

# Conditionally unbiased estimation in phase II/III clinical trials with early stopping for futility

Peter K. Kimani,<sup>a\*†</sup> Susan Todd<sup>b</sup> and Nigel Stallard<sup>a</sup>

Seamless phase II/III clinical trials combine traditional phases II and III into a single trial that is conducted in two stages, with stage 1 used to answer phase II objectives such as treatment selection and stage 2 used for the confirmatory analysis, which is a phase III objective. Although seamless phase II/III clinical trials are efficient because the confirmatory analysis includes phase II data from stage 1, inference can pose statistical challenges. In this paper, we consider point estimation following seamless phase II/III clinical trials in which stage 1 is used to select the most effective experimental treatment and to decide if, compared with a control, the trial should stop at stage 1 for futility. If the trial is not stopped, then the phase III confirmatory part of the trial involves evaluation of the selected most effective experimental treatment and the control. We have developed two new estimators for the treatment difference between these two treatments with the aim of reducing bias conditional on the treatment selection made and on the fact that the trial continues to stage 2. We have demonstrated the properties of these estimators using simulations. Copyright © 2013 John Wiley & Sons, Ltd.

**Keywords:** adaptive seamless designs; phase II/III clinical trials; estimation

## 1. Introduction

Modern innovations in clinical trial design have led to the availability of new approaches referred to as adaptive seamless designs (ASDs). Using an ASD, a clinical trial is conducted in 2 or more stages with interim analyses performed before the final stage to make adaptations. In this paper, we will consider two-stage ASDs where several doses or formulations of a drug, or several different treatments, are simultaneously compared with a standard/control with the poorly performing treatments dropped at stage 1 on the basis of interim analysis results. Such a trial is often termed a seamless phase II/III clinical trial. Unlike the traditional approach in which promising treatments are selected in a phase II trial separate to a confirmatory phase III trial, seamless phase II/III clinical trials combine aspects of both phases into a single trial with two or more stages. At the end of stage 1 of a two-stage seamless phase II/III clinical trial, an interim analysis is conducted to select the most promising treatment so that stage 1 resembles a phase II trial. The selected treatment together with the control treatment continues to stage 2 after which a confirmatory analysis is performed so that stage 2 resembles a fixed-sample-size phase III trial. The confirmatory analysis includes data from stages 1 and 2.

An ASD such as that described earlier poses a number of statistical challenges in both hypothesis testing and estimation of treatment effects because at the end of the trial, they use the data used in treatment selection to make inferences. An appropriate hypothesis testing method must be used to ensure that the overall type I error rate of the trial is not inflated. The evidence from the two stages can be combined using sufficient statistics from the accumulated data after each of the two stages or using the  $p$ -values from the two stages. Several authors [1–3] have proposed ASDs for which hypothesis testing is based on the sufficient statistics for the selected treatment effect, where the selected treatment is that which is seen to be most effective in the trial. If the selected treatment is not the most effective but testing is performed

<sup>a</sup>Warwick Medical School, The University of Warwick, Coventry, CV4 7AL, U.K.

<sup>b</sup>Department of Mathematics and Statistics, The University of Reading, Reading, RG6 6AX, U.K.

\*Correspondence to: Peter K. Kimani, Warwick Medical School, The University of Warwick, Coventry, CV4 7AL, U.K.

†E-mail: peter.k.u.kimani@warwick.ac.uk

as if the most effective treatment has been selected using the preceding methods, the test is conservative [4]. Hypothesis testing following an ASD can be carried out by combining the  $p$ -values from stages 1 and 2 [5,6]. This method of testing is very flexible with regard to the choice of the selection rule. The flexibility of this testing has been exploited to propose ASDs that use Bayesian techniques to make the selection but use frequentist methods for hypothesis testing [7–9]. A third technique of testing hypotheses after an adaptive trial is by using the conditional error principle [10] as in the adaptive Dunnett test [11].

The focus in this paper is estimation following an ASD. Estimation in this context is challenging because experimental treatments are retained in the trial precisely because they appear to be the most promising. Data suggesting that one treatment is superior may arise by chance even if the treatment is not truly superior to the other experimental treatments. Although the estimates may be biased, the bias can be quantified only if the rule for selecting the most promising treatment is specified in advance [12]. This is because bias is defined as an expectation and expectations are taken over all possible outcomes, requiring specification of the selection rule used. The most promising treatment may be chosen on the basis of effectiveness and other factors such as safety. In this paper, we will focus on selection where the most promising treatment is that which has the highest apparent effectiveness at the end of stage 1. For such a selection, the effectiveness of treatments chosen to remain in the clinical trial is likely to be overestimated.

Regulatory guidance [13,14] indicates that the bias of estimates obtained following an ASD should be considered. Cohen and Sackrowitz [15] and Shen [16] have proposed methods for estimating the mean of the selected treatment. The Cohen and Sackrowitz estimator is unbiased, whereas the Shen estimator reduces the bias relative to the naive estimator that ignores selecting the most effective treatment to continue to stage 2 based on the observed stage 1 data. Stallard and Todd [17] have proposed a method for estimating the mean of the selected treatment and also the means of the treatments that are dropped at stage 1. Cohen and Sackrowitz [15] and Shen [16] assume that the trial always continues to stage 2, whereas Stallard and Todd [17] assume that the trial may stop either for futility (when none of the experimental treatments are sufficiently effective on the basis of stage 1 data) or for efficacy. In this paper, we extend these methods to the setting where the trial can stop at stage 1 for futility, but not efficacy. This setting is common in practice. We will derive new estimators for the treatment difference for the selected treatment when estimation is unbiased conditional on continuing to stage 2. This differs from the Stallard and Todd estimator because the Stallard and Todd estimator is derived to be approximately unbiased conditional on the selected treatment whereas the estimators we will derive in this paper are obtained conditional on the selected treatment and the fact that the trial continues to stage 2. The two new conditional estimators (the word conditional is used to emphasize that estimation is unbiased conditional on continuing to stage 2) that we will derive extend the Cohen and Sackrowitz estimator and the Stallard and Todd estimator.

Like Koopmeiners *et al.* [18], we believe estimation unbiased conditional on continuing to stage 2 is of practical importance because when the trial cannot stop for efficacy at stage 1, it is reasonable to be interested in making a claim only when the trial continues to stage 2. Also, unconditionally unbiased estimators, that is estimators that do not condition on the stage at which the trial stops, may be conditionally biased [19]. Because of this, in [19], the authors proposed to obtain estimators unbiased conditional on the stage at which the trial stops.

We organized the remainder of the paper as follows. In Section 2.1, we describe the setting of interest while giving the notation. In Section 2.2, we derive an estimator that extends the Cohen and Sackrowitz estimator, and in Section 2.3, we derive expressions used to obtain an estimator that uses the Stallard and Todd principle. In Section 3, we present a worked example. We compare the various estimators using a simulation study in Section 4. The paper ends with a discussion in Section 5.

## 2. Estimating treatment difference after an adaptive seamless design

### 2.1. Setting and notation

As already mentioned, we will consider ASDs with two stages where stage 1 is used to select the most effective treatment and stage 2 is used for confirmatory analysis. Let  $k$  ( $\geq 2$ ) denote the number of experimental treatments available at stage 1 for comparison with the control treatment, with the experimental treatment showing the highest effectiveness based on stage 1 data selected to continue to stage 2 together with the control. Let the number of subjects allocated to each treatment at stage 1 be denoted by  $n_1$ .

We assume outcomes from treatment  $i$  ( $i = 0, 1, \dots, k$ ), with  $i = 0$  corresponding to the control treatment, are normally distributed with mean  $\mu_i$  and a known common variance  $\sigma^2$ , so that the stage 1 sample mean for treatment  $i$  follows a normal distribution  $N(\mu_i, \sigma_1^2)$ , where  $\sigma_1^2 = \sigma^2/n_1$ . We denote the stage 1 sample mean for treatment  $i$  by  $X_i$  and the observed sample mean by  $x_i$ . Let the selected treatment be denoted by  $S$  ( $S \in \{1, \dots, k\}$ ), noting that  $S$  is a random variable, and the sample mean from the stage 2 data for treatment  $i$  ( $i = 0, S$ ), with  $i = 0$  corresponding to the control treatment, be denoted by  $Y_i$  with observed sample mean denoted by  $y_i$  so that  $Y_i \sim N(\mu_i, \sigma_2^2)$ , where  $\sigma_2^2 = \sigma^2/n_2$ , with  $n_2$  the number of subjects allocated to each treatment at stage 2. We suppose that the trial continues to stage 2 if  $x_S - x_0 \geq b$ . We will refer to  $b$  as the futility boundary.

We define the selection time as the proportion  $n_1/(n_1 + n_2)$ . This is the proportion of stage 1 data for the control and the selected treatment. We denote the selection time by  $t$  so that the sample mean from the two stages for the control treatment is given by  $Z_{0,MLE} = tX_0 + (1-t)Y_0$  and the sample mean for the selected treatment is given by  $Z_{S,MLE} = tX_S + (1-t)Y_S$ . After completion of the trial, the objective is to estimate the treatment difference  $\theta_S = \mu_S - \mu_0$ . We can base the inference on the naive maximum likelihood estimator (MLE) for the difference between the selected and control treatments given by

$$D_{S,MLE} = Z_{S,MLE} - Z_{0,MLE}. \tag{1}$$

We will refer to this as the naive estimator. When there is no opportunity to stop at stage 1, the naive estimator is positively biased [12, 20, 21]. This is because the chosen experimental treatment is selected on the basis of having the maximum observed treatment difference compared with the control treatment.

In this paper, the objective is to seek estimators, which are unbiased conditional on continuing to stage 2, for the setting where a trial can stop at stage 1 for futility. For this setting, the naive MLE is also positively biased because of selecting the highest effective treatment and also requiring  $x_S - x_0$ , the observed difference between the selected and control treatments at stage 1, to exceed the critical value  $b$ . If estimation is conditional on continuing to stage 2,  $Y_S$  and  $Y_0$  are respectively unbiased estimators for  $\mu_S$  and  $\mu_0$  so that the stage 2 sample difference

$$D_{S,2} = Y_S - Y_0 \tag{2}$$

is an unbiased estimator for  $\theta_S$ . However, this estimator, which we will henceforth refer to as the stage 2 estimator, is likely to be inefficient compared with estimators that use both stage 1 and 2 data. In Sections 2.2 and 2.3, we will derive two new estimators that use both stages 1 and 2 data.

