The Stata Journal

The *Stata Journal* publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go "beyond the Stata manual" in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those who teach, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

For more information on the *Stata Journal*, including information for authors, see the webpage

http://www.stata-journal.com

The *Stata Journal* is indexed and abstracted in the following:

- CompuMath Citation Index®
- Current Contents/Social and Behavioral Sciences®
- RePEc: Research Papers in Economics
- Science Citation Index Expanded (also known as SciSearch®)
- Scopus™
- Social Sciences Citation Index®

The *Stata Journal* (ISSN 1536-867X) is a publication of Stata Press. Stata, Mata, NetCourse, and Stata Press are registered trademarks of StataCorp LP.

The Stata Journal (2011)
11, Number 2, pp. 240–254

Simon's minimax and optimal and Jung's admissible two-stage designs with or without curtailment

Cornelia U. Kunz
Institute of Medical Biometry and Informatics
University of Heidelberg, Germany
kunz@imbi.uni-heidelberg.de

Meinhard Kieser
Institute of Medical Biometry and Informatics
University of Heidelberg, Germany

Abstract. This article describes a new Stata command called `simontwostage`, which calculates the critical values and sample sizes for two-stage designs for phase II oncology trials. Options are provided to determine the minimax and optimal designs proposed by Simon (1989, *Controlled Clinical Trials* 10: 1–10) and admissible designs described by Jung et al. (2004, *Statistics in Medicine* 23: 561–569). Furthermore, nonstochastic and stochastic curtailment rules can be implemented in both stages of the trial, and the properties of the curtailed designs can be examined.

Keywords: st0227, simontwostage, Simon's two-stage design, Jung's admissible design, phase II trials, conditional power, curtailment

1 Introduction

In phase II clinical trials in oncology, the potential efficacy of a new treatment regimen is assessed in terms of anticancer activity. The standard approach consists of a single-arm design; that is, all patients are treated with the new regimen. The endpoint of interest is typically tumor response with the possible outcomes "yes" or "no". For ethical and economical reasons, the trials are performed with planned interim analyses allowing for an early decision about stopping or continuing the development of the treatment if the number of observed responses is too low or already high enough, respectively. In practice, most trials are based on one interim analysis dividing the trial into two stages.

The most popular of such designs are still those proposed by Simon (1989). The course of those trials can be described as follows: In the first stage of the trial, a predefined number of patients is enrolled. If the number of observed responses is too low regarding a prespecified critical value, the trial is stopped and the treatment is considered inefficient. Otherwise, another fixed number of patients is enrolled in the trial. Again the number of observed responses is examined, and the treatment is considered inefficient if the total number of observed responses is too low compared with a prespecified critical value. Because the actual sample size is a random variable and cannot be fixed in advance, different criteria can be defined regarding maximum sample size and expected sample size leading to different designs. Two criteria are given by Simon

(1989): determination of designs with either the minimal total sample size for the whole trial (minimax design) or the minimal expected sample size given that the null hypothesis is true (optimal design). Jung et al. (2004) introduced the admissible designs that are compromises between the minimax and the optimal designs.

Various methodological extensions of this design have been proposed, for example, by applying alternative optimality criteria (see, for example, Chang et al. [1987], Shuster [2002], and Mander and Thompson [2010]). Instead of two stages (that is, one interim analysis) as in Simon's design, Chen (1997) implemented three stages. Furthermore, designs are proposed that use two endpoints (see, for example, Bryant and Day [1995]; Lin and Chen [2000]; Panageas et al. [2002]; Lu, Jin, and Lamborn [2005]; Lin, Allred, and Andrews [2008]; and Kunz and Kieser [forthcoming]). Other extensions can be found, for example, in Lin and Shih (2004) or Green and Dahlberg (1992).

A limitation of minimax, optimal, and admissible designs is that early stopping can only occur at the interim analysis; stopping can occur neither before termination of the first stage nor during the second stage, even when it becomes evident that rejection of the null hypothesis can no longer be achieved. Based on the conditional power (CP)—that is, the probability of rejecting $H_0 : p \leq p_0$ after the second stage under the alternative $H_a : p = p_1$ given the results observed so far—Ayanlowo and Redden (2007) implemented stochastic curtailment rules in the second stage of the trial. Their approach was extended by Kunz and Kieser (2011), who implemented nonstochastic (that is, when the CP is zero) and stochastic curtailment rules in both stages of the trial.

A Stata command (`simon2stage`) for calculating the critical values and sample sizes for the minimax and the optimal Simon's two-stage design was written by Adrian Mander (Mander 2009).

This article introduces another command for sample-size calculation for two-stage designs. This command, called `simontwostage`, allows determination of the critical values and sample sizes for the minimax, optimal, and admissible designs with or without nonstochastic or stochastic curtailment rules.

