Interim decision-making strategies in adaptive designs for population selection considering post-progression survival magnitudes

Ryuji Uozumi*1 and Chikuma Hamada2

1Department of Biomedical Statistics and Bioinformatics, Kyoto University Graduate School of Medicine, Kyoto 606-8507, Japan.
2Department of Management Science, Graduate School of Engineering, Tokyo University of Science, Tokyo 162-8601, Japan.

Abstract

The development of targeted therapies, which benefit only a subgroup of patients treated for a given type of cancer, has been extremely attractive to many investigators. Adaptive seamless phase II/III designs in oncology clinical trials with interim analyses for subpopulation selection could be used if pre-defined biomarker hypothesis exists. We consider the interim analysis using time-to-event endpoints, e.g., overall survival (OS) and progression-free survival (PFS), to identify whether the whole population or only the biomarker-positive population should be continued into the subsequent stage, whereas a final decision is based on OS data. In this paper, we propose the interim decision-making strategies in adaptive designs with correlated endpoints, considering post-progression survival (PPS) magnitudes. In our approach, the interim decision is made on the basis of predictive power, by incorporating information on OS as well as PFS. We consider PFS data only in making interim decision in order to supplement the immature OS data. Simulation studies assuming a targeted therapy show that our interim decision procedure performs well in terms of selecting the proper population, especially under a scenario in which PPS affects the translation of the benefit from PFS to OS.

1 Introduction

The development of targeted therapies, which benefit only a subgroup of patients treated for a given type of cancer, has been extremely attractive to many investigators, especially in oncology. If a pre-defined biomarker hypothesis exists, clinical trial designs should be set up considering the heterogeneity of patient subgroups by using the biomarker at the planning stage [1]. Currently, most trials demonstrate the benefit for a given patient subgroup retrospectively. For instance, it was shown that progression-free survival (PFS) was significantly improved among patients with pulmonary adenocarcinoma who tested positive for epidermal growth factor receptor (EGFR) inhibitor and received gefitinib (hazard ratio [95% CI] for PFS: 0.48 [0.36 to 0.64], \( P < 0.001 \)) [2]. In contrast, PFS was notably worsened among patients who lacked an EGFR mutation (hazard ratio [95% CI] for PFS, 2.85 [2.05 to 3.98]; \( P < 0.001 \)). In another example, both overall survival (OS) and PFS were significantly improved among patients with advanced colorectal cancer who received cetuximab, and had the wild-type K-ras gene (hazard ratio [95% CI] for PFS, 0.40 [0.30 to 0.54]; \( P < 0.001 \); hazard ratio [95% CI] for OS, 0.55 [0.41 to 0.74]; \( P < 0.001 \)), whereas no improvement was seen among patients with mutated K-ras tumors (hazard ratio [95% CI] for PFS, 0.99 [0.73 to 1.35]; \( P = 0.96 \); hazard ratio [95% CI] for OS, 0.98 [0.70 to 1.37]; \( P = 0.89 \)) [3]. Note that the treatment effects for the subgroup in these clinical trials were evaluated retrospectively. For these reasons, adaptive designs to select the subpopulation by using an identified biomarker are needed to develop such a therapy.

Interest in the use of adaptive designs has been increasing among many biostatisticians, including staff members of the Food and Drug Administration from the CDER Office of Epidemiology and Biostatistics, in the past 10 to 15 years [4]. We describe an adaptive seamless phase II/III design with two stages although adaptive designs, in general, encompass every phases in clinical trials. The different types of adaptive seamless phase II/III designs are broadly divided into adaptive treatment selection design and adaptive subpopulation selection design. Moreover,
they can minimize the period, white space, between the analysis of phase II data and the recruitment of phase III patients [5]. Recently, many papers on adaptive seamless phase II/III designs have been published [5–19]. There are two methods to combine the stagewise data: the Fisher’s product combination method by Bauer and Köhne [20] and the weighted inverse normal combination method [21, 22].

In this paper, we discuss an adaptive design for subpopulation selection using correlated time-to-event endpoints. This design could be applied to the development of targeted therapies. Brannath et al. [13] presented an adaptive seamless phase II/III methodology for subpopulation selection by using a single time-to-event endpoint. Following this, Jenkins et al. [15] proposed the aforementioned methodology by using the correlated time-to-event endpoints: OS and PFS as a short-term endpoint. Subsequently, Friede et al. [19] demonstrated a more powerful method by using a conditional error function approach [23]. Under the methodology by Jenkins et al. [15], the interim analysis involves the use of PFS data from stage 1 patients only, whereas the final analysis is conducted based on OS data from each stage. We propose interim decision-making strategies in adaptive designs for subpopulation selection. We extend the previous methodologies [13, 15] in two aspects. First, the interim analysis is conducted by incorporating information on PFS as well as OS. Second, we consider a scenario in which OS is calculated based on post-progression survival (PPS), if the progression is observed before death.

The paper is structured as follows. We first introduce the approach for final decision-making in adaptive subpopulation selection design using the combination test method proposed by Jenkins et al. [15] in Section 2. We then discuss the interim decision-making strategies using correlated time-to-event endpoints in Section 3. Section 4 presents a simulation study. Finally, we conclude the paper in Section 5 with a discussion.

2 Final decision-making using combination test approach

This section reviews the combination test approach for final decision-making in adaptive subpopulation selection design.

We suppose that the pre-defined biomarker hypothesis exists, and the full population $F$ consists of the pre-defined biomarker-positive population $P$ and the biomarker-negative population $N$. In adaptive subpopulation selection designs, an interim analysis takes place to identify whether $F$ or $P$ benefit based on stage 1 only. Subsequently, the final analysis would conduct hypotheses testing in both $F$ and $P$, or only in $P$.

In the combination test for a single one-sided null hypothesis $H_0^{(g)}$ for each population $g \in \{F, P\}$, a stagewise p-value $p_k^{(g)}$ for $H_0^{(g)}$ is calculated from OS data only for stage $k \in \{1, 2\}$ patients based on log-rank tests. As Jenkins et al. [15] have illustrated in the solution for time-to-event endpoints in adaptive seamless design, the additional follow-up during stage 2 of patients accrued in stage 1 contributes to stage 1 p-values. The intersection hypothesis is $H_0^{(F, P)} = \cap_{g \in \{F, P\}} H_0^{(g)}$. Closed testing procedures [24] are used to strongly control the familywise error rate. We shall use Simes’ procedure [24] for $H_0^{(F, P)}$ with $p_k^{(F, P)} = \min[2 \cdot \min(p_k^{(F)}, p_k^{(P)}), \max(p_k^{(F)}, p_k^{(P)})]$.

We then use a weighted inverse normal combination method with pre-specified weights $w_k$, where $0 \leq w_k \leq 1$ and $\sum_{k=1}^{2} w_k^2 = 1$. For OS data, we set $w_k = \sqrt{D_k}/\sum_j D_j$ as the weights with the anticipated numbers of OS events $D_k$ in order to combine the p-value from each stage $k$ and the null hypothesis for each population is rejected if the weighted inverse normal combination function

$$C(p_1^{(g)}, p_2^{(g)}) = w_1\Phi^{-1}(1 - p_1^{(g)}) + w_2\Phi^{-1}(1 - p_2^{(g)}) > c$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution and $c \approx 1.96$ represents the critical value for a one-sided significance level of 0.025. Note that it is vital to specify the combination function and the design of stage 1 at the planning stage.

