

Two-stage re-estimation adaptive design: a simulation study

FRANCESCA GALLI⁽¹⁾, LUIGI MARIANI⁽¹⁾

ABSTRACT

BACKGROUND: adaptive clinical trial design has been proposed as a promising new approach to improve the drug discovery process. Among the many options available, adaptive sample size re-estimation is of great interest mainly because of its ability to avoid a large ‘up-front’ commitment of resources. In this simulation study, we investigate the statistical properties of two-stage sample size re-estimation designs in terms of type I error control, study power and sample size, in comparison with the fixed-sample study.

METHODS: we simulated a balanced two-arm trial aimed at comparing two means of normally distributed data, using the inverse normal method to combine the results of each stage, and considering scenarios jointly defined by the following factors: the sample size re-estimation method, the information fraction, the type of group sequential boundaries and the use of futility stopping. Calculations were performed using the statistical software SASTM (version 9.2).

RESULTS: under the null hypothesis, any type of adaptive design considered maintained the prefixed type I error rate, but futility stopping was required to avoid the unwanted increase in sample size. When deviating from the null hypothesis, the gain in power usually achieved with the adaptive design and its performance in terms of sample size were influenced by the specific design options considered.

CONCLUSIONS: we show that adaptive designs incorporating futility stopping, a sufficiently high information fraction (50-70%) and the conditional power method for sample size re-estimation have good statistical properties, which include a gain in power when trial results are less favourable than anticipated.

Key words: Sample size re-estimation design; Sample size re-estimation method; Group sequential boundaries; Information fraction; Futility stopping

1) Medical Statistics, Biometry and Bioinformatics - Unit of Clinical Epidemiology and Trial Organization; Fondazione IRCCS Istituto Nazionale Tumori, Milan, Italy

CORRESPONDING AUTHOR: Francesca Galli, Fondazione IRCCS Istituto Nazionale Tumori, Medical Statistics, Biometry and Bioinformatics, Unit of Clinical Epidemiology and Trial Organization; Via G. Venezian 1, 20133 Milan (Italy). e-mail: francesca.galli01@gmail.com

DOI: 10.2427/8862

Published as Online First on March 18, 2014

INTRODUCTION

It is being increasingly recognized that a growing expense for biomedical research does not lead to an increased number of successful therapies that enter in clinical practice.

Reasons identified are: (1) a diminished margin for therapeutic improvement, which escalates the level of difficulty in proving drug benefit; (2) genomics and other innovative scientific tools having not yet reached their full potential; (3) mergers

and other business arrangements that have decreased developmental candidates; (4) easy targets are the focus as chronic diseases are more difficult to study; (5) a persistently high failure rate in drug development; and (6) the rapidly escalating research costs and complexity, which decrease willingness/ability to bring many candidates forward into the clinic [1].

In 2004, taking into account the situation, the Food and Drug Administration (FDA) launched a project, called *Critical Path Initiative*, aimed at driving innovation in the scientific processes through which medical products are developed, evaluated, and manufactured [2]. Later, in 2006, the FDA released the *Critical Path Opportunities List*, a document collecting 76 projects aimed at filling the gap between the resources spent and the results obtained in clinical research. One of these projects focused on the development of innovative clinical trial methodologies, such as Bayesian approaches or adaptive designs, able to use prior experience or accumulated information in the effort to improve research efficiency [3].

Adaptive designs in particular are defined by the FDA as typical of those studies that incorporate a prospectively planned opportunity for modification of one or more protocol specified aspects regarding patient selection, treatment and assessment, or even the statistical hypotheses to be tested; this is done based on the analysis of accumulated data from the subjects enrolled up to a given point in time. Analyses of the accumulated data performed at prospectively planned time points within the study are also labelled as interim analyses, and can be performed in a fully blinded or an unblinded manner, as well as with or without formal statistical hypothesis testing [4].