2 Two-stage designs with and without curtailment rules

2.1 Minimax and optimal designs

The objective of Simon's two-stage design is to test whether a new treatment should be developed any further. To allow for early stopping if the treatment is ineffective—that is, if the response rate p is less than or equal to a prespecified value, p_0—an interim analysis is performed after n_1 enrolled patients. Only if more than r_1 responses are observed will the trial continue to the second stage with further $n - n_1$ patients. At the end of the second stage, the null hypothesis $H_0 : p \leq p_0$ will be rejected if more than r responses out of n enrolled patients are observed. The critical values r_1 and r, as well as the sample sizes n_1 and n, can be determined for a specific level p_1 of the response rate in the alternative hypothesis H_a.

Let $b(n, x, p)$ denote the binomial distribution, and let $B(n, x, p)$ denote the cumulative binomial distribution. For prespecified response levels p_0 and p_1, the actual type I error rate α' and the actual type II error rate β' can be calculated for any given parameter set (r_1, n_1, r, n) in the following way:

$$\alpha' = 1 - \left\{ B(n_1, r_1, p_0) + \sum_{x=r_1+1}^{\min(n_1,r)} b(n_1, x, p_0)\, B(n - n_1, r - x, p_0) \right\} \tag{1}$$

$$\beta' = B(n_1, r_1, p_1) + \sum_{x=r_1+1}^{\min(n_1,r)} b(n_1, x, p_1)\, B(n - n_1, r - x, p_1) \tag{2}$$

The probability of early termination (PET) and the expected sample size (EN) if the treatment is inefficient are given by

$$\text{PET}(p_0) = B(n_1, r_1, p_0) \tag{3}$$

$$\text{EN}(p_0) = n_1 + \{1 - \text{PET}(p_0)\}(n - n_1) \tag{4}$$

The minimax design is defined as the design with the smallest total sample size n. If there is more than one design with smallest total sample size n, then the one with the smallest $\text{EN}(p_0)$ is chosen within the possible minimax designs. The optimal design is defined as the design with the smallest $\text{EN}(p_0)$ (Simon 1989).

2.2 Admissible designs

Sometimes designs can be found to fulfill the constraints regarding type I and type II errors even when those designs are neither minimax nor optimal but lie between the two regarding total sample size and EN. Those designs have a higher total sample size than the minimax design but a smaller total sample size than the optimal design. Furthermore, the EN of those designs is less than the EN of the minimax design but still higher than the EN of the optimal design. Those designs are candidates for an admissible design as defined by Jung et al. (2004). A design is admissible when it minimizes the Bayes risk $\rho = q \times n + (1 - q) \times \text{EN}(p_0) = \{n - \text{EN}(p_0)\} \times q + \text{EN}(p_0)$ for a given weight $q \in [0, 1]$. In general, a design is admissible for a range of q. The range can be found by plotting the Bayes risk ρ against the weight q for every design found for prespecified values of p_0, p_1, α, and β (see figure 1).

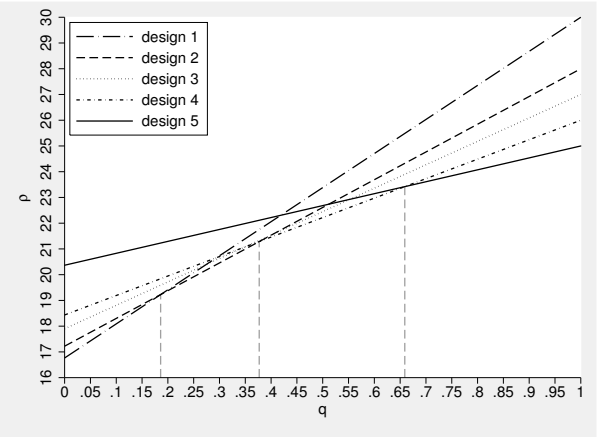

Figure 1. Bayes risk ρ against weight q for five different designs for $p_0 = 0.05$, $p_1 = 0.25$, $\alpha = 0.05$, and $\beta = 0.10$. Gray lines mark the borders for q.

For each design, the slope of the line is given by $n - \text{EN}(p_0)$, and the intercept is given by $\text{EN}(p_0)$. The range for q for every design can be found by determining the points of intersection of the lines. Between two points of intersection, the design with the lowest Bayes risk is admissible. For $q \in [0, 0.186]$, design 1 minimizes ρ and is, therefore, admissible in this range of q; for $q \in [0.186, 0.377]$, design 2 is admissible; for $q \in [0.377, 0.659]$, design 4 is admissible; and for $q \in [0.659, 1]$, design 5 is admissible. Although there is another design (design 3), it never minimizes ρ for any given q. Therefore, it is not an admissible design.

