Proceedings of the 2017 Winter Simulation Conference
W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, eds.

HYBRID RESEARCH SIMULATION MODELING FOR MAKING DECISIONS ON SAMPLE SIZE AND POWER OF RANDOMIZED CLINICAL TRIALS CONSIDERING EXPECTED NET BENEFITS

Ismail Abbas

Polytechnic University of Catalonia
Department of Statistics and Operations Research
Campus Nord, Jordi Girona, 1-3, Barcelona 08034, CA, SPAIN

ABSTRACT

This paper presents a framework aimed at making decisions on sample size and power that optimize expected net benefits of clinical trials that incorporate health and cost variables. Two-stages standard modeling and simulation was developed. The hybrid framework was populated with prior information on benefits of a clinical trial that compare magnetic resonance imaging vs. computerized axial tomography in a trial of acute ischemic stroke diagnostic, combined with an assumption on willingness to pay per health gain, correlated distributions, exclusivity and per patient-time cost. As results, an estimated optimal expected net benefits of €3.8M at optimal power of 34% and sample size of 1800 were calculated. Emphasis on main and sensitivity analysis results are also presented.

1 INTRODUCTION

Simulation modeling has recently made significant contributions designed for the calculation of sample size and power of clinical trials and, evaluating health and economic outcomes. (Wu 2015) generated data to estimate the difference between blind and non-blind estimation of means and variance of continuous endpoints for clinical trials, in terms of bias and mean square errors. (Stamey et al. 2013) formulated sample size calculations that account for continuous efficacy variable and binary safety variable simultaneously assuming correlations between them. Then, they generate data for a range of correlation coefficients searching for those coefficients that reduce the sample size of this type of clinical trials. (Lui et al. 2008) evaluated a clinical trial test statistic that incorporates noncompliance and missing data that may occur during the experiments. For a given sample size, simulation was used to calculate type I error of the null hypothesis and the power of the alternative hypothesis considering missing patients due to withdrawal or getting lost from the trial, and noncompliance due to side effects of the assigned treatment. (Laouenan et al. 2013) designed and performed a viral kinetic study to compare the antiviral effectiveness of two hepatitis-c protease inhibitors using clinical trial simulation. The power, type I error and parameters of the comparison test were calculated conducting 500 replication of the designed trial. The authors found that the simulated trial provide benefits over the most of empirical approaches. (Ciolino et al. 2012) studied the power of ignoring treatment allocation imbalance when continuous baseline covariates are involved in clinical trials of acute ischemic stroke. (Yang and Beerahee 2013) calculated the power of optimal trial design under pharmacokinetic modeling finding that the suggested approach is efficient and cost-effective when investigating a new ant-inflammatory drug in patients with rheumatoid arthritis. (Kraiczi and Frisen 2005) estimated the effects of parameters uncertainty on power calculations. (Bonate 2000) compare the type I error and the power of different metrics in the analysis of qtc with two dosing regimens, recommending not using the mean of qtc as a measure of drug effects. (Tsodikov et al. 1998) compared the power resulting from simulations and from bootstrap approach, giving data of a small trial for a chronic myelogenous leukemia. They found that the sampling data using simulation are more

accurate than data provided by bootstrap technique. (Henderson et al. 1996) calculated conditional power of three test statistics, log-rank, wilcoxon and kolmogorov-smirov tests comparing two treatments, making decision on the benefits of extending or not the trial. Using data from a clinical trial conducted after heart valve replacement, they recommend using the model for such decision. (Chabaud et al. 2002) investigated the best combination between safety and efficacy of Ivabradine, optimizing sample size needed to confirm the stability of patients with angina pectoris. (Jin et al. 2014) developed a dynamic simulation model to predict the clinical practice impact of controlling hypertension and its associated risk factors such as cardiovascular risk increasing the control rate and allowing a reduction in cardiovascular events. (Gold et al. 2014) selected a biomarker for chronic obstructive pulmonary disease. (Nixon et al. 2009) evaluated the probability of drug registration for treating rheumatoid arthritis considering parameters uncertainty. (Schackman et al. 2007) identified the benefits and risks of simplifying HIV treatment. (Hmelo et al. 2001) taught medical students using simulation to improve their required skills, and (Elashoff et al. 1988) used simulation to design a clinical trial to prevent ulcer recurrence. On clinical trials registration and simulation-based studies, (Cheng et al. 2014) encouraged on the needs for the inclusion of simulation-based studies when registering and publishing the results of clinical trials. (Nakai et al. 2015) conducted a clinical trial to compare two treatments for preventing osteoporosis, according to an optimized design proposed by modeling and simulation. In relation to the objective of this paper, dealing with health and economic outcomes of clinical trials, (Sanghera et al. 2014) developed modeling method for calculating long term impact of economic outcomes of health interventions, such as cost-effectiveness analysis. (Patel and Ankolekar 2007; Willan and Pinto 2005; and Burman et al. 2005) developed Bayesian models to incorporate economic variables, aimed at calculating sample size for an optimized expected net benefits of clinical trials. (Abbas et al. 2006, 2016) developed hybrid research simulation modeling of clinical trials that incorporate health and economic variables. In this paper, the main purpose is to make decision on optimal sample size and power of clinical trials for the optimization of their expected net benefits under uncertainty and variability.

2 METHODS

2.1 Clinical trials

Clinical trials are long health research experiments aimed at showing efficacy and safety of a medicine or health technology. Usually they are done to make assessment on a variety of health programs such as preventing population from a possible disease, making a diagnostic on a suspected disease, treating patients or promoting health care programs. The general purpose of conducting and analyzing such trials is to provide information on the benefits of these health programs compared with standards already assumed by patients (ICH 1998).

Often before designing a confirmatory trial, a small research study is conducted to identify the state of the clinical situation and economical implications of the proposed trial; in which all information around the disease in question and the proposed health intervention are collected, analyzed and used to make decisions on the final trial's design. Since these decisions will have direct impacts on the measurements of efficacy and benefits, researchers and decision makers may need additional measures to define the study population, response variables, follow-up periods, hypothesis of the trial, power and sample size, statistical design requirements, and strategies to manage the budget impact and expenditures on health care cost and benefits (Seidl et al. 2012). Hence, the development of clinical trials is a dynamic process where many variables are evaluated simultaneously to provide decision makers with their results efficiently, at a reasonable health benefit and cost. The pharmaceuticals usually make decisions on whether conducting a trial looking at its benefits in comparison with its total cost.

