

# %PT\_GSDesign: A SAS Macro for group sequential designs with time-to-event endpoint using the concept of proportional time

## Abstract

Sequential testing can be used to meet the specific needs of a clinical trial, all while adhering to the study's ethical, financial, and administrative considerations. When the assumption of proportional hazards or exponentially distributed lifetimes is not satisfied, the researcher can rely on the Proportional Time assumption for sample size calculation. The proportional time method has the advantage that previous study results can be used to bolster the current study design and provide an easier interpretation of the treatment benefit by reporting results as an improvement in longevity versus the more traditional interpretation of reduction in risk. This ease in interpretation of treatment benefit helps in raising interest in study participation. This novel method can be applied through a SAS macro and can be utilized for all distributions that belong to the generalized gamma family. The macro incorporates features specific to time-to-event data such as loss to follow-up, administrative censoring, differing accrual times and patterns, binding or non-binding futility rules with or without skips, and flexible alpha and beta spending functions. The macro includes validation for any parameters defined by the user, as well as suggestions for correcting erroneous input. This paper demonstrates two practical applications of the SAS macro with varying design inputs.

**Keywords:** efficacy, error spending, futility, proportional time, sample size, SAS software

Volume 11 Issue 2 - 2022

**Milind A Phadnis, Nadeesha Thewarapperuma**

Department of Biostatistics & Data Science, University of Kansas Medical Center, USA

**Correspondence:** Milind A. Phadnis, Department of Biostatistics & Data Science, University of Kansas Medical Center, Kansas City, USA, Tel 913-574-2029; Email [mphadnis@kumc.edu](mailto:mphadnis@kumc.edu)

**Received:** June 16, 2022 | **Published:** June 27, 2022

## Introduction

Two-arm randomized clinical trials (RCT) are considered the gold standard by biomedical researchers as they allow estimating how well a new treatment performs relative to a standard-of-care control. Treatments that are found to be promising in a Phase II trial are studied more comprehensively in a Phase III trial where, by enrolling a large number of patients (typically several hundred), researchers aim to investigate the effectiveness and safety of the new treatment against the current standard treatment. If such evidence is found in a Phase III trial, a new drug application (NDA) is submitted to the Food and Drug Administration (FDA) and on obtaining the FDA approval, the new drug becomes the *new* standard-of-care.

While traditional approaches require the calculation of a fixed sample size in advance of conducting a RCT depending on the type I error, power and clinically important treatment effect, in the medical setting they suffer from the limitation that patients are continually being accrued into a study which may be a time consuming process based on the accrual rate, the availability of qualified patients (based on inclusion/exclusion criteria) and the possibility of random dropouts among many factors. Thus, the primary outcome of interest is not available simultaneously on all patients and researchers may be interested to look at the early results on outcomes on the early enrollees and use that as a basis to decide whether the trial should be continued. This raises the concept of sequential testing in large-sized Phase III trials, where interim results can be used to – (i) stop the trial early for overwhelming evidence of efficacy, (ii) stop the trial early for overwhelming evidence for futility, and (iii) continue the trial for lack of evidence of efficacy or futility.

A Group Sequential Design (GSD) formalizes the above concept by providing a solid statistical framework under which either of the above three decisions can be taken after looking at results collected at interim points in the study. observation window. Ethical, financial,

and administrative requirements often guide the statistical designs of GSDs.<sup>1-3</sup> Such GSDs have been well developed for continuous and binary outcomes and have a long history starting with quality control application<sup>4</sup> and progressing to the medical setting.<sup>5</sup> Vast literature is available on this topic in many books<sup>6-10</sup> and overview articles.<sup>11-13</sup> When dealing with time-to-event outcome, a repeated significance testing approach incorporating a family of designs<sup>14-16</sup> can be combined with the error spending method<sup>17</sup> to implement a GSD using a log-rank test or by using the proportional hazards (PH) assumption. Popular statistical software often implement GSDs for time-to-event outcome using the weighted and unweighted versions of the log-rank test either explicitly assuming exponentially distributed survival times or with the PH assumption and are able to incorporate complexities of survival outcomes such as random dropouts, prespecified accrual and follow-up times, varying accrual patterns, equal/unequal spaced interim testing points (looks), efficacy-only designs, efficacy and futility designs, binding and non-binding futility rules, and many other flexible features specific to time-to-event outcomes.

