} \tag{7}$$

Otherwise: $\pi[\mu] = 0$.

The mode of density (7) is $\boldsymbol{\mu_0} = (\mu_0(x_1), \mu_0(x_2), \ldots, \mu_0(x_k))$, and the function $\boldsymbol{h_0} = (h_0(x_1), h(x_2), \ldots, h(x_k))$ determines the concentration of this density around the mode. For $\boldsymbol{h_0} = \mathbf{0}$ there is a uniform priori (note that a prior in a k-dimensional space is being considered).

Suppose that for each $x \in X$ the researcher has taken $m(x)$ observations. If $y_j(x)$ represents the j-th observation taken at x, the total sample may be written as the vector $y = \big(y_1(x_1), \ldots, y_{m(x_1)}(x_1); \ldots; y_1(x_k), \ldots, y_{m(x_k)}(x_k)\big)$. Considering the known precision vector $\boldsymbol{v} = (v(x_1), v(x_2), \ldots, v(x_k))$, the likelihood for this sample is:

$$\begin{aligned} L &= L(\mu;\, \boldsymbol{y}, \boldsymbol{v}) \\ &\propto \prod_{x \in X} \prod_{j=1}^{m(x)} exp\,\{[\Phi(\mu(x)) + [y_j(x) - \mu(x)]\varphi(\mu(x))]v(x)\} \\ &\propto \exp\left(\sum_{x \in X} \{[\Phi(\mu(x)) + [\bar{y}(x) - \mu(x)]\varphi(\mu(x))]v(x)m(x)\}\right). \end{aligned}$$

By Bayes' Theorem, the posterior distribution for the vector $\boldsymbol{\mu}$ is:

$$\begin{aligned} &\pi[\mu \mid \boldsymbol{y}] \propto \exp\left\{\sum_{x \in X}\left\{\left[\Phi\big(\mu(x)\big) + \big(\tilde{\mu}(x) - \mu(x)\big)\varphi\big(\mu(x)\big)\right]w(x)\right\}\right\}, \\ &\qquad \text{if } \mu(.) \text{ is isotonic on } X \\ &\pi[\mu \mid \boldsymbol{y}] = 0 \text{ otherwise,} \end{aligned} \tag{8}$$

where

$$\tilde{\mu}(x) = \frac{\bar{y}(x)v(x)m(x) + \mu_0(x)h_0(x)}{v(x)m(x) + h_0(x)},$$

and

$$w(x) = v(x)m(x) + h_0(x).$$

The mode of the posteriori (8) maximizes the sum in the exponent, in the class of the isotonic functions $\mu(x)$. Barlow et al. (1972) showed that one solution to this problem is the isotonic regression of $\tilde{\mu}(x)$, denoted by $\tilde{\mu}^*(x)$ with weights $w(x) = v(x)m(x) + h_0(x)$. The Theorem 4.2.2 summarizes the results of this section and will be fundamental to the design proposed in the current paper.

Theorem 4.2.2 Let $X = \{x_1, x_2, \ldots, x_k\}$ be a finite set with a partial order ($\preccurlyeq$). With the priori for the vector $\boldsymbol{\mu} = (\mu(x_1), \mu(x_2), \ldots, \mu(x_k))$ given by density (7), the posteriori distribution has density in same family. The mode of the posteriori is given by

$$\tilde{\boldsymbol{\mu}}^* = (\tilde{\mu}^*(x_1), \tilde{\mu}^*(x_2), \ldots, \tilde{\mu}^*(x_k)),$$

where, each $\tilde{\mu}^*(x)$ is the isotonic regression with weights $w(x) = v(x)m(x) + h_0(x)$, of the weighted mean of $\bar{y}(x)$ and $\mu_0(x)$:

$$\tilde{\mu}(x) = \frac{\bar{y}(x)v(x)m(x) + \mu_0(x)h_0(x)}{v(x)m(x) + h_0(x)}, \qquad x \in X.$$

In particular, if $\boldsymbol{h_0} = \mathbf{0}$ (flat prior) the mode of the posteriori coincides with the maximum likelihood estimates.

5 Methodology

To estimate quantiles of a cumulative distribution F, which represents the dose-response curve in several experimental situations, is a fact that usually occurs in clinical trials. In this work, an experimental design that is efficient in such situations is proposed.

5.1 Proposed design

Consider a set of doses $D = \{d_1, d_2, \ldots, d_k\}$, with $d_1 < d_2 < \ldots < d_k$, and $p_i = p(d_i)$ the probability of toxicity of dose d_i. In the Bayesian context the probabilities $p(d_i)$ are random variables having some priori distribution.

To each dose d_i associate a conditional distribution, with density in the exponential family as in (4), with unknown $\theta = p(d_i)$ and known $h = v(d_i)$, that is,

$$f(y \mid p(d_i), v(d_i)) \propto exp\{[\Phi(p(d_i)) + (y - p(d_i))\varphi(p(d_i))]\, v(d_i)\} \qquad (9)$$

Note that, for each dose d_i, it may be associated a Bernoulli random variable Y_i, with $E[Y_i \mid p(d_i), d_i] = p(d_i)$. The precision parameter $h = v(d_i)$ must be known. This parameter determines the variability of data around $p(d_i)$. For the Bernoulli distribution $h = 1$.