3 Interim decision-making

The purpose of this paper is to extend the recent methodologies [13, 15] for interim decision-making in two aspects. First, not only PFS data but also OS data are incorporated into the interim analysis in adaptive subpopulation selection design. PFS is defined as the time from randomization until objective tumor progression or death from any cause, namely, time-to-progression (TTP) or OS, whichever occurs first [26]. Several authors have discussed the issues in using PFS [27, 30]. Although the most commonly used endpoint required in phase III trials by regulatory agencies is OS. PFS is frequently used in phase II trials, especially for targeted therapies. OS generally requires
long follow-up durations after tumor progression. Therefore, a long study duration and a large number of patients is
required, making it expensive to conduct clinical trials to collect OS data. Hence, it would be practical to use PFS
data for interim decision-making only if PFS could be deemed as a short-term intermediate time-to-event endpoint,
because it is more quickly observed than OS. Second, we assume a scenario in which OS is calculated considering
PFS after tumor progression, if progression is observed before death. OS data frequently require long-term follow-up
periods. Therefore, it would also be pragmatic to consider the impact based on PPS data.

3.1 Procedures for interim decision-making
An interim analysis takes place to identify whether the full population $F$ or the pre-defined biomarker-positive
population $P$ benefit based on stage 1 only. In addition, a clinical trial can be discontinued early for futility, when
the success of the trials at the interim analysis is deemed unpromising. This enables sponsors and investigators to
optimize the investment of resources. In contrast, there is no consideration to stop the clinical trial early based on
efficacy. For interim decisions, the sponsor has to be blinded to any results at the interim stage and the Independent
Data Monitoring Committee (IDMC) makes the recommendation based on an interim decision rule [31].

3.2 Correlation model between OS and PFS
Several papers have been published to handle the correlation between time-to-event endpoints [32–36]. To measure
the correlation between two time-to-event variables such as PFS and OS, the correlation coefficient by Spearman [37],
which is nonparametric, is widely used in clinical trial reports (see, e.g., [38, 39]. However, these measurements do
not account for censoring.

One of the statistical models to handle censoring has been proposed by Fleischer et al. [36]. They use exponential
models for each time-to-event endpoint based on the assumption that TTP and OS are completely independent and
that PFS is given by the minimum of TTP and OS. Moreover, we shall consider the time-to-event endpoint $D$, say
the time to death without tumor progression. Then, OS is calculated as follows:

$$OS = \begin{cases} PFS & \text{if } PFS \neq TTP \\ TTP + PPS & \text{otherwise} \end{cases}$$

Suppose that each time-to-event endpoint $v \in \{TTP, PPS, D\}$ is exponentially distributed with parameter $\lambda_v$ where
$\lambda_v > 0$. Then the correlation between OS and PFS is given by

$$\rho = \text{Corr}(OS, PFS) = \frac{\lambda_{PPS}}{\sqrt{\lambda_{TTP}^2 + 2\lambda_{TTP}\lambda_D} + \lambda_{PPS}^2}$$

(2)

Note that $\rho = 1.0$ if no tumor progression occurs before death; in other words, PFS = OS.

3.3 Interim decision rule using predictive power
In this paper, the decision tool applied at the interim analysis relies on the predictive power approach [11, 40]. The
predictive power indicates how likely the various null hypotheses are to be rejected at the final analysis. For instance,
if the predictive power based on interim results is greater than the threshold for one hypothesis, then the possibility of
rejection is high for the corresponding hypothesis when continuing with the corresponding population. For simplicity,
let us assume non-informative priors for the treatment effect; the predictive power for each population $s \in \{F, P, N\}$
using the interim $e \in \{OS, PFS\}$ data is given as follows:

$$PP_{e}^{(s)} = 1 - \left[ \Phi \left(1 - \Pi_{e}^{(s)} \right)^{-1/2} \right]$$

where $\Pi_{e}^{(s)}$ represents the event fraction at the interim analysis and $z_{e}^{(s)}$ is the observed test statistic based on the
log-rank test using stage 1 data.

In the context of adaptive subpopulation selection designs, Brannath et al. [13] have demonstrated the decision
rule by using the predictive probability as well as the posterior probability at the interim analysis, whereas Jenkins
et al. [15] have proposed the use of the rule based on the estimated hazard ratios. For the clinical development
of targeted therapies, when we consider a scenario in which the experimental treatment is beneficial for $P$ but is
actually harmful for $N$, the problem of crossing hazard rates might be observed due to the violation of the proportional
hazard assumption. Hence, we provide the interim decision rule using predictive power for the sake of simplicity by modifying the rule demonstrated by Brannath et al. \cite{brannath} and by considering the use of multiple endpoints at the interim analysis. Figure 4 illustrates the decision rule at the interim stage.

Let $\pi^{(v)}_i$ denote the threshold of the predictive power for each population $s \in \{F, P, N\}$ in each case $i$ and let $\delta_s$ denote the relative importance that we assign to the corresponding endpoint $e \in \{OS, PFS\}$. Because of the nature that PFS data are more quickly observed than OS data, we should reflect the expected number of events accrued up to the interim analysis. For adaptive treatment selection design, Di Scala and Glimm \cite{scala} combined the predictive probability using weights similar to $\delta_s$ and have shown the simulation results. However, we evaluate multiple endpoints separately in considering the inconsistency as well as the correlation between OS and PFS due to the impact of PPS. A series of simulations are needed to set those thresholds to be used by the IDMC.

4 Simulation study

In this section, we describe a simulation study to show the operating characteristics of adaptive subpopulation selection designs based on the interim decision-making strategies presented in Section 3. The design assumption and simulation setting are presented in Section 4.1. Furthermore, each probability at the interim or the final decision is shown in Section 4.2.

4.1 Design assumption and simulation setting

We shall consider a randomized, parallel group clinical trial with two arms, experimental ($E$) and control ($C$). Assume that the median TTP for $C$ is 2 months and the median PPS for $C$ is set to (0.5, 0.5, 0.5) or (0.5, 2.0, 3.0) months, corresponding to the correlation between OS and PFS $\rho = (0.9, 0.7, 0.5)$. The clinical trial consists of two stages, namely, a randomized phase II trial and a confirmatory phase III trial.

Assume that the patients classified as being either biomarker-positive or biomarker-negative are included in the trial in order to consider a targeted therapy. Suppose that the prevalence of $P$ among $F$ is roughly set as $\gamma = 50\%$. Let $HR^{(v)}_i$ denote the hazard ratio for each population $t \in \{P, N\}$ using the interim $v \in \{TTP, PPS, D\}$ data. We consider several scenarios for the treatment effect:

\[
\begin{align*}
\text{(Scenario 1):} & \quad HR^{(P)}_1 = 0.50 \quad \text{and} \quad HR^{(N)}_1 = 0.90 \\
\text{(Scenario 2):} & \quad HR^{(P)}_2 = 0.50 \quad \text{and} \quad HR^{(N)}_2 = 1.00 \\
\text{(Scenario 3):} & \quad HR^{(P)}_3 = 0.50 \quad \text{and} \quad HR^{(N)}_3 = 1.11 \\
\text{(Scenario 4):} & \quad HR^{(P)}_4 = 0.50 \quad \text{and} \quad HR^{(N)}_4 = 1.43
\end{align*}
\]

where a hazard ratio less than 1 indicates an increased benefit from $E$. Scenarios 1 to 4 are those in which the experimental treatment is extremely beneficial for $P$, i.e., $HR^{(P)}_1 = 0.50$, and the hazard ratio for $P$ for TTP, PPS, and D is similar for the sake of simplicity. In scenario 1, $E$ is more beneficial for $P$ than for $N$. Scenario 2 is the scenario in which $E$ is beneficial for $P$ but not for $N$. In scenarios 3 and 4, $E$ is beneficial for $P$ but actually harmful for $N$. In calculating the predictive power in Section 3.3 it is required to take into account $N$ as well as $F$ and $P$ in these scenarios. Furthermore, the hazard ratio for $F$ is considered as $HR^{(F)}_v = \exp\{\gamma \cdot \log HR^{(P)}_v + (1 - \gamma) \cdot \log HR^{(N)}_v\}$.

