Original citation:

Permanent WRAP URL:
http://wrap.warwick.ac.uk/97283

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions.

This article is made available under the Creative Commons Attribution 4.0 International license (CC BY 4.0) and may be reused according to the conditions of the license. For more details see: http://creativecommons.org/licenses/by/4.0/

A note on versions:
The version presented in WRAP is the published version, or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk
Uniformly minimum variance conditionally unbiased estimation in multi-arm multi-stage clinical trials

BY NIGEL STALLARD AND PETER K. KIMANI

Warwick Medical School, University of Warwick, Coventry CV4 7AL, U.K.
n.stallard@warwick.ac.uk  peter.k-u.kimani@warwick.ac.uk

SUMMARY

Multi-arm multi-stage clinical trials compare several experimental treatments with a control treatment, with poorly performing treatments dropped at interim analyses. This leads to inferential challenges, including the construction of unbiased treatment effect estimators. A number of estimators which are unbiased conditional on treatment selection have been proposed, but are specific to certain selection rules, may ignore the comparison to the control and are not all minimum variance. We obtain estimators for treatment effects compared to the control that are uniformly minimum variance unbiased conditional on selection with any specified rule or stopping for futility.

Some key words: Adaptive seamless design; Drop-the-loser design; Point estimation; Treatment selection.

1. INTRODUCTION

Multi-arm multi-stage clinical trials compare several treatments to a common control treatment in a single trial with treatments dropped at interim analyses if, based on observed data, they are not sufficiently promising. Such designs have been used by, for example, MacArthur et al. (2013) and Barker et al. (2014). The approach can yield sample size reduction and administrative savings relative to running several two-arm trials, but presents challenges in terms of statistical analysis similar to those of post-model selection inference (Efron, 2014).

Proposed analysis methods have mainly focused on frequentist hypothesis tests (Thall et al., 1988, 1989; Stallard & Todd, 2003; Stallard & Friede, 2008; Magirr et al., 2012; Wason et al., 2017), with less work on estimation. Cohen & Sackrowitz (1989) consider two-stage designs with the treatment with the highest observed stage 1 average continuing to stage 2, with equal variances for the averages for each treatment in stage 1 and the stage 2 treatment. They derive an estimator of the stage 2 treatment mean that is uniformly minimum variance conditionally unbiased given the observed ordering of stage 1 treatment means. Bowden & Glimm (2008) extend the method to allow different variances for the treatment means and continuation to stage 2 of the s treatments with the largest observed stage 1 means. They provide expressions for uniformly minimum variance conditionally unbiased estimators for the means for these s treatment arms, again conditioning on the ordering of the stage 1 averages. Like Cohen & Sackrowitz (1989), they do not consider estimation relative to a control group.

Kimani et al. (2013), Bowden & Glimm (2014) and Robertson et al. (2016) derive conditionally unbiased estimators for the difference between selected treatments and a control in multi-arm multi-stage trials. Kimani et al. (2013) allow for stopping for futility at stage 1 of a two-stage trial assuming common variances for averages in different arms in stage 1, while Bowden & Glimm (2014) and Robertson et al. (2016) allow for different variances in different arms and correlation between these respectively. Bowden & Glimm (2014) also allow for more than two stages. As discussed below, the Kimani et al. (2013) and Robertson et al. (2016) estimators are conditionally unbiased but not generally minimum variance. We obtain minimum variance conditionally unbiased estimators in these settings.
Without loss of generality, label treatments such that specified rule for selection or stopping for futility. We show how uniformly minimum variance conditionally unbiased estimators may be obtained for comparisons with the control conditioning on a treatment not being dropped from the trial using any specified rule for selection or stopping for futility.