## 2.2. A new unbiased estimator for the treatment difference

Cohen and Sackrowitz [15], although not considering the control treatment, derived a uniformly minimum variance unbiased estimator (UMVUE) for  $\mu_S$  when the trial always continues to stage 2. When the trial always continues to stage 2, the bias of the naive estimator of  $\theta_S$  arises from using  $Z_{S,MLE}$  as an estimator for  $\mu_S$  [12]. Thus, replacing  $Z_{S,MLE}$  with the Cohen and Sackrowitz UMVUE for  $\mu_S$  in Equation (1) gives an unbiased estimator for  $\theta_S$  in the case where the trial always continues to stage 2. In this paper, we are interested in a setting where a trial can stop for futility and estimation is conditional on continuing to stage 2. For this setting,  $Z_{0,MLE}$  is biased for  $\mu_0$  and also the estimator for  $\mu_S$  derived by Cohen and Sackrowitz is no longer unbiased because it does not condition on continuing to stage 2. In the rest of this section, we will derive the UMVUE for  $\mu_S$  and the UMVUE for  $\mu_0$  when estimation is conditional on continuing to stage 2, and hence an unbiased estimator for  $\theta_S$ .

The UMVUEs are based on the Rao–Blackwell theorem (for example [22]). If estimation is conditional on continuing to stage 2,  $Y_S$  is an unbiased estimator of  $\mu_S$ . In the Rao–Blackwell theorem, a new estimator defined as the expected value of  $Y_S$  given a sufficient statistic for  $\mu_S$  is the UMVUE for  $\mu_S$ . Similarly,  $Y_0$  is an unbiased estimator of  $\mu_0$  so that the expected value of  $Y_0$  given a sufficient statistic for  $\mu_0$  is the UMVUE for  $\mu_0$ . Let  $X_{(1)} > X_{(2)} > \dots > X_{(k)}$  be the order statistics of stage 1 sample means so that  $X_S = X_{(1)}$ . For the selected treatment  $S$ , we show in Appendix A that the UMVUE for  $\mu_S$ , which we denote by  $Z_{S,CHN}$  with the notation chosen such that it reflects the fact that the estimator

extends the Cohen and Sackrowitz UMVUE for  $\mu_S$ , is given by

$$\begin{aligned} Z_{S,\text{CHN}} &= \frac{\sigma_2^2 X_S + \sigma_1^2 Y_S}{\sigma_1^2 + \sigma_2^2} - \frac{\sigma_2^2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \frac{\phi(W_B)}{\Phi(W_B)} \\ &= Z_{S,\text{MLE}} - \frac{\sigma_2^2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \frac{\phi(W_B)}{\Phi(W_B)}, \end{aligned} \quad (3)$$

where  $\phi(\cdot)$  and  $\Phi(\cdot)$  respectively denote the density and distribution functions of a standard normal,

$$W_B = \left( \sqrt{\sigma_1^2 + \sigma_2^2} / \sigma_1 \right) (Z_{S,\text{MLE}} - \max\{B, X_{(2)}\}) \quad (4)$$

and  $B = X_0 + b$ . For the control treatment, we show in Appendix B that the UMVUE for  $\mu_0$ , which we denote by  $Z_{0,\text{CHN}}$ , is given by

$$\begin{aligned} Z_{0,\text{CHN}} &= \frac{\sigma_2^2 X_0 + \sigma_1^2 Y_0}{\sigma_1^2 + \sigma_2^2} + \frac{\sigma_2^2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \frac{\phi(W_{B_1})}{\Phi(W_{B_1})} \\ &= Z_{0,\text{MLE}} + \frac{\sigma_2^2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \frac{\phi(W_{B_1})}{\Phi(W_{B_1})}, \end{aligned} \quad (5)$$

where

$$W_{B_1} = \left( \sqrt{\sigma_1^2 + \sigma_2^2} / \sigma_1 \right) (B_1 - Z_{0,\text{MLE}}) \quad \text{and} \quad B_1 = X_S - b.$$

Because  $Z_{S,\text{CHN}}$  and  $Z_{0,\text{CHN}}$  are unbiased estimators for  $\mu_S$  and  $\mu_0$ , then

$$D_{S,\text{CHN}} = Z_{S,\text{CHN}} - Z_{0,\text{CHN}} \quad (6)$$

is an unbiased estimator for  $\theta_S$ . We will refer to this estimator as the (new) unbiased estimator. If we set the futility boundary  $b = -\infty$  so that  $B = -\infty$ ,  $B_1 = \infty$  and  $\max\{B, X_{(2)}\} = X_{(2)}$ , then Equation (5) simplifies to

$$Z_{0,\text{CHN}} = Z_{0,\text{MLE}}$$

and Equation (4) simplifies to

$$W_B = \left( \sqrt{\sigma_1^2 + \sigma_2^2} / \sigma_1 \right) (Z_{S,\text{MLE}} - X_{(2)}). \quad (7)$$

The simplification of  $Z_{0,\text{CHN}}$  to  $Z_{0,\text{MLE}}$  supports the finding in [12] that when there is no opportunity to stop at stage 1, the bias when the naive estimator is used as an estimator of the treatment difference is only contributed to by using  $Z_{S,\text{MLE}}$  as an estimator for the selected treatment in Equation (1). If further we have  $\sigma^2 = 1$  and  $n_1 = n_2 = 1$  so that  $\sigma_1^2 = \sigma_2^2 = 1$ , formulae (3) and (7) reduce to the formulae given by Cohen and Sackrowitz.

Koopmeiners *et al.* [18] considered the setting with  $k = 1$ . Note that for a trial with no control arm, the UMVUE for this setting is given by Equation (3), but with  $W_B$  expressed as

$$W_B = \left( \sqrt{\sigma_1^2 + \sigma_2^2} / \sigma_1 \right) (Z_{S,\text{MLE}} - B), \quad (8)$$

where  $B$  is the futility boundary and  $Z_{S,\text{MLE}}$  is the sample mean for the experimental treatment. The same formula applies for the case with a control, replacing  $Z_{S,\text{MLE}}$  with  $D_{S,\text{MLE}}$ , the sample difference between the experimental treatment and the control, and appropriately defining  $\sigma_1^2$  and  $\sigma_2^2$  as the variances of stages 1 and 2 sample differences. Koopmeiners *et al.* also derived the UMVUE for the setting

with  $k = 1$ . The formula given by Koopmeiners *et al.* has a typological error. Defining  $\sigma_{\text{all}}^2 = \sigma^2/n$ , where  $n = n_1 + n_1$ , using our notation, they give the formula for UMVUE as

$$Z_{S,\text{MLE}} = \frac{\sigma_{\text{all}}^2}{\sqrt{\sigma_1^2 - \sigma_{\text{all}}^2}} \frac{\phi \left\{ (B - Z_{S,\text{MLE}}) \left( 1 / \sqrt{\sigma_1^2 - \sigma_{\text{all}}^2} \right) \right\}}{\Phi \left\{ (B - Z_{S,\text{MLE}}) \left( 1 / \sqrt{\sigma_1^2 - \sigma_{\text{all}}^2} \right) \right\}},$$

where  $B$  is the futility boundary, instead of

$$Z_{S,\text{MLE}} = \frac{\sigma_{\text{all}}^2}{\sqrt{\sigma_1^2 - \sigma_{\text{all}}^2}} \frac{\phi \left\{ (B - Z_{S,\text{MLE}}) \left( 1 / \sqrt{\sigma_1^2 - \sigma_{\text{all}}^2} \right) \right\}}{1 - \Phi \left\{ (B - Z_{S,\text{MLE}}) \left( 1 / \sqrt{\sigma_1^2 - \sigma_{\text{all}}^2} \right) \right\}}.$$

This formula can be shown to be equivalent to our formula. Thus, the Koopmeiners *et al.* estimator is a special case of our estimator.

### 2.3. A new bias-adjusted estimator for the treatment difference

Stallard and Todd [17] proposed a bias-adjusted estimator that involves estimating the bias of the naive estimator. A bias-adjusted estimate is then obtained by subtracting the estimate of the bias from the naive estimate. The bias-adjusted estimator is obtained as follows. Let  $\hat{\theta}_i$  ( $i = 1, \dots, k$ ) denote the naive maximum likelihood estimate for the treatment difference  $\theta_i = \mu_i - \mu_0$ . For  $i \neq S$ ,  $\hat{\theta}_i = x_i - x_0$ , and for  $i = S$ ,  $\hat{\theta}_i = D_{S,\text{MLE}}$ . If the true vector  $\theta = (\theta_1, \dots, \theta_k)'$  was known, the biases for the treatment differences of the naive maximum likelihood estimate  $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_k)'$  could be derived. Let the bias of  $\hat{\theta}_i$  be denoted by  $b_i(\theta)$  and the vector  $(b_1(\theta), \dots, b_k(\theta))'$  be denoted by  $\mathbf{b}(\theta)$ . Then, for example, the bias-adjusted estimator for  $\theta_S$  could be given by  $D_{S,\text{MLE}} - b_S(\theta)$ . However, the true mean vector  $\theta$  is unknown. Following [23], Stallard and Todd proposed estimating the bias vector by iteratively solving  $\tilde{\theta} = \hat{\theta} - \mathbf{b}(\tilde{\theta})$ , where  $\hat{\theta}$  is the naive maximum likelihood estimate of  $\theta$ . The initial value of  $\tilde{\theta}$  in the iteration procedure could be set to be  $\hat{\theta}$ . If the solution is achieved at iteration  $r$ , then the bias-adjusted estimator for  $\theta_S$  is given by

$$D_{S,\text{STL}} = D_{S,\text{MLE}} - b_S(\tilde{\theta}_r), \tag{9}$$

where the notation is chosen to reflect the fact that this estimator is obtained using the principles described by Stallard and Todd.