2.3 Algorithm for determining critical values and sample sizes

The critical values and sample sizes for the minimax, optimal, and admissible designs are determined using an algorithm that searches over n, n_1, r_1, and r. Using (2), the smallest possible values for n and n_1 can be calculated. Because both parts of the sum of (2) are positive, any given parameter set (r_1, n_1, r, n) must fulfill (5):

$$B(n_1, r_1, p_1) < \beta \tag{5}$$

Observing that $(1 - p_1)^{n_1} \leq B(n_1, r_1, p_1)$ leads to $n_1 > \log(\beta)/\log(1 - p_1)$. For $\beta \neq 1 - p_1$, the starting value for n_1 is $ceil\{\log(\beta)/\log(1 - p_1)\}$; and for $\beta = 1 - p_1$, the starting value for n_1 is 2. Because $n > n_1$, the starting value for the algorithm for n is $n_1 + 1$. The algorithm will, therefore, start with the smallest possible values for n_1 and n. For every r_1 in the range of $[0, n_1 - 1]$, we check whether $B(n_1, r_1, p_1) < \beta$ [see (5)]. If the inequality is not met, the algorithm will continue with the next n_1. Otherwise, it searches over r in the range of $[r_1 + 1, n - n_1 + r_1]$.

Because the type II error of the binomial test for a one-stage design is a lower boundary for the type II error of the Simon's design, we check whether $B(n, r, p_1) < \beta$

for every given parameter set (r_1, n_1, r, n). If the inequality does not hold, we stop the search over r and continue with the next r_1. Otherwise, type I and type II error rates are calculated. The algorithm continues to search over n_1 with a final value for n_1 of $n-1$ until the first design is found that fulfills the constraints for type I and type II errors. From the definitions of the minimax, optimal, and admissible designs, we know that if there is more than one design with minimal total sample size, the minimax design is the one with the smallest EN within those designs. Furthermore, we know that the EN of any admissible design must be smaller than the EN of the minimax design. And finally, the optimal design has a smaller EN than any admissible design. Therefore, we know that $\text{EN}(p_0)_{\text{optimal}} \leq \text{EN}(p_0)_{\text{admissible}} \leq \text{EN}(p_0)_{\text{minimax}}$. From (4), we know that for all designs $n_1 < \text{EN}(p_0)$. Combining both inequalities, we know that n_1 for the admissible design must be smaller than $\text{EN}(p_0)$ for the minimax design and that n_1 for the optimal design must be smaller than $\text{EN}(p_0)$ for the admissible design. Therefore, the algorithm searches over n_1 up to an end value of $n-1$ until the first design is found. The end value for n_1 is then replaced by $\text{EN}(p_0)$ of the first design. Whenever another design is found, the end value for n_1 is replaced with the smallest EN found so far.

2.4 Nonstochastic and stochastic curtailment

Based on the CP, nonstochastic and stochastic curtailment procedures allow stopping a trial whenever it becomes evident that rejection of the null hypothesis is impossible or that the probability of rejecting the null hypothesis is lower than a prespecified cutoff value, respectively. Ayanlowo and Redden (2007) implemented stochastic curtailment rules in the second stage of Simon's design. Their approach was extended by Kunz and Kieser (2011) who implemented nonstochastic and stochastic curtailment rules in both stages of the trial. The CP of rejecting the null hypothesis under the alternative $H_a : p = p_1$ if the trial would continue until termination of the second stage and if k responses are observed for the first $\widetilde{n}$ patients, $0 \leq \widetilde{n} \leq n$, is given by

$$\text{CP}_{k,\widetilde{n}} = \begin{cases} 0 & \text{if } 0 \leq n_1 - \widetilde{n} < r_1 - k + 1 \\ & \text{or } r - k + 1 > n - \widetilde{n} \\ 1 - \left\{ B(n_1 - \widetilde{n}, r_1 - k, p_1) + \sum\limits_{x=r_1-k+1}^{\substack{\min \\ r-k \\ n_1-\widetilde{n}}} b(n_1 - \widetilde{n}, x, p_1) \times \right. & \\ \qquad \left. B(n - n_1, r - x - k, p_1) \right\} & \text{if } 1 \leq r_1 - k + 1 \leq n_1 - \widetilde{n} \\ 1 - B(n - \widetilde{n}, r - k, p_1) & \text{if } r_1 < k \leq r \\ 1 & \text{if } k > r \end{cases}$$