2. Uniformly minimum variance conditionally unbiased estimation

2.1. A uniformly minimum variance conditionally unbiased estimator

Consider a multi-arm multi-stage clinical trial with up to \( r \) stages where, in stage \( j \), patients are randomized to a control treatment, treatment 0, or experimental treatments, labelled \( i \in \Omega_j \), with \( \Omega_1 \supseteq \cdots \supseteq \Omega_r \). Without loss of generality, label treatments such that \( \Omega_j = \{1, \ldots, s_j\} \) for some \( s_j \) \((j = 1, \ldots, r)\). Denoting the total number of experimental treatments, \( s_i \), by \( k \), let \( r_i = \max\{j: s_j \geq l\} \) \((i = 1, \ldots, k)\), so that treatment \( i \) is included in stages \( 1, \ldots, r_i \) and set \( r_0 = r \).

Let \( X_{ij} \) denote the stagewise average for treatment \( i \) in stage \( j \) \((i = 0, \ldots, k; j = 1, \ldots, r_i)\), with \( X_i = (X_{i1}, \ldots, X_{ip})^\top \), \( X_j = (X_{j1}, \ldots, X_{jp})^\top \) and \( \mathcal{X}_r = (X_1^\top, \ldots, X_r^\top)^\top \). Assume that the \( X_{ij} \) are jointly sufficient for \( \mu_0, \ldots, \mu_k \), with \( X_y \sim N(\mu, \tau_y^{-1}) \) independent with \( \tau_y \) known. Other cases, such as in §3 below, may use normal approximations or estimated variances. Experimental treatments are selected to continue, along with the control, to stage \( j+1 \) depending on \( \mathcal{X}_j \) according to some prespecified rule. Some possible rules are discussed below.

Define \( \theta_i = \mu_i - \mu_0 \) \((i = 1, \ldots, k)\). We wish to estimate \( \theta_i \) \((i = 1, \ldots, s_r)\). In particular, we will obtain uniformly minimum variance conditionally unbiased estimators of \( \theta_1, \ldots, \theta_{s_r} \) conditional on the event, \( Q \), that treatments \( 1, \ldots, s_r \) are selected to continue to the end of the trial according to the prespecified selection rule.

Let \( \tau_i = (\tau_{i1}, \ldots, \tau_{ip})^\top \), \( \tau_i = \tau_{i1} + \cdots + \tau_{ip} \), \( Z_i = X_i^\top \tau_i \tau_i^{-1} \) \((i = 0, \ldots, k)\), \( Z = (Z_0, \ldots, Z_k)^\top \) and \( \theta_{k+1} = \sum_{i=0}^k \tau_i \mu_i / \sum_{i=0}^k \tau_i \). The Appendix shows that \( \theta_{k+1} \) is orthogonal to \( \theta_i \), in that the information matrix term \( \phi_{i,j+1} = 0 \) \((i = 1, \ldots, k)\) and \( Z \) is complete and sufficient for \( \theta = (\theta_1, \ldots, \theta_{k+1})^\top \).

Let \( Y_j = X_{ij} - X_{i0} \) \((i = 1, \ldots, s_r)\); then \( Y_i \) is unbiased for \( \theta_i \) and, since \( Y_i \) is independent of \( \mathcal{X}_{r-1} \) and hence of \( Q \), is also conditionally unbiased for \( \theta_i \) given \( Q \). Thus, by the Rao–Blackwell theorem, \( E(Y_j \mid Z, Q) \) is a uniformly minimum variance conditionally unbiased estimator for \( \theta_i \) given \( Q \) \((i = 1, \ldots, s_r)\).

Since, given the specified selection rule, the event \( Q \) depends only on \( \mathcal{X}_{r-1} \),

\[
E(Y_i \mid Z, Q) = \frac{\int_Q E(Y_i \mid Z, \mathcal{X}_{r-1}) f(\mathcal{X}_{r-1} \mid Z) \, d\mathcal{X}_{r-1}}{\int_Q f(\mathcal{X}_{r-1} \mid Z) \, d\mathcal{X}_{r-1}},
\]

with the integrals taken over the region \( Q \) corresponding to those \( \mathcal{X}_{r-1} \) for which treatments \( 1, \ldots, s_r \) will be selected to continue to stage \( r \). The denominator is the probability of \( Q \) given \( Z \), which will be denoted by \( \Pr(Q \mid Z) \).