There are many proposals on how to conduct an adaptive study. The main design types can be classified according to the phase of clinical research (learning or confirmative phase) in which they are suitable. Designs indicated in the learning phase are (1) adaptive dose finding design and (2) adaptive seamless phase I/II design. Designs suitable in the confirmatory phase are (1) sample size re-estimation design; (2) biomarker adaptive (or enrichment) design; (3) adaptive randomization design; (4) adaptive treatment switching design; (5) adaptive group sequential design; (6) adaptive hypotheses design; (7) adaptive recursive design. Finally,

a type of adaptive design useful for both the learning and confirmatory phase is the adaptive seamless phase II/III design.

Among such a variety, the approach that has the largest applicability is the sample size re-estimation design, through which it is possible in the course of the study to increase or decrease the sample size initially planned based on interim results. The reasons for interest in this approach are obvious: one may avoid to oversize the study if intermediate results are favourable, or to miss clinically meaningful treatment effects if intermediate results are less optimistic than anticipated. At the same time, however, there is controversy in the statistical literature about the actual benefits implied by the re-estimation design in comparison with studies incorporating conventional interim analyses [5].

In this paper, we report the results of simulations used for investigating the statistical properties of two-stage sample size re-estimation designs in terms of type I error control, study power and sample size. Such an assessment was made in comparison with classic fixed-sample and group-sequential studies, and considering various options available for sample size re-estimation designs.

METHODS

Before describing simulations in detail, it is important to recall the process used to combine the results of each stage of an adaptive study according to the inverse normal method, as proposed by Lehmacher and Wassmer [6], and the available methods for sample size re-estimation.

In general, the test statistic resulting from the combination of independent p-values is given by

$$\frac{1}{\sqrt{k}} \sum_{\tilde{k}=1}^k \Phi^{-1}(1 - p_{\tilde{k}}), \tag{1}$$

where k is one of the K planned stages, $k = 1, 2, \dots, K$, $p_{\tilde{k}}$ is the p-value of the \tilde{k} -th stage and $\Phi^{-1}(\cdot)$ denotes the inverse cumulative standard normal distribution function.

An advantage of the inverse normal method, unlike the Fisher combination test [7], is that classical group sequential boundaries for early acceptance or rejection of the null hypothesis

may be used in conjunction with test statistic (1), thus simplifying study design. Since the $\Phi^{-1}(1 - p_k)$'s, $k = 1, 2, \dots, K$, are independent and standard normally distributed, the proposed approach maintains type I error rates α exactly for any (adaptive) choice of sample sizes [6].

Two methods may be used for sample-size re-estimation. The first one is based on the effect size ratio [8] which requires that re-estimation of the sample size after the interim analysis is given by

$$N = \left(\frac{E_0}{E} \right)^2 N_0,$$

where N is the newly estimated sample size for the entire study, N_0 is the initial sample size (estimated using the method of a classic study), E is the observed effect size and E_0 is the initial estimate of effect size (used for the estimation of N).

The second re-estimation method is based on the conditional power [1], i.e. the conditional probability of rejecting the null hypothesis during the rest of the trial based on the observed interim data. When using the inverse normal method for combining independent p-values, the re-estimation of sample size after the interim analysis in a two-stage adaptive trial is given by

$$N_2 = \frac{2\sigma^2}{\delta^2} \left[\frac{\Phi^{-1}(1 - \alpha_2) - \Phi^{-1}(1 - p_1)/\sqrt{2}}{1/\sqrt{2}} - \Phi^{-1}(1 - cP) \right]^2.$$

Where N_2 is the newly estimated sample size only for the second stage of the trial, σ^2 is the known variance, δ is the target difference between group means, $\Phi^{-1}(\cdot)$ denotes the inverse cumulative standard normal distribution function, α_2 is the critical value for the final analysis according to the chosen decision boundaries, p_1 is the p-value achieved at the interim analysis and cP is the pre-planned conditional power. For obvious practical reasons, sample size extension is typically capped to a suitably chosen maximum overall sample size.

As regards simulations, we assumed the situation of a balanced two-arm trial aimed at comparing two means of normally distributed data. Using a fixed-sample design, 233 observations per trial arm (466 overall) are required to test a standardized effect size $\delta=0.3$ with type I error rate and power set to 2.5%

and 90%, respectively. Various scenarios were considered based on a number of factors.