Regulators make decisions on the benefits in terms of patient health and cost of treatment under the intervention, in comparison with the standard. Hence, both departments are concerns on assessing safety and efficacy of the health interventions so obtaining regulatory approval. Obviously, regulators and pharmaceuticals seek to reach such approval at maximum health and economic benefits.

2.2 Modeling

Clinical trials are characterized by being risky experiments, complex and costly systems, and the benefits of its health and cost effects are uncertain. Modeling approach can be applied for producing their anticipated results, under which estimation and optimization of expected net benefits can be assessed.

Hence, the value of conducting these clinical trials under uncertainty requires the specification of key variables such as trial expenses, benefits, risks and probability of approvals. The expenses obviously depend on the sample size, recruitment rate, and number of centers and time of follow-up. For the trial risk it is considered to be the variability that arises from the included parameters. The expected power of the trial primary variable for a given sample size is considered to be the approval probability, which is the probability of the benefits that can be derived from the trial.

2.2.1 Sample size and power

The sample size required for a trial is usually calculated at the design stage, thus suppose that a protocol for a double blind randomized clinical trial is to be designed in order to study the efficacy of two treatments in patients with a given disease. The sample size calculation can be intended using formula (1.a, 1.b).

For binary response: $$n = \left(z_{1-\alpha/2} + z_{1-\beta}\right)^2 \sigma^2/\delta^2 \quad (1.a)$$

For continuous response: $$n = 2\left(z_{1-\alpha/2} + z_{1-\beta}\right)^2 \sigma^2/\delta^2 \quad (1.b)$$

Where, n is the sample size at each treatment assuming equal allocation, δ is the difference between the two treatments, σ^2 is common variance of the primary response variable and the $z_{1-\alpha/2}$ and $z_{1-\beta}$ are the values of the standardized normal distribution at the probabilities 1-α/2 and 1-β (α and β are the Type I and Type II errors). Formula (1.a and 1.b) can be inverted to obtain the corresponding formula (2.a and 2.b) for power calculation, giving a sample size n.

For binary response: $$1 - \beta = p\left(Z \le \left(\delta\sqrt{n} - \sigma z\right)/\sigma\right) \quad (2.a)$$

For continuous response: $$1 - \beta = p\left(Z \le \left(\delta\sqrt{n} - \sqrt{2}\sigma z\right)/\sqrt{2}\sigma\right) \quad (2.b)$$

When sample size is already recruited from the population and randomly assigned either to the new or to the standard intervention, where the patients will have a primary outcomes that are designated as u_2 and u_1, respectively. As soon as the data are available from the trial for evaluation, the a statistic such as z assuming normal distributions of difference can be applied to test the null hypothesis. Thus, using formula (3.a and 3.b), the observed value of such statistic can be obtained.

$$z = \frac{u_2 - u_1}{SE(u_2 - u_1)} \quad (3.a), \qquad SE(u_2 - u_1) = \sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}} \quad \text{where } n_1 = n_2 = n \quad (3.b)$$

When the resulting value of z is lower than $+z_{1-\alpha/2}$ and higher than $-z_{1-\alpha/2}$, the null hypothesis is not rejected concluding that there is not enough evidence to reject the null difference. When the resulting value of z is higher than $+z_{1-\alpha/2}$ (for two sided α), the null hypothesis is rejected advising that the new interventions is superior. When z is lower than $-z_{1-\alpha/2}$, the null hypothesis is rejected advising that the standard intervention remain superior, providing better health improvements. β is the probability of rejecting the alternative hypothesis when it is true, means that the observed value of the statistic z is lower compared with the nominal one (4.a). The complementary of this probability is known as the power, P = 1-β, of the hypotheses, which is the probability of rejecting the null hypothesis when the alternative is true (4.b1), where Z is the CDF-cumulative distribution function of the standard normal distribution, thus:

$$\beta = p\left(z \le z_{1-\alpha/2}\right) = p\left(Z \le z_{1-\alpha/2} - z\right) \quad (4.a)$$

$$P = 1 - p\left(Z \le z_{1-\alpha/2} - z\right) \quad (4.b1)$$

$$P = p\left(z > z_{1-\alpha/2}\right) = 1 - p\left(Z \le z_{1-\alpha/2} - z\right) \quad (4.b2)$$

The power also can be calculated directly as the probability of that the observed statistic is higher than the value of the reference one when the alternative hypothesis is true (4.b2). The reference value often is defined as the critical region of rejecting the null hypothesis. These formulas for β and 1-β can be used for both binary and continuous responses assuming equal or unequal allocation of subjects.

2.2.2 Cost and benefits

Multicentre trials are methods for increasing the recruitment of patients within a reasonable time-frame.

Thus, giving the length of clinical trials depends on the number of patients to be recruited, the number of centers that are available to participate, recruitment rate at each center and the suggested follow-up periods; the cost of trial can be reduced and the expected net benefits is optimized if the assigned budget depends on sample size and time. Hence, assuming that the total time of the clinical trial is the sum of the total time of patients recruitment (r) and the time of the follow-up (f), yield $\mathrm{r} + \mathrm{f}$. Assuming r depends on the recruitment rate, sample size n^* where $n^* = n_1 + n_2$ and the number of centers (S), yielding $\mathrm{r} = n^*/\lambda \mathrm{S}$ and then $(n^*/\lambda\ \mathrm{S}) + \mathrm{f}$. This formula assumes the patients are equally distributed between the centers according to the rate of enrolment (λ). Minimizing the total cost of randomized clinical trials of medicine or health care can be assessed by minimizing the total time of the trial and sample size. Therefore, this can be formulated as the total cost of the trial equals the cost per patient, denoted by c, multiplied by the total duration of the trial, $cn^*(\mathrm{r} + \mathrm{f})$. According to statistics theory, $\mathrm{V}(\sum_1^{\mathrm{n}} \mathrm{bX}) = \mathrm{nb}^2\mathrm{V}(\mathrm{X})$ where b is a constant. Therefore, assuming that the cost per patient is a random variable following a normal distribution, the variance of the total cost will be $\mathrm{V}\big(cn^*(\mathrm{r} + \mathrm{f})\big) = n^*(\mathrm{r} + \mathrm{f})^2\ \mathrm{V}(\mathrm{c})$. The total cost and its variance can be placed within the expected net benefit model, developed later on. When conducting clinical trials, economic evaluation studies are used to estimate the cost of treatments that occur beside the efficacy or effectiveness. The main objective is to compare the cost-effectiveness of treatments based on their incremental health benefits, measured as effectiveness or effectiveness adjusted by quality of life, divided by the incremental cost $\Delta\mathrm{E}/\Delta\mathrm{C}$. When the cost of the experimental intervention is higher than the standard one, there is an incremental cost of treatment, otherwise the experimental is dominant intervention, according to health economic theory. Due to statistical problems associated with ratios and their statistics, Benefits analysis was used to compare the effects of treatment on cost and benefits.