Stallard and Todd derived the bias vector conditional on the selected treatment. Let stage 1 treatment differences  $X_i - X_0$  ( $i = 1, \dots, k$ ) be denoted by  $D_i$  and the observed differences  $x_i - x_0$  by  $d_i$ . One of the densities Stallard and Todd need while deriving the bias vector, which we also need in this paper, is the joint density of  $S = i$  and  $d_i$  given by

$$f(d_i, S = i) = \int_{-\infty}^{\infty} \frac{1}{\sigma_1^2} \phi \left( \frac{w - d_i}{\sigma_1} \right) \phi \left( \frac{w - \theta_i}{\sigma_1} \right) \prod_{j \neq i} \Phi \left( \frac{w - \theta_j}{\sigma_1} \right) dw. \tag{10}$$

In the rest of this section, we will derive the bias vector when estimation is conditional on continuing to stage 2, where the trial continues to stage 2 if  $d_S \geq b$ . If the trial continues to stage 2, the expected value of the treatment difference for the selected treatment  $i$  is given by

$$E[D_i | S = i, D_i \geq b] = \frac{\int_b^{\infty} d_i f(d_i, S = i) dd_i}{\text{pr}(S = i, D_i \geq b)}, \tag{11}$$

where  $f(d_i, S = i)$  is the density given by Equation (10). The numerator and the denominator in Equation (11) are simplified to expressions with single integrals in Appendix C.1. The expected value of the naive estimator given by Equation (1) can be expressed as  $t(E[X_S - X_0] - \theta_S) + \theta_S$  so that the bias of the treatment difference for the selected treatment  $i$ , given that the trial continues to stage 2, may be written as

$$b_i(\theta) = t \cdot (E[D_i | S = i, D_i \geq b] - \theta_i). \tag{12}$$

Conditional on the trial continuing to stage 2, the expected value of the treatment difference between a dropped treatment  $i'$  and the control treatment is expressed by

$$E[D_{i'}|S = i, D_i \geq b] = \frac{E[D_{i'}, S = i, D_i \geq b]}{\text{pr}(S = i, D_i \geq b)}, \quad i' \neq i.$$

The expression for  $\text{pr}(S = i, D_i \geq b)$  is given earlier. The expression for  $E[D_{i'}, S = i, D_i \geq b]$  while using  $D_{i'}$  directly involves multidimensional integrals that cannot be simplified to fewer integrals. To overcome this, we define a new variable  $W_0$  that has a normal distribution  $N(0, \sigma_1^2)$  and its covariance with  $D_i (i = 1, \dots, k)$ ,  $\text{Cov}(W_0, D_i) = -\sigma_1^2$ . Then  $W_i = D_i + W_0$  is normally distributed with  $N(\theta_i, \sigma_1^2)$  and  $\text{Cov}(W_i, W_j) = 0$  for  $i \neq j = 0, 1, \dots, k$ . Note that

$$E[D_{i'}, S = i, D_i \geq b] = E[W_{i'}, S = i, D_i \geq b] - E[W_0, S = i, D_i \geq b].$$

The expressions for  $E[W_{i'}, S = i, D_i \geq b]$  and  $E[W_0, S = i, D_i \geq b]$  are simplified to single integrals in Appendix C.2. The bias of the treatment difference for a dropped treatment  $i'$  given that the trial continues to stage 2 may be written as

$$b_{i'}(\theta) = E[D_{i'}|S = i, D_i \geq b] - \theta_{i'}. \tag{13}$$

To obtain  $b_S(\tilde{\theta}_r)$  to substitute in Equation (9) and obtain a bias-adjusted estimate when estimation is conditional on continuing to stage 2, expressions (12) and (13) are used in the iteration procedure but with  $\theta_i$  and  $\theta_{i'}$  replaced by  $\tilde{\theta}_i$  and  $\tilde{\theta}_{i'}$ , respectively. We will refer to this estimator for  $\theta_S$  as the (new) bias-adjusted estimator. Koopmeiners *et al.* [18] derived a similar bias-adjusted estimator for the setting with  $k = 1$  so that their bias-adjusted estimator is a special case of our bias-adjusted estimator.

### 3. Example

In this section, using the two new estimators described in Sections 2.2 and 2.3, we compute estimates for an example constructed from the case study described in [21]. The case study is based on a comparison of three doses of an experimental drug for generalized anxiety disorder with a placebo. The primary endpoint is the change from baseline at 8 weeks of treatment in the total score on the Hamilton Rating Scale for Anxiety. The primary endpoint is taken to be normally distributed with a common standard deviation across the four treatment arms assumed to be 6 points. As in [21], we consider a two-stage ASD for the case study with  $n_1 = n_2 = 71$  so that  $t = 0.5$ .

Suppose that the true treatment means are the stage 1 estimates from [21], which we give in Table I (column 1), and that the observed stage 1 means from an adaptive trial are as given in column 2. We suppose the trial continues to stage 2 if the highest effective experimental dose is at least as effective as the placebo, that is, the observed difference between the highest effective experimental dose and the placebo is at least 0. On the basis of the observed stage 1 data, dose 2 and placebo would be tested further in stage 2. We suppose the results from stage 2 are as given in column 3.

With the results in Table I,  $z_{0,MLE} = (0.5 \times -0.082) + (0.5 \times 0.049) = -0.017$  and  $z_{2,MLE} = (0.5 \times 1.766) + (0.5 \times 1.451) = 1.609$  so that the naive maximum likelihood estimate for the difference between dose 2 and placebo is 1.626. The stage 2 estimate is 1.402 (column 5). For the new unbiased estimator, to compute  $z_{2,CHN}$ , we note that  $B = -0.082 + 0 = -0.082$  so that  $\max\{B, x_2\}$  is given by

Table I. Data from a seamless phase II/III clinical trial.					
Treatment	Treatment means			Treatment differences	
	True	Observed		Stage 1	Stage 2
		Stage 1	Stage 2		
Placebo	0	-0.082	0.049	—	—
Dose 1	0.8	0.413	—	0.495	—
Dose 2	1.5	1.766	1.451	1.848	1.402
Dose 3	2.6	1.567	—	1.649	—

$\max\{-0.082, 1.567\} = 1.567$  and that  $\sigma_1^2 = \sigma_2^2 = 6^2/71 = 0.507$  so that  $\sqrt{\sigma_1^2 + \sigma_2^2}/\sigma_1^2 = 1.986$  and  $\sigma_2^2/\sqrt{\sigma_1^2 + \sigma_2^2} = 0.504$ . By substituting the appropriate values in Equations (3) and (4),  $z_{2,\text{CHN}} = 1.261$ . For  $z_{0,\text{CHN}}$ , the only component we have not calculated is  $B_1$ , which is given by  $1.766 - 0$ . By substituting the appropriate values in Equation (5),  $z_{0,\text{CHN}} = -0.017$ . Therefore, the unbiased estimate for the difference between dose 2 and placebo is 1.278.

For the new bias-adjusted estimator, we note that the naive maximum likelihood estimate for  $(\theta_1, \theta_2, \theta_3)$  is (0.495, 1.626, 1.649). The bias function for doses 1 and 3 is given by expression (13), and the bias function for dose 2 is given by expression (12). Using a program written in the R statistical package, we obtain the value of  $\mathbf{b}(\hat{\theta})$  and hence  $\tilde{\theta} = \hat{\theta} - \mathbf{b}(\hat{\theta})$  at each iteration. The iteration procedure stops at iteration  $r$  if the Euclidean distance between  $\hat{\theta}_{r-1}$  and  $\hat{\theta}_r$  is less than or equal to 0.0005. The program is available at <https://files.warwick.ac.uk/nstallard/browse/adaptive>. We set  $\hat{\theta}_1 = (0.495, 1.626, 1.649)$ . The iteration procedure stopped at iteration 15, and the bias-adjusted estimate for the difference between dose 2 and placebo is 1.135.

Thus, the naive, stage 2, unbiased, and bias-adjusted estimates for the difference between dose 2 and the placebo are 1.626, 1.402, 1.278, and 1.135, respectively. The estimates are different with the naive estimate, as expected, having the highest value. The unbiased and bias-adjusted estimates correct for the bias, and their values are below both the stage 1 and 2 differences. The unbiased and bias-adjusted estimates are closer to the stage 2 difference, which is an unbiased estimate of the treatment difference. We explore the properties of the four estimators in the next section.

## 4. Simulation study

### 4.1. Simulation study settings

In this section, we describe a simulation study that was used to assess the bias and the mean squared error (MSE) of the estimators described in Section 2. Following expressions (10), (11), and (12), the bias of the naive estimator depends on the number of experimental treatments  $k$ , the selection time  $t$ , the value of the futility boundary, and the true parameter values. Therefore, we will consider several scenarios in the simulation study. We will consider scenarios where  $k$  is between 2 and 5. We believe this encompasses the majority of practical scenarios with  $k > 1$ . We will also consider different true parameter values for the means. In all simulations, we will take the variance of the outcomes  $\sigma^2$  to be 1. Hence, we will only consider small differences in true treatment means corresponding to the small standardized effect sizes that we might anticipate in clinical trials.

We will assess three different values for the futility boundary. In most simulations that we will describe, we will take the treatment difference between the most effective treatment(s) and the control treatment to be 0.05. The first futility boundary value is 0, so because it is below the highest treatment difference(s), this boundary will be used to assess the bias when some of the experimental treatments are more effective than minimally required. The second futility boundary value is 0.05, so it will be used to assess the bias when the highest treatment difference is on the futility boundary. The third futility boundary value is 0.10, so it will be used to assess the bias when none of the experimental treatments are as effective as is minimally desired. We will also describe simulation results for some scenarios where the treatment difference between the most effective treatments(s) is 0.1, 0.2, and 0.5 while using the same futility boundary values (0, 0.05, and 0.1). These simulations will be used to assess the bias and MSE when most bias is contributed by the selection of the most effective treatment and not because of the futility boundary.