For given $\widetilde{n} \leq n$, let $\widetilde{k}$ denote the maximum number of responses with $\text{CP}_{\widetilde{k},\widetilde{n}} = 0$. If $\widetilde{k}$ does not exist, the probability of early termination under the null hypothesis with $\widetilde{n}$ patients when applying nonstochastic curtailment $\text{PET}_{\widetilde{n}}(p0)$ is zero. If $\widetilde{k}$ exists, $\text{PET}_{\widetilde{n}}(p0)$ is given by

$$\text{PET}_{\widetilde{n}}(p_0) = \begin{cases} p_0^{\widetilde{k}} \times (1-p_0)^{\widetilde{n}-\widetilde{k}} \times \left\{ \sum_{x=0}^{\widetilde{k}} \binom{\widetilde{n}}{x} \times (-1)^{\widetilde{k}-x} \right\} & \text{if } \widetilde{n} \leq n_1 \\ p_0^{\widetilde{k}} \times (1-p_0)^{\widetilde{n}-\widetilde{k}} \times & \\ \quad \left\{ \sum_{x=r_1+1}^{\min(\widetilde{k},n_1)} \binom{n_1}{x} \times \sum_{y=0}^{\widetilde{k}-x} \binom{\widetilde{n}-n_1}{y} \times (-1)^{\widetilde{k}-x-y} \right\} & \text{if } \widetilde{n} > n_1 \end{cases}$$

The expected sample size under the null hypothesis is given by

$$\text{EN}(p_0) = \alpha' \times n + \sum_{\widetilde{n}=0}^{n} \text{PET}_{\widetilde{n}}(p_0) \times \widetilde{n}$$

where α' denotes the attained level of significance of the uncurtailed design.

While Simon's minimax and optimal designs and Jung's admissible designs are based on only two decision rules (r_1/n_1 and r/n), curtailed designs have $r+1$ decision rules, which can be determined as follows: for every $\widetilde{s}$ in the range of $[0, r]$ we search over m in the range of $[0, n]$ until the CP is less than a prespecified cutoff value c. The decision rule is given by $\widetilde{s}/\widetilde{m}$, where $\widetilde{m}$ denotes the smallest m for which $\text{CP}_{\widetilde{s},m} < c$.

3 The simontwostage command

The `simontwostage` command calculates the critical values and sample sizes for phase II oncology trials based on Simon's two-stage design. It allows choice among the minimax, optimal, and admissible designs—with or without curtailment rules. A scatterplot of those designs similar to the scatterplots of Jung et al. (2004) can be generated.

3.1 Syntax

```
simontwostage [ , p0(#) p1(#) alpha(#) beta(#)
   design(minimax|optimal|admissible|all) minn(#) maxn(#) plot
   curtailment(#| all [ , reps(#) ]) weight(#) ]
```

3.2 Options

`p0(#)` specifies the response rate under H_0. The response rate is a value between 0 and 1. The default is `p0(0.1)`.

`p1(#)` specifies the response rate under H_a. The response rate is a value between 0 and 1. Furthermore, `p1()` must be greater than `p0()`. The default is `p1(0.3)`.

`alpha(#)` specifies the significance level. The default is `alpha(0.05)`.

`beta(#)` specifies the limit for the type II error rate. The default is `beta(0.05)`.

`design(minimax|optimal|admissible|all)` specifies the design for which the sample size is determined. Simon's (1989) minimax and optimal designs can be calculated using `design(minimax)` and `design(optimal)`, respectively. Admissible design (Jung et al. 2004) can be obtained using `design(admissible)`. A table containing the results of all three designs can be generated using `design(all)`.

`minn(#)` specifies the start value for the total sample size. The default is `minn(-1)`. In this case, the minimal possible sample size will be calculated as described above.

`maxn(#)` specifies the end value for the total sample size. The default is `maxn(-1)`. In this case, an end value for the total sample size is calculated.

`plot` generates a scatterplot based on the approach of Jung et al. (2004). This option may be specified only if `design(all)` is also specified.

`curtailment(#|all [, reps(#)])` specifies the cutpoint for the CP. If `curtailment(0)` is specified, nonstochastic curtailment rules will be implemented in both stages of the trial. Type I error, type II error, EN ($p0$), and PET (p_0) for the end of the first stage and for every decision rule are calculated based on Kunz and Kieser (2011). Therefore, `reps()` could not be specified. For all other values of the cutpoint, stochastic curtailment rules will be implemented in both stages of the trial, and the values are determined using simulation studies. The number of replications can be specified using the `reps()` option; the default is `reps(1000)`. The decision rules are saved in a matrix. For the minimax, optimal, and admissible designs, the results are saved in `r(M)`, `r(O)`, and `r(A)`, respectively. If `curtailment(all [, reps(#)])` is specified, all cutpoints from 0 to 1 in steps of 0.01 will be implemented. Results will be shown only for every fifth step, but all results will be saved in the matrix `r(SC)`. Additionally, four plots will be generated. The plots show the results for the EN, the probability of early termination at the end of the first stage, the type I error, and the power, depending on the cutpoint. Gray bands show 95% confidence intervals (CIs).

`weight(#)` specifies the weight for the admissible design by Jung et al. (2004). The default is `weight(0.5)`.