The first one was the information fraction, i.e. the ratio between the sample size at which the interim analysis is conducted and the sample size fixed a priori for the entire study (obtained with conventional methods for the fixed-sample design). Most two-stage studies use an information fraction for the interim analysis equal to 50%, i.e. half of the sample size planned; in addition to this choice, we also considered information fractions of 30% or 70%.

The second factor was the type of group sequential boundaries. We decided to consider O'Brien and Fleming's and Pocock's boundaries, which are the most frequently used and cover two quite heterogeneous settings. While O'Brien and Fleming's boundaries imply monotonically decreasing critical values, making it difficult to stop the trial early but with almost no loss of statistical power, the opposite properties characterize Pocock's boundaries, in which critical values are constant at each stage [9].

Both O'Brien and Fleming's and Pocock's boundaries allow early stop of the study only for efficacy of treatment tested or, in other words, for the rejection of the null hypothesis at the interim analysis. Quite commonly, however, researchers are willing to stop the trial prematurely when results are not sufficiently promising, that is for futility. Therefore, we allowed for this by exploring the use of a futility threshold represented by a p-value at the interim analysis $p_1 \geq 0.5$.

Finally, we considered both re-estimation methods described before, with capping at one and a half the sample size fixed a priori (maximum 350 observations per trial arm) and cP set to 0.90.

For each scenario jointly defined by the above factors, we performed 1x10⁶ simulations and calculated the proportion of simulated studies yielding significant results: (1) under the null hypothesis, H_0 , to evaluate the type I error rates; (2) under the alternative hypothesis, H_1 , to assess the power of the study; (3) in between the null and alternative hypothesis, H_0/H_1 , as a way to show the gain achievable with sample size re-estimation. In particular, we assumed in this setting an effect size $\delta=0.2$, implying a drop in power of the fixed-sample design from 90% to 53%.

Additional simulation outputs were: (1) the futility stopping probability, i.e. the probability of early stopping of the study with acceptance

of the null hypothesis; (2) the efficacy stopping probability, i.e. the probability of early stopping of the study with rejection of the null hypothesis; (3) the average sample size.

Calculations were performed using two macros, written with the statistical software SASTM (version 9.2) by the authors, which are supplied in the Supplementary Materials.

RESULTS

Simulation results obtained using the effect size ratio and the conditional power re-estimation methods are shown in detail in Tables 1 and 2, respectively.

Under H_0 , any type of adaptive design considered was able to maintain the type I error rate equal to the nominal 2.5% level, with no difference over the distinct scenarios used for testing adaptive designs. A striking result, however, was the increase in sample size when the futility stopping rule was not adopted, compared to the size of the fixed-sample study. Such an increase was slightly affected by the chosen information fraction, the type of decision boundary and the method for sample size re-estimation. At worst (70% information fraction, O'Brien and Fleming's boundary, effect size re-estimation method, Table 1) the average increase amounted to 49% (348 subjects per trials arm compared to 233 for the fixed-sample design), which is close to the limit imposed by the capping. Such a drawback was no longer evident when adopting a futility stopping rule. Rather, a gain in efficiency was obtained in this case insofar, at no cost in terms of type I error probability, sample size was somewhat diminished, provided that the information fraction did not exceed 50%.

Results in terms of power and sample size under H_1 and H_0/H_1 are plotted in Figures 1 and 2. In these plots, power and sample size of the fixed-sample design are represented by horizontal reference lines. In power plots, the dots above or below the reference line indicate whether power is preserved (or even increased) or not in the specific simulation; similarly, dots in sample size plots above or below the reference line indicate whether sample size is increased (loss of efficiency) or decreased (gain in efficiency) compared to the fixed-sample design.

Under H_1 , and using the effect size

re-estimation method (Figure 1a), a gain in power beyond the nominal level was always achieved. However, contrary to what one normally expects from group sequential trials, sample size was increased when using the O'Brien and Fleming's boundary, even in the presence of a futility stopping rule. In contrast, a reduction in sample size was steadily obtained with the Pocock's boundary.

The picture was rather different when using the conditional power re-estimation method (Figure 1b). Power was again preserved, at least when adopting an information fraction of 50% or 70%, but sample size was steadily diminished by 7% to 20% compared to that required by the fixed-sample design, with no relevant effect of the type of decision boundary and the use or less of futility stopping.