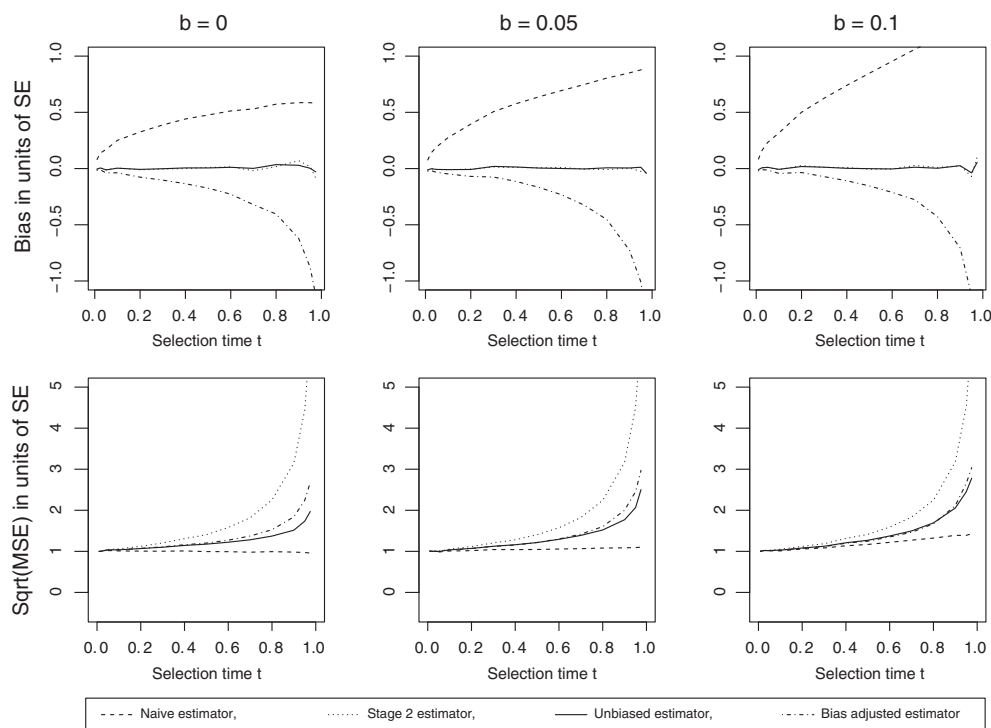
We perform simulations for 14 values of  $t$ , the selection time point, in the interval (0, 1). Because of the computations required, at each time point, we run 10,000 simulations that would continue to stage 2, that is, 10,000 simulations for which the simulated stage 1 treatment difference of the selected treatment is equal to or greater than the futility boundary value. For the treatment difference of the selected treatment  $S$ , in each simulation, we obtain the naive MLE  $d_{S,\text{MLE}}$  using Equation (1), the stage 2 estimate  $d_{S,2}$  using Equation (2), the unbiased estimate  $d_{S,\text{CHN}}$  using Equation (6), and the bias-adjusted estimate  $d_{S,\text{STL}}$  using Equation (9). We then calculate the differences  $(d_{S,\text{MLE}} - \theta_S)$ ,  $(d_{S,2} - \theta_S)$ ,  $(d_{S,\text{CHN}} - \theta_S)$ , and  $(d_{S,\text{STL}} - \theta_S)$  and the respective squares  $(d_{S,\text{MLE}} - \theta_S)^2$ ,  $(d_{S,2} - \theta_S)^2$ ,  $(d_{S,\text{CHN}} - \theta_S)^2$ , and  $(d_{S,\text{STL}} - \theta_S)^2$ . Then at each selection time point, for each estimator, the mean bias is obtained by taking the average of its corresponding 10,000 differences and the MSE by taking the average of its corresponding 10,000 square differences.

We will present the bias and the  $\sqrt{\text{MSE}}$  of the various estimators in units of the standard error (SE), the standard deviation for the estimator of the difference of a single experimental treatment–control comparison given by  $\sqrt{2/(n_1 + n_2)}$ . This makes the results invariant to changes in the sample sizes.

#### 4.2. Simulation results for $k = 2$ with $\theta_1 = \theta_2$

Figure 1 shows the bias and  $\sqrt{\text{MSE}}$  when two experimental treatments and a control are included in stage 1 with  $\theta_1 = \theta_2 = 0.05$ . Columns 1, 2, and 3 correspond to futility boundary values 0, 0.05, and 0.1, respectively. The dashed, dotted, continuous, and dash-dotted lines correspond to the naive, stage 2, unbiased, and bias-adjusted estimators, respectively. The naive estimator is biased, and the bias increases with selection time but not linearly and also as the futility boundary value increases. The stage 2 estimator, as expected, is mean unbiased for all selection time points and all futility boundary values. Because of the theoretical derivation of the unbiased estimator, this is also mean unbiased for all scenarios. The bias-adjusted estimator overcorrects for bias, and the overcorrection increases with selection time but decreases as the value of the futility boundary increases. The naive estimator has the lowest MSE at all selection times for all scenarios. The stage 2 estimator has the highest MSE. In all scenarios, up to selection time 0.7, the unbiased estimator and the bias-adjusted estimator have approximately equal MSE. Tables giving more details of the results in Figure 1 and of additional simulations mentioned in the following are available from the authors.

We also assessed the characteristics of the four estimators when  $\theta_1 = \theta_2 = 0.1$ ,  $\theta_1 = \theta_2 = 0.2$ , and  $\theta_1 = \theta_2 = 0.5$  for futility boundary values  $b = 0$ ,  $b = 0.05$ , and  $b = 0.1$ . We are able to clearly describe the findings from these scenarios without giving figures, and so we do not present figures for these scenarios. However, figures with these results and of additional simulations mentioned in the following, for which we have not presented figures, are available from the authors. For all the futility boundary values, the bias of the naive estimator decreases as the values of  $\theta_1$  and  $\theta_2$  increase. For  $\theta_1 = \theta_2 = 0.5$ , the biases of the naive estimator for  $b = 0$ ,  $b = 0.05$ , and  $b = 0.1$  are identical. This is because for this case, for the three futility boundary values, the stage 1 sample difference for experimental treatment 1



**Figure 1.** Bias (top row) and  $\sqrt{\text{MSE}}$  (bottom row) for various estimators. In all figures,  $\theta_1 = \theta_2 = 0.05$ . Dashed, dotted, continuous, and dash-dotted lines correspond to naive, stage 2, unbiased, and bias-adjusted estimators, respectively. The futility boundaries are given on top of each column. MSE, mean squared error; SE, standard error.



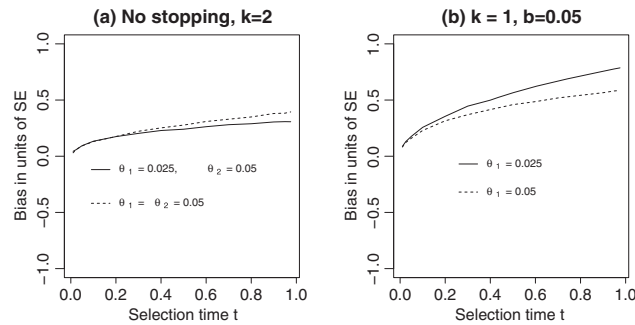
and the stage 1 sample difference for experimental treatment 2 beat the futility boundary values in most simulation runs so that the bias arises mostly because of the treatment selection and hence the similar bias of the naive estimator. As expected, the unbiased estimator and the stage 2 estimator are mean unbiased for all values of  $t$ ,  $b$ ,  $\theta_1$ , and  $\theta_2$ , whereas the bias-adjusted estimator is negatively biased, but the bias decreases as the values of  $\theta_1$  and  $\theta_2$  increase. For the three futility boundary values, as in the case where  $\theta_1 = \theta_2 = 0.05$ , for the cases where  $\theta_1 = \theta_2 = 0.1$ ,  $\theta_1 = \theta_2 = 0.2$ , and  $\theta_1 = \theta_2 = 0.5$ , the naive estimator has the lowest MSE at all selection times, the stage 2 estimator has the highest MSE at all selection times, and up to selection time 0.7, the unbiased estimator and the bias-adjusted estimator have approximately equal MSE.

4.3. Simulation results for  $k = 2$  with  $\theta_1 \neq \theta_2$

To assess the bias and the MSE when  $\theta_1 \neq \theta_2$ , we performed simulations with  $\theta_1 = 0.025$  and  $\theta_2 = 0.05$  for futility boundary values  $b = 0$ ,  $b = 0.05$ , and  $b = 0.1$ . Most findings from these scenarios are similar to findings in Figure 1, and so we describe the findings of these scenarios without presenting a figure summarizing the results. As in the case when  $\theta_1 = \theta_2 = 0.05$ , the bias of the naive estimator increases with futility boundary value and the selection time when  $\theta_1 = 0.025$  and  $\theta_2 = 0.05$ . As expected, the stage 2 and unbiased estimators are mean unbiased at all selection times and futility boundary values. In terms of bias, up to selection time 0.5, the bias-adjusted estimator performs almost as well as the stage 2 and unbiased estimators. The naive estimator has the lowest MSE whereas the stage 2 estimator has the highest MSE. The unbiased and bias-adjusted estimators have approximately equal MSE for all selection times and futility boundary values. For all the three boundaries considered, the bias of the naive estimator is slightly higher when  $\theta_1 = 0.025$  and  $\theta_2 = 0.05$  than when  $\theta_1 = \theta_2 = 0.05$ . The mean biases of the naive estimator for three values of  $t$  for the case where  $\theta_1 = \theta_2 = 0.05$  and for the case where  $\theta_1 = 0.025$  and  $\theta_2 = 0.05$  are given in Table II. This is unlike the setting in which the trial always continue to stage 2, where bias decreases as one of the experimental treatments becomes distinctly superior to the competing treatment [12]. To assess what may be causing this difference, in Table II, for both scenarios ( $\theta_1 = \theta_2 = 0.05$  and  $\theta_1 = 0.025$  and  $\theta_2 = 0.05$ ), we present the probabilities of continuing to stage 2 ( $\Pr[d_S \geq b]$ ) and, conditional on continuing to stage 2, the simulated probabilities of selecting treatment 1 ( $\Pr[S = 1|d_S \geq b]$ ), of selecting treatment 1 while treatment 2 also beats the futility boundary ( $\Pr[S_1, d_2 \geq b]$ ), and of selecting treatment 2 while treatment 1 also beats the futility boundary ( $\Pr[S_2, d_1 \geq b]$ ). As expected, for the case where  $\theta_1 = \theta_2 = 0.05$ , the simulated probabilities  $\Pr(S = 1|d_S \geq b)$  and  $\Pr(S = 1, d_2 \geq b)$  are respectively approximately equal to the simulated probabilities  $\Pr(S = 2|d_S \geq b)$  and  $\Pr(S = 2, d_1 \geq b)$ . For both scenarios, from  $\Pr(d_S \geq b)$ , we observe that as the selection is made later in the trial, it is more likely that a right decision of whether to continue to stage 2 or not will be made. However, we note that for the case where  $\theta_1 = 0.025$  and  $\theta_2 = 0.005$ , treatment 1 is still selected with relatively high probability (the minimum probability is 0.34

**Table II.** Bias of the naive estimator and probabilities of various outcomes when two experimental treatments are tested in stage 2 for the case where  $\theta_1 = \theta_2 = 0.05$  and the case where  $\theta_1 = 0.025$  and  $\theta_2 = 0.05$ .