For each dose d_i, consider independent samples with size $N(d_i)$. The researcher selects a previous probabilities vector $\boldsymbol{p_0} = (p_0(d_1), p_0(d_2), \ldots, p_0(d_k))$, where each $p_0(d_i)$ is the value considered plausible to represent the mode of distribution of the random variable $p(d_i)$. The natural assumption here is, if $d_1 < d_2 < \ldots < d_k$ then $p_0(d_1) \leq p_0(d_2) \leq \ldots \leq p_0(d_k)$. The vector of random components $\boldsymbol{p} = (p(d_1), p(d_2), \ldots, p(d_k))$ will have a priori distribution as in (7):

$$\pi[\boldsymbol{p}] \propto \prod_{i=1}^{k} exp\{[\Phi(p(d_i)) + (y - p_0(di))\varphi(p(d_i))]h_0(d_i)\},$$
if $p(d_i)$ is isotonic on D,

$$\pi[\boldsymbol{p}] = 0, \text{ otherwise.}$$

The Theorem 4.2.2 states that the posterior distribution of $\boldsymbol{p}$ has density in the same family, with mode $\tilde{\boldsymbol{p}}^* = (\tilde{p}^*(d_1),\ \tilde{p}^*(d_2),\ \ldots,\ \tilde{p}^*(d_k))$, where each $\tilde{p}^*(d_i)$ is the isotonic regression with weight $w(d_i) = v(d_i)\,N(d_i) + h_0(d_i)$, of the weighted mean of $\bar{y}(d_i)$ and $p_0(d_i)$,

$$\tilde{p}(d_i) = \frac{\bar{y}(d_i)\,v(d_i)N(d_i) + p_0(d_i)\,h_0(d_i)}{v(d_i)N(d_i) + h_0(d_i)} .$$

In a dose-response trial, it is natural to consider the random variable $Y_i \mid (p(d_i), n) \sim Ber(p(d_i))$. If $\bar{y}(d_i)$ is the proportion of positive responses for the dose d_i, that is,

$$\hat{p}(d_i) = \bar{y}(d_i) = \frac{R(d_i)}{N(d_i)} ,$$

where $R(d_i)$ is the number of positive responses to the dose d_i, as $v(d_i) = h = 1$, then

$$\tilde{p}(d_i) = \frac{\hat{p}(d_i)N(d_i) + p_0(d_i)h_0(d_i)}{N(d_i) + h_0(d_i)} .$$

The proposed design consists of the following four steps:

1. Specify in advance the vector of modes $\boldsymbol{p_0} = (p_0(d_1),\ p_0(d_2),\ \ldots,\ p_0(d_k))$ and the vector of precisions $\boldsymbol{h_0} = (h_0(d_1),\ h_0(d_2),\ \ldots,\ h_0(d_k))$, with $p_0(d_1) \leq p_0(d_2) \leq \ldots \leq p_0(d_k)$.

2. Starts the trial by testing a group of *M* patients in a dose level d_i, and observes the proportion of successes $\hat{p}(d_i)$. Considering $\hat{p}(d_l) = p_0(d_l)$ and $N(d_l) = 0$, for doses d_l not yet tested, weights $w(d_i) = N(d_i) + h_0(d_i)$ and

$$\tilde{p}(d_i) = \frac{\hat{p}(d_i)N(d_i) + p_0(d_i)h_0(d_i)}{N(d_i) + h_0(d_i)},$$

the isotonic regression $\tilde{p}^*(d_i)$ is calculated.

3. Next patient will receive the dose level:

   (a) d_{i-1}, if $\tilde{p}^*(d_i) > \Gamma$ and if, at least one among the s most recent patients on the same dose d_i, showed toxicity;
   (b) d_{i+1}, if $\tilde{p}^*(d_i) < \Gamma$ and if none among the s most recent patients on the same dose d_i showed toxicity;
   (c) d_i, otherwise.

4. Repeat steps (2) and (3) until a pre-defined stopping rule.

If, eventually, the random walk reaches doses d_1 or d_k, next step must be conveniently adjusted. This design allows the use of previously known information. Note that, actually, the only necessary specifications are the modes and the precision parameter.

Therefore, the calculations are fairly simple. This method will be called *Sequential Isotonic Bayesian Design - SIBD*.

5.2 Simulation study

The Sequential Isotonic Bayesian Design proposed was simulated and a nonparametric estimator was used. Several scenarios were analyzed and these are detailed below.

5.3 Evaluation of SIBD

To evaluate the Sequential Isotonic Bayesian Design we used the estimator $\hat{\theta}_I$ (Equation 2). The performance of this estimator on SIBD was compared with its performance on Modified Narayana Design - MND.