Under the intermediate hypothesis H_0/H_1 and using the effect size ratio re-estimation method (Figure 2a), the adaptive approach always improved study power compared to that obtained with a fixed-sample design; the best outcome was obtained with the O'Brien and Fleming's boundary and a 70% information fraction, a condition in which 74% power was achieved. On the other hand, the price to pay to improve study power was an increase in sample size varying between 8% and 32%.

With the conditional power re-estimation method (Figure 2b), study power showed considerable improvement only using an information fraction of 50% or 70%. Compared with the performance of the effect size ratio re-estimation method, such an improvement was slightly smaller, considering that study power reached a maximum of 67-68% (versus 74%), but the price in terms of sample size increase was also smaller, ranging between 4% and 17%.

Finally, the futility stopping probability (FSP) and the efficacy stopping probability (ESP), obviously not affected by the re-estimation method, were influenced by the information fraction and, when considering ESP, also by the type of decision boundaries. In particular, FSP (when allowed for by design) tended to decrease with increasing information fraction under H_1 and H_0/H_1 . ESP, while stable under H_0 , tended to increase with increasing information fraction under H_1 and H_0/H_1 ; furthermore, under otherwise similar conditions, ESP was always larger with Pocock's compared to O'Brien and Fleming's boundaries, compatibly with the conservative nature of the latter.

TABLE 1

TWO-STAGE ADAPTIVE DESIGN WITH THE EFFECT SIZE RATIO RE-ESTIMATION METHOD*								
INFORMATION FRACTION	DECISION BOUNDARIES	HYPOTHESIS	FUTILITY STOPPING	FSP (%)	ESP (%)	AVERAGE SAMPLE SIZE (PER GROUP)	PROBABILITY OF REJECTING H_0 (%)	SIMULATION NUMBER
30%	O'Brien/Fleming	H_0	No	0	0.3	336	2.5	1
			Yes	50.0	0.3	203	2.5	2
		H_1	No	0	15.4	258	95.4	3
			Yes	3.8	15.4	247	93.1	4
		H_0/H_1	No	0	5.3	301	67.7	5
			Yes	11.9	5.3	268	65.5	6
	Pocock	H_0	No	0	1.5	334	2.5	7
			Yes	50.0	1.5	201	2.5	8
		H_1	No	0	34.4	227	93.2	9
			Yes	3.8	34.4	216	91.2	10
		H_0/H_1	No	0	16.0	284	60.8	11
			Yes	11.8	16.0	251	59.2	12
50%	O'Brien/Fleming	H_0	No	0	0.3	345	2.5	13
			Yes	50.0	0.3	231	2.5	14
		H_1	No	0	30.9	245	96.7	15
			Yes	1.1	30.9	242	96.1	16
		H_0/H_1	No	0	10.3	303	71.9	17
			Yes	6.3	10.3	288	71.1	18
	Pocock	H_0	No	0	1.5	344	2.5	19
			Yes	50.0	1.5	229	2.5	20
		H_1	No	0	54.7	217	95.7	21
			Yes	1.1	54.7	214	95.2	22
		H_0/H_1	No	0	25.9	284	67.2	23
			Yes	6.3	25.9	270	66.5	24
70%	O'Brien/Fleming	H_0	No	0	0.3	348	2.5	25
			Yes	50.0	0.3	255	2.5	26
		H_1	No	0	46.6	246	97.2	27
			Yes	0.3	46.6	246	97.0	28
		H_0/H_1	No	0	16.1	307	73.8	29
			Yes	3.6	16.1	301	73.6	30
	Pocock	H_0	No	0	1.5	346	2.5	31
			Yes	50.0	1.5	254	2.5	32
		H_1	No	0	70.2	219	96.6	33
			Yes	0.3	70.2	218	96.5	34
		H_0/H_1	No	0	35.5	284	70.2	35
			Yes	3.6	35.5	277	69.9	36

*Futility stopping probability (FSP), efficacy stopping probability (ESP), average sample size and probability of rejecting the null hypothesis by varying information fraction, decision boundaries and futility stopping rule