Characteristic	$t = 0.2$			$t = 0.5$			$t = 0.8$		
	$b = 0$	$b = 0.05$	$b = 0.1$	$b = 0$	$b = 0.05$	$b = 0.1$	$b = 0$	$b = 0.05$	$b = 0.1$
$\theta_1 = 0.05, \theta_2 = 0.05$									
Mean bias	0.0239	0.0291	0.0336	0.0326	0.0453	0.0608	0.0386	0.0568	0.0807
$\Pr(d_S \geq b)$	0.78	0.67	0.53	0.84	0.67	0.45	0.87	0.67	0.40
$\Pr(S = 1 d_S \geq b)$	0.49	0.50	0.50	0.50	0.49	0.50	0.50	0.50	0.51
$\Pr(S = 1, d_2 \geq b)$	0.29	0.25	0.21	0.33	0.25	0.19	0.34	0.25	0.16
$\Pr(S = 2, d_1 \geq b)$	0.30	0.24	0.21	0.32	0.26	0.18	0.34	0.25	0.16
$\theta_1 = 0.025, \theta_2 = 0.05$									
Mean bias	0.0244	0.0293	0.0354	0.0360	0.0467	0.0631	0.0411	0.0615	0.0856
$\Pr(d_S \geq b)$	0.76	0.64	0.50	0.80	0.62	0.41	0.83	0.61	0.34
$\Pr(S = 1 d_S \geq b)$	0.44	0.43	0.41	0.40	0.39	0.38	0.37	0.36	0.34
$\Pr(S = 1, d_2 \geq b)$	0.26	0.22	0.17	0.26	0.19	0.14	0.26	0.17	0.11
$\Pr(S = 2, d_1 \geq b)$	0.30	0.26	0.21	0.36	0.26	0.18	0.39	0.26	0.16



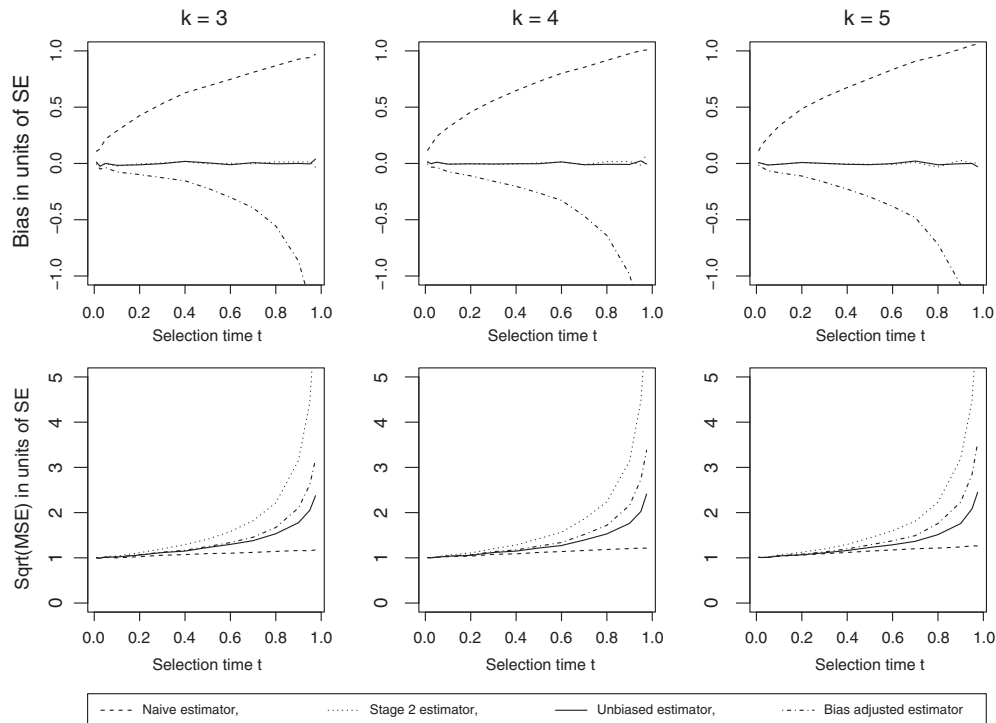
**Figure 2.** Panel (a) shows bias of the naive estimator when  $k = 2$ , and the trial always continues to stage 2 when  $\theta_1 = \theta_2 = 0.05$  (dashed line) and when  $\theta_1 = 0.025$  and  $\theta_2 = 0.05$  (continuous line). Panel (b) shows bias of the naive estimator when  $k = 1$  and futility boundary  $b = 0.05$ . The continuous and dashed lines correspond to  $\theta_1 = 0.025$  and  $\theta_1 = 0.05$ , respectively.

when  $t = 0.8$  and boundary value  $b = 0.1$ ). Also, when treatment 1 is selected to continue to stage 2, the treatment difference for treatment 2 is usually below the boundary ( $\Pr(S = 1, d_2 \geq b)$  is small). We use Figure 2 to assess whether the instances where treatment 1 is selected are the ones that make bias higher when  $\theta_1 = 0.025$  and  $\theta_2 = 0.05$  than when  $\theta_1 = \theta_2 = 0.05$ . Figure 2(a) shows the bias of the naive estimator when  $\theta_1 = \theta_2 = 0.05$  (dashed line) and when  $\theta_1 = 0.025$  and  $\theta_2 = 0.05$  (continuous line) in the case where the trial always continues to stage 2, and as expected, following [12], bias is higher when  $\theta_1 = \theta_2 = 0.05$ . The proof that for the case where the trial always continues to stage 2, the naive estimator is maximally biased when all experimental treatments are equally effective is given in [24]. Figure 2(b) shows the bias of the naive estimator when  $k = 1$  and the futility boundary value  $b = 0.05$ . The continuous and dashed lines correspond to  $\theta_1 = 0.025$  and  $\theta_1 = 0.05$ , respectively. The bias is higher when  $\theta_1 = 0.025$  than when  $\theta_1 = 0.05$ . Comparing Figure 2(a and b), we see that the futility boundary seems to contribute more to the bias. This may explain why in the case where there is a futility boundary and  $k = 2$ , the bias of the naive estimator is higher when  $\theta_1 = 0.025$  and  $\theta_2 = 0.05$  than when  $\theta_1 = \theta_2 = 0.05$ . Although the selected treatment may be the most promising because the treatment effects are distinct and hence reduce the selection bias, whenever the least effective treatment is selected, the bias is higher because we have a futility boundary.

We also performed simulations when  $(\theta_1, \theta_2) = (0.075, 0.1)$ ,  $(\theta_1, \theta_2) = (0.175, 0.2)$ , and  $(\theta_1, \theta_2) = (0.475, 0.5)$  using the futility boundary values  $b = 0$ ,  $b = 0.05$ , and  $b = 0.1$ . Note that for these parameter vectors, as for the case considered earlier where  $\theta_1 = 0.025$  and  $\theta_2 = 0.05$ ,  $\theta_2 - \theta_1 = 0.025$ . We describe the findings from these scenarios without giving the figures. For the three futility boundary values, the bias of the naive estimator decreases as the values of  $\theta_1$  and  $\theta_2$  increase. For  $(\theta_1, \theta_2) = (0.475, 0.5)$ , the biases of the naive estimator for  $b = 0$ ,  $b = 0.05$ , and  $b = 0.1$  are identical. This is because for this case, for the three futility boundary values, the stage 1 sample difference for experimental treatment 1 and the stage 1 sample difference for experimental treatment 2 beat the futility boundary values in most simulation runs so that the bias arises mostly because of the treatment selection and hence the similar bias of the naive estimator. For the three futility boundary values, compared with the case where  $(\theta_1, \theta_2) = (0.5, 0.5)$ , the biases of the naive estimator when  $(\theta_1, \theta_2) = (0.475, 0.5)$  are lower. This is because, for futility boundary values 0, 0.05 and 0.1, stage 1 sample differences for treatments 1 and 2 beat the futility boundary values in most simulations for the cases where  $(\theta_1, \theta_2) = (0.5, 0.5)$  and  $(\theta_1, \theta_2) = (0.475, 0.5)$  so that most bias arises from treatment selection and selection bias is maximal when experimental treatments are equally effective [12, 24].

#### 4.4. Simulation results for $k \geq 3$

When three or more experimental treatments are tested in stage 1, there are several possible configurations of the treatment differences, and this leads to several scenarios. Therefore, we will first describe general findings for such scenarios without presenting figures and then describe results of a few specific scenarios using a figure. On the basis of results that are not presented here, as in the case when two treatments are tested in stage 1, when three or more experimental treatments are tested in stage 1, the bias of the naive estimator increases with the futility boundary value, and estimation using the bias-adjusted



**Figure 3.** Bias (top row) and  $\sqrt{\text{MSE}}$  (bottom row) for various estimators when fertility boundary  $b = 0.05$ . Dashed, dotted, continuous, and dash-dotted lines correspond to naive, stage 2, unbiased, and bias-adjusted estimators, respectively. The number of treatments is given on top of each column, and the treatment differences are all equal to 0.05.

estimator improves with higher fertility boundary value whereas the stage 2 and unbiased estimators, as expected, provide unbiased estimators for all fertility boundary values. Figure 3 shows results when treatment differences are all equal to 0.05 and the fertility boundary is 0.05. Columns 1 to 3 give results when three, four, and five experimental treatments, respectively, are tested in stage 1. For the naive estimator, we observe that the bias increases slightly as the number of treatments increases. The stage 2 and unbiased estimators, as expected, are mean unbiased at all selection times and when three, four, or five experimental treatments are tested in stage 1. The bias-adjusted estimator again overcorrects for bias, and the overcorrection increases with selection time. Also, as the number of treatments increases, the overcorrection of the bias-adjusted estimator increases slightly. The naive estimator has the least MSE whereas the unbiased and bias-adjusted estimators have similar MSE for selection times up to 0.6. The stage 2 estimator has the highest MSE, and the difference between the MSE for the stage 2 estimator and the other estimators increases with selection time.

#### 4.5. Summary of findings from the simulation study

From the simulation study, we have observed that the bias of the naive estimator increases with the selection time, the number of experimental treatments, and the fertility boundary value. The treatment differences affect the bias of the naive estimator, but this also depends on the fertility boundary value so that it is not possible to generalize the bias on the basis of treatment differences only. The stage 2 and unbiased estimators, as expected, provide mean unbiased estimates. The bias-adjusted estimator overcorrects for bias, but under some configurations of treatment differences, if selection is carried out up to selection time 0.4, it performs fairly well. For MSE, the unbiased and bias-adjusted estimators perform similarly up to time 0.6, whereas unsurprisingly, the stage 2 estimator performs worst. Regulation guidelines [13] suggest that methods for estimating treatment effect and confidence intervals with appropriate coverage should be provided as well as for controlling the prespecified type I error, whereas in [14], the importance of controlling the bias of the point estimate is emphasized. Hence, from the simulation findings and the importance of not overestimating treatment effect as described in [13, 14], we recommend the unbiased estimator.

## 5. Discussion

In drug development, the need to reduce the cost and time taken to test new treatments has led to the use of ASDs. ASDs combine several phases of a clinical development program into a single trial. However, compared with traditional testing strategies, ASDs pose additional challenges in statistical analysis. In this paper, we have considered point estimation following an ASD where, on the basis of observed data at stage 1, the experimental treatment that is superior to the competing experimental treatments at stage 1 continues to stage 2 together with the control. Cohen and Sackrowitz [15] and Shen [16] have considered this setting in the case where the trial always continues to stage 2 and proposed estimators for the treatment difference. Stallard and Todd [17] have also proposed an estimator that can be applied in this setting.