$$\mathrm{B} = \mathrm{k}\big(\mu_{\mathrm{e}2} - \mu_{\mathrm{e}1}\big) - \big(\mu_{\mathrm{c}2} - \mu_{\mathrm{c}1}\big) \tag{5}$$

Thus, the cost-benefits of comparing two treatments expressed in formula (5) estimate the benefits of conducting a trial where μ_{e2} and μ_{e1} are, respectively, the mean health benefits for patients under the new and the standard treatment; μ_{C2} and μ_{C1} are the mean cost for patients under the new and the standard treatment. In this paper it is regarded as the benefits of clinical trials, comparing two health interventions. The benefits requires the specification of k, which is the aim of an advanced paper, regarded as the amount of willing to pay for increasing health benefits by one unit. As shown later on, the statistical problems associated with four planes of the incremental cost-effectiveness and benefits analysis can be recognized by the standardized expected net benefits measures and their probabilities.

2.2.3 Optimal expected net benefits with uncorrelated distributions

Expected Net Benefits is the key approach for assessing the global value of conducting clinical trials by converting available data into information and conclusions to assist making decisions on whether clinical trials provide net benefits for health systems. Maximizing the expected net benefits depends on the power calculated for an expected difference on primary health outcome, cost and benefits. In mathematical term, defining ENB as the expected net benefit from the trial; P denotes the probability of rejecting the null hypotheses of the trial; B as the total benefits (for example, the benefits derived from a trial comparing two interventions using formula (5)); D denotes the years of exclusivity; C the total cost of the trial that

$$\mathrm{ENB} = \mathrm{PBD} - \mathrm{C} \tag{6.a}$$

$$\mathrm{STD(ENB)} = \sqrt{\mathrm{V[PBD - C]}} \tag{6.b}$$

includes the design cost, managements, follow-up period, data analysis and other administrative costs; yield formula 6.a and 6.b. for its standard deviation. So assuming predefined parameters on sample size, power and benefits, follow-up and period of exclusivity, we rewrite formula 6.a as formula 7.a for the total ENB. Again, according to statistics theory, $E(\sum_1^n bX) = nbX, \mathrm{V}(\sum_1^{\mathrm{n}} \mathrm{bX}) = \mathrm{nb}^2\mathrm{V}(\mathrm{X})$ where b is a constant, we rewrite the variance of ENB in 6.b as formula 7.b.

$$\mathrm{ENB} = \left(1 - p\left(\mathrm{Z} \le \mathrm{z}_{1-\alpha/2} - \mathrm{z}\right)\right)\left(\mathrm{k}\left(\mathrm{n}_2\mu_{\mathrm{e}2} - \mathrm{n}_1\mu_{\mathrm{e}1}\right) - \left(\mathrm{n}_2\mu_{\mathrm{c}2} - \mathrm{n}_1\mu_{\mathrm{c}1}\right)\right)\mathrm{D/f} - cn^*\ (n^*/\lambda\ \ \mathrm{S} + \mathrm{f}) \quad (7.a)$$

$$\mathrm{V(ENB)} = \left[\left(1 - p\left(\mathrm{Z} \le \mathrm{z}_{1-\alpha/2} - \mathrm{z}\right)\right)^2\left(\left(\mathrm{k}^2\sigma_{\mathrm{e}1}^2 + \sigma_{\mathrm{c}1}^2\right)\mathrm{n}_1 + \left(\mathrm{k}^2\sigma_{\mathrm{e}2}^2 + \sigma_{\mathrm{c}2}^2\right)\mathrm{n}_2\right)\right]\mathrm{D}^2/\mathrm{f}^2 + (n^*/\lambda\ \ \mathrm{S} + \mathrm{f})^2\,\sigma_{\mathrm{C}}^2 n^* \quad (7.b)$$

Where, σ_{e2} the standard deviation of health benefits for patients under the new treatment, σ_{e1} the standard deviation of health benefits for patients under the standard treatment, σ_{C2} the standard deviation of health cost for patients under the new treatment, σ_{C1} the standard deviation of health cost for patients under the standard treatment, μ_C the mean cost of the trial under standard or new treatment, σ_C the standard deviation of the mean cost of the trial between patients, k willingness to pay per health gain, f the follow-up and D the exclusivity time, n_1 and n_2 are the number of subjects at the standard and the new intervention, $n^* = n_1 + n_2$. Notice that, the resulting benefits from the trial is extrapolated to real treatment practice by a number of periods D/f, including the benefits resulting from the follow-up period and the subsequent periods of exclusivity, each of which is equal to the follow-up period until the expiration of exclusivity. The resulting standardized normal statistical test of ENB (7.a), for known variance (7.b) is:

$$\mathrm{z} = \mathrm{ENB}/\sqrt{\mathrm{V(ENB)}} \quad (8.a)$$

Then, the power of net benefits can be assessed using formula (8.b). For known variance, the standardized normal statistical test may provide evidence if there is difference between the benefits and the cost of conducting the trial.

$$P = 1 - p\left(Z \le z_{1-\alpha/2} - z\right) \quad (8.b)$$

Where Z is the CDF-cumulative distribution function of the standard normal distribution. The model expressed in formula (8.a) or its probabilities in formula (8.b) can be used to calculate the optimal sample size and power of clinical trials that consider their expected net benefits.