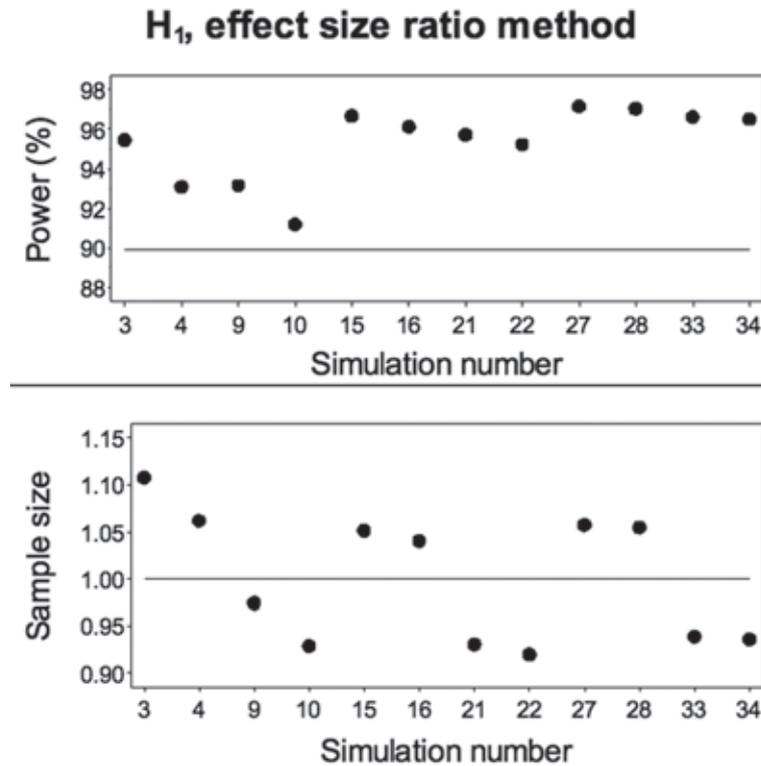
TABLE 2

TWO-STAGE ADAPTIVE DESIGN WITH THE CONDITIONAL POWER RE-ESTIMATION METHOD*								
INFORMATION FRACTION	DECISION BOUNDARIES	HYPOTHESIS	FUTILITY STOPPING	FSP (%)	ESP (%)	AVERAGE SAMPLE SIZE (PER GROUP)	PROBABILITY OF REJECTING H_0 (%)	SIMULATION NUMBER
30%	O'Brien/Fleming	H_0	No	0	0.3	325	2.5	37
			Yes	50.0	0.3	184	2.5	38
		H_1	No	0	15.4	198	90.2	39
			Yes	3.8	15.4	187	87.8	40
		H_0/H_1	No	0	5.3	250	59.3	41
			Yes	11.9	5.3	216	57.1	42
	Pocock	H_0	No	0	1.5	333	2.5	43
			Yes	50.0	1.5	193	2.5	44
		H_1	No	0	34.4	205	91.0	45
			Yes	3.8	34.4	195	89.0	46
		H_0/H_1	No	0	16.0	262	57.0	47
			Yes	11.8	16.0	229	55.5	48
50%	O'Brien/Fleming	H_0	No	0	0.3	336	2.5	49
			Yes	50.0	0.3	220	2.5	50
		H_1	No	0	30.9	200	92.2	51
			Yes	1.1	30.9	197	91.6	52
		H_0/H_1	No	0	10.3	258	64.1	53
			Yes	6.3	10.3	242	63.3	54
	Pocock	H_0	No	0	1.5	341	2.5	55
			Yes	50.0	1.5	224	2.5	56
		H_1	No	0	54.7	198	93.8	57
			Yes	1.1	54.7	195	93.4	58
		H_0/H_1	No	0	25.9	263	63.4	59
			Yes	6.3	25.9	248	62.8	60
70%	O'Brien/Fleming	H_0	No	0	0.3	344	2.5	61
			Yes	50.0	0.3	250	2.5	62
		H_1	No	0	46.6	217	93.9	63
			Yes	0.3	46.6	216	93.8	64
		H_0/H_1	No	0	16.1	273	67.7	65
			Yes	3.6	16.1	266	67.4	66
	Pocock	H_0	No	0	1.5	345	2.5	67
			Yes	50.0	1.5	252	2.5	68
		H_1	No	0	70.2	210	95.7	69
			Yes	0.3	70.2	209	95.5	70
		H_0/H_1	No	0	35.5	272	68.0	71
			Yes	3.6	35.5	265	67.7	72