When the underlying assumptions that drive the analytical and simulation-based approaches using the framework of the log-rank test are not valid, hardly any alternate methods are available in literature or in standard statistical software. Recent developments in this field have considered relaxing the PH assumption in favor of a 'proportionality of time' (PT) assumption leading to development of GSDs in the context of an accelerated failure time (AFT) model.<sup>18</sup> The authors have described various scenarios in the biomedical setting where their approach could be advantageous compared to the standard methods with the help of real-life examples. Their proposed GSD method provides an alternate approach when the PH assumption is not appropriate and allows various hazard shapes (increasing/decreasing monotonically over time, bathtub shaped, arc-shaped) using the generalized gamma ratio distribution.<sup>19</sup> The purpose of this paper is to present a fully functional SAS macro that can be used to implement

their GSD method. The SAS macro incorporates multitude of design features specific to a two-arm GSD for time-to-event outcome and the accompanying discussion of results provide information on how this macro can be implemented.

## Materials and Methods

### Statistical methods for GSD using the proportional time (PT) framework

The statistical framework for the method proposed based on the PT assumption<sup>18</sup> assumes that the survival times follow a generalized gamma (GG) distribution.<sup>20</sup> The probability density function of the GG distribution is given as:

$$f(t) = \frac{\beta}{\Gamma(\kappa)\theta} \left(\frac{t}{\theta}\right)^{\kappa\beta-1} e^{-\left(\frac{t}{\theta}\right)^\beta} \quad (1)$$

where  $\beta > 0$  and  $\kappa > 0$  are the shape parameters,  $\theta > 0$  is the scale parameter and  $\Gamma(k)$  is the gamma function defined as  $\Gamma(k) = \int_0^\infty x^{k-1} e^{-x} dx$ . For model fitting purposes a re-parametrization  $GG(\mu, \sigma, \lambda)$  is used to avoid convergence problems using location parameter  $\mu$ , scale parameter  $\sigma$ , and shape parameter  $\lambda$  that generalizes the two-parameter gamma distribution. The density function is given by:

$$f_{GG}(t) = \frac{|\lambda|}{\sigma \Gamma(\lambda^{-2})} \left[ \lambda^{-2} \{ \exp(-\mu) t \}^{\lambda/\sigma} \right]^{\lambda^{-2}} \exp \left[ -\lambda^{-2} \{ \exp(-\mu) t \}^{\lambda/\sigma} \right] \quad (2)$$

A complete taxonomy of the various hazard functions for the GG family is explained in literature.<sup>21</sup> Briefly, the GG family allows the flexibility of modeling different hazard shapes such as increasing from 0 to  $\infty$  or from a constant to  $\infty$  decreasing from  $\infty$  to 0, or from  $\infty$  to a constant, arc shaped hazards, and bathtub shaped hazards. Special cases of the GG family are (i) two parameter gamma:  $\lambda = \sigma$  (ii) standard gamma:  $\mu = 0; \sigma = 1$  for fixed values of  $\lambda$  (iii) Weibull:  $\lambda = 1$  (iv) exponential:  $\lambda = \sigma = 1$  (v) lognormal:  $\lambda = 0$  (vi) inverse Weibull:  $\lambda = -1$  (vii) inverse gamma:  $\lambda = -\sigma$  (viii) ammag:  $\lambda = 1/\sigma$  (ix) inverse ammag:  $\lambda = -1/\sigma$ .