In this paper, we have considered the setting where the trial can stop for futility and estimation is unbiased conditional on continuing to stage 2. We have extended the Cohen and Sackrowitz method to construct an unbiased estimator for this setting. We have referred to this estimator as the (new) unbiased estimator. Carreras and Brannath [24] compared the Cohen and Sackrowitz estimator and the Stallard and Todd estimator when the trial always continues to stage 2. Their findings show that although the Cohen and Sackrowitz estimator is unbiased, it is not the best in terms of MSE. Thus, although the estimator we derive by extending the Cohen and Sackrowitz estimator to the setting where the trial can stop at stage 1 for futility is unbiased by construction, it is of interest to compare it with other estimators in terms of bias and MSE. Therefore, we have also developed a new bias-adjusted estimator that extends the Stallard and Todd estimator to our setting.

We also considered extending the Shen [16] estimator. The Shen estimator was proposed when the trial always continues to stage 2 and adjusts for bias by proposing a step function. When the trial always continues to stage 2, the step function depends on the absolute differences between the experimental treatment means and a tuning parameter. The best value for the tuning parameter depends on the unknown true values of the treatment means. With the possibility of early stopping, the bias depends not only on the absolute differences between the means of the experimental treatments but also on the values of observed differences between these and the mean of the control because of the futility boundary. This makes it challenging to propose a step function, and because we also know it will depend on a tuning parameter whose best value depends on the unknown true treatment means, we did not pursue this estimator further.

In terms of MSE, if treatment selection and the decision whether to continue to stage 2 are made at a selection time  $t < 0.6$ , the unbiased and bias-adjusted estimators perform similarly. The stage 2 estimator performs worst in terms of MSE, and the naive estimator (unadjusted for the possibility of stopping and for selection) performs the best. In terms of bias, the unbiased and stage 2 estimators are unbiased, and the naive estimator is positively biased whereas the bias-adjusted estimator is negatively biased. From this finding, we propose using the new unbiased estimator we have derived in this paper by extending the Cohen and Sackrowitz estimator [15] when a trial can stop for futility and estimation is performed conditional on continuing to stage 2. We emphasize that, although in the simulation study, we averaged over all simulations and the selected treatments, by derivation, the new unbiased estimator fulfills a stronger condition of unbiasedness in that it is unbiased with respect to each treatment whenever it is selected.

In this paper, we have considered point estimation following a two-stage adaptive seamless trial in which at stage 1, there is treatment selection and the possibility of early stopping for futility and estimation is conditional on the trial continuing to stage 2. As mentioned in Section 4.5, methods for interval estimation (confidence intervals) that adjust for the adaptation so that the right coverage is achieved are also important. There exist methods for constructing confidence intervals that can be used for the setting considered in this paper [17, 20, 25]. However, the confidence intervals following these methods are not based on the principle used to develop the estimators in this paper. For further research, we are considering confidence intervals based on the principle used to derive the unbiased estimator.

## Appendix A. Deriving the uniformly minimum variance unbiased estimator for $\mu_S$

For ease of notation in the derivation of UMVUEs for  $\mu_S$  and  $\mu_0$ , without loss of generality, we let  $X_1 > \dots > X_k$  so that  $X_{(i)} = X_i$  ( $i = 1, \dots, k$ ) and  $X_S = X_1$  and  $Y_S = Y_1$ . For the mean of the selected

treatment, we are seeking the UMVUE for  $\mu_1$ . We will skip details of steps that are similar to steps given in [15] and [26]. Denote by  $Q_B(X)$  the event  $\{X_0, X_1 > \dots > X_k, X_1 \geq B\}$ , where  $B = X_0 + b$ . We will first write and re-express the density of  $(Y_0, Y_1, X)$ , where  $X = (X_0, X_1, \dots, X_k)$  given  $Q_B$  in order to deduce the sufficient statistics for estimating  $\mu_1$  that combine stage 1 and 2 means for treatment 1 into a single quantity. The density of  $(Y_0, Y_1, X)$  given  $Q_B$ , denoted by  $f(y_1, y_0, x|Q_B)$ , is given by

$$K^{-1}(\mu) \frac{1}{\sigma_2} \phi\left(\frac{y_1 - \mu_1}{\sigma_2}\right) \frac{1}{\sigma_1} \phi\left(\frac{x_1 - \mu_1}{\sigma_1}\right) \mathbf{1}_{[Q_B(x)]} \psi(y_0, x_0, x'),$$

where  $\sigma_1$  and  $\sigma_2$  are as defined in Section 2.1,  $\mathbf{1}_{[Q_B]}$  is the indicator for  $Q_B(x)$ ,  $K(\mu) = \text{Prob}_\mu(\mathbf{1}_{[Q_B(x)]} = 1)$ , and

$$\psi(y_0, x_0, x') = \frac{1}{\sigma_2} \phi\left(\frac{y_0 - \mu_0}{\sigma_2}\right) \prod_{\substack{i=0 \\ i \neq 1}}^k \frac{1}{\sigma_1} \phi\left(\frac{x_i - \mu_i}{\sigma_1}\right).$$

The preceding density can be re-expressed as

$$K^{-1}(\mu) \frac{1}{\sigma_2} \phi\left(\frac{\frac{\sigma_2}{\sigma_1} x_1 + \frac{\sigma_1}{\sigma_2} y_1 - \mu_1 \alpha_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) \frac{1}{\sigma_1} \phi\left(\frac{x_1 - \left(\frac{\sigma_2}{\sigma_1} x_1 + \frac{\sigma_1}{\sigma_2} y_1\right) / \alpha_1}{\frac{\sigma_1}{\sigma_2} \alpha_2}\right) \mathbf{1}_{[Q_B(x)]} \psi(y_0, x_0, x'),$$

where

$$\alpha_1 = \frac{\sigma_1}{\sigma_2} + \frac{\sigma_2}{\sigma_1} \quad \text{and} \quad \alpha_2 = \frac{\sigma_2^2}{\sqrt{\sigma_1^2 + \sigma_2^2}}.$$

Let  $z_1 = (\sigma_2/\sigma_1)x_1 + (\sigma_1/\sigma_2)y_1$ ; then from the preceding density,  $(X_0, X_2, \dots, X_k, Y_0, Z_1)$  is sufficient and complete for the problem of seeking an estimate for  $\mu_1$  given  $Q_B$ . Therefore, conditional on  $Q_B$ , the UMVUE for  $\mu_1$  is given by  $E[Y_1|X_0, X_2, \dots, X_k, Y_0, Z_1, Q_B]$ . We obtain the expression for this by deriving the density  $f(y_1|x_0, x_2, \dots, x_k, y_0, z_1, Q_B)$  and using it to get the expected value we are seeking.

Transforming the density  $f(y_1, y_0, x|Q_B)$  into the density  $f(x, y_0, z_1|Q_B)$  gives

$$K^{-1}(\mu) \phi\left(\frac{z_1 - \mu_1 \alpha_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) \frac{1}{\sigma_1^2} \phi\left(\frac{x_1 - \frac{z_1}{\alpha_1}}{\frac{\sigma_1}{\sigma_2} \alpha_2}\right) \mathbf{1}_{[Q_B(x)]} \psi(y_0, x_0, x'),$$

and transforming the density  $f(y_1, y_0, x|Q_B)$  into the density  $f(y_1, x_0, x_2, \dots, x_k, y_0, z_1|Q_B)$  gives

$$K^{-1}(\mu) \phi\left(\frac{z_1 - \mu_1 \alpha_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) \frac{1}{\sigma_2^2} \phi\left(\frac{y_1 - \frac{z_1}{\alpha_1}}{\alpha_2}\right) \mathbf{1}_{[Q_B(x)]} \psi(y_0, x_0, x').$$

The density  $f(x_0, x_2, \dots, x_k, y_0, z_1|Q_B)$  is obtained from  $f(x, y_0, z_1|Q_B)$  by integrating out  $x_1$  as follows

$$\begin{aligned} & K^{-1}(\mu) \frac{1}{\sigma_1^2} \phi\left(\frac{z_1 - \mu_1 \alpha_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) \mathbf{1}_{[Q_B(x)]} \psi(y_0, x_0, x') \int_{\max\{B, x_2\}}^{\infty} \phi\left(\frac{x_1 - \frac{z_1}{\alpha_1}}{\frac{\sigma_1}{\sigma_2} \alpha_2}\right) dx_1 \\ & = K^{-1}(\mu) \frac{\alpha_2}{\sigma_2^2} \phi\left(\frac{z_1 - \mu_1 \alpha_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) \mathbf{1}_{[Q_B(x)]} \psi(y_0, x_0, x') [1 - \Phi(-W_B(1, 2))], \end{aligned}$$

where

$$\begin{aligned} W_B(1, 2) &= \frac{1}{\sigma_1^2} \left( \frac{\sigma_2^2 x_1 + \sigma_1^2 y_1}{\sqrt{\sigma_1^2 + \sigma_2^2}} - \max\{B, x_2\} \sqrt{\sigma_1^2 + \sigma_2^2} \right) \\ &= \left( \sqrt{\sigma_1^2 + \sigma_2^2} / \sigma_1^2 \right) (z_{1, \text{MLE}} - \max\{B, x_2\}). \end{aligned}$$

We are seeking the density of  $(Y_1|X_0, X_2, \dots, X_k, Y_0, Z_1, Q_B)$ , which is expressed as

$$\begin{aligned} f(y_1|x_0, x_2, \dots, x_k, y_0, z_1; Q_B) &= \frac{f(y_1, x_0, x_2, \dots, x_k, y_0, z_1|Q_B)}{f(x_0, x_2, \dots, x_k, y_0, z_1|Q_B)} \\ &= \frac{\frac{1}{\alpha_2} \phi\left(\frac{y_1 - \frac{z_1}{\alpha_1}}{\alpha_2}\right)}{1 - \Phi(-W_B(1, 2))} \cdot I\left[y_1 < \frac{\sigma_2}{\sigma_1} \left(z_1 - \frac{\sigma_2}{\sigma_1} \max\{B, x_2\}\right)\right]. \end{aligned}$$

Therefore,  $E[Y_1|X_0, X_2, \dots, X_k, Y_0, Z_1, Q_B]$  is equal to

$$\frac{\frac{1}{\alpha_2} \int_{-\infty}^{\frac{\sigma_2}{\sigma_1} \left(z_1 - \frac{\sigma_2}{\sigma_1} \max\{B, x_2\}\right)} y_1 \phi\left(\frac{y_1 - \frac{z_1}{\alpha_1}}{\alpha_2}\right) dy_1}{1 - \Phi(-W_B(1, 2))} = \frac{\int_{-\infty}^{\frac{\sigma_2}{\sigma_1} \left(z_1 - \frac{\sigma_2}{\sigma_1} \max\{B, x_2\}\right)} \frac{y_1}{\alpha_2} \phi\left(\frac{y_1 - \frac{z_1}{\alpha_1}}{\alpha_2}\right) dy_1}{\Phi(W_B(1, 2))}.$$

Following [26], the numerator of the preceding expectation simplifies to

$$-\alpha_2 \phi(W_B(1, 2)) + \frac{z_1}{\alpha_1} \Phi(W_B(1, 2))$$

so that  $E[Y_1|X_0, X_2, \dots, X_k, Y_0, Z_1, Q_B]$  reduces to

$$\frac{z_1}{\alpha_1} - \alpha_2 \frac{\phi(W_B(1, 2))}{\Phi(W_B(1, 2))} = \frac{\sigma_2^2 x_1 + \sigma_1^2 y_1}{\sigma_1^2 + \sigma_2^2} - \frac{\sigma_2^2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \frac{\phi(W_B(1, 2))}{\Phi(W_B(1, 2))}$$

and formula (3) gives the estimator when there is the possibility of stopping for futility and estimation is unbiased conditional on continuing to stage 2.