2.2.4 Optimal expected net benefits with correlated distributions

Previously, the first response variable measures are assumed to be independent. This is typically true for the clinical outcomes. However, more complicated measures for cost-benefit analysis are typically correlated. Many times in cost-benefit analysis, a more expensive intervention will produce better results. These correlations will influence the cost-benefit analysis. Assuming normality, the difference of two correlated quantities has a variance of $\sigma^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$ where σ_1 is the variance of the first quantity and σ_2 is the variance of the second quantity, ρ is the correlation coefficient. Consider $\mathrm{B}_2 = \mathrm{k}\,\mathrm{U}_2 - \mathrm{C}_2$ and $\mathrm{B}_1 = \mathrm{k}\,\mathrm{U}_1 - \mathrm{C}_1$, where B_2 and B_1 are the total benefits per patient of the new and the standard intervention; k is the willingness to pay to a unit of health benefits per patient; U_2 and U_1 are the health utilities of the new and standard intervention; and C_2 and C_1 are the associated costs with the new and standard intervention. By rewriting (7.a) as (9.a), correlations can be explicitly considered:

$$\mathrm{ENB} = \left(1 - p\left(\mathrm{Z} \le \mathrm{z}_{1-\alpha/2} - \mathrm{z}\right)\right)\ (\ \mathrm{n}_2\mathrm{B}_2 - \ \mathrm{n}_1\mathrm{B}_1)\,\mathrm{D/f} - \mathrm{cn}^*\ (\mathrm{n}^*/\lambda\ \ \mathrm{S} + \mathrm{f}) \quad (9.a)$$

Notice that the data for U_2 and C_2 are paired data. Similarly, so are the data for U_1 and C_1. Again, according to statistics theory $E\left(\sum_1^n bX\right) = nbX, \mathrm{V}\left(\sum_1^{\mathrm{n}} \mathrm{bX}\right) = \mathrm{nb}^2\mathrm{V(X)}$ where b is a constant and assuming correlation between health benefits an d cost of treatment as in (9.a), we rewrite (7.b) as 9.b.

$$\mathrm{V(ENB)} = \left(1 - p\left(\mathrm{Z} \le \mathrm{z}_{1-\alpha/2} - \mathrm{z}\right)\right)^2\left(\sigma_{\mathrm{B}2}^2\mathrm{n}_2 + \sigma_{\mathrm{B}1}^2\mathrm{n}_1\right)\mathrm{D}^2/\mathrm{f}^2 + (\mathrm{n}^*/\lambda\ \ \mathrm{S} + \mathrm{f})^2\,\sigma_{\mathrm{C}}^2\mathrm{n}^* \quad (9.b)$$

Hence

$$\mathrm{z} = \mathrm{ENB}/\sqrt{\mathrm{V(ENB)}} \quad (9.c)$$

In (9.b), $\sigma_{\mathrm{B}2}^2 = \left(\mathrm{k}^2\sigma_{\mathrm{e}2}^2 + \sigma_{\mathrm{c}2}^2 - 2\mathrm{k}\rho\sigma_{\mathrm{e}2}\sigma_{\mathrm{c}2}\right)$ and $\sigma_{\mathrm{B}1}^2 = \left(\mathrm{k}^2\sigma_{\mathrm{e}1}^2 + \sigma_{\mathrm{c}1}^2 — 2\mathrm{k}\rho\sigma_{\mathrm{e}1}\sigma_{\mathrm{c}1}\right)$ are the variances for B_2 and B_1 respectively, ρ is the correlation coefficient between health benefits and cost of the treatment within each intervention. All subsequent analysis follows the same as before. Formula (8.b) should still be used, but for *z* coming from equation (9.c).

2.3 Simulation

In this paper, simulation incorporates mathematical equations and statistical distributions to draw random data using computing random numbers generators. The random numbers and data were generated using a common modeling and simulation software, allowing for creating data that are comparable to real data

generated in clinical trials. The simulation starts at recognizing the clinical trial parameters including sample size, number of centers, rate of recruitment, follow-up period, the number of evaluations, probabilities or the values of reaching the first and the second clinical response whether binary variable or continuous variable, respectively, cost of treating each patient following the administration of each compared health intervention, covariance, and the cost of enrolling each patient. The simulation is run in two stages, the first is to simulate the power and sample size considering the first response variable of the trial, the second simulates the power and sample size considering the expected net benefits. Both stages can be done assuming non-correlation and correlation between health benefits and cost of treatment. The two stages should be repeated for a range of sample size, thus to obtain optimal ENBs, n^* and power, as shown in the results of the applications.

2.3.1 Stage 1. Simulating sample size and power of the trial

This stage of simulation starts at enrolling the subjects of the trial as the following:

1. The first patient from the population is recruited to any centre, and then randomly will be chosen to be allocated either to standard or experimental intervention according to the probability of randomization. Assuming equal allocation if $p = 0.5 \leq \text{rand}(0,1)$ then this patient will be allocated to the standard health intervention, otherwise will be at the experimental intervention.
2. Subsequently, this patient randomly will remain in the actual health state or to have an updated health state according to the probabilities or the values of reaching the first clinical response in the case of binary variable or continuous variable.
3. The cost of trialing this patient will be calculated depending on the time of recruitment and follow-up of the trial. This is calculated according to formula $cn^*(n^*/\lambda\ S + f)$, where: c is the cost per patient sampled from normal distribution with mean u_c and standard deviation σc multiplied by the sum of actual time of recruitment $n^*/\lambda\ S$ and the trial follow-up time f.
4. The three steps will be repeated as many as the required number patients, n_1, n_2.
5. Then, calculations are done for the proportion of responses under the new intervention p_2 and proportion of responses under the standard intervention p_1, in the case of binary response; and values of responses under the new intervention u_2 and values of responses under the standard intervention u_1, in the case of continuous response.
6. The previous steps are repeated for 10000 trials on the same parameters and sample size.
7. The difference between the two proportions p_2-p_1 or values u_2-u_1 are calculated for each repetition, and the standard deviations of difference is calculated as the overall standard deviation between the simulated trials.
8. Each difference of each simulated trial is divided by the overall standard deviation, obtaining the statistic to test the hypothesis of no difference between the two interventions. If the resulting statistic is higher than the reference one, we conclude that there is an evidence of difference in favor the new intervention. If the resulting statistic is lower than the reference one, we conclude that there is evidence of difference in favor the standard intervention. If the resulting statistic is between the higher and lower value of the reference one, we conclude that there is no evidence of difference in favor the new intervention. The power is calculated as the proportion of the number of the trials that showed evidence of difference out of the total number of simulated trials. The mean of the total cost is calculated.

2.3.2 Stage 2. Simulating sample size and power of ENB with uncorrelated distributions

In this scenario the simulation will treat the random variables distributions as independent, so the covariance is not considered and consequently there are no variability reduction within or between subjects.