Two sets of doses were used: $D_1 = \{0.5; 1.0; 1.5; 2.0; 2.5; 3.0; 3.5; 4.0\}$ and $D_2 = \{0.5; 0.75; 1.0; 1.25; 1.5; 1.75; 2.0; 2.25; 2.5; 2.75; 3.0; 3.25; 3.5; 3.75; 4.0\}$. We analyzed 21 scenarios with set D_1 and 18 scenarios with D_2. For D_1, two priori information scenarios were used, with different population parameters for different scenarios (Table 1). For D_2 only a single priori mode vector with variation in parameters was considered, according to Table 2.

Table 1 - Priori vectors of modes ($\boldsymbol{p_{01}}$ and $\boldsymbol{p_{02}}$) and precision ($\boldsymbol{h_{01}}$ and $\boldsymbol{h_{02}}$) used with D_1. The vectors $\boldsymbol{p_1}$, $\boldsymbol{p_2}$ and $\boldsymbol{p_3}$ represents population probabilities

Vectors
$p_{01} = (0.02;\ 0.07;\ 0.15;\ 0.25;\ 0.40;\ 0.60;\ 0.70;\ 0.80)$
$p_{02} = (0.10;\ 0.15;\ 0.30;\ 0.60;\ 0.80;\ 0.90;\ 0.95;\ 0.99)$
$p_1 = (0.04;\ 0.11;\ 0.23;\ 0.40;\ 0.60;\ 0.77;\ 0.89;\ 0.96)$
$p_2 = (0.14;\ 0.26;\ 0.36;\ 0.45;\ 0.53;\ 0.59;\ 0.65;\ 0.70)$
$p_3 = (0.08;\ 0.32;\ 0.58;\ 0.76;\ 0.88;\ 0.94;\ 0.97;\ 0.99)$
$h_{01} = (2.00;\ 2.00;\ 2.00;\ 2.00;\ 2.00;\ 2.00;\ 2.00;\ 2.00)$
$h_{02} = (3.00;\ 3.00;\ 3.00;\ 3.00;\ 3.00;\ 3.00;\ 3.00;\ 3.00)$

Table 2 - Priori vector of modes ($\boldsymbol{p_{03}}$) and precision ($\boldsymbol{h_{03}}$ and $\boldsymbol{h_{04}}$) used with D_2. The vectors $\boldsymbol{p_4}$, $\boldsymbol{p_5}$ and $\boldsymbol{p_6}$ represent population probabilities

Vectors
$p_{03} = (0.02; 0.04; 0.11; 0.18; 0.25; 0.30; 0.35; 0.42; 0.49; 0.57; 0.70; 0.84; 0.87; 0.92; 0.94)$
$p_4 = (0.04; 0.07; 0.11; 0.16; 0.23; 0.31; 0.40; 0.50; 0.60; 0.69; 0.77; 0.84; 0.89; 0.93; 0.96)$
$p_5 = (0.14; 0.20; 0.26; 0.31; 0.36; 0.41; 0.45; 0.49; 0.53; 0.56; 0.59; 0.62; 0.65; 0.68; 0.70)$
$p_6 = (0.08; 0.19; 0.32; 0.46; 0.58; 0.68; 0.76; 0.83; 0.88; 0.91; 0.94; 0.96; 0.97; 0.98; 0.99)$
$h_{03} = (2.00; 2.00; 2.00; 2.00; 2.00; 2.00; 2.00; 2.00; 2.00; 2.00; 2.00; 2.00; 2.00; 2.00; 2.00)$
$h_{04} = (20.0; 20.0; 20.0; 20.0; 20.0; 20.0; 20.0; 20.0; 20.0; 20.0; 20.0; 20.0; 20.0; 20.0; 20.0)$

5.4 Software

Routines were developed in R language, using the *isotone* package and the *gpava* function.

6 Results and discussion

The design SIBD proposed was evaluated for two sets of doses, with equal amplitudes and two different levels of spacing. All scenarios were evaluated considering the same estimator (Equation (2)), calculating their Mean Square Error (MSE) and bias. The results for the various scenarios were similar and, for simplicity, only a few specific cases, where the Tolerance is modeled by a normal distribution, will be presented and discussed.

6.1 First set of doses

The evaluation of SIBD for the doses $D_1 = \{0.5; 1.0; 1.5; 2.0; 2.5; 3.0; 3.5; 4.0\}$ and the various scenarios (Table 3) presented similar results, and only one case will be displayed and discussed: Tolerance $normal(2.25,1)$.