## Appendix B. Deriving the uniformly minimum variance unbiased estimator for $\mu_0$

Let  $Z_0 = (\sigma_2/\sigma_1)X_0 + (\sigma_1/\sigma_2)Y_0$ . If we write and re-express the density  $f(x_0, x_1, \dots, x_k, y_0, y_1)$  as in Appendix A, we can deduce that  $(Z_0, Y_1, X_1, \dots, X_k)$  is sufficient and complete for estimating  $\mu_0$ . Thus, we are seeking  $E[Y_0|Z_0, Y_1, X_1, \dots, X_k]$  so that if we denote by  $X$  the vector  $(X_1, \dots, X_k)'$ , we need the density

$$f(y_0|z_0, y_1, x) = \frac{f(y_0, z_0, y_1, x)}{f(z_0, y_1, x)}.$$

Similar to derivation in Appendix A, it can be shown that the density  $f(z_0, y_1, x_0, x)$  is given by

$$K^{-1}(\mu) \phi\left(\frac{z_0 - \mu_0 \alpha_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) \frac{1}{\sigma_1^2} \phi\left(\frac{x_0 - z_0/\alpha_1}{\frac{\sigma_1^2}{\sigma_2^2} \alpha_2}\right) \mathbf{1}_{[Q_B(x)]} \psi(y_1, x),$$

where

$$\psi(y_1, x) = \frac{1}{\sigma_2} \phi\left(\frac{y_1 - \mu_1}{\sigma_2}\right) \prod_{i=1}^k \frac{1}{\sigma_1} \phi\left(\frac{x_i - \mu_i}{\sigma_1}\right)$$

and  $\mathbf{1}_{[Q_B(x)]}$  and  $K(\mu)$  are as defined in Appendix A whereas the density  $f(y_0, z_0, y_1, x)$  is given by

$$K^{-1}(\mu)\phi\left(\frac{z_0 - \mu_0\alpha_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)\frac{1}{\sigma_2^2}\phi\left(\frac{y_0 - z_0/\alpha_1}{\alpha_2}\right)\mathbf{1}_{[Q_B(x)]}\psi(y_1, x). \quad (14)$$

The density  $f(z_0, y_1, x)$  is obtained by integrating out  $x_0$  in the density  $f(z_0, y_1, x_0, x)$ . The range of  $x_0$  is  $-\infty$  to  $B_1 = (x_1 - b)$  because the trial continues to stage 2 if  $(x_1 - x_0) \geq b$ . The density  $f(z_0, y_1, x)$  becomes

$$K^{-1}(\mu)\phi\left(\frac{z_0 - \mu_0\alpha_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)\frac{\alpha_2}{\sigma_2^2}\cdot\Phi(W_{B_1})\mathbf{1}_{[Q_B(x)]}\psi(y_1, x), \quad (15)$$

where

$$\begin{aligned} W_{B_1} &= \frac{1}{\sigma_1^2}\left(B_1\cdot\sqrt{\sigma_1^2 + \sigma_2^2} - \frac{\sigma_2^2x_0 + \sigma_1^2y_0}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) \\ &= \left(\sqrt{\sigma_1^2 + \sigma_2^2}/\sigma_1^2\right)(B_1 - z_{0,\text{MLE}}). \end{aligned}$$

Using (14) and (15), we obtain

$$f(y_0, z_0, y_1, x_1, \dots, x_k) = \frac{\frac{1}{\alpha_2}\phi\left(\frac{y_0 - z_0/\alpha_1}{\alpha_2}\right)}{\Phi(W_{B_1})}\cdot I\left[y_0 > \frac{\sigma_2}{\sigma_1}\left(z_0 - \frac{\sigma_2}{\sigma_1}B_1\right)\right]$$

so that the expression for  $E[Y_0|Z_0, Y_1, X_1, \dots, X_k]$  is

$$\frac{\frac{1}{\alpha_2}\int_{\frac{\sigma_2}{\sigma_1}(z_0 - \frac{\sigma_2}{\sigma_1}B_1)}^{\infty} y_0\phi\left(\frac{y_0 - z_0/\alpha_1}{\alpha_2}\right)dy_0}{\Phi(W_{B_1})}. \quad (16)$$

The numerator can be re-expressed as

$$\int_{\frac{\sigma_2}{\sigma_1}(z_0 - \frac{\sigma_2}{\sigma_1}B_1)}^{\infty} \frac{y_0}{\alpha_2}\phi\left(\frac{y_0 - z_0/\alpha_1}{\alpha_2}\right)dy_0 = \frac{Z_0}{\alpha_1} - \int_{\infty}^{\frac{\sigma_2}{\sigma_1}(z_0 - \frac{\sigma_2}{\sigma_1}B_1)} \frac{y_0}{\alpha_2}\phi\left(\frac{y_0 - z_0/\alpha_1}{\alpha_2}\right)dy_0.$$

Similar to derivation in Appendix A, it can be shown that

$$\int_{\infty}^{\frac{\sigma_2}{\sigma_1}(z_0 - \frac{\sigma_2}{\sigma_1}B_1)} \frac{y_0}{\alpha_2}\phi\left(\frac{y_0 - z_0/\alpha_1}{\alpha_2}\right)dy_0 = -\alpha_2\phi(-W_{B_1}) + \frac{Z_0}{\alpha_1}\Phi(-W_{B_1})$$

so that the numerator in Equation (16) becomes

$$\begin{aligned} \frac{Z_0}{\alpha_1} + \alpha_2\phi(-W_{B_1}) - \frac{Z_0}{\alpha_1}\Phi(-W_{B_1}) &= \frac{Z_0}{\alpha_1}(1 - \Phi(-W_{B_1})) + \alpha_2\phi(-W_{B_1}) \\ &= \frac{Z_0}{\alpha_1}\Phi(W_{B_1}) + \alpha_2\phi(W_{B_1}). \end{aligned} \quad (17)$$

By substituting (17) in (16),  $E[Y_0|Z_0, Y_1, X_1, \dots, X_k]$  simplifies to

$$\frac{\frac{Z_0}{\alpha_1}\Phi(W_{B_1}) + \alpha_2\phi(W_{B_1})}{\Phi(W_{B_1})} = \frac{Z_0}{\alpha_1} + \alpha_2\frac{\phi(W_{B_1})}{\Phi(W_{B_1})},$$

so that the UMVUE for  $\mu_0$  is given by expression (5).

## Appendix C. Expressions for the expected values

### C.1. Probability of selecting a treatment and its expected value

The expression for the numerator in Equation (11) is given by

$$\int_b^\infty d_i \int_{-\infty}^\infty \frac{1}{\sigma_1^2} \phi\left(\frac{w-d_i}{\sigma_1}\right) \phi\left(\frac{w-\theta_i}{\sigma_1}\right) \prod_{j \neq i} \Phi\left(\frac{w-\theta_j}{\sigma_1}\right) dw dd_i,$$

which by changing the order of integration becomes

$$\int_{-\infty}^\infty \frac{1}{\sigma_1} \phi\left(\frac{w-\theta_i}{\sigma_1}\right) \prod_{j \neq i} \Phi\left(\frac{w-\theta_j}{\sigma_1}\right) \int_b^\infty \frac{d_i}{\sigma_1} \phi\left(\frac{d_i-w}{\sigma_1}\right) dd_i dw.$$

As described in Appendices A and B, integration with respect to  $d_i$  is given by

$$w - \int_{-\infty}^b \frac{d_i}{\sigma_1} \phi\left(\frac{d_i-w}{\sigma_1}\right) dd_i = w - \left[ w \Phi\left(\frac{b-w}{\sigma_1}\right) - \sigma_1 \phi\left(\frac{b-w}{\sigma_1}\right) \right]$$

so that the numerator in Equation (11) is given by

$$\int_{-\infty}^\infty \frac{1}{\sigma_1} \phi\left(\frac{w-\theta_i}{\sigma_1}\right) \prod_{j \neq i} \Phi\left(\frac{w-\theta_j}{\sigma_1}\right) \left[ \sigma_1 \phi\left(\frac{b-w}{\sigma_1}\right) + w \Phi\left(\frac{w-b}{\sigma_1}\right) \right] dw.$$

The  $\text{pr}(S = i, D_i \geq b)$ , the denominator in Equation (11), is obtained by solving the integral  $\int_b^\infty f(d_i, S = i) dd_i$ , which can be simplified as follows

$$\begin{aligned} & \int_b^\infty \int_{-\infty}^\infty \frac{1}{\sigma_1^2} \phi\left(\frac{w-d_i}{\sigma_1}\right) \phi\left(\frac{w-\theta_i}{\sigma_1}\right) \prod_{j \neq i} \Phi\left(\frac{w-\theta_j}{\sigma_1}\right) dw dd_i \\ &= \int_{-\infty}^\infty \frac{1}{\sigma_1} \phi\left(\frac{w-\theta_i}{\sigma_1}\right) \prod_{j \neq i} \Phi\left(\frac{w-\theta_j}{\sigma_1}\right) \int_b^\infty \frac{1}{\sigma_1} \phi\left(\frac{d_i-w}{\sigma_1}\right) dd_i dw \\ &= \int_{-\infty}^\infty \frac{1}{\sigma_1} \phi\left(\frac{w-\theta_i}{\sigma_1}\right) \prod_{j \neq i} \Phi\left(\frac{w-\theta_j}{\sigma_1}\right) \left[ 1 - \Phi\left(\frac{b-w}{\sigma_1}\right) \right] dw. \end{aligned}$$