Note that the scenario 3 above is roughly based on a real trial for patients with advanced colorectal cancer receiving cetuximab \cite{example}. That result shows that the hazard ratios among patients with the wild-type K-ras gene are $HR_{PFS}^{(P)} = 0.40$ and $HR_{OS}^{(P)} = 0.55$ while the hazard ratios among patients with mutated K-ras tumors are $HR_{PFS}^{(N)} = 0.99$ and $HR_{OS}^{(N)} = 0.98$, respectively.

Moreover, we assume that the final analysis is performed after 300 OS events occur in reference to the example \cite{example}. An interim analysis is conducted after $\tau = (25\%, 33\%, 50\%)$ PFS events of pre-planned OS events are observed. Here, we assume that the overall number of patients is 400.

In terms of interim decision-making, the thresholds that needed to be pre-specified at the planning stage are roughly set as $(\pi^{(P)}_1, \pi^{(N)}_1) = (10\%, 5\%)$ and $\pi^{(P)}_2 = 20\%$ with weights of $\delta_{OS} = 1/2$ and $\delta_{PFS} = 1$, respectively, because OS is a primary endpoint whereas PFS data are more rapidly observed than OS data.

With respect to data generation within the simulation, each time-to-event endpoint is assumed to arise from the exponential distribution considering the correlation between OS and PFS as well as the impact of PPS after tumor...
progression using Equation (2) as we demonstrated in Section 3.2. Selected results for the simulation are presented in Section 4.2. Furthermore, the comparison of approaches that use OS data only or PFS data only at the interim analysis is also given.

4.2 Simulation results

All subsequently reported results are obtained based on 10,000 simulation replications per scenario.

First, Table 1 presents the probability of rejecting at least one null hypothesis for $F$ or $P$ for $HR_{OS}^{(P)} = HR_{PFS}^{(P)} = HR_{OS}^{(N)} = HR_{PFS}^{(N)} = 1.00$, under the assumption that the median PPS for $C$ is set to $(0.5, 0.5, 0.5)$ months. It demonstrates that the familywise type I error rate is controlled at less than 2.5% across all scenarios. The correlation of the p-values between $F$ and $P$ and the independence of stagewise p-values across stages are provided to confirm that the assumption of the Simes’ procedure and the weighted inverse normal combination method are met. The correlation of p-values between $F$ and $P$ is positive and the independence between stage 1 and stage 2 is also illustrated under all scenarios.

Second, Tables 2 and 3 show the each probability of selecting each corresponding population at the interim analysis. The lower probability is better in Table 2, whereas the higher is better in Table 3 because of the simulation setting in which the targeted therapy is considered. It is expected to see a greater probability of selecting each population when using OS and PFS, under the assumption that the median PPS for $C$ is set to $(0.5, 2.0, 3.0)$ months, particularly where the correlation between OS and PFS is not very strong, i.e., $\rho = (0.7, 0.5)$, in scenarios 1 to 3. These indicate that the impact of PPS leads to the misspecification of the population rather than the correlation between OS and PFS. On the other hand, each probability is similar under the assumption that median PPS for $C$ is set to $(0.5, 0.5, 0.5)$ months for every scenario. In addition, the probability of discontinuing the trial early for futility is less than 10% under every scenario in Table 4, and this would be valid in terms of the targeted therapy setting for simulations, although it depends on the pre-specified thresholds.

Lastly, Tables 5 and 6 show the each probability of rejecting the null hypothesis, i.e., power, for each population at the final analysis. The weighted inverse normal combination method is used here (see, Equation (1) in Section 2). Note that the lower power is better because it is beneficial only for $P$ in Table 5. Therefore, incorporation of information from OS and PFS results in good performance in the same way as described in Tables 2 and 3. Furthermore, each overall power gets lower when the timing of the interim analysis $\tau$ is earlier, because of the immature OS and PFS data. However, an interim decision conducted after the observation of PFS events constituting 50% of the pre-planned OS events would be unrealistic, due to the adaptive seamless phase II/III clinical trial design considered. More patients may be needed in order to construct the adaptive seamless phase II/III design.

5 Discussion

The aim of this paper was to propose the interim decision-making strategies in adaptive designs for subpopulation selection. Interim decision-making strategies using OS as well as PFS perform well in terms of selecting the most appropriate population, especially under a scenario in which PPS affects the translation of the benefit from PFS to OS, although there are concerns about bias with respect to operational facets.

This paper has been limited to a situation in which the source of the pre-defined subgroup regarded as biomarker-positive for efficacy is known. Moreover, a restriction in our simulation results is that sample size calculation is not considered for the sake of simplicity. In practice, this is determined based on the expected treatment effect for both $F$ and $P$. However, a large number of patients would be required when the experimental treatment benefits $P$ but has a negative effect on $N$, especially in scenario 4 of Section 4.1. This is a realistic scenario referring to an actual trial for gefitinib [2]. Alternative clinical trial designs that enroll only $P$, i.e., enrichment designs [1], could be considered, particularly in scenarios in which the prevalence of $P$ is small. Meanwhile, note that the aforementioned approach meets the regulatory requirement in that it provides the efficacy results for $N$.

In oncology, whether PFS instead of OS as a primary endpoint is acceptable for a given treatment evaluation for marketing approval will depend on the disease setting. Nevertheless, we consider PFS data only in making interim decision to supplement the immature OS data. Furthermore, we also consider PPS lengths in using the correlation between OS and PFS. Consequently, the probability of selecting the proper population at the interim analysis is improved under the circumstance in which a relatively long PPS is expected. However, as Zhang et al. [12] have mentioned, the OS benefit given the PFS benefit also depends largely on the crossover rate, because treatment crossover from the control to experimental groups frequently occurs right after tumor progression in real trials.
Regarding the correlation of multiple time-to-event endpoints, it would be worthwhile to consider the impact of crossover rates in addition to the magnitude of PPS.

Regarding multiplicity issues, we use the Simes’ procedure to control the familywise type I error rate. As presented in Section 4.2, however, these results are conservative because of the asymptotic bivariate normal distribution and positive correlation. As Friede et al. [19] have proposed, the method based on the conditional error function approach would be more powerful.

As shown in Section 3.3, we assume non-informative priors at the time of interim decision-making for simplicity in this paper. A possible future investigation is to consider how external information or any other knowledge prior to the phase II trial might be incorporated into the use of informative priors and setting the thresholds for the predictive power in the interim decision.

References

[1] Simon R. Genomic Clinical Trials and Predictive Medicine. Cambridge University Press: New York, 2013.

[2] Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, Saijo N, Sunpaweravong P, Han B, Margono B, Ichinose Y, Nishiwaki Y, Ohe Y, Yang JJ, Chewaskulyong B, Jiang H, Duffield EL, Watkins CL, Armour AA, Fukuoka M. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. New England Journal of Medicine 2009; 361:947-957.

[3] Karapetis CS, Khambata-Ford S, Jonker DJ, O'Callaghan CJ, Tu D, Tebbutt NC, Simes RJ, Chalchal H, Shapiro JD, Robitaille S, Price TJ, Shepherd L, Au HJ, Langer C, Moore MJ, Zalcberg JR. K-ras mutations and benefit from cetuximab in advanced colorectal cancer. New England Journal of Medicine 2008; 359:1757-1765.