As

\[
\begin{pmatrix}
X_{i1} \\
\vdots \\
X_{i,r-1} \\
Z_i
\end{pmatrix} \sim N
\begin{pmatrix}
\mu_i \\
\vdots \\
\mu_i
\end{pmatrix},
\begin{pmatrix}
\tau_{i1}^{-1} & 0 & \cdots & 0 & \tau_{i1}^{-1} \\
0 & \tau_{i2}^{-1} & \cdots & \vdots & \vdots \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & \cdots & 0 & \tau_{i,r-1}^{-1} & \tau_{i,r-1}^{-1} \\
\tau_{i1}^{-1} & \cdots & \cdots & \cdots & \tau_{i,r-1}^{-1}
\end{pmatrix}
\]

and, for \( i \neq i' \) and any \( j,j' \), \( X_{ij} \) is independent of \( X_{i'j'} \) and \( Z_i \), we have that

\[
f(\mathcal{X}_{r-1} \mid Z) = \prod_{i=0}^k \phi_{i,j-1}(\langle X_{i1}, \ldots, X_{i,r-1} \rangle; \mu_i, \Sigma_i),
\]

as in §3.
where \( \phi_n(X; \mu, \Sigma) \) denotes the \( n \)-dimensional multivariate normal density with mean \( \mu \) and variance matrix \( \Sigma \), evaluated at \( X \), with \( \mu_i = (Z_0, \ldots, Z_i)^\top \) and \( \Sigma_i \) having diagonal elements \( (\tau_{i1} - \tau_{i1}^{-1}, \ldots, \tau_{ir} - \tau_{ir}^{-1}) \) and all off-diagonal elements equal to \( -\tau_{ij}^{-1} \). If \( r_i = 1 \), \( X_{i1} \) given \( Z \) has a degenerate distribution, since \( X_{i1} = Z_i \) and \( \tau_{i1} = \tau_i \).

Since \( Y_i = X_{ir} - X_{0r} \) and \( X_{ir} = \tau_{ir}^{-1}(\tau_iZ_i - \sum_{j=1}^{r-1}X_{ij}\tau_{ij}) \) (\( i = 0, \ldots, s_r \)),

\[
E(Y_i \mid Z, \mathcal{R}_{r-1}) = \tau_{ir}^{-1}\tau_iZ_i - \tau_{0r}^{-1}r_0Z_0 - \sum_{j=1}^{r-1}(\tau_{ir}^{-1}\tau_jX_{ij} - \tau_{0r}^{-1}\tau_0X_{0j}).
\]

Thus (1) gives

\[
E(Y_i \mid Z, Q) = \tau_{ir}^{-1}\tau_iZ_i - \tau_{0r}^{-1}r_0Z_0 - \text{pr}(Q \mid Z)^{-1}\int_\mathcal{R}_{r-1} \sum_{j=1}^{r-1}(\tau_{ir}^{-1}\tau_jX_{ij} - \tau_{0r}^{-1}\tau_0X_{0j}) f(\mathcal{R}_{r-1} \mid Z) d\mathcal{R}_{r-1}. \tag{3}
\]

An important special case arises when the selection rule does not depend on \( X_0 \), the observed averages for the control arm. In the integral over the event \( Q \), the range of integration with respect to \( X_i \) (\( i = 1, \ldots, k; j = 1, \ldots, r_i \)) then does not depend on \( X_0 \) and the integration with respect to the elements of \( X_0 \), i.e., \( X_{01}, \ldots, X_{0r-1} \), is over the whole real line. Thus

\[
\text{pr}(Q \mid Z)^{-1}\int_\mathcal{R}_{r-1} \tau_{0r}^{-1}\sum_{j=1}^{r-1}X_{0j}\tau_0 f(\mathcal{R}_{r-1} \mid Z) d\mathcal{R}_{r-1} = \tau_{0r}^{-1}\sum_{j=1}^{r-1}\tau_0 E(X_{0j} \mid Z),
\]

and since \( E(X_{0j} \mid Z) = Z_j \) (\( j = 1, \ldots, r - 1 \)), we have

\[
E(Y_i \mid Z, Q) = \tau_{ir}^{-1}\tau_iZ_i - \text{pr}(Q \mid Z)^{-1}\int_\mathcal{R}_{r-1} \tau_{ir}^{-1}\sum_{j=1}^{r-1}X_{ij}\tau_j f(\mathcal{R}_{r-1} \mid Z) d\mathcal{R}_{r-1} - Z_0. \tag{4}
\]