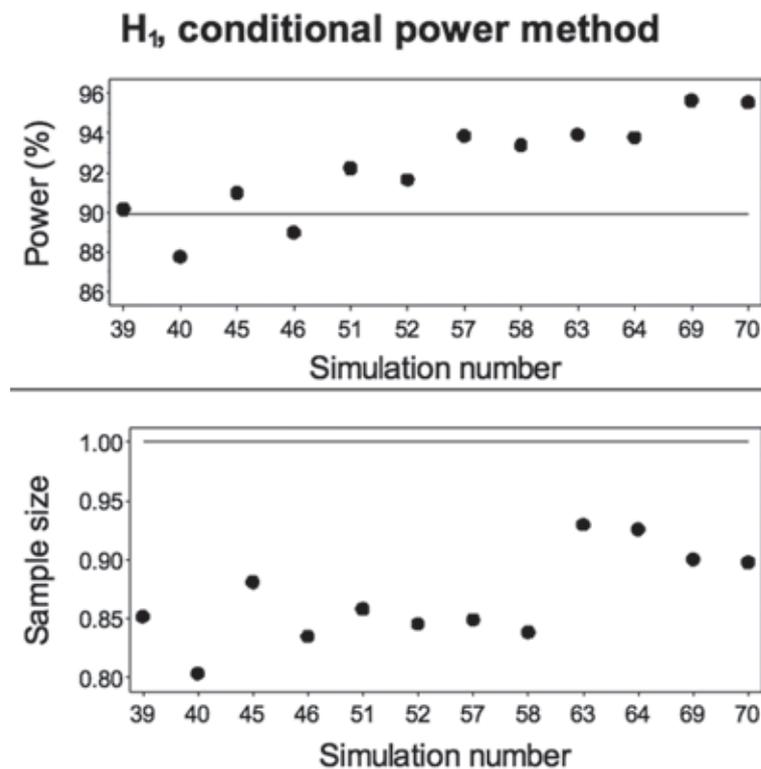
*Futility stopping probability (FSP), efficacy stopping probability (ESP), average sample size and probability of rejecting the null hypothesis by varying information fraction, decision boundaries and futility stopping rule

FIGURE 1

a)



b)

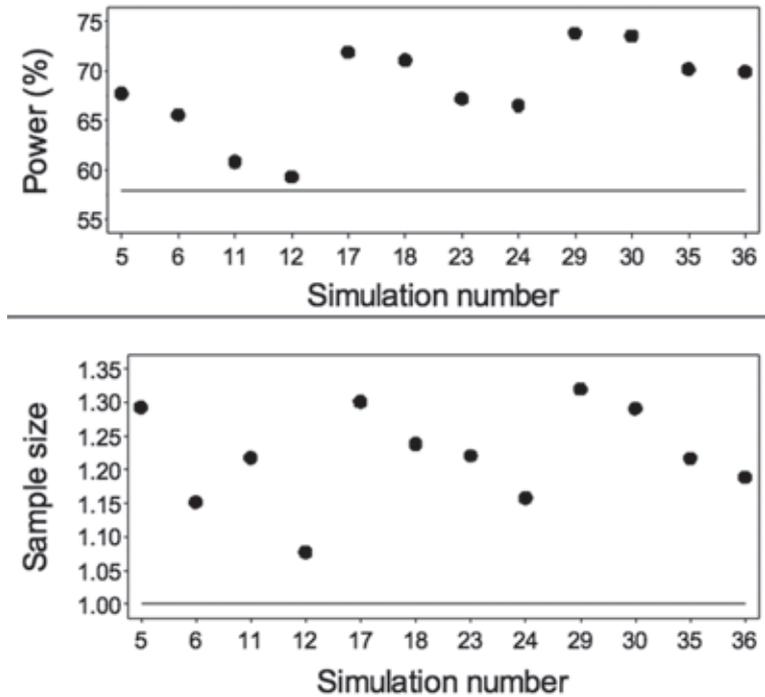


Power and average sample size under the alternative hypothesis (H_1) according to simulation number (see Tables 1-2 for identifying associated scenarios) using the effect size ratio re-estimation method (Figure 1a) or the conditional power method (Figure 1b)