3.3 Saved results

`simontwostage` saves the following in `r()`:

Scalars

`r(minn)`	minimum for total sample size as finally used by Mata
`r(maxn)`	maximum for total sample size as finally used by Mata
minimax	
`r(r1_m)`	critical value for the first stage
`r(n1_m)`	sample size for the first stage
`r(r_m)`	critical value for the end of the trial
`r(n_m)`	total sample size
`r(en0_m)`	EN($p0$)
`r(pet0_m)`	PET($p0$) at the end of the first stage
`r(tie_m)`	type I error rate
`r(tiie_m)`	type II error rate
optimal	
`r(r1_o)`	critical value for the first stage
`r(n1_o)`	sample size for the first stage
`r(r_o)`	critical value for the end of the trial
`r(n_o)`	total sample size
`r(en0_o)`	EN($p0$)
`r(pet0_o)`	PET($p0$) at the end of the first stage
`r(tie_o)`	type I error rate
`r(tiie_o)`	type II error rate
admissible	
`r(r1_a)`	critical value for the first stage
`r(n1_a)`	sample size for the first stage
`r(r_a)`	critical value for the end of the trial
`r(n_a)`	total sample size
`r(en0_a)`	EN($p0$)
`r(pet0_a)`	PET($p0$) at the end of the first stage
`r(tie_a)`	type I error rate
`r(tiie_a)`	type II error rate
if option `curtailment()` is specified	
`r(cp_min)`	minimal possible CP for the specified design
`r(en_nsc)`	EN($p0$) for the nonstochastic curtailed design
`r(en_sc)`	EN($p0$) for the stochastic curtailed design
`r(en_sc_l)`	lower boundary of 95% CI for EN($p0$) for the stochastic curtailed design
`r(en_sc_u)`	upper boundary of 95% CI for EN($p0$) for the stochastic curtailed design
`r(pet_sc)`	PET($p0$) at the end of the first stage for the stochastic curtailed design
`r(pet_sc_l)`	lower boundary of 95% CI for PET($p0$) at the end of the first stage for the stochastic curtailed design
`r(pet_sc_u)`	upper boundary of 95% CI for PET($p0$) at the end of the first stage for the stochastic curtailed design
`r(tie_sc)`	type I error rate for the stochastic curtailed design
`r(tie_sc_l)`	lower boundary of 95% CI for type I error rate for the stochastic curtailed design
`r(tie_sc_u)`	upper boundary of 95% CI for type I error rate for the stochastic curtailed design
`r(tiie_sc)`	type II error rate for the stochastic curtailed design
`r(tiie_sc_l)`	lower boundary of 95% CI for type II error rate for the stochastic curtailed design
`r(tiie_sc_u)`	upper boundary of 95% CI for type II error rate for the stochastic curtailed design

Matrices

`r(M)`	stopping rules and PET($p0$) for the minimax design
`r(O)`	stopping rules and PET($p0$) for the optimal design
`r(A)`	stopping rules and PET($p0$) for the admissible design
`r(C)`	CP for every possible outcome that can be observed during the study
`r(R)`	critical values, sample sizes, EN, PET($p0$), type I error, type II error, interval for weight
`r(SC)`	cutpoint for CP, EN, PET($p0$), type I error, type II error, lower and upper limits for the 95% CI

4 Use of simontwostage

In this section, the use of `simontwostage` will be explained. Let us consider that we want to plan a phase II trial in oncology. We assume that the response rate under H_0 is $p_0 \leq 0.05$. The new treatment is considered efficient if the response rate under H_a is $p_1 > 0.25$. Furthermore, we choose a significance level of $\alpha = 0.05$ and a power of $1 - \beta = 0.90$. As a first step, we use the following command:

```
. simontwostage, p0(0.05) p1(0.25) alpha(0.05) beta(0.10) design(all) plot
```

```
Simon´s Two Stage Design
```

```
This algorithm is searching for the minimax, optimal, and admissible design
for the following parameter specifications

Type I error = 0.05
Power        = 0.90
H0: p <= 0.05
H1: p >  0.25
NOTE: the program searches many designs
It can take a while to find the best design
An occassional design is displayed
```

Design	type I	type II	PET(p0)	EN(p0)
(0/15,3/25)	0.034	0.099	0.463	20.367
(0/12,3/26)	0.036	0.095	0.540	18.435
(0/11,3/27)	0.040	0.092	0.569	17.899
(0/10,3/28)	0.043	0.094	0.599	17.223
(0/9,3/30)	0.049	0.098	0.630	16.765

	r1/n1	r/n	EN_0	PET_0	t_I_e*	t_II_e*	weight
minimax	0/15	3/25	20.37	0.463	0.0336	0.0992	[0.659;1.000]
admissible	0/12	3/26	18.43	0.540	0.0365	0.0953	[0.377;0.659]
not admissible	0/11	3/27	17.90	0.569	0.0398	0.0917	[-;-]
admissible	0/10	3/28	17.22	0.599	0.0426	0.0942	[0.186;0.377]
optimal	0/ 9	3/30	16.76	0.630	0.0489	0.0981	[0.000;0.186]

```
* t_I_e:  type I error
  t_II_e: type II error
```

