Smoothing Corrections for Improving Sample Size Recalculation Rules in Adaptive Group Sequential Study Designs

Carolin Herrmann1  Geraldine Rauch1

1 Charité—Universitätsmedizin Berlin, Corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Institute of Biometry and Clinical Epidemiology, Berlin, Germany

Address for correspondence Carolin Herrmann, MSc, Charité—Universitätsmedizin Berlin, Corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Institute of Biometry and Clinical Epidemiology, Charitéplatz 1, Berlin 10117, Germany (e-mail: carolin.herrmann@charite.de).

Abstract

Background    An adequate sample size calculation is essential for designing a successful clinical trial. One way to tackle planning difficulties regarding parameter assumptions required for sample size calculation is to adapt the sample size during the ongoing trial. This can be attained by adaptive group sequential study designs. At a predefined timepoint, the interim effect is tested for significance. Based on the interim test result, the trial is either stopped or continued with the possibility of a sample size recalculation.

Objectives    Sample size recalculation rules have different limitations in application like a high variability of the recalculated sample size. Hence, the goal is to provide a tool to counteract this performance limitation.

Methods     Sample size recalculation rules can be interpreted as functions of the observed interim effect. Often, a "jump" from the first stage’s sample size to the maximal sample size at a rather arbitrarily chosen interim effect size is implemented and the curve decreases monotonically afterwards. This jump is one reason for a high variability of the sample size. In this work, we investigate how the shape of the recalculation function can be improved by implementing a smoother increase of the sample size. The design options are evaluated by means of Monte Carlo simulations. Evaluation criteria are univariate performance measures such as the conditional power and sample size as well as a conditional performance score which combines these components.

Results    We demonstrate that smoothing corrections can reduce variability in conditional power and sample size as well as they increase the performance with respect to a recently published conditional performance score for medium and large standardized effect sizes.

Conclusion    Based on the simulation study, we present a tool that is easily implemented to improve sample size recalculation rules. The approach can be combined with existing sample size recalculation rules described in the literature.
Introduction

A reliable sample size calculation is an important determinant for the success of a clinical trial. However, even with thorough literature research and solid medical expertise, it is not always possible to make reasonable planning assumptions. Thus, changing the sample size during an ongoing trial seems appealing to incorporate evidence from the data collected so far. This is the key idea of adaptive study designs with unblinded sample size recalculation.1,2 Sample size recalculation rules can be interpreted as functions of the observed interim effect. Current rules proposed in the literature often suffer from a high variability of the recalculated sample size,3 which is a random variable. Recalculation rules have in common that they take values between the first stage’s sample size $n_1$ (no additional sample size) and a predefined maximal sample size $n_{\text{max}}$. Usually, a single “jump” from $n_1$ to $n_{\text{max}}$ is implemented. However, a medical researcher would probably not understand that for an observed interim effect of 0.22 the study must be stopped early for futility and for an observed effect of 0.23 the study continues with $n_{\text{max}}$. This “jump” is one reason for a high variability of the recalculated sample size.

Objectives

In this work, we investigate how the shape of the recalculation function can be improved by a smoother increase of the sample size. These smoothing corrections are evaluated by means of Monte-Carlo simulations. Performance indicators are the conditional power and sample size of the second stage as well as a conditional performance score which incorporates sample size and power components.