In this case the uniformly minimum variance conditionally unbiased estimator for \( \theta_i \) is the difference between the uniformly minimum variance conditionally unbiased estimator of \( \mu_i \), which can be calculated ignoring the observed value of \( X_0 \), and \( Z_0 \), the usual uniformly minimum variance unbiased estimator of \( \mu_0 \), which does not depend on the selection.

The integrals in (3) generally cannot be evaluated using standard functions. A numerical approach is Monte Carlo integration with rejection sampling, simulating \( \mathcal{R}_{r-1} \) from (2), its conditional distribution given \( Z \), accepting those in \( Q \) for the specified selection rule. This approach can be used for any selection rule with treatments proceeding to stage \( j + 1 \) dependent on \( \mathcal{R}_j \).

2.2. Selection of the best-performing treatment in a two-stage trial

Much previous work on conditionally unbiased estimation in clinical trials with treatment selection has focused on the special case of \( r = 2 \) and \( s_2 = 1 \), with the experimental treatment having the largest observed stage 1 mean continuing along with the control to stage 2. Thus \( Q \) is the event \( X_{11} > m \) with \( m = \max(X_{31}, \ldots, X_{k1}) \). As \( Q \) is independent of \( X_0 \), the uniformly minimum variance conditionally unbiased estimator for \( \theta \) is (4). Since treatments 2, \ldots, \( k \) are observed in stage 1 only, we have \( Z_i = X_{i1} \) (\( i = 2, \ldots, k \)) so that the numerator and denominator in the fractional term in (4) are respectively \( \tau_{11}^{-1}\int_0^\infty X_{11} f(X_{11} \mid Z_1) dX_{11} \) and \( \text{pr}(X_{11} > m \mid Z_1) \).

As \( X_{11} \mid Z_1 \sim N(Z_1, \upsilon_1) \) with \( \upsilon_1 = \tau_{11}^{-1} - \tau_{11}^{-1} \), this integral and probability are \( Z_1[1 - \Phi(\tilde{Z}_1)] + \upsilon_1^{1/2}\phi(\tilde{Z}_1) \) (Todd et al., 1996) and \( 1 - \Phi(\tilde{Z}_1) \) respectively, where \( \tilde{Z}_1 = (m - Z_1)\upsilon_1^{1/2} \). Thus \( E(Y_i \mid Z, Q) = Z_i - Z_0 - \tau_{11}^{-1}\upsilon_1^{1/2}\phi(\tilde{Z}_1)/[1 - \Phi(\tilde{Z}_1)] \) confirming this to be the difference between the uniformly minimum variance conditionally unbiased estimator for \( \mu_1 \) ignoring the control treatment given by Bowden & Glimm (2008) and \( Z_0 \).

Kimani et al. (2013) consider a similar two-stage trial in which treatment 1 is selected if \( X_{11} > m \) but continues to the second stage only if \( X_{11} > X_{01} + c \) for some specified \( c \), the trial otherwise stopping.

\[
}\text{Downloaded from https://academic.oup.com/biomet/article-abstract/105/2/495/4913762 by guest on 05 June 2018}\]
The primary endpoint was clinical response, defined as at most two watery stools per week during at least two percentage points above the placebo. The trial was early. In this case, \( Q \) is the event \( X_{11} > \max(m, X_{01} + c) \). Since this depends on \( X_{01} \), the form (3) must be used rather than (4) to give the uniformly minimum variance conditionally unbiased estimator of \( \theta_1 \). The integral in (3) is taken over \( X_{11} \) and \( X_{01} \), but may be rewritten in terms of \( X_{11} \) and \( (X_{11} - X_{01}) \), noting that