### Concept of proportional time (PT) as a special case of relative time (RT)

For a  $GG(\mu, \sigma, \lambda)$  distribution, we have

$$\log\{t_{GG(\mu, \sigma, \lambda)}(p)\} = \mu + \sigma \cdot \log\{t_{GG(0, 1, \lambda)}(p)\} \\ = \mu + \sigma \cdot g_\lambda(p) \quad (3)$$

$g_\lambda(p)$  is the logarithm of the  $p^{\text{th}}$  quantile from the GG (0,1, $\lambda$ ) distribution. The location parameter  $\mu$  acts as a time-multiplier and governs the values of the median for fixed values of  $\sigma$  and  $\lambda$  resulting in the accelerated failure time (AFT) model. The scale parameter  $\sigma$  determines the interquartile ratio for fixed values of  $\lambda$  and independently of  $\mu$ . The shape parameter  $\lambda$  determines the GG (0,1, $\lambda$ ) distribution. Together,  $\sigma$  and  $\lambda$  describe the type of hazard function for the GG(0,1, $\lambda$ ) distribution.

The time by which  $p\%$  of the population experience an event can lead to a statistic called ‘relative times RT(p),’ which can be used to compare survival profiles of patients in different treatment arms (new treatment versus standard treatment). Thus,

$$RT(p) = \frac{t_1(p)}{t_0(p)} = \frac{S_1^{-1}(1-p)}{S_0^{-1}(1-p)} \quad (4)$$

The interpretation of RT(p) is that the time required for  $p\%$  of individuals in one study arm to experience an event is RT(p) times

the time required for  $p\%$  of individuals in the second study arm. Thus if  $(\mu_0, \sigma_0, \lambda_0)$  and  $(\mu_1, \sigma_1, \lambda_1)$  denote two different sets of GG parameter values, then

$$RT(p) = \exp\left((\mu_1 - \mu_0) + \sigma_1 \cdot g_{\lambda_1}(p) - \sigma_0 \cdot g_{\lambda_0}(p)\right) \quad (5)$$

The manner in which covariates affect RT(p) can be summarized as:

- 1) If  $\lambda_1 = \lambda_0$  and  $\sigma_1 = \sigma_0$ , then we have a conventional AFT model resulting in non-PH, but proportional RT or simply “proportional times (PT) assumption”. Then, covariates affect  $\mu$  only.  
 $RT(p) = \exp(\mu_1 - \mu_0) = \Delta_{PT} \equiv PT \text{ assumption}$
- 2) If only  $\lambda_1 = \lambda_0$ , then we have a model that results in non-PH and nonproportional RT(p). Then, covariates affect both  $\mu$  and  $\sigma$ .
- 3) Full generalization is obtained by having covariates affect all 3 parameters.

### Test Statistic

Based on the discussion above, a test statistic that follows the four-parameter generalized gamma ratio (GGR) distribution can be developed.<sup>20</sup> That is, the parameters of the GG distribution can be used to express RT(p) as:

$$RT(p) = \exp(\mu_1 - \mu_0) = \exp\left\{\ln(\theta_1) + \frac{1}{\beta} \ln(\lambda^{-2}) - \ln(\theta_0) - \frac{1}{\beta} \ln(\lambda^{-2})\right\} = \frac{\theta_1}{\theta_0} \quad (6)$$

Thus, for new treatment to standard treatment allocation ratio  $r = n_1/n_0$  we get a test statistic Q that follows the GGR distribution.

$$Q = \frac{\bar{\theta}_1}{\bar{\theta}_0} \sim GGR\left(\frac{1}{r} \left[\frac{\theta_1}{\theta_0}\right]^\beta, \frac{n_1 k}{r}, n_1 k, \beta\right) \quad (7)$$

Although this test statistic can be used to calculate the sample size for a two-arm RCT in the case of a *fixed* study design (a design without any interim testing), when designing a more complex study incorporating all the desired features of a GSD, calculations become more complicated and have to be conducted using a simulation-based approach. The remainder of the paper discusses how the simulation-based GSD method of Phadnis et al.<sup>18</sup> can be implemented for a two-arm phase III trial using a SAS macro.