Table 3 - Scenarios used in the simulation study to evaluate the SIBD under D_1 (**PP** = Population Parameters)

Scenario	Design	Distribution	Γ	p_0	h_0	j	K	PP
(a)	SIBD	$N(2.25; 1)$	0.10	$\boldsymbol{p_{01}}$	$\boldsymbol{h_{01}}$	2	7	$\boldsymbol{p_1}$
(b)	SIBD	$N(2.25; 1)$	0.10	$\boldsymbol{p_{02}}$	$\boldsymbol{h_{02}}$	1	7	$\boldsymbol{p_1}$
(c)	MND	$N(2.25; 1)$	0.10	-	-	1	7	$\boldsymbol{p_1}$
(d)	SIBD	$N(2.25; 1)$	0.33	$\boldsymbol{p_{01}}$	$\boldsymbol{h_{01}}$	4	2	$\boldsymbol{p_1}$
(e)	SIBD	$N(2.25; 1)$	0.33	$\boldsymbol{p_{02}}$	$\boldsymbol{h_{02}}$	3	2	$\boldsymbol{p_1}$
(f)	MND	$N(2.25; 1)$	0.33	-	-	1	2	$\boldsymbol{p_1}$
(g)	SIBD	$N(2.25; 1)$	0.50	$\boldsymbol{p_{01}}$	$\boldsymbol{h_{01}}$	5	1	$\boldsymbol{p_1}$
(h)	SIBD	$N(2.25; 1)$	0.50	$\boldsymbol{p_{02}}$	$\boldsymbol{h_{02}}$	4	1	$\boldsymbol{p_1}$
(i)	MND	$N(2.25; 1)$	0.50	-	-	1	1	$\boldsymbol{p_1}$
(j)	SIBD	$Exp(0.3)$	0.10	$\boldsymbol{p_{01}}$	$\boldsymbol{h_{01}}$	2	7	$\boldsymbol{p_2}$
(l)	MND	$Exp(0.3)$	0.10	-	-	1	7	$\boldsymbol{p_2}$
(m)	SIBD	$Exp(0.3)$	0.33	$\boldsymbol{p_{01}}$	$\boldsymbol{h_{01}}$	4	2	$\boldsymbol{p_2}$
(n)	MND	$Exp(0.3)$	0.33	-	-	1	2	$\boldsymbol{p_2}$
(o)	SIBD	$Exp(0.3)$	0.50	$\boldsymbol{p_{01}}$	$\boldsymbol{h_{01}}$	5	1	$\boldsymbol{p_2}$
(p)	MND	$Exp(0.3)$	0.50	-	-	1	1	$\boldsymbol{p_2}$
(q)	SIBD	$Gamma(3; 0.5)$	0.10	$\boldsymbol{p_{01}}$	$\boldsymbol{h_{01}}$	2	7	$\boldsymbol{p_3}$
(r)	MND	$Gamma(3; 0.5)$	0.10	-	-	1	7	$\boldsymbol{p_3}$
(s)	SIBD	$Gamma(3; 0.5)$	0.33	$\boldsymbol{p_{01}}$	$\boldsymbol{h_{01}}$	4	2	$\boldsymbol{p_3}$
(t)	MND	$Gamma(3; 0.5)$	0.33	-	-	1	2	$\boldsymbol{p_3}$
(u)	SIBD	$Gamma(3; 0.5)$	0.50	$\boldsymbol{p_{01}}$	$\boldsymbol{h_{01}}$	5	1	$\boldsymbol{p_3}$
(v)	MND	$Gamma(3; 0.5)$	0.50	-	-	1	1	$\boldsymbol{p_3}$

For estimation of the quantiles $\theta_{0.10}$, $\theta_{0.33}$ and $\theta_{0.50}$ both the MSE and the bias for SIBD tend to approach zero faster than for MND (Figures 2, 3 and 4).

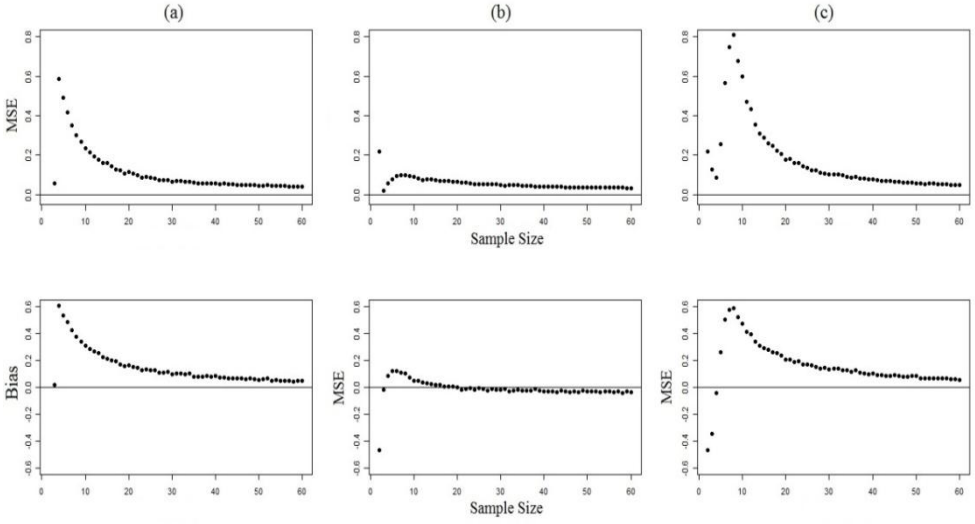

Figure 2 - MSE and bias for estimator $\hat{\theta}_{\Gamma}$, generated assuming a $normal(2.25,1)$ and $\Gamma = 0.10$, for each sample size. Scenarios (a), (b) and (c) are shown in Table 3: the scenarios (a) and (b) are related to SIBD, with different prior, and scenario (c) is related to MND.