\[
\begin{pmatrix}
X_{11} \\
X_{11} - X_{01}
\end{pmatrix}
\sim N
\left\{
\begin{pmatrix}
Z_1 \\
Z_1 - Z_0
\end{pmatrix},
\begin{pmatrix}
v_1 & v_1 \\
v_1 & v_2
\end{pmatrix}
\right\}
\]

with \( v_1 \) as above and \( v_2 = v_1 + \tau_0^{-1} - \tau_0^{-1} \). Since \( Q \) is the rectangular region \( \{ X_{11} > m, (X_{11} - X_{01}) > c \} \), (3) is a truncated bivariate normal expectation and can be evaluated using Rosenbaum (1961) to obtain the uniformly minimum variance conditionally unbiased estimator

\[
\tau_1 \tau_2^{-1} Z_1 - \tau_{11} \tau_2^{-1} E(X_{11} \mid Z, Q) - \tau_0 \tau_{02}^{-1} Z_0 - \tau_0 \tau_{02}^{-1} E(X_{01} \mid Z, Q),
\]

where \( E(X_{11} \mid Z, Q) = v_1^{1/2} A + Z_1 \) and \( E(X_{01} \mid Z, Q) = v_1^{1/2} A - v_2^{1/2} B + Z_0 \) with

\[
A = \frac{\phi(a)(1 - \Phi(\tilde{b})) + \rho \phi(b)(1 - \Phi(\tilde{a}))}{U(a; b; \rho)}, \quad B = \frac{\rho \phi(a)(1 - \Phi(\tilde{b})) + \phi(b)(1 - \Phi(\tilde{a}))}{U(a; b; \rho)};
\]

where \( a = (m - Z_1) v_1^{-1/2}, b = (c - Z_1 + Z_0) v_2^{-1/2}, \tilde{a} = (a - \rho b)(1 - \rho^2)^{-1/2}, \tilde{b} = (b - \rho a)(1 - \rho^2)^{-1/2}, \rho^2 = v_1/v_2 \) and \( U(a; b; \rho) = \Phi(u_1 > a, u_2 > b) \) for \((u_1, u_2)^T\) standard bivariate normal with correlation \( \rho \). This is not the estimator proposed by Kimani et al. (2013), which is conditionally unbiased but not minimum variance.

Robertson et al. (2016) also consider two-stage trials with early stopping for futility, though they do not assume equal variances and consider ranking by standardized observed stage 1 treatment effect estimates, \( R_i = (X_{1i} - X_{0i})/(\tau_i^{-1} + \tau_0^{-1})^{-1/2} \), assuming treatment 1 is selected if \( R_i > \max(R_2, \ldots, R_k, c) \) for specified \( c \). In constructing their estimators, they condition on statistics based on the observed treatment effects, \( X_{ij} - X_{0j} \) (\( i = 1, \ldots, k; j = 1, 2 \)). Since these may not be sufficient, for example in the case of normal data with a common control, their estimator is also conditionally unbiased but not minimum variance; indeed they show that in the common variance case it has larger variance than the estimator proposed by Kimani et al. (2013).

3. Numerical Example

The ADVENT trial (MacArthur et al., 2013) was a two-stage study comparing 125 mg, 250 mg and 500 mg doses of crofelemer with a placebo in noninfectious chronic diarrhea in HIV-seropositive patients. The primary endpoint was clinical response, defined as at most two watery stools per week during at least two weeks of treatment. At the end of stage 1, based on data from 200 patients randomized equally between the four treatment arms, a single dose of crofelemer would continue with the control to stage 2, with a further 150 patients randomized equally between these two groups. In the absence of safety concerns, the dose selected would be the lowest dose with an observed clinical response rate within two percentage points of the best-performing dose. Although not explicitly stated by MacArthur et al. (2013), we assume that the trial would have stopped at the first stage if the best-performing dose did not have observed clinical response rate at least two percentage points above the placebo. The trial was analysed using the method of Posch et al. (2005) to control the familywise Type I error rate, but apparently no attempt was made to obtain unbiased estimators of the treatment effect for the selected dose. The U.S. Food and Drug Administration (2012) report gives results of the two stages of the study. In stage 1, 50, 44, 54 and 46 patients received the placebo and the three doses respectively, with 1, 9, 5 and 9 patients in these groups showing clinical response. In stage 2, 88 and 92 further patients received the placebo and 125 mg crofelemer respectively, with 10 and 15 demonstrating clinical response. A naive estimate of the effect of the 125 mg dose relative to placebo is thus 24/136 − 11/138 = 0.097.
Table 1. Simulated probability of selection and bias (with root mean squared error in parentheses), with all values multiplied by 100, for the estimator given by (3) and the naive estimator for $\theta_i$ conditional on selection of treatment $i$ ($i = 1, \ldots, 3$) for a range of $\mu$ values ($\mu_0 = 0.02$ in all cases)