FIGURE 2

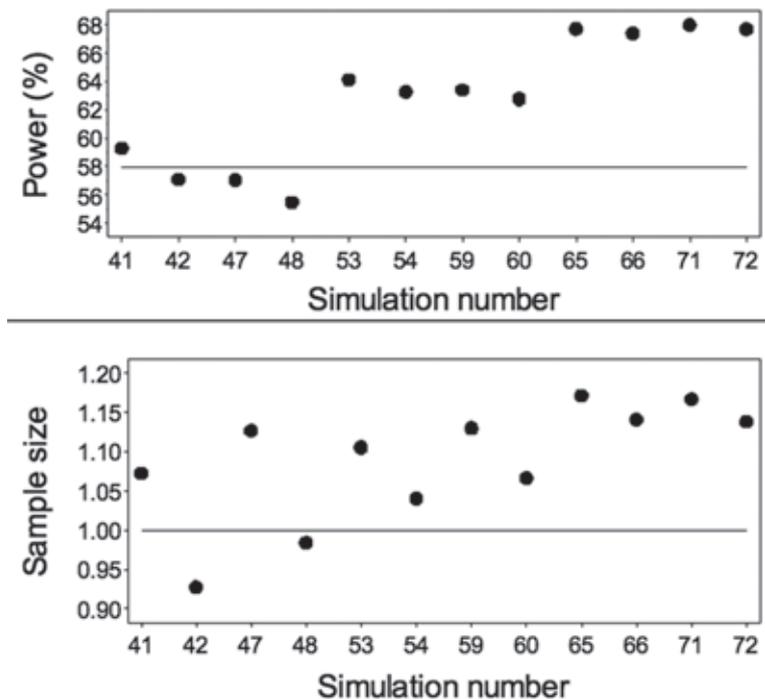
a)

H_0/H_1 , effect size ratio method



b)

H_0/H_1 , conditional power method



Power and average sample size under the intermediate hypothesis (H_0/H_1) according to simulation number (see Tables 1-2 for identifying associated scenarios) using the effect size ratio re-estimation method (Figure 2a) or the conditional power method (Figure 2b)

DISCUSSION

The possibility of adapting some of the characteristics of a clinical trial, such as sample size, based on the results of one or more interim analysis is definitely a great opportunity. It must be underlined, however, that this kind of adaptation is not always applicable or ideal: for example, it is more easily manageable when the response to treatment can be measured and recorded in a short period of time. Furthermore, any type of interim analysis adds complexity to the logistics management of the trial.

These problems aside, from the statistical viewpoint it is important that changes made on the basis of interim results do not adversely affect the operative characteristics of the study. Basically, it is fundamental that type I error rate is maintained as close as possible to the pre-established level under the null hypothesis, and that study power is preserved under a reasonably wide range of conditions not covered by the null hypothesis, with the least possible price to pay in terms of sample size.

Results of our simulations are very instructive in showing that a gain in power is generally obtained with re-estimation adaptive designs, but that at the same time not all choices about the various options available are equally able to ensure good trial statistical properties. Therefore, when designing a study which allows for sample size re-estimation, it is of utmost importance to carefully check the implications of the options specifically chosen. This task is typically performed by simulation, using for instance the macros supplied in the Supplementary Materials or one of the commercially available ad hoc packages. There are however some guidelines that may be drawn from our investigation useful to guide applied statisticians.

First of all, one very obvious and unwanted drawback of re-estimation adaptive designs is the remarkable increase in sample size under the null hypothesis. This disadvantage, however, can be totally prevented by incorporating an early stopping rule for futility, which use is thus to be considered mandatory in this setting.