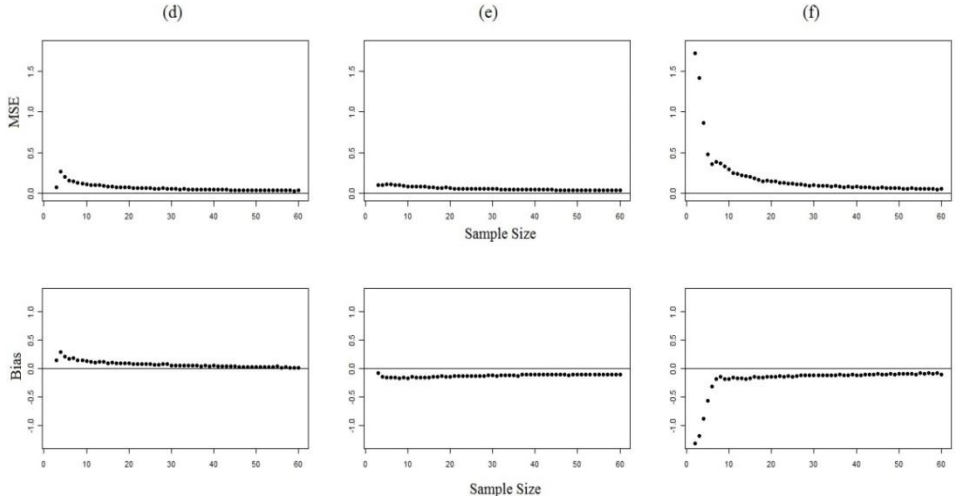

Figure 3 - MSE and bias for estimator $\hat{\theta}_{\Gamma}$, generated assuming a $normal(2.25,1)$ and $\Gamma = 0.33$, for each sample size. Scenarios (d), (e) and (f) are shown in Table 3: the scenarios (d) and (e) are related to SIBD, with different prior, and scenario (f) related to MND.

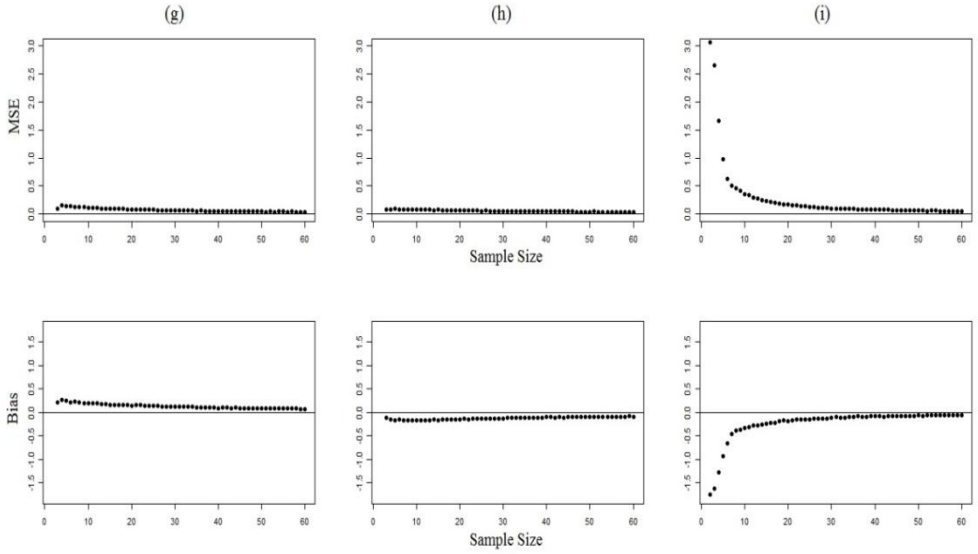

Figure 4 - MSE and bias for estimator $\hat{\theta}_{\Gamma}$, generated assuming a $normal(2.25,1)$ and $\Gamma = 0.50$, for each sample size. Scenarios (g), (h) and (i) are shown in Table 3: the scenarios (g) and (h) are related to SIBD, with different prior, and scenario (i) related to MND.

For quantile $\theta_{0.10}$ the MSE and the bias are stabilized near zero for sample sizes near 20 for SIBD and for MND for sample size near 30 (Figure 2). In the case of the other two quantiles ($\theta_{0.33}$ and $\theta_{0.50}$), stabilization occurred faster for both designs. However, SIBD shows better results (Figure 3 and 4).

When the transition probabilities were obtained from an exponential distribution, the MND and SIBD showed to be similar. A slight advantage had been presented by MND in contrast with SIBD in the Gama case.

Starting with sample size more than 20, Stylianou, Proschan e Flournoy (2003) found MSE near zero, using others estimators and the biased coin design. However, their method did not use priori information. In that sense, besides the fact that SIBD had shown a good performance related to bias and MSE, it uses prior information as usual with Bayesian approach.

Another detail to be considered, when comparing SIBD with other methods, is its easy computational implementation, with no Monte Carlo, no Gibbs sampler, no Metropolis-Hasting and no complex closed posteriori. Researcher only needs the priori information about probability for toxicity to perform the four steps.