Methods

Test Problem and Trial Design

We consider a 1:1 randomized, controlled clinical trial with a normally distributed primary end point with means $\mu_i$ in the control group, $\mu_c$ in the intervention group, and known variance $\sigma^2$ for both groups. The hypotheses for the one-sided test problem are formulated as

$$H_0 : \mu_I - \mu_C \leq 0 \quad \text{and} \quad H_1 : \mu_I - \mu_C > 0.$$ 

The study is conducted as a two-stage adaptive group-sequential clinical trial design with $n_1$ patients per group for the first stage and a function $n_2(\cdot) \leq n_{\text{max}} - n_1$ for the number of patients per group for the second stage, where $n_{\text{max}}$ is the maximal total sample size. The null hypothesis is tested with the common Z-test, where $Z_1$ defines the interim test statistic and the test statistic for the final analysis $Z_{1,2}$ is obtained by means of the inverse normal combination test.4 If $Z_1 \geq c_{\text{eff}}$ or $Z_{1,2} \geq c_{\text{final}}$, the null hypothesis is rejected. The trial is continued to the second stage if $Z_1$ falls within the so-called “recalculation area” $c_{\text{fit}} \leq Z_1 < C_{\text{eff}}$, where $c_{\text{fit}}$ is the futility stopping boundary, $c_{\text{eff}}$ is the futility stopping boundary after the first, and $c_{\text{final}}$ after the second stage.

Sample Size Recalculation

There exist many ways of recalculating the sample size at the interim analysis. Most established recalculation rules are based on conditional power arguments.2,5 For illustrative purposes, we focus here on the commonly used “restricted observed conditional power approach.” At the interim analysis, the sample size for the second stage is calculated such that a predefined conditional power value $1 - \beta$ can be reached. Thereby, the observed interim effect is used as an estimator for the true underlying effect. If the recalculated sample size exceeds the maximally feasible sample size $n_{\text{max}}$, the sample size is restricted to $n_{\text{max}}$. Furthermore, a minimal conditional power $1 - \beta_{\text{min}}$ must be reached to justify the increase to $n_{\text{max}}$. This results in a sample size recalculation function starting with a plateau at $n_1$ (no increase), then jumping to a plateau at $n_{\text{max}}$, and then decreasing monotonically.

Smoothing Correction

To reduce the variability of the recalculated sample size, we propose a smoothing correction to increase the sample size from $n_1$ to $n_{\text{max}}$ within the interval $[c_{\text{fit}} : c_{\text{max}}]$, where $c_{\text{max}} < c_{\text{eff}}$ is the smallest interim test statistic suggesting $n_{\text{max}}$ according to the selected sample size recalculation rule.

We consider five classes of simple smoothing functions to do so, as graphically illustrated in –Fig. 1 and described mathematically in –Appendix A:

- A linear increase,
- A stepwise increase,
- A sigmoid increase,
- A concave increase and
- A convex increase.

Note that these five function classes represent different general approaches for smoothing of which we aim to identify the most promising. We do not aim at optimizing a specific function shape within this work.

Performance Evaluation

Whereas in a one-stage design, the performance measures are simply given by power and sample size, in an adaptive design, both the conditional power and the second stage sample size are random variables. A good performance is therefore given if the average conditional power meets its target, the average sample size is neither too high nor too low, and the corresponding variances are reasonably small. Recently, Herrmann et al.6 proposed a conditional performance score $CPS$ averaging these indicators (location of conditional power and sample size, variation of conditional power and sample size) within a single performance measure. The location components are constructed as follows

$$\text{location}_X = 1 - \frac{E[X] - X_{\text{target}}}{(X_{\text{max}} - X_{\text{min}})},$$

where $X$ refers either to conditional power or sample size. The expectation can be estimated via the corresponding average, all other values in the formula are fixed quantities. From the location formula, it can be seen that the idea is to compare the expected value $E[X]$ to a predefined target value $X_{\text{target}}$ in relation to the
the maximally possible variance. Here, observed variance is seen in relation to components are formulated as performance. between 0 and 1 and higher values refer to a better well as its components are constructed such that they range values of the interim test statistics. Based on this set of variations were performed with the software R.

Simulation Setup
We conduct a Monte-Carlo simulation study to assess the potential performance improvement when adding the smoothing corrections presented above. We set the global one-sided significance level to \( \alpha = 0.025 \) and the binding futility stopping bound to \( c_{fu} = 0.0 \). The inverse normal parameter 0.25.