[4] Food and Drug Administration. Guidance for industry: adaptive design clinical trials for drugs and biologics. U.S. Food and Drug Administration 2010, draft.

[5] Schmidli H, Bretz F, Racine A, Maurer W. Confirmatory seamless phase II/III clinical trials with hypotheses selection at interim: applications and practical considerations. Biometrical Journal 2006; 48:635-643.

[6] Todd S, Stallard N. A new clinical trial design combining phases 2 and 3: sequential designs with treatment selection and a change of endpoint. Drug Information Journal 2005; 39:109-118.

[7] Kelly PJ, Stallard N, Todd S. An adaptive group sequential design for phase II/III clinical trials that select a single treatment from several. Journal of Biopharmaceutical Statistics 2005; 15:641-658.

[8] Bretz F, Schmidli H, Koenig F, Racine A, Maurer W. Confirmatory seamless phase II/III clinical trials with hypotheses selection at interim: general concepts. Biometrical Journal 2006; 48:623-634.

[9] Jennison C, Turnbull BW. Confirmatory seamless phase II/III clinical trials with hypotheses selection at interim: opportunities and limitations. Biometrical Journal 2006; 48:650-655.

[10] Jennison C, Turnbull BW. Adaptive seamless designs: selection and prospective testing of hypotheses. Journal of Biopharmaceutical Statistics 2007; 17:1135-1161.

[11] Schmidli H, Bretz F, Racine-Poon A. Bayesian predictive power for interim adaptation in seamless phase II/III trials where the endpoint is survival up to some specified timepoint. Statistics in Medicine 2007; 27:4925-4938.

[12] Bischoff W, Miller F. A seamless phase II/III design with sample-size re-estimation. Journal of Biopharmaceutical Statistics 2009; 19:595-609.

[13] Brannath W, Zuber E, Branson M, Bretz F, Gallo P, Posch M, Racine-Poon A. Confirmatory adaptive designs with Bayesian decision tools for targeted therapy in oncology. Statistics in Medicine 2009; 28:1445-1463.

[14] Stallard N. A confirmatory seamless phase II/III clinical trial design incorporating short-term endpoint information. Statistics in Medicine 2010; 30:959-971.

[15] Jenkins M, Stone A, Jennison C. An adaptive seamless phase II/III design for oncology trials with subpopulation selection using correlated survival endpoints. Pharmaceutical Statistics 2011; 10:347-356.
[16] Friede T, Parsons N, Stallard N, Todd S, Valdes Marquez E, Chataway J, Nicholas R. Designing a seamless phase II/III clinical trial using early outcomes for treatment selection: an application in multiple sclerosis. *Statistics in Medicine* 2011; **30**:1528-1540.

[17] Uozumi R, Hamada C. An Adaptive subpopulation selection design using time-to-event endpoints. *Abstracts for the XXVIth International Biometric Conference*, Kobe, Japan, 26-31 August 2012.

[18] Kimani PK, Glimm E, Maurer W, Hutton JL, Stallard N. Practical guidelines for adaptive seamless phase II/III clinical trials that use Bayesian methods. *Statistics in Medicine* 2012; **31**:2068-2085.

[19] Friede T, Parsons N, Stallard N. A conditional error function approach for subgroup selection in adaptive clinical trials. *Statistics in Medicine* 2012; **31**:4309-4320.

[20] Bauer P, Köhne K. Evaluation of experiments with adaptive interim analyses. *Biometrics* 1994; **50**:1029-1041.

[21] Lehmacher W, Wassmer G. Adaptive sample size calculation in group sequential trials. *Biometrics* 1999; **55**:853-857.

[22] Müller HH, Schäfer H. Adaptive group sequential designs for clinical trials: combining the advantage of adaptive and of classical group sequential approaches. *Biometrics* 2001; **57**:886-891.

[23] Marcus R, Peritz E, Gabriel KR. On closed testing procedures with special reference to ordered analysis of variance. *Biometrika* 1976; **63**:655-660.

[24] Lehmacher W, Wassmer G. Adaptive sample size calculation in group sequential trials. *Biometrics* 1999; **55**:1286-1290.

[25] Simes RJ. An improved Bonferroni procedure for multiple tests of significance. *Biometrika* 1986; **73**:751-754.

[26] Food and Drug Administration. Guidance for industry: clinical trial endpoints for the approval of cancer drugs and biologics. *Biotechnology Law Report* 2007; **26**:375-386.

[27] Fleming TR, Rothmann MD, Lu HL. Issues in using progression-free survival when evaluating oncology products. *Journal of Clinical Oncology* 2009; **27**:2874-2880.

[28] Saad ED, Katz A, Hoff PM, Buyse M. Progression-free survival as surrogate and as true end point: insights from the breast and colorectal cancer literature. *Annals of Oncology* 2010; **21**:7-12.

[29] Kay R, Wu J, Wittes J. On assessing the presence of evaluation-time bias in progression-free survival in randomized trials. *Pharmaceutical Statistics* 2011; **10**:213-217.

[30] Booth CM, Eisenhauer EA. Progression-free survival: meaningful or simply measurable? *Journal of Clinical Oncology* 2012; **30**:1030-1033.

[31] Food and Drug Administration. Guidance for clinical trial sponsors on the establishment and operation of clinical trial data monitoring committees. *U.S. Food and Drug Administration* 2006.

[32] Gumbel EJ. Bivariate exponential distributions. *Journal of the American Statistical Association* 1960; **55**:698-707.

[33] Marshall AW, Olkin I. A generalized bivariate exponential distribution. *Journal of Applied Probability* 1967; **4**:291-302.

[34] Clayton DG. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease. *Biometrika* 1978; **65**:141-151.

[35] Frank MJ. On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$. *Aequationes Mathematicae* 1979; **19**:194-226.

[36] Fleischer F, Gaschler-Markefski B, Bluhmki E. A statistical model for the dependence between progression-free survival and overall survival. *Statistics in Medicine* 2009; **28**:2669-2686.

[37] Spearman C. The proof and measurement of association between two things. *American Journal of Psychology* 1904; **15**:72-101.
[38] Tang PA, Bentzen SM, Chen EX, Siu LL. Surrogate end points for median overall survival in metastatic colorectal cancer: literature-based analysis from 39 randomized controlled trials of first-line chemotherapy. *Journal of Clinical Oncology* 2007; 25:4562-4568.

[39] Petrelli F, Barni S. Correlation of progression-free and post-progression survival with overall survival in advanced colorectal cancer. *Annals of Oncology* 2013; 24:186-192.

[40] Dmitrienko A, Wang MD. Bayesian predictive approach to interim monitoring in clinical trials. *Statistics in Medicine* 2006; 25:2178-2195.

[41] Di Scala L, Glimm E. Time-to-event analysis with treatment arm selection at interim. *Statistics in Medicine* 2011; 30:3067-3081.