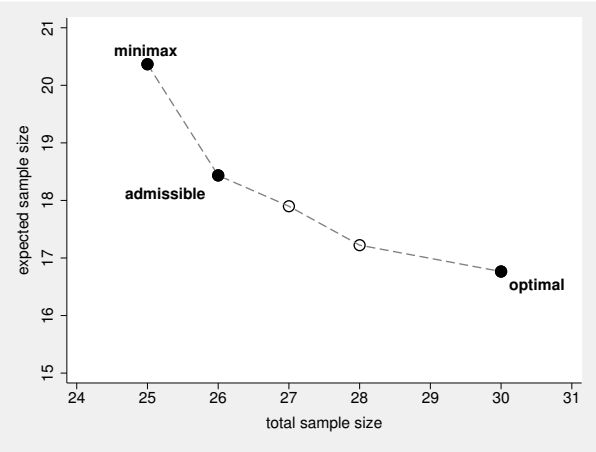

Figure 2. Scatterplot based on Jung et al. (2004) generated by the `plot` option

As can be seen in the output, five designs fulfilling the prespecified constraints for p_0, p_1, α, and β are found. The first one is Simon's minimax design, the second and the fourth ones are admissible designs according to Jung et al. (2004), and the last one is Simon's optimal design (see figure 2). The third design is neither minimax nor admissible nor optimal.

The first column gives the critical value and the sample size for the first stage for each design. For example, the critical value and the sample size for the first stage for the minimax design are 0 and 15, respectively. The second column gives the critical value and the sample size for the whole trial. For the minimax designs, these are 3 and 25. The third and the fourth columns give the EN under H_0 [see (4)] and the probability of early termination under H_0 [see (3)]. The next two columns give the actual type I and type II errors [see (1) and (2)]. And the last column gives the range for the weight q as defined by Jung et al. (2004). In the plot, the first admissible design with $n = 26$ and $\text{EN}(p_0) = 18.43$ is marked. The default value for the `weight()` option is 0.5, which is in the interval of $[0.377, 0.659]$. Because no other value was specified in the command, the corresponding admissible design is marked in the plot.

Let us assume that we want to use the second admissible design with $n = 28$ and $\text{EN}(p_0) = 17.22$ for our study. We therefore have to specify the `weight()` option with any value between $[0.186, 0.377]$. In the next step, we use the following command to determine the effect of implementing curtailment rules (also see figure 3):

```
. simontwostage, p0(0.05) p1(0.25) alpha(0.05) beta(0.10) design(admissible)
> weight(0.3) curtailment(all, reps(10000))
```

Simon´s Two Stage Design

This algorithm is searching for the admissible design
for the following parameter specifications

Type I error = 0.05
Power   = 0.90
H0: p <= 0.05
H1: p > 0.25
NOTE: the program searches many designs
It can take a while to find the best design
An occassional design is displayed

Design	type I	type II	PET(p0)	EN(p0)
(0/15,3/25)	0.034	0.099	0.463	20.367
(0/12,3/26)	0.036	0.095	0.540	18.435
(0/11,3/27)	0.040	0.092	0.569	17.899
(0/10,3/28)	0.043	0.094	0.599	17.223
(0/9,3/30)	0.049	0.098	0.630	16.765

admissible	r1/n1	r/n	EN(p0)	PET(p0)	t_I_e	t_II_e
original	0/10	3/28	17.22	0.5987	0.0426	0.0942
NSC			16.80	0.5987	0.0426	0.0942
SC(0.05)			16.53	0.6060	0.0406	0.0982
SC(0.10)			16.39	0.5951	0.0449	0.0947
SC(0.15)			16.09	0.6033	0.0421	0.0948
SC(0.20)			15.72	0.6048	0.0389	0.1013
SC(0.25)			14.58	0.6293	0.0398	0.1146
SC(0.30)			14.47	0.6238	0.0362	0.1162
SC(0.35)			14.09	0.6357	0.0349	0.1231
SC(0.40)			12.69	0.6689	0.0307	0.1481
SC(0.45)			12.69	0.6640	0.0327	0.1526
SC(0.50)			12.49	0.6566	0.0295	0.1609
SC(0.55)			11.11	0.6961	0.0282	0.1832
SC(0.60)			10.84	0.7016	0.0225	0.1996
SC(0.65)			9.59	0.7278	0.0237	0.2378
SC(0.70)			7.97	0.7734	0.0164	0.3082
SC(0.75)			7.82	0.7700	0.0159	0.3149
SC(0.80)			6.28	0.8137	0.0121	0.4000
SC(0.85)			4.53	0.8609	0.0069	0.5063
SC(0.90)			1.54	0.9835	0.0018	0.7998
SC(0.95)			0.00	1.0000	0.0000	1.0000
SC(1.00)			0.00	1.0000	0.0000	1.0000