For each scenario, we draw 10,000 replications from a normal distribution \( N(\Delta/\sqrt{n}, 2, 1) \) to generate the observed values of the interim test statistics. Based on this set of observed values, we recalculated the sample size according to the restricted observed conditional power approach with \( 1 - \beta_{min} = 0.6 \) and an anticipated conditional power \( 1 - \beta = 0.8 \) combined with and without the smoothing corrections presented above. We evaluated the scenarios with respect to average conditional power and sample size, their variances, and the conditional performance score. Simulations were performed with the software R.

Results
Within this section, we discuss all five classes of smoothing functions in all five scenarios presented above. For the sake of readability, we restrict our tabulated results to Scenario 1 in the main manuscript (Table 1). The tabulated results for Scenarios 2 to 5 can be found in the -Supplementary Tables S1 to S4 (available online only).

The results of Scenario 1 show that all smoothing corrections result in slightly larger average sample sizes, as the smoothing correction implies an increase in sample size (cf. -Table 1 Column 3). It can be seen that stepwise...
Table 1 Estimated pointwise conditional performance score and related conditional performance measures with $n_1 = 50$, $n_{\text{max}} = 200$, $\alpha = 0.025$, $c_{\text{final}} = 1.973$, $c_{\text{decr}} = 1.332$, $c_{\text{off}} = 2.790$, $c_{\text{inccr}} = 1.116$ (multiplicity adjustment according to O'Brien and Fleming⁷) and weights $1/\sqrt{2}$ for the inverse normal combination test (Scenario 1)