1. In simulating health benefits of each patient, suppose $z_1 \sim N(0,1)$ and $z_2 \sim N(0,1)$ are two standard normal distributions, the health benefits of the new (10) and the standard (11) interventions are the sum of the mean and two values sampled from these two distributions:

$$U_2 = u_{e2} + Z_1\sigma_{e21} + Z_2\sigma_{e22}, \text{ where } \sigma_{e2} = \sqrt{\sigma_{e21}^2 + \sigma_{e22}^2} \quad (10)$$

$$U_1 = u_{e1} + Z_1\sigma_{e11} + Z_2\sigma_{e12}, \text{ where } \sigma_{e1} = \sqrt{\sigma_{e11}^2 + \sigma_{e12}^2} \quad (11)$$

2. In simulating the cost of treatment, suppose $Z_1 \sim N(0,1)$ and $Z_2 \sim N(0,1)$ are the same two standard normal distributions, the cost of treating patients with the new (12) and the standard (13) interventions are the sum of their mean cost and two values generated from these two distributions:

$$C_2 = u_{c2} + Z_1\sigma_{c21} + Z_2\sigma_{c22}, \text{ where } \sigma_{c2} = \sqrt{\sigma_{c21}^2 + \sigma_{c22}^2} \quad (12)$$

$$C_1 = u_{c1} + Z_1\sigma_{c11} + Z_2\sigma_{c12}, \text{ where } \sigma_{c1} = \sqrt{\sigma_{c11}^2 + \sigma_{c12}^2} \quad (13)$$

3. Once the health benefits and cost of treatment are sampled using these statistical models whether the patient is allocated to the new or the standard intervention, the sampled value of health benefits will be multiplied by monetary health value (k) that assumed to be the monetary value per Qalys.
4. The cost of treating will be subtracted to obtain the benefits of the new intervention, $B_2 = k\,U_2 - C_2$ and the standard intervention, $B_1 = k\,U_1 - C_1$, thus to obtain the overall benefits of the trial $B = B_2 - B_1$.
5. The four steps will be repeated for all patients simulated at stage 1 (sample size of the trial), thus obtaining the total benefits.
6. Then the expected net benefits of this trial can be calculated according to formula (9.c) introduced in the modeling section, where the power and the cost of the trial are results of stage 1.
7. The six steps are repeated for the 10000 trials, thus to obtain 10000 ENBs. Their overall standard deviations is calculated, $STD(ENB) = \sqrt{\frac{1}{N}\sum_1^N (ENB - E(ENB))^2}$, as the deviation of the mean expected net benefits from all simulated ENBs, where N=10000 is the number of simulated trials under the same settings.
8. Dividing each trial's ENB by the overall standard deviation will result in obtaining the statistic used to test the null hypothesis, and comparing the resulting statistic with a reference one (z or t).
9. Then the results of the comparison is used to assess if the expected net benefits of the trial is statistically significant or not. The power is calculated as the proportion of the number of significant expected net benefits divided by the total number of repeated trials.

2.3.3 Stage 2. Simulating sample size and power of ENB with correlated distributions

In this scenario, the simulation considers variability reduction for within and between subjects by incorporating covariance between the health benefits and cost of treatment, within each health intervention. As commented in the modeling section, commonly health benefits are positively correlated with the cost of the treatment. Thus, modeling the correlation requests that the variance of health benefits and costs should be discomposed into two variances; one for the correlated part between the health benefits and costs and one for the uncorrelated variance. The procedure to discompose the variances and generate correlated values between benefits and costs can be as the following:

1. In simulating health benefits of each patient, suppose $z_1 \sim N(0,1)$ and $z_2 \sim N(0,1)$ are two standard normal distributions, the health benefits of the new (14) and the standard (15) interventions are the sum of the mean and two values sampled from these two distributions:

$$U_2 = u_{e2} + Z_1\sigma_{e21} + Z_2\sigma_{e22}, \text{ where } \sigma_{e2} = \sqrt{\sigma_{e21}^2 + \sigma_{e22}^2} \quad (14)$$

$$U_1 = u_{e1} + Z_1\sigma_{e11} + Z_2\sigma_{e12}, \text{ where } \sigma_{e1} = \sqrt{\sigma_{e11}^2 + \sigma_{e12}^2} \quad (15)$$

2. In simulating the cost of each patient, suppose $Z_1 \sim N(0,1)$ and $Z_2 \sim N(0,1)$ are two standard normal distributions, the cost of treating patients with the new (16) and the standard (17) interventions are the sum of their mean cost and two values generated from these two distributions. However, when assuming correlations, Cholesky decomposition of variability was used to generate the data, thus:

$$C_2 = u_{c2} + Z_1\rho\sigma_{c2} + Z_2\sigma_{c2}\sqrt{(1-\rho^2)} \quad (16)$$

$$C_1 = u_{c1} + Z_1\rho\sigma_{c1} + Z_2\sigma_{c1}\sqrt{(1-\rho^2)} \quad (17)$$

3. Then to complete the simulation, steps 3-9 in section 2.3.2 should be repeated in this section.

3 RESULTS

The modeling and simulation methods were applied to real data that come from a clinical trial. The parameters on primary response and benefits were obtained from their published results. An assumption was made on cost of the trial, recruitment rate, number of centers, the monetary contribution and the time of exclusivity. Since these parameters are rather health policy parameters, they can be adjusted to reflect values coming from real data. The clinical trial was conducted for comparing magnetic resonance imaging (MRI) and computerized axial tomography scanner (CT) in the diagnostic of patients with acute ischemic stroke. The information obtained from the trial, are the follow-up years where the evaluation of effects will be done, f = 1 year, the rate of primary event (stroke-independent) per year provided that patients are under MRI ($p_1 = 0.429$) and under CT ($p_2 = 0.506$), the cost of health care per patient under MRI (μ_{C2} = €6184, standard deviation σ_{C2} = €2413) and CT (μ_{C1} = €6129, standard deviation σ_{C1} = €2546), the effectiveness adjusted by quality of life for a patient under MRI ($\mu_{e2} = 0.1123$, standard deviation $\sigma_{e2} = 0.2815$) and CT ($\mu_{e1} = 0.1638$, standard deviation $\sigma_{e1} = 0.2286$). Suppose that the rate of recruitment per year, $\lambda = 150$ patients at each centre, the number of centers, S = 2, the cost per patient of the trial is assumed to be μ_C = €300, standard deviation σ_C = €20; the monetary value per Qalys, k = €8000, and the supposed total time of patent and exclusivity, D = 20 years.