* t_I_e: type I error
  t_II_e: type II error
  Based on 10000 replications

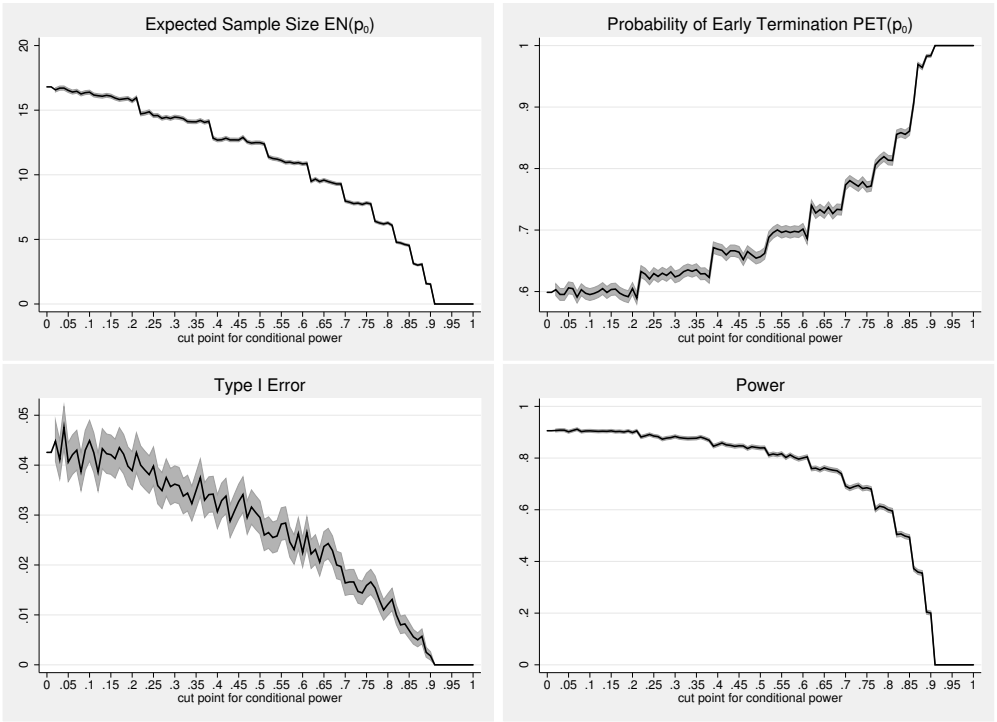

Figure 3. Plots based on 10,000 replications generated with the `curtailment(all, reps(10000))` option. Gray areas mark 95% CIs.

The first row in the output table gives the results for the original uncurtailed design with EN $(p_0) = 17.22$ and a type II error of $\beta' = 0.942$. The next row, labeled NSC, gives the values for the nonstochastic curtailed design with EN $(p_0) = 16.80$. The type II error for this design is exactly the same as for the original uncurtailed design. The next rows correspond to stochastic curtailed designs with cutoff values for the CP of 0.05 up to 1.00 in steps of 0.05. As described above, results are calculated for cutpoints from 0 to 1 in steps of 0.01 but are only shown for every fifth cutpoint. All other results can be obtained by typing

```
. matrix list r(SC)
```
(*output omitted*)

We decide to take a cutpoint of 0.20 because the EN is now EN $(p_0) = 15.72$ (95% CI $[15.58, 15.86]$) compared with EN $(p_0) = 17.22$ for the uncurtailed design. The type II error is now $\beta' = 0.1013$ (95% CI $[0.0954, 0.1072]$), and therefore is higher compared with $\beta' = 0.0942$ for the uncurtailed design, but we regard it as acceptable. Finally, we want to get the new decision rules for the stochastic curtailed design with a cutpoint for the CP of 0.20:

```
. simontwostage, p0(0.05) p1(0.25) alpha(0.05) beta(0.10) design(admissible)
> weight(0.3) curtailment(0.20, reps(10000))

Simon´s Two Stage Design

This algorithm is searching for the admissible design
for the following parameter specifications
Type I error = 0.05
Power        = 0.90
H0: p <= 0.05
H1: p >  0.25
NOTE: the program searches many designs
It can take a while to find the best design
An occassional design is displayed
Design          type I    type II  PET(p0)   EN(p0)
(0/15,3/25)     0.034     0.099    0.463     20.367
(0/12,3/26)     0.036     0.095    0.540     18.435
(0/11,3/27)     0.040     0.092    0.569     17.899
(0/10,3/28)     0.043     0.094    0.599     17.223
(0/9,3/30)      0.049     0.098    0.630     16.765

admissible  r1/n1   r/n     EN(p0)       PET(p0)      t_I_e        t_II_e
original     0/10   3/28    17.22        0.5987       0.0426       0.0942
SC(0.20)    mat list r(A)   15.78        0.6017       0.0377       0.0992
95%-CI                      [15.6;15.9]  [0.59;0.61]  [0.03;0.04]  [0.09;0.11]

* t_I_e:  type I error
  t_II_e: type II error
  Based on 10000 replications
```