| $\Delta$ | Smoothing | Average sample size second stage | Variance of sample size second stage | Average conditional power | Variance of conditional power | Conditional performance score⁸ |
|----------|------------|---------------------------------|-------------------------------------|---------------------------|-------------------------------|---------------------------------|
| 0.0      | Without    | 75.873                          | 2,500.657                           | 0.204                     | 0.116                         | 0.574                           |
|          | Linear     | 126.114                         | 2,117.314                           | 0.292                     | 0.101                         | 0.493                           |
|          | Stepwise   | 108.195                         | 2,377.465                           | 0.274                     | 0.101                         | 0.518                           |
|          | Sigmoid    | 117.582                         | 3,403.994                           | 0.295                     | 0.102                         | 0.464                           |
|          | Concave    | 144.750                         | 2,143.337                           | 0.305                     | 0.099                         | 0.459                           |
|          | Convex     | 107.477                         | 2,520.527                           | 0.278                     | 0.103                         | 0.511                           |
| 0.1      | Without    | 83.687                          | 2,806.324                           | 0.300                     | 0.144                         | 0.507                           |
|          | Linear     | 128.908                         | 2,152.401                           | 0.385                     | 0.114                         | 0.453                           |
|          | Stepwise   | 113.744                         | 2,343.830                           | 0.368                     | 0.117                         | 0.473                           |
|          | Sigmoid    | 122.781                         | 3,247.434                           | 0.388                     | 0.115                         | 0.426                           |
|          | Concave    | 144.544                         | 2,239.472                           | 0.397                     | 0.110                         | 0.423                           |
|          | Convex     | 113.273                         | 2,523.429                           | 0.373                     | 0.119                         | 0.466                           |
| 0.2      | Without    | 89.223                          | 2,803.946                           | 0.407                     | 0.153                         | 0.464                           |
|          | Linear     | 128.110                         | 2,241.270                           | 0.486                     | 0.111                         | 0.427                           |
|          | Stepwise   | 115.975                         | 2,236.044                           | 0.470                     | 0.116                         | 0.448                           |
|          | Sigmoid    | 124.451                         | 3,059.403                           | 0.490                     | 0.111                         | 0.406                           |
|          | Concave    | 140.330                         | 2,479.820                           | 0.497                     | 0.106                         | 0.400                           |
|          | Convex     | 115.890                         | 2,444.768                           | 0.475                     | 0.117                         | 0.439                           |
| 0.3      | Without    | 93.038                          | 2,645.027                           | 0.522                     | 0.139                         | 0.432                           |
|          | Linear     | 122.799                         | 2,413.852                           | 0.588                     | 0.090                         | 0.543                           |
|          | Stepwise   | 113.843                         | 2,223.251                           | 0.575                     | 0.097                         | 0.526                           |
|          | Sigmoid    | 121.177                         | 2,948.012                           | 0.591                     | 0.089                         | 0.524                           |
|          | Concave    | 131.118                         | 2,766.602                           | 0.596                     | 0.085                         | 0.551                           |
|          | Convex     | 114.480                         | 2,419.475                           | 0.580                     | 0.096                         | 0.522                           |
| 0.4      | Without    | 92.842                          | 2,289.227                           | 0.622                     | 0.106                         | 0.620                           |
|          | Linear     | 112.624                         | 2,390.235                           | 0.667                     | 0.064                         | 0.655                           |
|          | Stepwise   | 106.905                         | 2,122.581                           | 0.659                     | 0.070                         | 0.666                           |
|          | Sigmoid    | 111.960                         | 2,707.405                           | 0.670                     | 0.063                         | 0.648                           |
|          | Concave    | 117.824                         | 2,761.486                           | 0.673                     | 0.060                         | 0.640                           |
|          | Convex     | 107.424                         | 2,273.953                           | 0.662                     | 0.069                         | 0.662                           |
| 0.5      | Without    | 88.376                          | 1,870.067                           | 0.694                     | 0.070                         | 0.656                           |
|          | Linear     | 101.010                         | 2,181.853                           | 0.724                     | 0.039                         | 0.665                           |
|          | Stepwise   | 97.490                          | 1,896.476                           | 0.718                     | 0.043                         | 0.675                           |
|          | Sigmoid    | 100.912                         | 2,364.080                           | 0.726                     | 0.038                         | 0.661                           |
|          | Concave    | 104.038                         | 2,502.587                           | 0.727                     | 0.036                         | 0.654                           |
|          | Convex     | 97.983                          | 2,022.049                           | 0.721                     | 0.042                         | 0.671                           |
| 0.6      | Without    | 83.047                          | 1,495.924                           | 0.740                     | 0.042                         | 0.689                           |
|          | Linear     | 90.506                          | 1,833.702                           | 0.759                     | 0.021                         | 0.699                           |
|          | Stepwise   | 88.523                          | 1,609.744                           | 0.756                     | 0.024                         | 0.705                           |
|          | Sigmoid    | 90.556                          | 1,944.184                           | 0.760                     | 0.020                         | 0.696                           |
|          | Concave    | 92.056                          | 2,039.396                           | 0.761                     | 0.019                         | 0.692                           |
Smoothing reduces the variability in sample size for small and medium true effect sizes ($\Delta = 0.0 - 0.4$, cf. Table 1 Column 4). Within that effect range, often the linear, concave, and convex smoothing lead to a reduction in variability in sample size as well. In contrast, the sigmoid smoothing approach, however, adds to an increase in variance in sample size for all considered effect sizes. The reason is that for a null effect or very small effects, the target second stage sample size within the conditional performance score is 0 and the optimal conditional power is given by the local significance level. Thus, any smoothing correction that increases the sample size has a negative impact on the score. If the interim effect is larger, the target second stage sample size is different from 0 and the target conditional power is $1 - \beta$. Here, the reduced variability caused by smoothing as well as the increased conditional power has a positive impact on the score. However, for larger effect sizes, that improvement becomes smaller or is no longer apparent ($\Delta = 1.0$). This is due to the fact that the target sample size of the score is