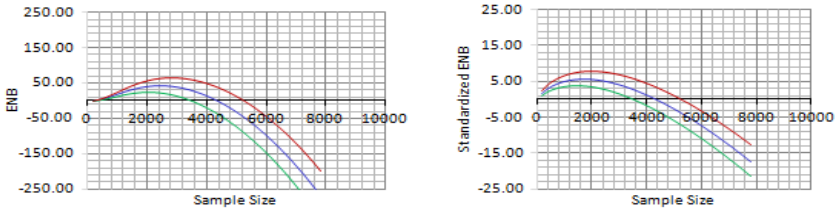

Figure 1: Absolute (left panel) and standardized (right panel) ENB as a function of n^* (ENB in €100.000).

Left Figure 1 shows three curves of expected net benefits of three parameterizations for a range of sample sizes. All results are calculated using formula 7.b and simulation without considering variability. The blue one is the results of the model with the real parameters of the trial that are described above. The expected net benefits increases with sample size reaching its maximum of €4.3M at optimal sample size of 2400 subjects, and declines smoothly. The power at these optimal results is 0.96. The sample size is 1300 for a power of 0.8, yielding to an expected net benefits of €2.8M, which is 45% of the optimal one at 45% reduction of sample size. The other two curves shows the sensitivity of expected net benefits to health benefits resulting from diagnosing patients either with MRI or with CT. At the green curve, the expected net benefits increases with sample size reaching its maximum of €2.5M at optimal sample size of 2000 subjects, and declines smoothly. The power at these optimal results is 0.93. The sample size is 1300 for a power of 0.8, yielding to an expected net benefits of €1.8M, which is 75% of the optimal one at 35% reduction of sample size. At the red curve, the expected net benefits increases with sample size reaching its maximum of €6.7M at optimal sample size of 2800 subjects, and declines smoothly. The power at these optimal results is around 98%. The sample size is 1300 for a power of 0.8, yielding to an expected net benefits of €3.7M, which is 50% of the optimal one at 55% reduction of sample size. Right Figure 1 shows the three curves for the standardized SENB giving sample sizes. All results are calculated considering risk, which is the variability that ENB can suffer from uncertain variables. For the blue curve (middle), the SENB increases with sample size reaching its maximum of 1.55 (€3.8M) at optimal sample size of 1800 subjects, and declines smoothly. The power at these optimal results is 34%. The other two curves shows the sensitivity of expected net benefits to change to the health benefits resulting from diagnosing patients either with MRI or with CT. At the green curve (bottom), the expected net benefits increases with sample size reaching its maximum of 1.024 (€2.0M) at optimal sample size of 1400 subjects, and declines subsequently. The power at these optimal results is 0.17. At the red curve, the expected net benefits increases with sample size reaching its maximum of 2.15 (€5.8M) at optimal sample size of 2000 subjects, and declines smoothly. The power at these optimal results is around 58%.

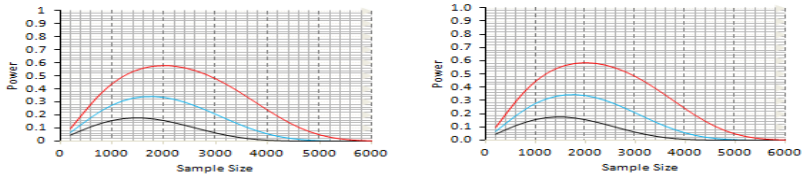

Figure 2: Power of ENB as a function of n^* for non-correlated distributions by modeling and simulation.

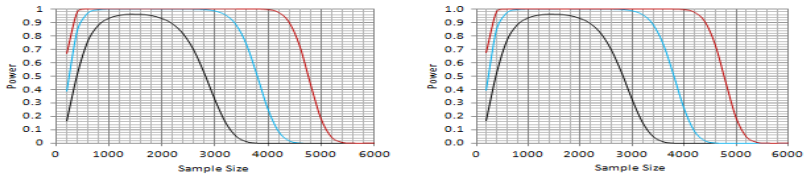

Figure 3: Power of ENB as a function of n^* for correlated distributions by modeling and simulation.

When looking at Figure 2 and Figure 3, we see that there are three uncorrelated and correlated curves, respectively, for the power of the expected net benefits of three parameterizations for a range of sample sizes. The calculation is done by modeling (left curves) and simulation (right curves), showing identical results. The blue curve (middle) of represents the results of the model with the real parameters of the trial that are described above. The power increases with sample size reaching its maximum of 34% (99% for ρ = 0.95 correlated distributions) at optimal sample size of 1800 subjects, and declines smoothly. The expected net benefits at these optimal results is €3.8M. The other two curves shows the sensitivity of expected net benefits to change to the health benefits resulting from diagnosing patients either with MRI or with CT. At the black curve (bottom), the power increases with sample size reaching its maximum of 18% (96% for ρ = 0.95 correlated distributions) at optimal sample size of 1400 subjects, and declines smoothly. The expected net benefits at these optimal results is €2.0M. In the case of the red curve (top), the power increases with sample size reaching its maximum of 58% (100% for ρ = 0.95 correlated distributions) at optimal sample size of 2000 subjects, and declines smoothly. The expected net benefits at these optimal results is €5.8M.

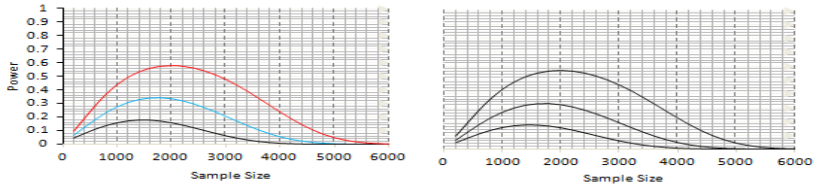

Figure 4: Power of ENB as a function of n^* for uncorrelated distributions by modeling and simulation.

Even considering the calculation for uncorrelated distributions, modeling still rely on strong assumption that the diagonal part of the non-self correlation matrix is 1. However, when facing situations in which all variables of the matrix are uncorrelated, the simulation can be run with independent random numbers generator for each variable of the matrix, leading to the same optimal results on sample size and expected benefits but at smaller optimal power as shown at Figure 4 (right curves, 17%, 32%, 0.56%), compared with the results assuming the previous uncorrelated matrix (left curves, 18%, 34%, 58%).