### Description of SAS Macro

A ten-step algorithm has been detailed in the GSD method of Phadnis et al.<sup>18</sup> along with the appropriate formulas for performing sample size calculations. The proposed SAS macro titled PT\_GS Design fully implements this algorithm and is written in base SAS and SAS/STAT.<sup>22</sup> The various design features available in our macro are summarized below:

- NumSimul: Number of simulated samples for the given sample size
- alpha: Type I error
- sides: 1-sided or 2-sided test
- lambda: Shape parameter of the Control Arm using GG distribution
- sigma: Scale parameter of the Control Arm using GG distribution
- med: User entered Median of the Control Arm using GG distribution
- evt\_rate: Anticipated event rate for loss-to-follow-up (right censoring)
- seed: A random seed is chosen
- r: Allocation Ratio: (number in Treatment arm)/ (number in Standard Arm)

Delta\_PT\_Ha: Under the alternative, PT is greater than 1

a: Accrual time for the study

a\_type: Accrual pattern: «1» = Uniform, «2» = Truncated Exponential (parameter omega)

a\_omega: Parameter of «2» = truncated exponential distribution: >0 (convex) or <0 (concave); input will only be used with truncated exponential

t: Total time for the study = Accrual time + Follow-up time

bind: Binding futility = 1; Non-binding futility=0

num\_look: Total number of looks (including the look at the end-of-study)

look\_points: equally spaced looks = 1, unequally spaced looks = 2

alpha\_spend: Type of Alpha spending function: 1 = Jennison-Turnbull, 2 = Hwang-Shih-DeCani, 3 = User defined spending.

The following datasets are needed to take advantage of the macro's user defined options.

- If unequal look points (look\_points=2) are selected then UserDefTime will need to include the number of look points (numuser) followed by the time points. The time points need to be in cumulative, ascending order and the last time point must equal the total study time, t.
- If user defined alpha spending (alpha\_spend=3) or beta spending (beta\_spend=3) are selected then UserDefAlpha and UserDefBeta values must be split in cumulative, ascending order and the last entry must match the inputs for alpha and beta.

Default values for the macro parameters have been provided in the text description. Error checks in the code prohibit a user to input impossible values for the macro parameters. For example, where numerical input is required, character values cannot be entered. Likewise, numerical input outside the natural range of a macro parameters are not allowed. If such impossible values are entered, the macro will stop executing and display an error message in the log window suggesting corrections to the input values.

In addition to the above, the following extra features are provided in the macro:

1. The user can create his/her own user-defined inputs for: (i) Times at which the interim looks occur using the dataset UserDefTime (ii) The type-I error to be spent at each interim look using the dataset UserDefAlpha (iii) The type-II error to be spent at each interim look using the dataset UserDefBeta. If the user-defined options are selected and these datasets are left empty or contain incorrect (or trash) values, the macro generates an error message asking the user to enter correct input. If the default options of Pocock or O'Brien-Fleming and equally spaced look points are chosen, then the user-defined datasets are ignored.
2. Warning messages have been added to the code wherever necessary. If the first interim is conducted very early (small value of time relative to the total study time) with only few events, the program may run into convergence problems with LIFEREG procedure, and the resulting output may not be reliable for the first interim. On such occasions, a warning message will be displayed in the SAS log file. This may also cause delay for full

execution of the program. On such occasions, the user may want to re-run the program ensuring that the first look is conducted late enough to allow more events to be observed thereby ensuring convergence is obtained with the LIFEREG procedure.

3. The SAS log file is saved as a separate text file, Mydoc.log, under the user-defined file path, while output tables are generated as an ods listing file. The first fifty errors will be printed at the bottom of the ods listing file.
4. The macro will not run if the number of simulations < 1000. To make sure the macro parameters are valid, the user can comment out this statement and check the ods listing file for any errors before submitting a full simulation.

We have also provided a "README.pdf" file detailing a step-by-step procedure to help users navigate through the process of entering input values. This, along with the full code, is available at [https://github.com/thewan05/GSD\\_SAS\\_Macro](https://github.com/thewan05/GSD_SAS_Macro).