[42] Zhang L, Ko CW, Tang S, Sridhara R. Relationship between progression-free survival and overall survival benefit: a simulation study. *Therapeutic Innovation & Regulatory Science* 2013; 47:95-100.
Table 1: The familywise type I error rate, the correlation of the p-values between $F$ and $P$, and the independence of stagewise p-values across stages under the assumption that the median PPS for $C$ is set to $(0.5,0.5,0.5)$ months based on 10,000 simulation replications per scenario.

| $\tau$ | $\rho$ | $p_{1}^{(F)}$ and $p_{2}^{(F)}$ | $p_{1}^{(P)}$ and $p_{2}^{(P)}$ | $p_{1}^{(F,P)}$ and $p_{2}^{(F,P)}$ | $p_{1}^{(F)}$ and $p_{1}^{(P)}$ | $p_{2}^{(F)}$ and $p_{2}^{(P)}$ | Familywise type I error rate |
|--------|--------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| 25%    | 0.90   | -0.0001           | 0.0079            | 0.0039            | 0.6261            | 0.3351            | 0.0159            |
|        | 0.70   | 0.0004            | 0.0034            | 0.0051            | 0.6330            | 0.3240            | 0.0142            |
|        | 0.50   | 0.0069            | 0.0038            | 0.0072            | 0.6277            | 0.3325            | 0.0155            |
| 33%    | 0.90   | 0.0157            | 0.0003            | 0.0016            | 0.6207            | 0.4220            | 0.0125            |
|        | 0.70   | 0.0169            | 0.0041            | 0.0088            | 0.6128            | 0.4121            | 0.0124            |
|        | 0.50   | 0.0088            | -0.0043           | -0.0009           | 0.6095            | 0.4180            | 0.0134            |
| 50%    | 0.90   | 0.0094            | 0.0162            | 0.0111            | 0.6303            | 0.6269            | 0.0179            |
|        | 0.70   | 0.0022            | -0.0002           | -0.0002           | 0.6355            | 0.6278            | 0.0154            |
|        | 0.50   | 0.0069            | -0.0013           | -0.0020           | 0.6213            | 0.6295            | 0.0189            |
Table 2: Probabilities of selecting each population at the interim analysis using the corresponding data under the assumption that the median PPS for $C$ is set to (0.5, 0.5, 0.5) months based on 10,000 simulation replications per scenario.

| $\tau$ | $HR_{0.5}^{[N]}$ | $\rho$ | Probabilities of selecting $F$ using PFS only | Probabilities of selecting OS only | Probabilities of selecting OS and PFS | Probabilities of selecting $P$ using PFS only | Probabilities of selecting OS only | Probabilities of selecting OS and PFS |
|--------|-----------------|--------|-------------------------------------------|-------------------------------|-----------------------------------|-------------------------------------------|-------------------------------|-----------------------------------|
| 25%    | 0.90            | 0.90   | 0.7718                                    | 0.7838                        | 0.7463                            | 0.2120                                    | 0.2032                        | 0.2360                            |
|        | 0.70            | 0.90   | 0.7703                                    | 0.7606                        | 0.7407                            | 0.2127                                    | 0.2159                        | 0.2386                            |
|        | 0.50            | 0.90   | 0.7687                                    | 0.7609                        | 0.7523                            | 0.2140                                    | 0.2168                        | 0.2275                            |
| 1.00   | 0.90            | 0.6611 | 0.6658                                    | 0.6246                        | 0.3182                            | 0.3166                                    | 0.3524                        |                                    |
|        | 0.70            | 0.6597 | 0.6614                                    | 0.6258                        | 0.3197                            | 0.3091                                    | 0.3487                        |                                    |
|        | 0.50            | 0.6596 | 0.6559                                    | 0.6436                        | 0.3182                            | 0.3166                                    | 0.3317                        |                                    |
| 1.11   | 0.90            | 0.5379 | 0.5348                                    | 0.4919                        | 0.4371                            | 0.4441                                    | 0.4800                        |                                    |
|        | 0.70            | 0.5321 | 0.5462                                    | 0.4985                        | 0.4405                            | 0.4194                                    | 0.4688                        |                                    |
|        | 0.50            | 0.5334 | 0.5405                                    | 0.5169                        | 0.4395                            | 0.4277                                    | 0.4537                        |                                    |
| 1.43   | 0.90            | 0.2338 | 0.2076                                    | 0.1900                        | 0.7302                            | 0.7613                                    | 0.7702                        |                                    |
|        | 0.70            | 0.2355 | 0.2585                                    | 0.2096                        | 0.7282                            | 0.6935                                    | 0.7474                        |                                    |
|        | 0.50            | 0.2390 | 0.2576                                    | 0.2224                        | 0.7238                            | 0.6989                                    | 0.7372                        |                                    |
| 33%    | 0.90            | 0.90   | 0.7172                                    | 0.7375                        | 0.6879                            | 0.2704                                    | 0.2526                        | 0.2973                            |
|        | 0.70            | 0.7188 | 0.7095                                    | 0.6853                        | 0.2678                            | 0.2714                                    | 0.2977                        |                                    |
|        | 0.50            | 0.7185 | 0.7138                                    | 0.7038                        | 0.2687                            | 0.2686                                    | 0.2819                        |                                    |
| 1.00   | 0.90            | 0.5777 | 0.5795                                    | 0.5394                        | 0.4048                            | 0.4068                                    | 0.4407                        |                                    |
|        | 0.70            | 0.5884 | 0.5827                                    | 0.5498                        | 0.3950                            | 0.3933                                    | 0.4294                        |                                    |
|        | 0.50            | 0.5829 | 0.5842                                    | 0.5639                        | 0.4008                            | 0.3933                                    | 0.4179                        |                                    |
| 1.11   | 0.90            | 0.4269 | 0.4101                                    | 0.3764                        | 0.5508                            | 0.5733                                    | 0.5986                        |                                    |
|        | 0.70            | 0.4285 | 0.4402                                    | 0.3931                        | 0.5505                            | 0.5302                                    | 0.5803                        |                                    |
|        | 0.50            | 0.4325 | 0.4373                                    | 0.4132                        | 0.5470                            | 0.5344                                    | 0.5640                        |                                    |
| 1.43   | 0.90            | 0.1274 | 0.1049                                    | 0.0959                        | 0.8420                            | 0.8686                                    | 0.8699                        |                                    |
|        | 0.70            | 0.1300 | 0.1496                                    | 0.1103                        | 0.8388                            | 0.8074                                    | 0.8521                        |                                    |
|        | 0.50            | 0.1291 | 0.1483                                    | 0.1183                        | 0.8412                            | 0.8150                                    | 0.8489                        |                                    |
| 50%    | 0.90            | 0.90   | 0.5978                                    | 0.6179                        | 0.5545                            | 0.3915                                    | 0.3757                        | 0.4337                            |
|        | 0.70            | 0.5953 | 0.5846                                    | 0.5479                        | 0.3936                            | 0.3958                                    | 0.4358                        |                                    |
|        | 0.50            | 0.5923 | 0.5871                                    | 0.5660                        | 0.3955                            | 0.3964                                    | 0.4197                        |                                    |
| 1.00   | 0.90            | 0.4067 | 0.4104                                    | 0.3595                        | 0.5795                            | 0.5804                                    | 0.6251                        |                                    |
|        | 0.70            | 0.4110 | 0.4115                                    | 0.3643                        | 0.5744                            | 0.5639                                    | 0.6158                        |                                    |
|        | 0.50            | 0.4139 | 0.4125                                    | 0.3867                        | 0.5701                            | 0.5655                                    | 0.5956                        |                                    |
| 1.11   | 0.90            | 0.2408 | 0.2293                                    | 0.1973                        | 0.7428                            | 0.7587                                    | 0.7838                        |                                    |
|        | 0.70            | 0.2439 | 0.2518                                    | 0.2087                        | 0.7375                            | 0.7171                                    | 0.7669                        |                                    |
|        | 0.50            | 0.2435 | 0.2509                                    | 0.2242                        | 0.7367                            | 0.7226                                    | 0.7539                        |                                    |
| 1.43   | 0.90            | 0.0310 | 0.0204                                    | 0.0183                        | 0.9496                            | 0.9626                                    | 0.9587                        |                                    |
|        | 0.70            | 0.0295 | 0.0388                                    | 0.0227                        | 0.9469                            | 0.9218                                    | 0.9469                        |                                    |
|        | 0.50            | 0.0298 | 0.0391                                    | 0.0258                        | 0.9454                            | 0.9274                                    | 0.9474                        |                                    |
Table 3: Probabilities of selecting each population at the interim analysis using the corresponding data under the assumption that the median PPS for $C$ is set to (0.5, 2.0, 3.0) months based on 10,000 simulation replications per scenario.