| $\Delta$ | Smoothing | Average sample size second stage | Variance of sample size second stage | Average conditional power | Variance of conditional power | Conditional performance score
|----------|------------|----------------------------------|--------------------------------------|---------------------------|-------------------------------|-------------------------------|
| 0.7      | Without    | 78.031                           | 1,105.555                           | 0.770                     | 0.022                         | 0.735                         |
|          | Linear     | 81.977                           | 1,346.081                           | 0.781                     | 0.010                         | 0.744                         |
|          | Stepwise   | 81.023                           | 1,209.807                           | 0.779                     | 0.011                         | 0.747                         |
|          | Sigmoid    | 82.230                           | 1,414.946                           | 0.781                     | 0.009                         | 0.742                         |
|          | Concave    | 82.741                           | 1,477.070                           | 0.782                     | 0.009                         | 0.740                         |
|          |Convex      | 81.212                           | 1,256.071                           | 0.780                     | 0.011                         | 0.746                         |
| 0.8      | Without    | 73.463                           | 858.014                             | 0.790                     | 0.007                         | 0.779                         |
|          | Linear     | 74.731                           | 956.005                             | 0.794                     | 0.003                         | 0.788                         |
|          | Stepwise   | 74.404                           | 900.420                             | 0.793                     | 0.003                         | 0.789                         |
|          | Sigmoid    | 74.737                           | 969.232                             | 0.794                     | 0.003                         | 0.788                         |
|          | Concave    | 74.939                           | 992.560                             | 0.794                     | 0.003                         | 0.787                         |
|          |Convex      | 74.523                           | 931.230                             | 0.793                     | 0.003                         | 0.788                         |
| 0.9      | Without    | 67.541                           | 402.875                             | 0.794                     | 0.004                         | 0.820                         |
|          | Linear     | 68.238                           | 462.443                             | 0.796                     | 0.002                         | 0.825                         |
|          | Stepwise   | 67.999                           | 432.512                             | 0.796                     | 0.003                         | 0.826                         |
|          | Sigmoid    | 68.221                           | 476.467                             | 0.796                     | 0.002                         | 0.824                         |
|          | Concave    | 68.372                           | 483.860                             | 0.796                     | 0.002                         | 0.824                         |
|          |Convex      | 68.103                           | 446.870                             | 0.796                     | 0.002                         | 0.825                         |
| 1.0      | Without    | 64.458                           | 241.188                             | 0.793                     | 0.005                         | 0.831                         |
|          | Linear     | 65.316                           | 306.793                             | 0.795                     | 0.003                         | 0.832                         |
|          | Stepwise   | 65.271                           | 298.788                             | 0.795                     | 0.003                         | 0.833                         |
|          | Sigmoid    | 65.469                           | 337.396                             | 0.795                     | 0.003                         | 0.830                         |
|          | Concave    | 65.570                           | 361.005                             | 0.795                     | 0.003                         | 0.829                         |
|          |Convex      | 65.062                           | 268.475                             | 0.795                     | 0.003                         | 0.834                         |

$\Delta$, true standardized treatment effect.

*Conditional performance score with an equal weighting of the components and target values as suggested in Herrmann et al.6
smaller for higher effect sizes. The sampling probability to observe a small interim effect when the true effect is large intuitively becomes smaller and consequently the smoothing effect is less prominent. Among all smoothing corrections for Scenario 1, the stepwise approach usually outperforms the other four smoothing approaches with respect to variation in sample size and the overall conditional performance score over the range of different effect sizes (cf. \textit{Table 1} Columns 4 and 7). The convex approach shows a reasonable overall performance as well.