## Results and discussion

These examples were first published under the methodology paper.<sup>18</sup> The examples are presented once more so the reader can easily reproduce them. There may be some minor variations, depending on the seed used.

These macro parameters are used to obtain the results for example one:

```
NumSimul=10000, alpha=0.025, sides=1, lambda=0.5,
sigma=0.75, med=20, evt_rate=0.7, seed=1729, r=1, Delta_PT_
Ha=1.4, a=12, a_type=1, a_omega=1, t=60, bind=1, num_look=3,
look_points=2, alpha_spend=1, rho=1, beta=0.10, beta_spend=1,
rho_f=1, num_skip=0, maxiter=200, convg=1E3, direct=C:\Users\
user1\Desktop
```

UserDefTime dataset: 3 24 36 60 (Table 1).

These macro parameters are used to obtain the results for example two:

```
NumSimul=10000, alpha=0.025, sides=1, lambda=0.5,
sigma=0.75, med=20, evt_rate=0.7, seed=1729, r=1, Delta_PT_
Ha=1.4, a=12, a_type=1, a_omega=1, t=60, bind=1, num_look=3,
look_points=2, alpha_spend=1, rho=3, beta=0.10, beta_spend=1,
rho_f=3, num_skip=0, maxiter=200, convg=1E3, direct=C:\Users\
user1\Desktop
```

UserDefTime dataset: 3 24 36 60 (Table 2).

These macro parameters are used to obtain the results for example three:

```
NumSimul=10000, alpha=0.025, sides=1, lambda=0.5,
sigma=0.75, med=20, evt_rate=0.7, seed=1729, r=1, Delta_PT_
Ha=1.4, a=12, a_type=1, a_omega=1, t=60, bind=1, num_look=3,
look_points=2, alpha_spend=3, rho=3, beta=0.10, beta_spend=3,
rho_f=3, num_skip=0, maxiter=200, convg=1E3, direct=C:\Users\
user1\Desktop
```

UserDefTime dataset: 3 24 36 60

UserDefAlpha dataset: 0.0050 0.0125 0.0250

UserDefBeta dataset: 0.0100 0.0350 0.1000 (Table 3).

These macro parameters are used to obtain the results for example four:

NumSimul=10000, alpha=0.025, sides=1, lambda=1, sigma=1, med=1, evt\_rate=1, seed=1729, r =1, Delta\_PT\_Ha=1.75, a=1,

a\_type=1, a\_omega=1, t=4, bind=1, num\_look=4, look\_points=1, alpha\_spend=2, rho=1, beta=0.20, beta\_spend=2, rho\_f=1, num\_skip=2, maxiter=200, convg=1E3, direct=C:\Users\user1\Desktop (Table 4), (Figure 1), (Figure 2).

**Table 1** GSD - Ovarian CT using proposed method with 10,000 simulations; Pocock plans (efficacy and futility at all looks)

Look no.	Look times	No. events– H <sub>0</sub> control arm	No. events– H <sub>0</sub> treatment arm	Alpha spent	Cumul. alpha spent	Upper Significance boundary (efficacy) GGR Test statistic	Stop probability under H <sub>0</sub>	Cumul. stop probability under H <sub>0</sub>	Cumul. subject time under H <sub>0</sub>
1	24	58.21	58.25	0.00883	0.00833	1.333	0.8053	0.8053	1992.86
2	36	87.74	87.74	0.00833	0.01667	1.259	0.1448	0.9501	2559.65
3	60	109.09	109.05	0.00833	0.025	1.219	0.0499	1	2940.79
Look no.	Look times	No. events – H <sub>A</sub> control arm	No. events – H <sub>A</sub> treatment arm	Beta spent	Cumul. beta spent	Lower significance boundary (efficacy) GGR Test statistic	Stop probability under H <sub>A</sub>	Cumul. stop probability under H <sub>A</sub>	Cumul. subject time under H <sub>A</sub>
1	24	58.21	40.03	0.03173	0.03173	1.105	0.6792	0.6792	2244.29
2	36	87.74	69.12	0.03173	0.06347	1.174	0.2136	0.8928	3093.34
3	60	109.09	99.41	0.03173	0.0952	1.219	0.1072	1	3875.05