<table>
<thead>
<tr>
<th>$(\mu_1, \mu_2, \mu_3)$</th>
<th>pr(selected)</th>
<th>Estimator from (3)</th>
<th>Naive estimator</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3</td>
<td>$\theta_1$</td>
<td>$\theta_2$</td>
</tr>
<tr>
<td>(0.02, 0.02, 0.02)</td>
<td>8 5 4</td>
<td>$-0.2 (1.5)$</td>
<td>$-0.2 (1.5)$</td>
</tr>
<tr>
<td>(0.02, 0.02, 0.2)</td>
<td>&lt;1 &lt;1 99</td>
<td>$-0.4 (1.3)$</td>
<td>$-0.4 (1.4)$</td>
</tr>
<tr>
<td>(0.02, 0.1, 0.2)</td>
<td>&lt;1 14 86</td>
<td>$-0.2 (1.5)$</td>
<td>$-0.7 (3.3)$</td>
</tr>
<tr>
<td>(0.02, 0.2, 0.2)</td>
<td>&lt;1 61 39</td>
<td>$-0.2 (0.9)$</td>
<td>$-0.7 (4.1)$</td>
</tr>
<tr>
<td>(0.1, 0.2, 0.2)</td>
<td>5 8 37</td>
<td>$-0.6 (3.3)$</td>
<td>$-0.7 (4.1)$</td>
</tr>
<tr>
<td>(0.2, 0.2, 0.2)</td>
<td>46 31 23</td>
<td>$-0.7 (4.2)$</td>
<td>$-0.5 (4.2)$</td>
</tr>
</tbody>
</table>

For illustrative purposes, we treat estimated event rates as asymptotically normally distributed with variances based on the observed responses in each group and stage and set $(X_{01}, \ldots, X_{11}) = (0.02, 0.2045, 0.0926, 0.1956)$, $(\tau_{01}, \ldots, \tau_{11}) = (2551, 270.4, 642.1, 292.3)$, $(X_{02}, X_{12}) = (0.1136, 0.1630)$ and $(\tau_{02}, \tau_{12}) = (873.7, 674.2)$. Based on these results, we obtained an estimate from (3), calculating the integrals using 100,000 simulations conditional on 125 mg being the lowest dose with observed clinical response rate within two percentage points of the best-performing dose, giving an estimate of 0.114.

Properties of the estimator (3) were assessed in a realistic setting via simulations based on the ADVENT trial. For a range of $\mu$ values, for each of 100,000 simulated trials, we simulated $B_1 \sim B(50, \mu)$ ($i = 0, \ldots, 3$), adding or subtracting 1 if $B_{31} = 0$ or 50 to enable variance estimation. We then set $X_1 = B_{11}/50$ and $\tau^{-1}_1 = X_1(1 - X_1)/50$ ($i = 0, \ldots, 3$). For $I = \min\{i \in \{0, \ldots, 3\} : X_1 \geq \max_{r=0,\ldots,3}|X_{r1} - 0.02|\}$, the trial was assumed to stop at the end of stage 1 if $I = 0$; otherwise, we simulated $B_{32} \sim B(75, \mu)$ ($i = 0, I$), again adding or subtracting 1 if $B_{32} = 0$ or 75, and set $X_2 = B_{32}/75$ and $\tau^{-1}_2 = X_2(1 - X_2)/75$. The estimate from (3) for $\theta_1$, conditional on treatment $I$ being selected, was obtained using 10,000 Monte Carlo simulations to evaluate the integrals, and was compared with the naive estimator $(B_{11} + B_{12})/125 - (B_{01} + B_{02})/125$. Table 1 gives simulated selection probabilities and the bias and root mean squared error for the naive estimator and estimator (3). Bias estimates have standard error of at most 0.0002, with the exception of those for $\theta_1$ when $\mu = (0.02, 0.02, 0.2, 0.2)^T$ and for $\theta_3$ when $\mu = (0.02, 0.2, 0.2, 0.2)^T$, with standard errors of 0.04 and 0.03 respectively, and estimators for treatments selected with a probability of less than 0.01. Recall that if observed differences between treatments are small, the selection rule favours treatments with the lowest indices, explaining the differences in selection probabilities for treatments with the same mean.