| $\tau$ | $HR_{3}^{(N)}$ | $\rho$ | Probabilities of selecting $F$ using | Probabilities of selecting $P$ using |
|--------|-----------------|--------|-------------------------------------|-------------------------------------|
|        |                 |        | PFS only | OS only | OS and PFS | PFS only | OS only | OS and PFS |
| 25%    | 0.90            | 0.90   | 0.7718   | 0.7838  | 0.7463     | 0.2120   | 0.2032  | 0.2360     |
|        | 0.70            | 0.7794 | 0.7970   | 0.7213  |             | 0.2043   | 0.1953  | 0.2589     |
|        | 0.50            | 0.7790 | 0.7635   | 0.6950  |             | 0.2059   | 0.2089  | 0.2766     |
| 1.00   | 0.90            | 0.6611 | 0.6658   | 0.6246  |             | 0.3182   | 0.3166  | 0.3524     |
|        | 0.70            | 0.6684 | 0.6673   | 0.5830  |             | 0.3117   | 0.3208  | 0.3923     |
|        | 0.50            | 0.6710 | 0.6642   | 0.5689  |             | 0.3094   | 0.3009  | 0.3943     |
| 1.11   | 0.90            | 0.5379 | 0.5348   | 0.4919  |             | 0.4371   | 0.4441  | 0.4800     |
|        | 0.70            | 0.5461 | 0.5140   | 0.4373  |             | 0.4293   | 0.4706  | 0.5323     |
|        | 0.50            | 0.5397 | 0.5394   | 0.4225  |             | 0.4362   | 0.4179  | 0.5327     |
| 1.43   | 0.90            | 0.2338 | 0.2076   | 0.1900  |             | 0.7302   | 0.7613  | 0.7702     |
|        | 0.70            | 0.2343 | 0.1648   | 0.1256  |             | 0.7315   | 0.8091  | 0.8319     |
|        | 0.50            | 0.2294 | 0.2390   | 0.1371  |             | 0.7363   | 0.6983  | 0.8020     |
| 33%    | 0.90            | 0.7172 | 0.7375   | 0.6879  |             | 0.2704   | 0.2526  | 0.2973     |
|        | 0.70            | 0.7251 | 0.7392   | 0.6563  |             | 0.2602   | 0.2520  | 0.3248     |
|        | 0.50            | 0.7214 | 0.7022   | 0.6220  |             | 0.2648   | 0.2674  | 0.3470     |
| 1.00   | 0.90            | 0.5777 | 0.5795   | 0.5394  |             | 0.4048   | 0.4068  | 0.4407     |
|        | 0.70            | 0.5811 | 0.5774   | 0.4848  |             | 0.3997   | 0.4096  | 0.4899     |
|        | 0.50            | 0.5750 | 0.5701   | 0.4603  |             | 0.4066   | 0.3923  | 0.5007     |
| 1.11   | 0.90            | 0.4269 | 0.4101   | 0.3764  |             | 0.5508   | 0.5733  | 0.5986     |
|        | 0.70            | 0.4317 | 0.3935   | 0.3169  |             | 0.5447   | 0.5895  | 0.6529     |
|        | 0.50            | 0.4201 | 0.4267   | 0.3013  |             | 0.5566   | 0.5298  | 0.6508     |
| 1.43   | 0.90            | 0.1274 | 0.1049   | 0.0959  |             | 0.8420   | 0.8686  | 0.8699     |
|        | 0.70            | 0.1231 | 0.0811   | 0.0541  |             | 0.8429   | 0.8919  | 0.9029     |
|        | 0.50            | 0.1241 | 0.1295   | 0.0605  |             | 0.8446   | 0.8021  | 0.8792     |
| 50%    | 0.90            | 0.5978 | 0.6179   | 0.5545  |             | 0.3915   | 0.3757  | 0.4337     |
|        | 0.70            | 0.6014 | 0.6380   | 0.5186  |             | 0.3893   | 0.3577  | 0.4687     |
|        | 0.50            | 0.5985 | 0.5873   | 0.4722  |             | 0.3927   | 0.3885  | 0.5040     |
| 1.00   | 0.90            | 0.4067 | 0.4104   | 0.3595  |             | 0.5795   | 0.5804  | 0.6251     |
|        | 0.70            | 0.4135 | 0.4137   | 0.3086  |             | 0.5743   | 0.5784  | 0.6745     |
|        | 0.50            | 0.4122 | 0.4149   | 0.2872  |             | 0.5757   | 0.5515  | 0.6833     |
| 1.11   | 0.90            | 0.2408 | 0.2293   | 0.1973  |             | 0.7428   | 0.7587  | 0.7838     |
|        | 0.70            | 0.2411 | 0.2159   | 0.1441  |             | 0.7427   | 0.7745  | 0.8346     |
|        | 0.50            | 0.2425 | 0.2475   | 0.1394  |             | 0.7428   | 0.7112  | 0.8237     |
| 1.43   | 0.90            | 0.0310 | 0.0204   | 0.0183  |             | 0.9496   | 0.9626  | 0.9587     |
|        | 0.70            | 0.0339 | 0.0147   | 0.0076  |             | 0.9458   | 0.9719  | 0.9680     |
|        | 0.50            | 0.0345 | 0.0324   | 0.0102  |             | 0.9463   | 0.9121  | 0.9440     |
Table 4: Probabilities of discontinuing the trial early for futility at the interim analysis using the corresponding data based on 10,000 simulation replications per scenario.