**Table 2** GSD - Ovarian CT using proposed method with 10,000 simulations; O'Brien-Fleming plan (efficacy and futility at all looks)

Look no.	Look Times	No. events – H <sub>0</sub> control arm	No. events – H <sub>0</sub> treatment arm	Alpha spent	Cumul. alpha spent	Upper significance boundary (efficacy) GGR test statistic	Stop probability under H <sub>0</sub>	Cumul. stop probability under H <sub>0</sub>	Cumul. subject time under H <sub>0</sub>
1	24	52.57	52.63	0.00093	0.00093	1.457	0.4309	0.4309	1797.58
2	36	79.19	79.21	0.00648	0.00741	1.312	0.4378	0.8687	2308.8
3	60	98.42	98.4	0.01759	0.025	1.222	0.1313	1	2652.04
Look no.	Look times	No. events – H <sub>A</sub> control arm	No. events – H <sub>A</sub> treatment arm	Beta spent	Cumul. beta spent	Lower significance boundary (efficacy) GGR test statistic	Stop probability under H <sub>A</sub>	Cumul. stop probability under H <sub>A</sub>	Cumul. subject time under H <sub>A</sub>
1	24	52.57	36.08	0.00352	0.00352	0.978	0.3887	0.3887	2026.69
2	36	79.19	62.34	0.02463	0.02815	1.127	0.356	0.7447	2793.83
3	60	98.42	89.7	0.06685	0.095	1.222	0.2557	1.0004	3501.2

**Table 3** GSD - Ovarian CT using proposed method with 10,000 simulations; user-defined alpha and beta spending (efficacy and futility at all looks)

Look no.	Look times	No. events – H <sub>0</sub> control arm	No. events – H <sub>0</sub> treatment arm	Alpha spent	Cumul. alpha spent	Upper significance boundary (efficacy) GGR Test statistic	Stop probability under H <sub>0</sub>	Cumul. stop Probability under H <sub>0</sub>	Cumul. subject time under H <sub>0</sub>
1	24	54.32	54.32	0.005	0.005	1.384	0.5758	0.5758	1859.49
2	36	81.82	81.83	0.0075	0.0125	1.279	0.3191	0.8949	2388.16
3	60	101.73	101.73	0.0125	0.025	1.224	0.1043	0.9992	2743.29
Look no.	Look times	No. events – H <sub>A</sub> control arm	No. events – H <sub>A</sub> treatment arm	Beta spent	Cumul. beta spent	Lower significance boundary (efficacy) GGR Test statistic	Stop probability under H <sub>A</sub>	Cumul. stop probability under H <sub>A</sub>	Cumul. subject time under H <sub>A</sub>
1	24	54.32	37.35	0.00922	0.00922	1.022	0.5416	0.5416	2094.91
2	36	81.82	64.53	0.02305	0.03227	1.135	0.2776	0.8192	2887.1
3	60	101.73	92.82	0.05933	0.0922	1.224	0.1798	0.999	3615.75

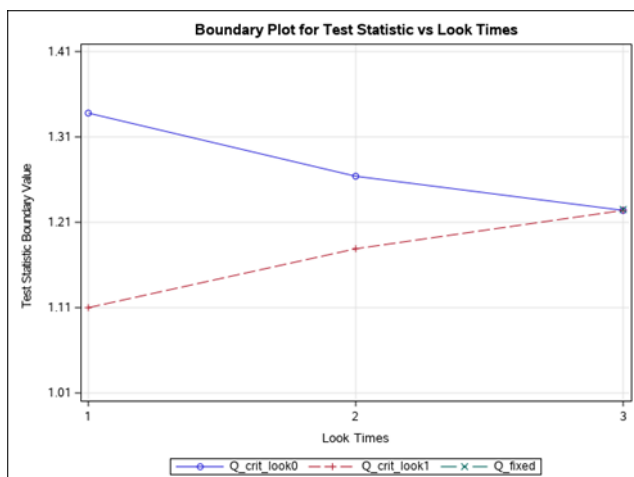


**Table 4** GSD output for exponential distributed data using proposed method with 10,000 simulation (two futility skips)