The choice of information fraction is also crucial. Only if it is sufficiently high (50% or 70% in our simulations), study power shows considerable improvement when intermediate results are less favourable than anticipated. Theoretically, the nominal power might be

fully achieved with sample size re-estimation. However, since a clear upfront commitment in terms of sample size is typically required by both the sponsors and ethical committees, limitation in the maximum sample size is commonly adopted, which limits as well maximum power recovery. Again, the “capping threshold” (1.5 in our simulations) is one of the study design parameters that need to be assessed in the planning phase under reasonable scenarios.

Regarding re-estimation methods, we showed more favourable properties for the conditional power method. Such a procedure, compared with the effect size re-estimation method, is more efficient (sample size is steadily diminished) under the alternative hypothesis, as well as in conditions intermediate between the null and alternative hypotheses.

Finally, when considering the joint use of futility stopping, 50-70% information fraction and conditional power re-estimation method, only trivial differences emerged in the statistical properties of O'Brien and Fleming's or Pocock's decision boundaries. It must be considered, however, that researchers are usually willing not to stop the trial too early because of favourable results, which is a justification for preferring the more conservative O'Brien and Fleming's boundaries.

As mentioned in the Introduction, there is controversy about the actual benefits implied by the re-estimation design in comparison with studies incorporating conventional interim analyses. In such a comparison, the most critical aspect from the statistical viewpoint is the loss of statistical efficiency resulting from the weighting scheme used to combine the results of each stage of an adaptive study according to the inverse normal method, as well explained by Fleming [10]. This author supports his argument on the basis of an example in which first and second stage sample sizes, 200 and 1 100 respectively, are greatly imbalanced. The imbalance, however, is unlikely to be so relevant in most practical conditions. The consequent loss of efficiency might thus be trivial and easily offset by the advantage of a smaller up-front investment of sample size resources, followed by a larger subsequent investment contingent on seeing promising results from the interim analysis, whenever a futility stopping rule (as done in our simulations) or conditional power calculation are incorporated into study monitoring.

In conclusion, although the presented simulations are limited to a restricted set of possible scenarios, our findings are in agreement with theory and may be regarded as sufficiently generalizable to other settings. In any case, as underlined before, applied statistician willing to adopt the type of adaptive designs here

discussed must carefully check their properties in the specific situations through simulation of a number of possible scenarios. In this case, as we have shown with our investigation, re-estimation adaptive designs may actually improve the statistical quality of clinical trials.

References

- [1] Chang M. Adaptive design theory and implementation using SAS and R. New York: Chapman & Hall/CRC, 2008
- [2] The Food and Drug Administration (FDA). FDA's Critical Path Initiative. Available from: <http://www.fda.gov/ScienceResearch/SpecialTopics/CriticalPathInitiative/ucm076689.htm> [Accessed January 10, 2013]
- [3] The Food and Drug Administration (FDA). Critical Path Opportunities List. Available from: <http://www.fda.gov/downloads/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/UCM077258.pdf> [Accessed January 10, 2013]
- [4] The United States Food and Drug Administration (FDA). Draft Guidance for Industry – Adaptive Design Clinical Trials for Drugs and Biologics. Available from: www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM201790.pdf [Accessed January 10, 2013]
- [5] Tsiatis AA, Mehta C. On the inefficiency of the adaptive design for monitoring clinical trials. *Biometrika* 2003; 20: 367-78
- [6] Lehmacher W, Wassmer G. Adaptive sample size calculations in group sequential trials. *Biometrics* 1999;55:1286-90
- [7] Bauer P, Köhne K. Evaluation of Experiments with Adaptive Interim Analyses. *Biometrics* 1994; 50: 1029-41
- [8] Cui L, Hung HMJ, Wang SJ. Modification of sample-size in group sequential trials. *Biometrics* 1999, 55: 853-7
- [9] Wang SK, Tsiatis AA. Approximately optimal one-parameter boundaries for a sequential trials. *Biometrics* 1987; 43: 193-200
- [10] Fleming TR. Standard versus adaptive monitoring procedures: A commentary. *Statistics in Medicine* 2006; 25: 3305-12