| $\tau$ | $HR_v^{(N)}$ | $\rho$ | Probabilities of discontinuing the trial early for futility under the assumption that median PPS for $C = (0.5, 0.5, 0.5)$ months | median PPS for $C = (0.5, 2.0, 3.0)$ months |
|-------|---------------|-------|-------------------------------------------------|--------------------------------------------|
|       |               |       | PFS only | OS only | OS and PFS | PFS only | OS only | OS and PFS |
| 25%   | 0.90          | 0.90  | 0.0162   | 0.0130  | 0.0177     | 0.0162   | 0.0130  | 0.0177     |
|       | 0.70          | 0.0170 | 0.0235   | 0.0207  | 0.0163     | 0.0163   | 0.0077  | 0.0198     |
|       | 0.50          | 0.0173 | 0.0223   | 0.0202  | 0.0151     | 0.0196   | 0.0349  | 0.0284     |
| 1.00  | 0.90          | 0.0207 | 0.0176   | 0.0230  | 0.0207     | 0.0199   | 0.0119  | 0.0247     |
|       | 0.70          | 0.0206 | 0.0295   | 0.0255  | 0.0199     | 0.0250   | 0.0211  | 0.0281     |
|       | 0.50          | 0.0222 | 0.0275   | 0.0247  | 0.0196     | 0.0196   | 0.0349  | 0.0368     |
| 1.11  | 0.90          | 0.0250 | 0.0211   | 0.0281  | 0.0250     | 0.0250   | 0.0211  | 0.0281     |
|       | 0.70          | 0.0274 | 0.0344   | 0.0327  | 0.0246     | 0.0246   | 0.0154  | 0.0304     |
|       | 0.50          | 0.0271 | 0.0318   | 0.0294  | 0.0241     | 0.0241   | 0.0427  | 0.0448     |
| 1.43  | 0.90          | 0.0360 | 0.0311   | 0.0398  | 0.0360     | 0.0360   | 0.0311  | 0.0398     |
|       | 0.70          | 0.0363 | 0.0480   | 0.0430  | 0.0342     | 0.0342   | 0.0261  | 0.0425     |
|       | 0.50          | 0.0372 | 0.0435   | 0.0404  | 0.0343     | 0.0343   | 0.0627  | 0.0609     |
| 33%   | 0.90          | 0.0124 | 0.0099   | 0.0148  | 0.0124     | 0.0124   | 0.0099  | 0.0148     |
|       | 0.70          | 0.0134 | 0.0191   | 0.0170  | 0.0147     | 0.0147   | 0.0088  | 0.0189     |
|       | 0.50          | 0.0128 | 0.0176   | 0.0143  | 0.0138     | 0.0138   | 0.0304  | 0.0310     |
| 1.00  | 0.90          | 0.0175 | 0.0137   | 0.0199  | 0.0175     | 0.0175   | 0.0137  | 0.0199     |
|       | 0.70          | 0.0166 | 0.0240   | 0.0208  | 0.0192     | 0.0192   | 0.0130  | 0.0253     |
|       | 0.50          | 0.0163 | 0.0225   | 0.0182  | 0.0184     | 0.0184   | 0.0376  | 0.0390     |
| 1.11  | 0.90          | 0.0223 | 0.0166   | 0.0250  | 0.0223     | 0.0223   | 0.0166  | 0.0250     |
|       | 0.70          | 0.0210 | 0.0296   | 0.0266  | 0.0236     | 0.0236   | 0.0170  | 0.0302     |
|       | 0.50          | 0.0205 | 0.0283   | 0.0228  | 0.0233     | 0.0233   | 0.0435  | 0.0479     |
| 1.43  | 0.90          | 0.0306 | 0.0265   | 0.0342  | 0.0306     | 0.0306   | 0.0265  | 0.0342     |
|       | 0.70          | 0.0312 | 0.0430   | 0.0376  | 0.0340     | 0.0340   | 0.0270  | 0.0430     |
|       | 0.50          | 0.0297 | 0.0367   | 0.0328  | 0.0313     | 0.0313   | 0.0684  | 0.0603     |
| 50%   | 0.90          | 0.0107 | 0.0064   | 0.0118  | 0.0107     | 0.0107   | 0.0064  | 0.0118     |
|       | 0.70          | 0.0111 | 0.0196   | 0.0163  | 0.0093     | 0.0093   | 0.0043  | 0.0127     |
|       | 0.50          | 0.0122 | 0.0165   | 0.0143  | 0.0088     | 0.0088   | 0.0242  | 0.0238     |
| 1.00  | 0.90          | 0.0138 | 0.0092   | 0.0154  | 0.0138     | 0.0138   | 0.0092  | 0.0154     |
|       | 0.70          | 0.0146 | 0.0246   | 0.0199  | 0.0122     | 0.0122   | 0.0079  | 0.0169     |
|       | 0.50          | 0.0160 | 0.0220   | 0.0177  | 0.0121     | 0.0121   | 0.0336  | 0.0295     |
| 1.11  | 0.90          | 0.0164 | 0.0120   | 0.0189  | 0.0164     | 0.0164   | 0.0120  | 0.0189     |
|       | 0.70          | 0.0186 | 0.0311   | 0.0244  | 0.0162     | 0.0162   | 0.0096  | 0.0213     |
|       | 0.50          | 0.0198 | 0.0265   | 0.0219  | 0.0147     | 0.0147   | 0.0413  | 0.0369     |
| 1.43  | 0.90          | 0.0194 | 0.0170   | 0.0230  | 0.0194     | 0.0194   | 0.0170  | 0.0230     |
|       | 0.70          | 0.0236 | 0.0394   | 0.0304  | 0.0203     | 0.0203   | 0.0134  | 0.0244     |
|       | 0.50          | 0.0248 | 0.0335   | 0.0268  | 0.0192     | 0.0192   | 0.0555  | 0.0458     |
Table 5: Probabilities of rejecting the null hypothesis for each population at the final analysis using the corresponding data under the assumption that the median PPS for $C$ is set to $(0.5, 0.5, 0.5)$ months based on 10,000 simulation replications per scenario.

| $\tau$ | $HR_c^{(x)}$ | $\rho$ | Probabilities of rejecting $H_0^{(P)}$ using PFS only | Probabilities of rejecting $H_0^{(P)}$ using OS only | Probabilities of rejecting $H_0^{(P)}$ using OS and PFS |
|--------|---------------|--------|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|
| 25%    | 0.90          | 0.90   | 0.5136                                          | 0.5224                                          | 0.5020                                          |
|        | 0.70          | 0.3989 | 0.4012                                          | 0.3893                                          | 0.1799                                          |
|        | 0.50          | 0.4027 | 0.4037                                          | 0.3985                                          | 0.1491                                          |
| 1.00   | 0.90          | 0.2348 | 0.2399                                          | 0.2283                                          | 0.2687                                          |
|        | 0.70          | 0.1863 | 0.1921                                          | 0.1822                                          | 0.2320                                          |
|        | 0.50          | 0.1923 | 0.1951                                          | 0.1897                                          | 0.2399                                          |
| 1.11   | 0.90          | 0.0649 | 0.0657                                          | 0.0620                                          | 0.3615                                          |
|        | 0.70          | 0.0620 | 0.0675                                          | 0.0614                                          | 0.3072                                          |
|        | 0.50          | 0.0619 | 0.0651                                          | 0.0616                                          | 0.3127                                          |
| 1.43   | 0.90          | 0.0003 | 0.0002                                          | 0.0002                                          | 0.5952                                          |
|        | 0.70          | 0.0009 | 0.0010                                          | 0.0009                                          | 0.4942                                          |
|        | 0.50          | 0.0014 | 0.0015                                          | 0.0014                                          | 0.4954                                          |
| 33%    | 0.90          | 0.5969 | 0.6154                                          | 0.5977                                          | 0.2530                                          |
|        | 0.70          | 0.6194 | 0.5689                                          | 0.5759                                          | 0.2373                                          |
|        | 0.50          | 0.6037 | 0.5703                                          | 0.5778                                          | 0.2373                                          |
| 1.00   | 0.90          | 0.3443 | 0.3514                                          | 0.3309                                          | 0.3770                                          |
|        | 0.70          | 0.2836 | 0.2902                                          | 0.2745                                          | 0.3302                                          |
|        | 0.50          | 0.2927 | 0.2989                                          | 0.2879                                          | 0.3373                                          |
| 1.11   | 0.90          | 0.1374 | 0.1387                                          | 0.1297                                          | 0.5107                                          |
|        | 0.70          | 0.1190 | 0.1273                                          | 0.1144                                          | 0.4585                                          |
|        | 0.50          | 0.1229 | 0.1282                                          | 0.1197                                          | 0.4585                                          |
| 1.43   | 0.90          | 0.0018 | 0.0016                                          | 0.0018                                          | 0.7738                                          |
|        | 0.70          | 0.0042 | 0.0049                                          | 0.0041                                          | 0.6868                                          |
|        | 0.50          | 0.0044 | 0.0044                                          | 0.0043                                          | 0.6974                                          |
| 50%    | 0.90          | 0.5875 | 0.6085                                          | 0.5467                                          | 0.3886                                          |
|        | 0.70          | 0.5664 | 0.5799                                          | 0.5192                                          | 0.3800                                          |
|        | 0.50          | 0.5602 | 0.5590                                          | 0.5378                                          | 0.3831                                          |
| 1.00   | 0.90          | 0.3864 | 0.3925                                          | 0.3453                                          | 0.5765                                          |
|        | 0.70          | 0.3625 | 0.3710                                          | 0.3281                                          | 0.5528                                          |
|        | 0.50          | 0.3715 | 0.3761                                          | 0.3509                                          | 0.5504                                          |
| 1.11   | 0.90          | 0.2121 | 0.2083                                          | 0.1785                                          | 0.7377                                          |
|        | 0.70          | 0.1920 | 0.2088                                          | 0.1714                                          | 0.7077                                          |
|        | 0.50          | 0.2003 | 0.2103                                          | 0.1882                                          | 0.7100                                          |
| 1.43   | 0.90          | 0.0171 | 0.0226                                          | 0.0117                                          | 0.9410                                          |
|        | 0.70          | 0.0147 | 0.0226                                          | 0.0124                                          | 0.9028                                          |
|        | 0.50          | 0.0153 | 0.0230                                          | 0.0144                                          | 0.9079                                          |
Table 6: Probabilities of rejecting the null hypothesis for each population at the final analysis using the corresponding data under the assumption that the median PPS for $C$ is set to $(0.5, 2.0, 3.0)$ months based on 10,000 simulation replications per scenario.