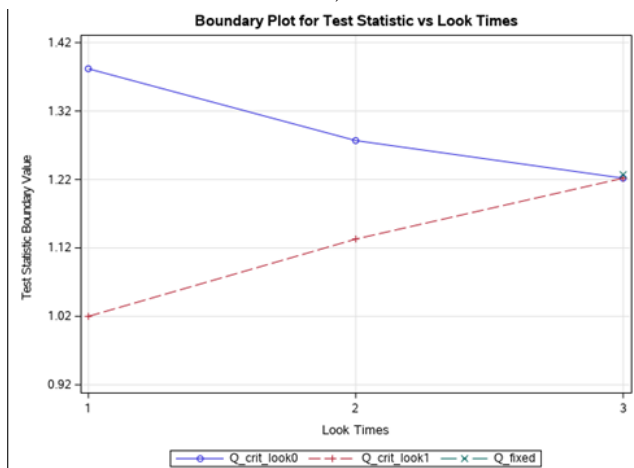
Look no.	Look times	No. events – $H_0$ control arm	No. events – $H_0$ treatment arm	Alpha spent	Cumul. alpha spent	Upper significance boundary (efficacy) GGR Test statistic	Stop probability under $H_0$	Cumul. stop probability under $H_0$	Cumul. subject time under $H_0$
1	1	22.3	22.3	0.00875	0.00875	2.218	0.0088	0.0088	32.156
2	2	51.23	51.17	0.00681	0.01556	1.632	<0.0001	0.0088	73.726
3	3	65.58	65.61	0.00531	0.02087	1.522	0.9837	0.9922	94.496
4	4	72.79	72.84	0.00413	0.025	1.452	0.0077	0.9999	104.891

Look no.	Look times	No. events – $H_A$ control arm	No. events – $H_A$ treatment arm	Beta spent	Cumul. beta spent	Lower significance boundary (efficacy) GGR Test statistic	Stop probability under $H_A$	Cumul. stop probability under $H_A$	Cumul. subject time under $H_A$
1	1	22.3	13.93	0	0	-	0.2674	0.2674	35.198
2	2	51.23	35.59	0	0	-	0.3524	0.6198	89.716
3	3	65.58	50.14	0.16268	0.16268	1.45	0.3082	0.928	126.351
4	4	72.79	59.94	0.03222	0.1949	1.453	0.0722	1.0002	150.989



**Figure 1** Ovarian CT efficacy and futility boundaries using Pocock plan under Test-statistic scale. Look times are at 24, 36 and 60 months.



**Figure 2** Ovarian CT efficacy and futility boundaries using user-defined values under Test-statistic scale. Look times are at 24, 36 and 60 months.

## Conclusion

A GSD is generally implemented as a large sample Phase III trial and therefore provides an opportunity to incorporate information obtained

from a preceding moderate-sized Phase II study. In our paper, we have built a SAS macro that implements a GSD incorporating various design features specific to time-to-event outcome in the case of non-proportional hazards. While earlier methods using the nonparametric log-rank test or the PH assumption are available in standard statistics software, our macro is the first of its kind in implementing a GSD in the non-PH case using a three-parameter GG distribution. The macro fully executes the method based on the PT assumption<sup>18</sup> and thereby offers researchers an additional option in designing Phase III trials for the non-PH case. Some of the advantages of using the macro are - it handles different types of hazard shapes, utilizes Phase II data to ensure that early interims are not conducted with too few events, is simulation based and does not depend on any asymptotic normality of the test statistic, and most importantly provides clinical meaningful and easy-to-interpret efficacy and/or futility boundaries based on the concept of improvement in longevity. Due to this direct interpretation of “treatment effect” as an improvement in survival time, we hope that researchers working in this area will find our SAS macro to be of practical value in implementing a GSD for Phase III time-to-event trials.