Using `matrix list r(A)`, we obtain the new decision rules. We now have four decision rules—(0/10), (1/22), (2/25), and (3/28)—instead of the two decision rules for the uncurtailed admissible design. But we also have a smaller EN, and furthermore, the curtailed design allows for early stopping after the end of the first stage but before the end of the second stage. For example, the cumulative probability of early termination if the treatment is ineffective with 22 enrolled patients is $0.5913 + 0.1737 = 0.765$.

```
. matrix list r(A)
r(A)[4,3]
        r      n    PET
r1      0     10  .5913
r2      1     22  .1737
r3      2     25  .1276
r4      3     28  .0613
```

5 Summary

This article described a new Stata command for calculating the critical values and sample sizes for Simon's minimax, Simon's optimal, and Jung's admissible designs. The command also implements nonstochastic and stochastic curtailment rules in both stages of the trial and allows comparison of the resulting designs with those of the uncurtailed designs.

6 Acknowledgments

We would like to thank Adrian Mander, Ulrich Abel, and Tom Palmer for their helpful suggestions.

7 References

Ayanlowo, A. O., and D. T. Redden. 2007. Stochastically curtailed phase II clinical trials. *Statistics in Medicine* 26: 1462–1472.

Bryant, J., and R. Day. 1995. Incorporating toxicity considerations into the design of two-stage phase II clinical trials. *Biometrics* 51: 1372–1383.

Chang, M. N., T. M. Therneau, H. S. Wieand, and S. S. Cha. 1987. Designs for group sequential phase II clinical trials. *Biometrics* 43: 865–874.

Chen, T. T. 1997. Optimal three-stage designs for phase II cancer clinical trials. *Statistics in Medicine* 16: 2701–2711.

Green, S. J., and S. Dahlberg. 1992. Planned versus attained design in phase II clinical trials. *Statistics in Medicine* 11: 853–862.

Jung, S.-H., T. Lee, K. Kim, and S. L. George. 2004. Admissible two-stage designs for phase II cancer clinical trials. *Statistics in Medicine* 23: 561–569.

Kunz, C. U., and M. Kieser. 2011. Non-stochastic and stochastic curtailment in phase II oncology trials. Working paper.

———. Forthcoming. Optimal two-stage designs for single-arm phase II oncology trials with two binary endpoints. *Methods of Information in Medicine*. http://www.ncbi.nlm.nih.gov/pubmed/21057719.

Lin, S. P., and T. T. Chen. 2000. Optimal two-stage designs for phase II clinical trials with differentiation of complete and partial responses. *Communications in Statistics—Theory and Methods* 29: 923–940.

Lin, X., R. Allred, and G. Andrews. 2008. A two-stage phase II trial design utilizing both primary and secondary endpoints. *Pharmaceutical Statistics* 7: 88–92.

Lin, Y., and W. J. Shih. 2004. Adaptive two-stage designs for single-arm phase IIA cancer clinical trials. *Biometrics* 60: 482–490.

Lu, Y., H. Jin, and K. R. Lamborn. 2005. A design of phase II cancer trials using total and complete response endpoints. *Statistics in Medicine* 24: 3155–3170.

Mander, A. 2009. simon2stage: Stata module to compute the design parameters for a Simon two-stage design. Statistical Software Components S457081, Department of Economics, Boston College. http://ideas.repec.org/c/boc/bocode/s457081.html.

Mander, A. P., and S. G. Thompson. 2010. Two-stage designs optimal under the alternative hypothesis for phase II cancer clinical trials. *Contemporary Clinical Trials* 31: 572–578.

Panageas, K. S., A. Smith, M. Gönen, and P. B. Chapman. 2002. An optimal two-stage phase II design utilizing complete and partial response information separately. *Controlled Clinical Trials* 23: 367–379.

Shuster, J. 2002. Optimal two-stage designs for single arm phase II cancer trials. *Journal of Biopharmaceutical Statistics* 12: 39–51.

Simon, R. 1989. Optimal two-stage designs for phase II clinical trials. *Controlled Clinical Trials* 10: 1–10.

About the authors

Cornelia Ursula Kunz is studying for a PhD in medical biometry and has currently submitted her PhD thesis, "Two-Stage Designs for Phase-II-Trials with One or Two Endpoints".

Meinhard Kieser is the director of the Institute of Medical Biometry and Informatics, Heidelberg University, Germany.