| $\tau$ | $HR_0^{(T)}$ | $\rho$ | Probabilities of rejecting $H_0^{(T)}$ using $PFS$ only | Probabilities of rejecting $H_0^{(T)}$ using $OS$ only | Probabilities of rejecting $H_0^{(T)}$ using $OS$ and $PFS$ |
|-------|-------------|-------|----------------|----------------|----------------|
| 25%   | 0.90        | 0.90  | 0.5136         | 0.5224         | 0.5020         |
|       |             |       | 0.1799         | 0.1728         | 0.2006         |
|       | 0.70        | 0.6583| 0.6779         | 0.6162         | 0.1976         |
|       |             |       | 0.1890         | 0.2505         |                |
|       | 0.50        | 0.5825| 0.5865         | 0.5330         | 0.1908         |
|       |             |       | 0.1925         | 0.2555         |                |
| 1.00  | 0.90        | 0.2348| 0.2399         | 0.2283         | 0.2687         |
|       |             |       | 0.2658         | 0.2962         |                |
|       | 0.70        | 0.3258| 0.3333         | 0.2969         | 0.3007         |
|       |             |       | 0.3093         | 0.3792         |                |
|       | 0.50        | 0.2691| 0.2836         | 0.2432         | 0.2842         |
|       |             |       | 0.2765         | 0.3629         |                |
| 1.11  | 0.90        | 0.0649| 0.0657         | 0.0620         | 0.3615         |
|       |             |       | 0.3685         | 0.3979         |                |
|       | 0.70        | 0.0797| 0.0809         | 0.0716         | 0.4126         |
|       |             |       | 0.4529         | 0.5124         |                |
|       | 0.50        | 0.0682| 0.0752         | 0.0609         | 0.3967         |
|       |             |       | 0.3800         | 0.4864         |                |
| 1.43  | 0.90        | 0.0003| 0.0002         | 0.0002         | 0.5952         |
|       |             |       | 0.6203         | 0.6299         |                |
|       | 0.70        | 0.0001| 0.0001         | 0.0001         | 0.6965         |
|       |             |       | 0.7677         | 0.7923         |                |
|       | 0.50        | 0.0002| 0.0003         | 0.0002         | 0.6591         |
|       |             |       | 0.6280         | 0.7216         |                |
| 33%   | 0.90        | 0.5969| 0.6154         | 0.5797         | 0.2530         |
|       |             |       | 0.2373         | 0.2785         |                |
|       | 0.70        | 0.6862| 0.7038         | 0.6261         | 0.2591         |
|       |             |       | 0.2509         | 0.3234         |                |
|       | 0.50        | 0.6369| 0.6309         | 0.5595         | 0.2604         |
|       |             |       | 0.2635         | 0.3420         |                |
| 1.00  | 0.90        | 0.3443| 0.3514         | 0.3309         | 0.3770         |
|       |             |       | 0.3789         | 0.4104         |                |
|       | 0.70        | 0.4356| 0.4409         | 0.3798         | 0.3977         |
|       |             |       | 0.4077         | 0.4876         |                |
|       | 0.50        | 0.3715| 0.3883         | 0.3176         | 0.3991         |
|       |             |       | 0.3861         | 0.4924         |                |
| 1.11  | 0.90        | 0.1374| 0.1387         | 0.1297         | 0.5107         |
|       |             |       | 0.5318         | 0.5563         |                |
|       | 0.70        | 0.1684| 0.1646         | 0.1397         | 0.5414         |
|       |             |       | 0.5861         | 0.6491         |                |
|       | 0.50        | 0.1323| 0.1505         | 0.1110         | 0.5450         |
|       |             |       | 0.5212         | 0.6392         |                |
| 1.43  | 0.90        | 0.0018| 0.0016         | 0.0018         | 0.7738         |
|       |             |       | 0.7987         | 0.8006         |                |
|       | 0.70        | 0.0006| 0.0011         | 0.0006         | 0.8362         |
|       |             |       | 0.8851         | 0.8962         |                |
|       | 0.50        | 0.0008| 0.0018         | 0.0008         | 0.8208         |
|       |             |       | 0.7814         | 0.8560         |                |
| 50%   | 0.90        | 0.5875| 0.6085         | 0.5467         | 0.3896         |
|       |             |       | 0.3738         | 0.4316         |                |
|       | 0.70        | 0.5999| 0.6368         | 0.5177         | 0.3892         |
|       |             |       | 0.3575         | 0.4686         |                |
|       | 0.50        | 0.5930| 0.5835         | 0.4690         | 0.3922         |
|       |             |       | 0.3881         | 0.5035         |                |
| 1.00  | 0.90        | 0.3864| 0.3925         | 0.3453         | 0.5765         |
|       |             |       | 0.5770         | 0.6215         |                |
|       | 0.70        | 0.4097| 0.4113         | 0.3075         | 0.5741         |
|       |             |       | 0.5781         | 0.6743         |                |
|       | 0.50        | 0.3966| 0.4038         | 0.2799         | 0.5745         |
|       |             |       | 0.5510         | 0.6824         |                |
| 1.11  | 0.90        | 0.2121| 0.2083         | 0.1785         | 0.7377         |
|       |             |       | 0.7539         | 0.7789         |                |
|       | 0.70        | 0.2278| 0.2088         | 0.1390         | 0.7426         |
|       |             |       | 0.7444         | 0.8345         |                |
|       | 0.50        | 0.2117| 0.2279         | 0.1290         | 0.7414         |
|       |             |       | 0.7104         | 0.8225         |                |
| 1.43  | 0.90        | 0.0171| 0.0122         | 0.0117         | 0.9410         |
|       |             |       | 0.9554         | 0.9512         |                |
|       | 0.70        | 0.0152| 0.0084         | 0.0043         | 0.9456         |
|       |             |       | 0.9717         | 0.9678         |                |
|       | 0.50        | 0.0122| 0.0152         | 0.0054         | 0.9446         |
|       |             |       | 0.9110         | 0.9425         |                |
Case 1: \[ PP_e^{(s)} > \delta_e \cdot \pi_1^{(s)} \rightarrow \text{Go with } F \]

Case 2: if Case 1. is not met,

\[ PP_e^{(P)} > \delta_e \cdot \pi_2^{(P)} \rightarrow \text{Go with } P \]

Case 3: if both Case 1. and Case 2. are not met,

Otherwise. \[ \rightarrow \text{Stop for futility} \]

Figure 1: Interim decision rule