Similar results can be found in Scenarios 2 and 3 for the other two \( n_1 \) and \( n_{\text{max}} \) combinations. Again, the average sample size is increased for sample size recalculation with smoothing corrections (cf. \textit{Supplementary Tables S1} and S2 Column 3, available online only). The variance in sample size is reduced for the stepwise smoothing approach for effect sizes up to \( \Delta = 0.5 \) and the sigmoid smoothing approach records an increase in the variance in sample size for all considered effect sizes (cf. \textit{Supplementary Tables S1} and S2 Column 4, available online only). In line with Scenario 1, the average conditional power is increased and the variance of the conditional power is reduced when comparing sample size recalculation with and without smoothing corrections (cf. \textit{Supplementary Tables S1} and S2 Columns 5 and 6, available online only). Moreover, the conditional performance score declares a benefit for the smoothing corrections for effect sizes from \( \Delta = 0.4 \) since the target second stage sample size of the conditional performance score equals 0 for a broader effect size range than in Scenario 1 (cf. \textit{Supplementary Tables S1} and S2 Column 7, available online only). Among the smoothing approaches, the stepwise smoothing performs usually better than the other ones with respect to variability in sample size and the conditional performance score throughout the different effect sizes.

Scenarios 4 and 5 behave similarly. Again, we observe an increase in the average sample size (cf. \textit{Supplementary Tables S3} and S4 Column 3, available online only). Note that the different multiplicity adjustments have an impact on the width of the recalculation area with \( c_{\text{eff}} = 2.176 \) for the adjustment according to Pocock\textsuperscript{8} and \( c_{\text{eff}} = 2.420 \) for the adjustment according to Wang and Tsiatis\textsuperscript{9} while it was \( c_{\text{eff}} = 2.790 \) for O’Brien and Fleming.\textsuperscript{7} Thus, in these two Scenarios 4 and 5, the recalculation area becomes smaller compared with Scenarios 1 to 3 and as a consequence, the smoothing corrections also lead to a variance reduction in sample size for effect sizes above 0.4 (cf. \textit{Supplementary Tables S3} and S4 Column 4, available online only), in particular the stepwise approach. Same as in Scenario 1, we observe an increase in the average conditional power, a reduction in the variance of the conditional power, and a conditional score benefit for effect sizes from \( \Delta = 0.3 \) (cf. \textit{Supplementary Tables S3} and S4 Columns 5–7, available online only).

Throughout the different scenarios, all smoothing corrections result in a slightly larger average conditional power as well as they all reduce the variance in conditional power for all considered effect sizes. Moreover, it can be seen that the average sample size is increased. Sample size recalculation with a smoothing correction decreases the variance in sample size for a small selection of smoothing corrections and effect sizes up to \( \Delta = 0.5 \) with multiplicity adjustment according to O’Brien and Fleming.\textsuperscript{7} For multiplicity adjustments according to Pocock\textsuperscript{8} or the selected Wang and Tsiatis\textsuperscript{9} boundaries, the recalculation area becomes smaller and the variance in sample size is also decreased for higher effect sizes. For effect sizes below \( \Delta = 0.3 \) (for \( n_{\text{max}} = 200 \)) or \( \Delta = 0.4 \) (for \( n_{\text{max}} = 150 \)), the reduction of variance in sample size is outweighed by the increase in average sample size, which results in better conditional performance scores without smoothing correction. For larger effect sizes, all five smoothing corrections result usually in better conditional performance scores, but also for smaller effect sizes most smoothing corrections show a considerable benefit with respect to variance reduction in sample size. Among the smoothing corrections, overall, the stepwise smoothing correction turns out to be performing well or pointwise even best with respect to average sample size and the conditional performance score throughout the different effect sizes and scenarios.