4 DISCUSION AND CONCLUSIONS

This paper describes a modeling and simulation framework for making decisions on optimal sample size and power considering expected net benefits of clinical trials. The main difference between the two applied methods, modeling and simulation, resides in the calculations and analyses of the results.
Analyses and decisions can be made on sample size and power based on empirical distributions of the expected net benefits means which is not treated in this paper. It can be done by, for example, the widely

used curve of acceptability for their empirical distributions. In this case, the described simulation will cover up the limitations of modeling. However, analyses should be made without considering variability in expected net benefits, concluding weaker decisions. The framework was kept simple so it can be applied, reproduced or updated by related researchers. It considers two methods of calculations assuming normality. In sample size section, the statistical models expressed in formula 4.b1 and 4.b2 provides a method of power calculation which was the base for formulating the power calculation of expected net benefits as expressed in formula 8.b. The results on sample size, power and expected net benefits were presented without considering variability for a range of sample sizes. However, including variability or correlations of some of parameters imply that the standardized expected net benefits might be used to calculate probabilities, power, sample sizes and optimal expected net benefits. The framework was populated with real data on benefits obtained from two clinical trials on stroke disease, combined with an assumption about health gain value, exclusivity and per patient-time trial cost. Using the developed modeling and simulation framework, an optimal expected net benefits can be calculated for an optimal sample sizes and powers. Modeling first uses the statistical-mathematical model (7.a) to calculate optimal sample size and expected net benefits without considering variability for a range of power that was obtained for the first response variable of the trial. Then modeling again uses the statistical models in formula 8.a and 8.b, by hand, calculates the optimal results considering variability. These hand calculations were repeated for the correlated distributions using formula 9.c, and then the optimal result were calculated using formula 8.b. Simulation relied on the method of repetitions to calculate the optimal results. It calculates, in random fashion for each patient, the effects on health and costs of the health interventions. This fashion is repeated for a given sample size of the trial and repeated many times to calculate the power and expected net benefits. The expected net benefits is calculated as the sum patient-based ENB, and the resulting standardized ENBs are used to calculate the power comparing the standardized values with reference value from the standard normal distribution. To find optimal results, these are repeated for a range of sample sizes. Sensitivity analysis was applied by simulating different settings of the parameters, to quantify uncertainty on the optimal expected net benefits, sample size and power. Uncertainty was considered in the health benefits, and accordingly three sets of results were generated. Still, although this hybrid research describes a research methods based on two-stages modeling and simulation, the results of these clinical trials are likely to arise until a complete real data are available, allowing for an updated and validated optimal results. Nonetheless, further developments can be assumed based on multiple-testing (for example, superiority, equivalence and non-inferiority simultaneously) or multi-two-stages. n-two-stages of modeling and simulation of a clinical trial means that the results of the applied modeling and simulation can be updated each time there are additional data, and consequently sample size and power of an optimal expected net benefits will be modified depending on previous updates. The regulatory approval of a prevention strategy, diagnostic technology, pharmaceutical or heath care program requires the development of clinical trials ensuring that benefits are derived. Since that these trials are long, costly and complex to be manage and their result are subject to variability and uncertainty, the anticipated analysis of their expected data on cost-benefit for health system and society can bring invaluable contributions. Modeling and simulation not only could make such analysis possible. They also make available research methods on how a particular clinical trial should be designed, conducted and extrapolated considering their expected net benefits. The modeling and simulation framework is based on scientific methods, making decisions on whether conducting clinical trials provides net benefits for health system, and contributing to the transfer of the discovered knowledge at experimental environments to medicine applications.

ACKNOWLEDGMENTS

We would like to thank anonymous reviewers for their invaluable comments and suggestions that improved a previous version of this paper. We also appreciate Dr. Navonil Mustafee from the University of Exeter for editing the manuscript.

REFERENCES

Abbas, I. 2016. "Modeling and Simulation in Clinical Trials". SpringSim-MSM 2016, April 3-6 2016, Pasadena, CA, USA. The Society of Modeling and Simulation International (SCS), 788-793.

Abbas, I. 2015. "Simulating The Expected Net Benefits of Health Research: The Case of Randomized Clinical Trials for Comparing two Health Interventions". SpringSim-MSM 2015 April 12-15, 2015 Alexandria, VA, USA. The Society of Modeling and Simulation International (SCS), 640-648.

Abbas, I., J. Rovira, and J. Casanovas. 2006a. "Validation by Simulation of a Clinical Trial Model Using Standardized Mean and Variance Criteria". *Journal of Biomedical Informatics* 39:687-696.

Abbas, I., J. Rovira, and J. Casanovas. 2007. "Clinical Trial Optimization: Monte Carlo Simulation Markov Model for Planning Clinical Trials Recruitment". *Contemporary Clinical Trials* 28(3):220-231.

Abbas, I., J. Rovira, Casanovas J, and T. Greenfield. 2008b. "Optimal Design of Clinical Trials with Computer Simulation Based on Results of Earlier Trials, Illustrated with a Lipodystrophy Trial in HIV Patients". *Journal of Biomedical Informatics* 41:1053–1061.

Abbas, I., J. Rovira, E. Cobo, and J. Casanovas. 2006b. "Simulation Models for Optimizing The Design of Clinical Trials". *Quality and Reliability Engineering International* 22:683-691.

Abbas, I., J. Rovira, E. Cobo, J. Romeu, and J. Casanovas. 2008a. "Disease and Clinical Trials Modeling". *Encyclopedia of Statistics in Quality and Reliability*, ed1: 558-561.

Abbas I., J. Rovira, and J. Casanovas. 2013. "Hypothesis Testing by Simulation of a Medical Study Model Using the Expected Net Benefits Criteria". *Health* 5: 364-374.

Bonate, P. L. 2000. "Rank Power of Metrics Used to Assess QTC Interval Prolongation by Clinical Trial Simulation". *Journal of Clinical Pharmacology,* 40(5):468-474.

Burman, F. C., B. Hamren, and P. Olsson 2005. "Modelling and Simulation to Improve Decision-Making in Clinical Development". *Pharmaceutical Statistics* , 4:47-58.

Chabaud, S., P. Girard, P. Nony, and J. P. Boissel 2002. "Clinical Trial Simulation Using Therapeutic Effect Modeling: Application to Ivabradine Efficacy in Patients with Angina Pectoris". *Journal of Pharmacokinetics and Pharmacodynamics,* 29 (4): 339-363.