6.2 Second doses set

The evaluation of SIBD for the doses on D_2 = {0.5; 0.75; 1.0; 1.25; 1.5; 1.75; 2.0; 2.25; 2.5; 2.75; 3.0; 3.25; 3.5; 3.75; 4.0} and the various scenarios presented in Table 4 showed similar results, and a single case will be presented: the Normal Distribution with mean $\mu = 2.25$ and variance $\sigma^2 = 1.0$.

Table 4 - Scenarios used in the study simulation to evaluate the SIBD under D_2 (**PP** = Population Parameters).

Scenario	Design	Distribution	Γ	p_0	h_0	j	K	PP
(A)	SIBD	$N(2.25;1)$	0.10	$\boldsymbol{p_{03}}$	$\boldsymbol{h_{03}}$	2	7	$\boldsymbol{p_4}$
(B)	MND	$N(2.25;1)$	0.10	-	-	1	7	$\boldsymbol{p_4}$
(C)	SIBD	$N(2.25;1)$	0.33	$\boldsymbol{p_{03}}$	$\boldsymbol{h_{03}}$	6	2	$\boldsymbol{p_4}$
(D)	MND	$N(2.25;1)$	0.33	-	-	1	2	$\boldsymbol{p_4}$
(E)	SIBD	$N(2.25;1)$	0.50	$\boldsymbol{p_{03}}$	$\boldsymbol{h_{03}}$	9	1	$\boldsymbol{p_4}$
(F)	MND	$N(2.25;1)$	0.50	-	-	1	1	$\boldsymbol{p_4}$
(G)	SIBD	$Exp(0.3)$	0.10	$\boldsymbol{p_{03}}$	$\boldsymbol{h_{03}}$	2	7	$\boldsymbol{p_5}$
(H)	MND	$Exp(0.3)$	0.10	-	-	1	7	$\boldsymbol{p_5}$
(I)	SIBD	$Exp(0.3)$	0.33	$\boldsymbol{p_{03}}$	$\boldsymbol{h_{03}}$	6	2	$\boldsymbol{p_5}$
(J)	MND	$Exp(0.3)$	0.33	-	-	1	2	$\boldsymbol{p_5}$
(L)	SIBD	$Exp(0.3)$	0.50	$\boldsymbol{p_{03}}$	$\boldsymbol{h_{03}}$	9	1	$\boldsymbol{p_5}$
(M)	MND	$Exp(0.3)$	0.50	-	-	1	1	$\boldsymbol{p_5}$
(N)	SIBD	$Gamma(3;0.5)$	0.10	$\boldsymbol{p_{03}}$	$\boldsymbol{h_{03}}$	2	7	$\boldsymbol{p_6}$
(O)	MND	$Gamma(3;0.5)$	0.10	-	-	1	7	$\boldsymbol{p_6}$
(P)	SIBD	$Gamma(3;0.5)$	0.33	$\boldsymbol{p_{03}}$	$\boldsymbol{h_{03}}$	6	2	$\boldsymbol{p_6}$
(Q)	MND	$Gamma(3;0.5)$	0.33	-	-	1	2	$\boldsymbol{p_6}$
(R)	SIBD	$Gamma(3;0.5)$	0.50	$\boldsymbol{p_{03}}$	$\boldsymbol{h_{03}}$	9	1	$\boldsymbol{p_6}$
(S)	MND	$Gamma(3;0.5)$	0.50	-	-	1	1	$\boldsymbol{p_6}$

The results show that, decreasing the spacing of the doses, the performance of SIBD improves in relation to MND. The performance of the MND was not better in any of the scenarios for this set of doses (Figures 5, 6 e 7). In this sense, SIBD shown interesting aspects: use of researcher priori information, maximum likelihood estimation with order restriction (isotonic regression) and easy computational implementation.

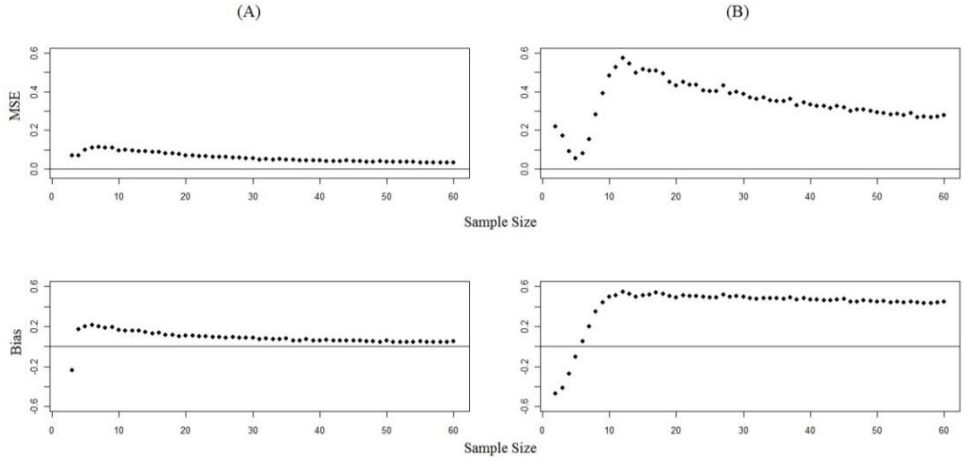

Figure 5 - MSE and bias for estimator $\hat{\theta}_{\Gamma}$, samples generated assuming a $normal(2.25,1)$ and Γ = 0.10. Scenarios (A) and (B), shown in Table 4, refers to SIBD and MND, respectively.