Application of Smoothing Corrections to a Medical Example

To illustrate the presented methodology, we consider a clinical study example. Bowden and Mander discussed clinical trials where an adaption of the planned study design may become necessary as the assumed effect gained from a pilot study might correspond to a low level of evidence.\textsuperscript{11} As an example, they present a clinical trial scenario where the aim is to compare a new versus standard treatment with respect to the end point pain relief in osteoarthritis patients.\textsuperscript{11} Let us assume that pain relief is measured on the McGill pain scale\textsuperscript{12} ranging from 0 to 50 where higher values indicate a worse pain and we assume the endpoint to be normally distributed. We are interested in the difference between the two groups with respect to short-term pain reduction from baseline to 2 weeks of treatment. We assume that there exists a small pilot study which supports the superiority of the new intervention over the standard treatment with an observed standardized effect of 0.4, which should be confirmed now. Therefore, we formulate the hypotheses as

\[
H_0: \mu_I - \mu_C \leq 0 \quad \text{and} \quad H_1: \mu_I - \mu_C > 0,
\]

where \( \mu_I := \mu_{I, \text{baseline}} - \mu_{I, \text{2 weeks}} \) refers to the pain reduction within the 2 weeks in the intervention group and \( \mu_C := \mu_{C, \text{baseline}} - \mu_{C, \text{2 weeks}} \) respectively in the control group. As the pilot study was rather small, we decide on an adaptive study design with one interim analysis after \( n_1 = 50 \) patients per group (half of the fixed sample size per group at \( \Delta = 0.4 \)). Thereby, we choose the inverse normal combination test\textsuperscript{4} with equal weights, a global one-sided significance level of 0.025 and locally adjusted significance levels according to O’Brien and Fleming.\textsuperscript{7} The binding futility stopping bound is set to \( c_{\text{eff}} = 0.0 \). At interim, if the study is neither stopped for futility nor for efficacy, the sample size is recalculated based on the “restricted observed conditional power approach”
combined with stepwise smoothing with a maximal sample size of \( n_{\text{max}} = 200 \) per group. Note that this refers to simulation Scenario 1 where the performance results are given in Table 1.

At the interim analysis, the observed interim effect turns out to be 0.2. Without the smoothing correction, the trial would have been stopped for futility in this case. However, for an observed interim effect of 0.228 and higher, the trial would have continued with the maximal sample size, so the result would be difficult to communicate to the investigator. When applying sample size recalculation with stepwise smoothing correction as anticipated here, the study is continued with a total sample size of 150 per group, and hence offers the possibility of still showing a possibly clinically relevant difference between the two treatments after the second stage. This increase in sample size corresponds to a value midway of no increase (as suggested without smoothing) and the maximal increase (as suggested for an effect 0.228 or higher) and is thus relatively easy to communicate to non-statisticians.

### Discussion

When using an adaptive design with sample size recalculation, it seems intuitive that for small interim effects, no increase of sample size is recommended (early stop for futility). It also seems plausible that starting from a certain value of the interim test statistic, an increase of sample size is justified. However, the decision on this boundary \( c_{\text{incr}} \) is somehow arbitrary. To overcome this problem, we presented five classes of simple smoothing corrections (linear, stepwise, sigmoid, concave, and convex increase) to be combined with existing recalculation rules to decrease the variability of sample size and conditional power. These variability reductions were applied to different sample sizes as well as different multiplicity adjustments (Scenarios 1–5). Moreover, a clinical study example was provided for illustration. Our main motivation to choose the smoothing functions was to propose a simple approach that does not need any analytical derivations. However, even when focusing on these five simple smoothing classes, there remain many possibilities of adaption for the specific function shapes. The intention of our work is to highlight the general impact of a smoothing function in different scenarios, whereas the optimization of a specific function shape will be addressed in future work.

### Conclusion

Our findings generally support the application of a smoothing correction, in particular the stepwise smoothing approach, to achieve the aim of reducing the variability in sample size and conditional power. These variability reductions are only one aspect of the performance for adaptive sample size recalculation, whereas a correct target sample size and power define the other perspective. The performance score by Herrmann et al.\(^6\) assessing both—variability and location of power and sample size—shows an overall benefit of smoothing corrections for medium and large effect sizes. Generally, there is no globally optimal approach across all effect sizes. The R code underlying the simulations of this paper is available on https://github.com/shareCH/SSR-smoothing-corrections.