Cheng, A., and D. B. Raeme. 2014. "Is Clinical Trial Registration for Simulation-Based Research Necessary?" *Simulation in Healthcare,* 9(6):350-352.

Ciolino, J., W. Zhao, R. Martin, Y. Palesch 2010. "Quantifying the Cost in Power of Ignoring Continuous Covariant Imbalances in Clinical Trial Randomization". *Contemporary Clinical Trials*, 32(2):250-259.

Elashoff, J. D., G. G. Koch, G. Y. Chi 1988. "Designing a Clinical Trial to Demonstrate Prevention of Ulcer Recurrence: Modeling Simulation Approaches". *Statistics in Medicine*, 7(8):877-888.

Gold, D. L., M. Dawson, H. Yang, J. Parker, and D. L. Gossage. 2014. "Clinical Trial Simulation to Assist in COPD Trial Planning and Design with a Biomarker-Based Diagnostic: When to Pull the Trigger?" *COPD: Journal of Chronic Obstructive Pulmonary Disease*, 11(2):226-235.

Henderson, W. G., S. G. Fisher, L. Weber, K. E. Hammermeister, and G. Sethi. 1991. "Conditional Power for Arbitrary Survival Curves to Decide Whether to Extend a Clinical Trial". *Control Clinical Trials*, 12(2):304-313.

Hmelo, C. E., S. Ramakrishnan, R. S. Day, W. E. Shirey, A. Brufsky, C. Johnson, J. Baar, and Q. Huang. 2001. "Oncology Thinking Cap: Scaffolded Use of a Simulation to Learn Clinical Trial Design". *Teaching Learning Medicine*, 13(3):183-191.

Jin, Y., R. Bies, M. R. Gastonguay, Y. Wang, N. Stockbridge, J. Gobburu, R. Madabushi. 2014. "Predicted Impact of Various Clinical Practice Strategies on Cardiovascular Risk for the Treatment of Hypertension: a Clinical Trial Simulation Study". *Jouranl of Pharmacokinetics and Pharmacodynamics*, 41(6):693-704.

Kraiczi, H., M. Frisen. 2005. "Effect of Uncertainty about Population Parameters on Pharmacodynamics-Based Prediction of Clinical Trial Power". *Contemporary Clinical Trials*, 26(2):118-130.

Laouenan, C., J. Guedj, F. Mentre. 2013. "Clinical trial simulation to evaluate power to compare the antiviral effectiveness of two hepatitis C protease inhibitors using nonlinear mixed effect models: a viral kinetic approach". *BMC Medical Research Methodology*, 25: 13-60.

Lui, K. J., and K. C. Chang. 2008. "Test Equality and Sample Size Calculation Based on Risk Difference in a Randomized Clinical Trial with Noncompliance and Missing Outcomes". *Biometrical Journal*. 50(2):224-236.

Nakai, K., S Iida, M. Tobinai, J. Hashimoto, and T. Kawanishi. 2015. "Application of Modeling and Simulation to a Long-Term Clinical Trial: a Direct Comparison of Simulated Data and Data Actually Observed in Japanese Osteoporosis Patients Following 3-year Ibandronate Treatment". *Clinical Pharmacokinetics*, 54(3):295-304.

Nixon, R. M., A. O'Hagan, J. Oakley, J. Madan, W. Stevens, N. Bansback, and A. Brennan. 2009. "The Rheumatoid Arthritis Drug Development Model: a Case Study in Bayesian Clinical Trial Simulation". *Pharmaceutical Statistics*, 8(4): 371-389.

Patel, N. R., S. Ankolekar. 2007. A Bayesian approach for incorporating economic factors in sample size design for clinical studies of individual drugs and portfolios of drugs. *Statistics in Medicine*, 26:4976-4988.

Sanghera, S., T. E. Roberts, P. Barton, E. Frew, J. Daniels, L. Middleton, L. Gennard, J. Kai, J. K. Gupta. 2014. "Levonorgestrel-Releasing Intrauterine System vs. Usual Medical Treatment for Menorrhagia: an Economic Evaluation Alongside a Randomised Controlled Trial". *PLoS One*, 17;9(3):e91891.

Schackman, B. R., C. A. Scott, P. E. Sax, E. Losina, T. J. Wilkin, J. E. McKinnon, S. Swindells, W. C. Weinstein, and K. A. Freedberg. 2007. "Potential risks and benefits of HIV Treatment Simplification: a Simulation Model of a Proposed Clinical Trial". *Clinical Infection Disease*, 45(8): 1062-1072.

Seidl, H., C. Meisinger, R. Wende, and R. Holle. 2012. "Empirical Analysis Shows Reduced Cost Data Collection May be an Efficient Method in Economic Clinical Trials". *BMC Health Services Research*, 12:318.

Stamey, J. D., F. Natanegara, and J. W. Seaman. 2013. "Bayesian Sample Size Determination for a Clinical Trial with Correlated Continuous and Binary Outcomes". *Journal of Biopharmaceutical Stat*istics, 23(4):790-803.

Statistical Principals for Clinical Trials 1998. "International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceutical for Human Use".

Tsodikov, A., D. Hasenclever, and M. Loeffler. 1998. "Regression with Bounded Outcome Score: Evaluation of Power by Bootstrap and Simulation in a Chronic Myelogenous Leukemia Clinical Trial". *Statistics in Medicine*, 17(17):1909-1922.

Willan A. R. and E. M. Pinto. 2005. "The Value of Information and Optimal Clinical Trial Design". *Statistics in Medicine*; 24(12): 1791-1806.

Wu, C. H. 2015. "Sample Size Re-Estimation without Breaking the Blind in Clinical Trial". *Journal of Biopharmaceutical Statistics*, 25(5):1114-1129.

Yang, S. and M. Beerahee. 2011. "Power Estimation Using a Population Pharmacokinetics Model with Optimal Design by Clinical Trial Simulations: Application in Pharmacokinetics Drug-Drug Interaction Studies". *European Journal of Clinical Pharmacology,* 67 (3): 225-233.

AUTHOR BIOGRAPHY

ISMAIL ABBAS is a professor of the Department of Statistics and Operations Research of the Polytechnic University of Catalonia. He holds a PhD in Statistics and Operation Research from Polytechnic University of Catalonia. His main research interests include, modeling and simulation of clinical trials, probability and statistics. His email address is ismail.i.abbas@gmail.com.