### C.2. Expected value for the dropped treatments

Without loss of generality, suppose treatment 1 is selected to continue to stage 2. Then for a dropped treatment  $i'$ , the density of  $W_{i'}$  when  $d_1 \geq b$  is given by

$$\int_{w_{i'} < w_1} \int_{w_1 - w_0 \geq b} \frac{1}{\sigma_1^3} \phi\left(\frac{w_0}{\sigma_1}\right) \phi\left(\frac{w_1 - \theta_1}{\sigma_1}\right) \phi\left(\frac{w_{i'} - \theta_{i'}}{\sigma_1}\right) \prod_{\substack{j \neq 1 \\ j \neq i'}} \Phi\left(\frac{w_1 - \theta_j}{\sigma_1}\right) dw_0 dw_1,$$

which simplifies to

$$\int_{w_{i'}}^\infty \frac{1}{\sigma_1^2} \phi\left(\frac{w_{i'} - \theta_{i'}}{\sigma_1}\right) \phi\left(\frac{w_1 - \theta_1}{\sigma_1}\right) \Phi\left(\frac{w_1 - b}{\sigma_1}\right) \prod_{\substack{j \neq 1 \\ j \neq i'}} \Phi\left(\frac{w_1 - \theta_j}{\sigma_1}\right) dw_1$$

so that  $E[W_{i'}, S = 1, D_1 \geq b]$  is

$$\int_{-\infty}^\infty \int_{w_{i'}}^\infty \frac{w_{i'}}{\sigma_1^2} \phi\left(\frac{w_{i'} - \theta_{i'}}{\sigma_1}\right) \phi\left(\frac{w_1 - \theta_1}{\sigma_1}\right) \Phi\left(\frac{w_1 - b}{\sigma_1}\right) \prod_{\substack{j \neq 1 \\ j \neq i'}} \Phi\left(\frac{w_1 - \theta_j}{\sigma_1}\right) dw_1 dw_{i'}.$$

Changing the order of integration gives

$$\int_{-\infty}^\infty \int_{-\infty}^{w_1} \frac{w_{i'}}{\sigma_1} \phi\left(\frac{w_{i'} - \theta_{i'}}{\sigma_1}\right) \frac{1}{\sigma_1} \phi\left(\frac{w_1 - \theta_1}{\sigma_1}\right) \Phi\left(\frac{w_1 - b}{\sigma_1}\right) \prod_{\substack{j \neq 1 \\ j \neq i'}} \Phi\left(\frac{w_1 - \theta_j}{\sigma_1}\right) dw_{i'} dw_1.$$



From Appendices A and B, we can write the closed-form solution of the integration with respect to  $w_{i'}$  so that the expression for  $E[W_{i'}, S = i, D_i \geq b]$  is

$$\int_{-\infty}^{\infty} \psi(w_i, \theta_{i'}, \sigma_1) \frac{1}{\sigma_1} \phi\left(\frac{w_i - \theta_i}{\sigma_1}\right) \Phi\left(\frac{w_i - b}{\sigma_1}\right) \prod_{\substack{j \neq i \\ j \neq i'}} \Phi\left(\frac{w_i - \theta_j}{\sigma_1}\right) dw_i,$$

where

$$\psi(w_i, \theta_{i'}, \sigma_1) = \theta_{i'} \Phi\left(\frac{w_i - \theta_{i'}}{\sigma_1}\right) - \sigma_1 \phi\left(\frac{w_i - \theta_{i'}}{\sigma_1}\right).$$

The density for  $W_0$  when treatment 1 is selected and  $d_1 \geq b$  is given by

$$\begin{aligned} f(w_0, d_1 \geq b) &= \int_{w_1 - w_0 \geq b} \int_{-\infty}^{w_1} \dots \int_{-\infty}^{w_1} f(w_0) f(w_1) f(w_2) \dots f(w_k) dw_2 \dots dw_k dw_1 \\ &= \int_{w_0 + b}^{\infty} \frac{1}{\sigma_1} \phi\left(\frac{w_0}{\sigma_1}\right) \frac{1}{\sigma_1} \left(\frac{w_1 - \theta_1}{\sigma_1}\right) \prod_{j \neq 1} \Phi\left(\frac{w_1 - \theta_j}{\sigma_1}\right) dw_1 \end{aligned}$$

so that  $E[W_0, S = 1, D_1 \geq b]$  is given by

$$\int_{-\infty}^{\infty} \int_{w_0 + b}^{\infty} \frac{w_0}{\sigma_1} \phi\left(\frac{w_0}{\sigma_1}\right) \frac{1}{\sigma_1} \phi\left(\frac{w_1 - \theta_1}{\sigma_1}\right) \prod_{j \neq 1} \Phi\left(\frac{w_1 - \theta_j}{\sigma_1}\right) dw_1 dw_0,$$

which by changing the order of integration becomes

$$\int_{-\infty}^{\infty} \int_{-\infty}^{w_1 - b} \frac{w_0}{\sigma_1} \phi\left(\frac{w_0}{\sigma_1}\right) \frac{1}{\sigma_1} \phi\left(\frac{w_1 - \theta_1}{\sigma_1}\right) \prod_{j \neq 1} \Phi\left(\frac{w_1 - \theta_j}{\sigma_1}\right) dw_0 dw_1.$$

When we use the results in earlier parts of the appendices, the solution for the integral with respect to  $w_0$  is  $-\sigma_1 \phi\{(w_1 - b)/\sigma_1\}$  so that  $E[W_0, S = 1, D_1 \geq b]$  is given by

$$E[W_0, S = i, D_i \geq b] = - \int_{-\infty}^{\infty} \phi\left(\frac{w_i - b}{\sigma_1}\right) \phi\left(\frac{w_i - \theta_i}{\sigma_1}\right) \prod_{j \neq i} \Phi\left(\frac{w_i - \theta_j}{\sigma_1}\right) dw_i.$$

## Acknowledgements

The authors would like to thank the associate editor and two reviewers for comments that helped improve the paper. This work was supported in part by the Targeted Research Programme of the Medicines and Healthcare products Regulatory Agency (MHRA).

## References

1. Thall PF, Simon R, Ellenberg SS. Two-stage selection and testing designs for comparative clinical trials. *Biometrika* 1988; **75**:303–310.
2. Schaid DJ, Wieand S, Therneau TM. Optimal two-stage screening designs for survival comparison. *Biometrika* 1990; **77**:507–513.
3. Stallard N, Todd S. Sequential designs for phase III clinical trials incorporating treatment selection. *Statistics in Medicine* 2003; **22**:689–703.
4. Jennison C, Turnbull BW. Adaptive seamless designs: selection and prospective testing of hypotheses. *Journal of Biopharmaceutical Statistics* 2007; **17**:1135–1161.
5. Hommel G. Adaptive modifications of hypotheses after an interim analysis. *Biometrical Journal* 2001; **43**:581–589.
6. Bretz F, König 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.
7. 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; **26**:4925–4938.
8. Brannath W, Zuber E, Branson M, Bretz F, Gallo P, Posch M, Racine-Poon A. Confirmatory adaptive designs with Bayesian decision tools for a targeted therapy in oncology. *Statistics in Medicine* 2009; **28**:1445–1463.
9. Kimani PK, Stallard N, Hutton JL. Dose selection in seamless phase II/III clinical trials based on efficacy and safety. *Statistics in Medicine* 2009; **28**:917–936.

10. Müller H, Schäffer H. Adaptive group sequential designs for clinical trials: combining the advantages of adaptive and of classical group sequential approaches. *Biometrics* 2001; **57**:886–891.
11. Koenig F, Brannath W, Bretz F, Posch M. Adaptive dunnett tests for treatment selection. *Statistics in Medicine* 2008; **27**:1612–1625.
12. Bauer P, Koenig F, Brannath W, Posch M. Selection and bias—two hostile brothers. *Statistics in Medicine* 2010; **29**:1–13.
13. CHMP, EMEA. Reflection paper on methodological issues in confirmatory clinical trials planned with an adaptive design @ONLINE, 11 February 2013. (Available from: [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Scientific\\_guideline/2009/09/WC500003616.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003616.pdf)).
14. Food and Drug Administration. Guidance for industry: adaptive design clinical trials for drugs and biologics @ONLINE, 11 February 2013. (Available from: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM201790.pdf>).
15. Cohen A, Sackrowitz HB. Two stage conditionally unbiased estimators of the selected mean. *Statistics and Probability Letters* 1989; **8**:273–278.
16. Shen L. Estimation following selection of the largest of two normal means. *Statistics in Medicine* 2001; **20**:1913–1929.
17. Stallard N, Todd S. Point estimates and confidence regions for sequential trials involving selection. *Journal of Statistical Planning and Inference* 2005; **135**:402–419.
18. Koopmeiners JS, Feng Z, Pepe MS. Conditional estimation after a two-stage diagnostic biomarker study that allows early termination for futility. *Statistics in Medicine* 2012; **31**:420–435.
19. Troendle JF, Yu KF. Conditional estimation following a group sequential clinical trial. *Communication in Statistics—Theory and Methods* 1999; **28**(7):1617–1634.
20. Posch M, Koenig F, Branson M, Brannath W, Dunger-Baldauf C, Bauer P. Testing and estimation in flexible group sequential designs with adaptive treatment selection. *Statistics in Medicine* 2005; **24**:3697–3714.
21. Bretz F, Koenig F, Brannath W, Glimm E, Posch M. Adaptive designs for confirmatory clinical trials. *Statistics in Medicine* 2009; **28**:1181–1217.
22. Lindgren BW. *Statistical Theory*, (3rd edn). Collier Macmillan International Editions: London, 1976.
23. Whitehead J. On the bias of maximum likelihood estimation following a sequential test. *Biometrika* 1986; **73**(3):573–581.
24. Carreras M, Brannath W. Shrinkage estimation in two-stage adaptive designs with mid-trial treatment selection. *Statistics in Medicine* 2013. DOI: 10.1002/sim.5463.
25. Wu SS, Wang W, Yang MCK. Interval estimation for drop-the-losers design. *Biometrika* 2010; **97**(2):405–418.
26. Bowden J, Glimm E. Unbiased estimation of selected treatment means in two-stage trials. *Biometrical Journal* 2008; **50**(4):515–527.