### Ethical Approval

This research is exclusively based on simulations and does not involve any human subject data.

### Funding

This work was supported by the German Research Foundation (grant RA 2347/4-1).

### Conflict of Interest

C.H. reports a grant from Deutsche Forschungsgesellschaft/German Research Foundation (cf. funding declaration), during the conduct of the study. G.R. reports grants from null, during the conduct of the study.

### References

1. Friede T, Kieser M. A comparison of methods for adaptive sample size adjustment. Stat Med 2001;20(24):3861–3873
2. Chen YH, DeMets DL, Lan KK. Increasing the sample size when the unblinded interim result is promising. Stat Med 2004;23(07):1023–1038
3. Levin GP, Emerson SC, Emerson SS. Adaptive clinical trial designs with pre-specified rules for modifying the sample size: understanding efficient types of adaptation. Stat Med 2013;32(08):1259–1275, discussion 1280–1282
4. Lehmann W, Wassmer G. Adaptive sample size calculations in group sequential trials. Biometrics 1999;55(04):1286–1290
5. Posch M, Bauer P. Adaptive two stage designs and the conditional error function. Biometrical J 1999;41:689–696
6. Herrmann C, Pilz M, Kieser M, Rauch G. A new conditional performance score for the evaluation of adaptive group sequential designs with sample size recalculation. Stat Med 2020;39(15):2067–2100
7. O’Brien PC, Fleming TR. A multiple testing procedure for clinical trials. Biometrics 1979;35(03):549–556
8. Pocock SJ. Group sequential methods in the design and analysis of clinical trials. Biometrika 1977;64:191–199
9. Wang SK, Tsatis AA. Approximately optimal one-parameter boundaries for group sequential trials. Biometrics 1987;43(01):193–199
10. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; 2020
11. Bowden J, Mander A. A review and re-interpretation of a group-sequential approach to sample size re-estimation in two-stage trials. Pharm Stat 2014;13(03):163–172
12. Melzack R, Torgerson WS. On the language of pain. Anesthesiology 1971;34(01):50–59
Appendix A

The five simple smoothing options illustrated in Fig. 1 can be described mathematically by

- \( n_{\text{total}}(z_1) = (n_{\text{max}} - n_1)/c_{\text{incr}} \cdot z_1 + n_1 \) for \( z_1 \in [c_{\text{fut}}; c_{\text{incr}}] \),

for the linear function,

- \( n_{\text{total}}(z_1) = \begin{cases} 
\frac{n_1}{3} + n_1 & \text{for } z_1 \in \left( c_{\text{fut}}/2; 1 \right) \\
2 \cdot \left( n_{\text{max}} - n_1 \right)/3 + n_1 & \text{for } z_1 \in \left( 1; c_{\text{incr}} \right) \\
\end{cases} \)

for the step function,

- \( n_{\text{total}}(z_1) = 0.5 \cdot \frac{n_{\text{max}} - n_1}{0.5 + e^{10 \cdot (z_1 - 1)}} + n_1 \) for \( z_1 \in [c_{\text{fut}}; c_{\text{incr}}] \),

for the sigmoid function,

- \( n_{\text{total}}(z_1) = \frac{n_1 - n_{\text{max}}}{c_{\text{incr}}} \cdot (z_1 - c_{\text{incr}})^2 + n_{\text{max}} \) for \( z_1 \in [c_{\text{fut}}; c_{\text{incr}}] \),

for the concave function and

- \( n_{\text{total}}(z_1) = \frac{n_{\text{max}} - n_1}{c_{\text{incr}}} \cdot z_1^2 + n_1 \) for \( z_1 \in [c_{\text{fut}}; c_{\text{incr}}] \),

for the convex function.

The total sample size \( n_{\text{total}}(z_1) \) for \( z_1 \in [c_{\text{incr}}; c_{\text{off}}] \) is determined according to the initially proposed sample size recalculation rule.