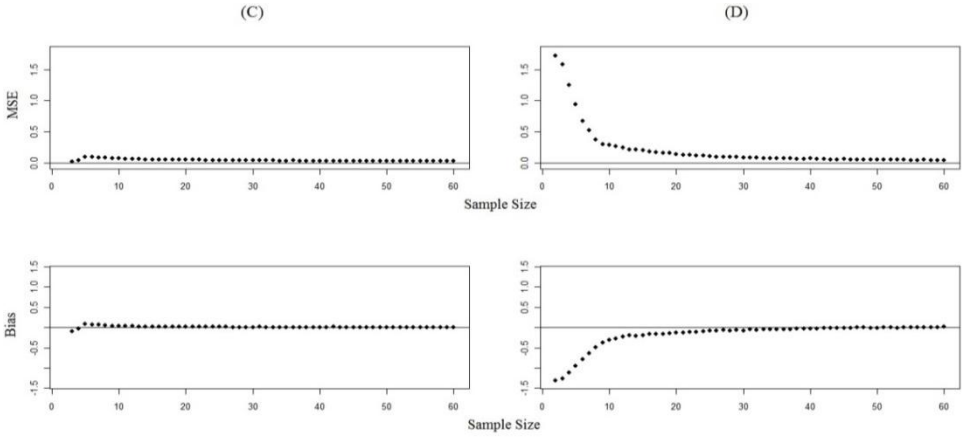

Figure 6 - MSE and bias for estimator $\widehat{\theta}_{\Gamma}$, samples generated assuming a $normal(2.25,1)$ and Γ = 0.33. Scenarios (C) and (D), shown in Table 4, refers to SIBD and MND, respectively.

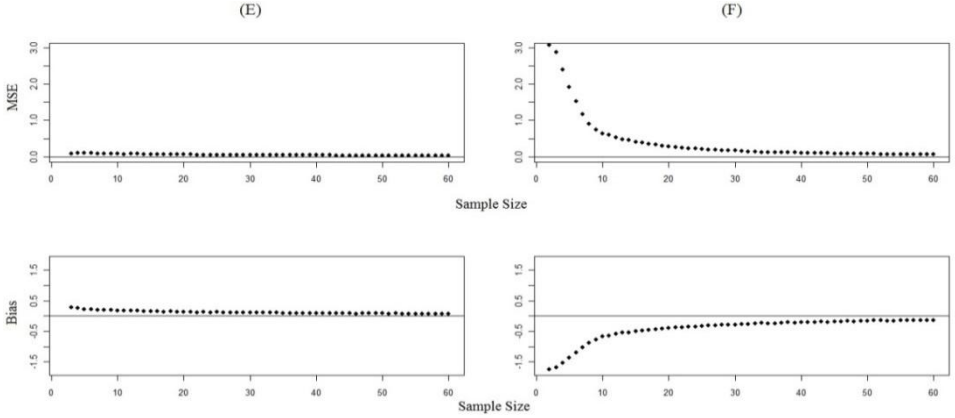

Figure 7 - MSE and bias estimator $\hat{\theta}_{\Gamma}$, samples generated assuming a $normal(2.25,1)$ for Γ = 0.50. Scenarios (E) and (F) shown in Table 4, refers to SIBD and MND, respectively.

The reason for the use of lower time intervals relies in the fact that the presupposition that the proportions increases with the increasing of doses can be more easily violated with small time gaps between doses, in reason of the randomness of the process. Things like that justify furthermore the use of isotonic regression, as in SIBD, to correct eventual distortions.

The SIBD is quite different compared to those Bayesian methods presented by O'Quigley, Pepe and Fisher (1990) and Ji, Li and Yin (2007). It is interesting to note that prior for the transition probability may be any member of the exponential family (6). This is important because the theory can be used with a wide family of probability distributions, as Normal, Exponencial, t-Student, Chi-Square. Researcher only needs no

suppose a unimodal distribution in the exponential family and use its previous knowledge to start the estimation process.

Conclusion

The Sequential Isotonic Bayesian Design - SIBD showed better results than MND for the scenarios studied, with reduction of MSE and bias when using $\hat{\theta}_{\Gamma}$, Weighted Average estimator. It is easy to manipulate and incorporates previous information about the transition probabilities from one dose to another.

Acknowledgements

To CAPES for the PhD scholarship granted to the first author and to CNPq for the Research Productivity scholarship granted to the third author.