## Acknowledgements

The High performance computing capabilities, which were used to conduct some of the analyses described in this paper, were supported in part by the National Cancer Institute (NCI) Cancer Center Support Grant P30 CA168524; the Kansas IDeA Network of Biomedical Research Excellence Bioinformatics Core, supported by the National Institute of General Medical Science award P20 GM103418; and the Kansas Institute for Precision Medicine COBRE, supported by the National Institute of General Medical Science award P20 GM130423.

## Conflicts of interest

The authors declare no conflicts of interest.

## References

- Enas G, Dornseif B, Sampson C, Rochhol F, et al. Monitoring versus Interim Analysis of Clinical Trials: Perspective from the Pharmaceutical Industry. *Controlled Clinical Trials*. 1989;10(1):57–70.
- Jennison C, Turnbull BW. Statistical Approaches to Interim Monitoring of Medical Trials: A Review and Commentary. *Statistical Science*. 1990;5(3):299–317.

3. Ellenberg SS, Fleming TR, DeMets DL. Data Monitoring Committees in Clinical Trials: A Practical Perspective. *John Wiley & Sons*. 2019.
4. Wald A. *Sequential analysis*. Sequential analysis. John Wiley; 1947:xii, 212-xii, 212.
5. Armitage P. *Sequential Medical Trials*. 1st ed. Springfield; 1960.
6. Whitehead J. *The Design and Analysis of Sequential Clinical Trials*. 2nd ed. John Wiley & Sons; 1997.
7. Proschan MA, Lan KKG, Wittes JT. *Statistical Monitoring of Clinical Trials. Statistics for Biology and Health*. Springer-Verlag; 2006.
8. Dmitrienko A, Molenberghs G, Chuang-Stein C, et al. *Analysis of Clinical Trials Using SAS: A Practical Guide*. SAS Institute; 2005.
9. Wassmer G, Brannath W. *Group Sequential and Confirmatory Adaptive Designs in Clinical Trials*. 1 ed. Springer Series in Pharmaceutical Statistics. Springer Cham; 2016.
10. Jennison & Turnbull, CRC/Chapman & Hall, Boca Raton. Group Sequential Methods with Applications to Clinical Trials. 2000.
11. Whitehead J. A unified theory for sequential clinical trials. *Stat Med*. 1999;18(17-18):2271–2286.
12. Todd SA. 25-year review of sequential methodology in clinical studies. *Statistics in Medicine*. 2007;26(2):237–252.
13. Mazumdar M, Bang H. 16 *Sequential and Group Sequential Designs in Clinical Trials: Guidelines for Practitioners*. Handbook of Statistics. Elsevier; 2007.
14. Pocock SJ. Group Sequential Methods in the Design and Analysis of Clinical Trials. *Biometrika*. 1977;64(2):191–199.
15. O'Brien PC, Fleming TR. A Multiple Testing Procedure for Clinical Trials. *Biometrics*. 1979;35(3):549–556.
16. Wang SK, Tsiatis AA. Approximately Optimal One-Parameter Boundaries for Group Sequential Trials. *Biometrics*. 1987;43(1):193–199.
17. Lan KKG, Demets DL. Discrete Sequential Boundaries for Clinical Trials. *Biometrika*. 1983;70(3):659–663.
18. Phadnis MA, Mayo MS. Group sequential design for time-to-event data using the concept of proportional time. *Statistical methods in medical research*. 2020;29(7):1867–1890.
19. Phadnis MA, Wetmore JB, Mayo MS. A clinical trial design using the concept of proportional time using the generalized gamma ratio distribution. *Statistics in medicine*. 2017;36(26):4121–4140.
20. Stacy EW, Mihram GA. Parameter Estimation for a Generalized Gamma Distribution. *Technometrics*. 1965;7(3):349–358.
21. Cox C, Chu H, Schneider MF, et al. Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution. *Statistics in Medicine*. 2007;26(23):4352–4374.
22. SAS Institute Inc. *SAS Software, Version 9.4*. Cary, NC, USA. 2017.