SOUZA, E. M., CHAVES, L. M., MUNIZ, J. A., SOUZA, D. J. Um delineamento Bayesino para testar ensaios sequenciais do tipo dose-resposta utilizando regressão isotônica. *Rev. Bras. Biom*. Lavras, v.34, n.2, p.347-364, 2016.

- *ABSTRACT: O presente trabalho é relacionado à obtenção de uma dose com uma taxa de toxicidade pré-especificada, em uma população alvo. Propõe, utilizando métodos Bayesianos e regressão isotônica, um delineamento sequencial que permita aos pesquisadores atualizar a informação a priori. A teoria por detrás deste delineamento não é de fácil assimilação para não estatísticos. Entretanto, o delineamento proposto reduz-se a quatro passos de fácil manipulação, sem a necessidade de esforço computacional. Procedimentos de simulação confirmaram a eficácia da metodologia proposta.*

- *KEYWORDS: Delineamento sequencial; método up-and-down; delineamento da moeda viciada; delineamento de Narayana.*

References

BARLOW, R. E.; BARTHOLOMEW, D. J.; BREMNER, J. M.; BRUNK, H. D. *Statistical Inference under Order Restrictions: The Theory and Application of Isotonic Regression*. 1.ed. New York: John Wiley & Sons, 1972. 388p.

BERKSON, J. Application to the logistic function to bio-trial. *Journal of the American Statistical Association*, Washington, v.39, n. 227, p.357-365, 1944.

BLISS, C. I. The calculator of the dosage-mortality curve. *Annals of Applied Biology*, Warwick, v.22, n. 1, p.134-167, 1935.

DERMAN, C. Nonparametric up and down experimentation. *The Annals of Mathematical Statistics*, Beachwood, v.28, n. 3, p.795-798, 1957.

DIXON, W. J.; MOOD, A. M. A method for obtaining and analyzing sensitivity data. *Journal of the American Statistical Association*, Washington, v.43, n. 241, p.109-126, 1948.

DURHAM, S. D.; FLOURNOY, N.; ROSENBERGER, W. F. A random walk rule for phase I clinical trials. *Biometrics***,** Arlington, v.53, n. 2, p.745-760, 1997.

GERKE, O.; SIEDENTOP, H. Optimal phase I dose-escalation trial designs in oncology: a simulation study. *Statistics in Medicine*, Malden, v.27, n. 28, p.5329-5344, 2007.

GIOVAGNOLI, A.; PINTACUDA, N. Properties of frequency distribution induced by general up-and-dow methods for estimating quantiles. *Journal of Statistical Planning and Inference*, Cupertino, v.74, n. 1, p.51-63, 1998.

IVANOVA, A.; HAGHIGHI, A. M.; MUHANTY, S. G.; DURHAM, S. D. Improved up-and-down designs for phase I trials. *Statistics in Medicine*, Malden, v.22, n. 1, p.69-82, 2003.

LEEUW, J.; HORNIK, K.; MAIR, P. Isotone optimization in R: Pool - adjacent - violators algorithm (pava) and active set methods. *Journal of Statistical Software*, Los Angeles, v.32, n. 5, p.1-24, 2009.

JI, Y.; LI, Y.; YIN, G. Bayesian dose finding in phase I clinical trials based on a new statistical framework. *Statistica Sinica*, Taiwan, v.17, n. 2, p.531-547, 2007.

O'QUIGLEY, J.; PEPE, M.; FISHER, L. Continual reassessment method: a practical design for phase I clinical trials in cancer. *Biometrics*, Arlington, v.46, n. 1, p.33-48, 1990.

PACE, N. L.; STYLIANOU, M. P. Advances in and limitations of up-and-down methodology. *Anesthesiology*, Bethesda, v.107, n. 1, p.144-152, 2007.

PAUL, R. K.; ROSENBERGER, W. F.; FLOURNOY, N. Quantile estimation following non-parametric phase I clinical trials with ordinal response. *Statistics in Medicine*, Malden, v.23, n. 16, p.2483-2495, 2004.

STYLIANOU, M.; FLOURNOY, N. Dose finding using the biased coin up-and-down design and isotonic regression. *Biometrics*, Arlington, v.58, n. 1, p.171-177, 2002.

STYLIANOU, M.; PROSCHAN, M.; FLOURNOY, N. Estimating the probability of toxicity at the target dose following an up-and-down design. *Statistics in Medicine*, Malden, v.22, n. 4, p.535-543, 2003.

TOURNEAU, C. L.; LEE, J. J.; SIU, L. L. Dose escalation methods in phase I cancer clinical trials. *Journal of the National Cancer Institute*, Bethesda, v.101, n.10, p.708-720, 2009.

SOUZA, E. M. *Theory of the isotonic regression and sequential designs applied to dose-response essays*, 2011. 371p. Thesis - Federal University of Lavras, MG, Brazil, 2011.

SOUZA, E. M.; CHAVES, L. M.; MUNIZ, J. A. Isotonic regression geometrical aspects and an application to the estimation for maximum likelihood to the exponential family. *Biometric Brazilian Journal*, Sao Paulo, v.30, n. 1, p.76-105, 2012.

Received in 03.09.2015

Approved after revised in 24.02.2016