The naive estimator is conditionally biased, overestimating the true effect. The bias is relatively small, but in some cases it is, like the difference between the naive and new estimates for the ADVENT data reported above, close to the difference in clinical response rates of two percentage points considered important in the trial design. Only settings with $\mu = (0.02, 0.02, 0.02, 0.2)^T$ and $(0.02, 0.02, 0.1, 0.2)^T$, where the highest dose is nearly always selected, have bias near zero. The estimator (3) has bias near zero, though given that the derivation is based on assumed normality with known variance, it may not be unbiased; the simulated bias is negative in all cases, suggesting that the true treatment effect is slightly underestimated. The root mean squared errors of the two estimators are similar, though with some suggestion that the naive estimator has slightly smaller mean squared error in situations in which it has larger bias, consistent with theoretical results in settings in which the single most promising treatment is selected (Stallard et al., 2008; Bauer et al., 2010). Given the simulation results, it is interesting that the estimate from (3) using the observed ADVENT trial data is larger than the naive estimate. This appears to be due to the large difference between $X_{01}$ and $X_{02}$, the placebo arm means in the two stages. A full analysis might involve investigation of possible causes for this difference.

The estimators proposed by Kimani et al. (2013) and Robertson et al. (2016) were not included in the simulations as they are not applicable for the selection rule used. We conducted 10^6 additional simulations.
with \( r = k = 2 \) and \( X_i \sim N(0,1) \) (\( i = 0, 1, 2; j = 1, 2 \)), selecting the treatment with the largest \( X_i \) to continue to stage 2 provided \( X_i - X_0 > c \), when the Kimani et al. (2013) estimator is conditionally unbiased and has smaller mean squared error than that of Robertson et al. (2016). If \( c \to -\infty \), so there is no stopping for futility, the Kimani et al. (2013) estimator and (3) correspond, with root mean squared error of 1.08. For \( c = 0 \), the two estimators have root mean squared error of 1.23 and 1.21 respectively, with these increasing to 1.34 and 1.31 respectively for \( c = 2 \). The probabilities of stopping for futility in these two cases are 0.64 and 0.87.

4. DISCUSSION

The method proposed gives uniformly minimum variance conditionally unbiased estimators in multi-arm multi-stage clinical trials. Estimators are obtained for treatment effects relative to the control for treatments continuing to the end of the trial conditional on the event, \( Q \), that these treatments are selected to do so. The conditioning event is weaker than that considered by Bowden & Glimm (2008) and Cohen & Sackrowitz (1989), who condition on the ordering of stage 1 treatment means. We consider conditioning on \( Q \) to be appropriate since decisions regarding effectiveness of treatments are likely to depend on treatments continuing to the end of the trial but not on the ordering of continuing treatments at earlier analyses. The conditionally unbiased estimators are identical for a two-stage design with only the best-performing treatment continuing along with the control to stage 2, as shown in §2.2 above.

We have considered the total number of treatments and selection rule to be specified in advance. Others have proposed more flexible approaches, such as in the STAMPEDE trial (Sydes et al., 2012), where treatment arms were added during the trial. Our method might not yield conditionally unbiased estimators in such a trial. In a fully flexible approach, it is not clear that it is even possible to define the bias, as this would be an expectation over an unspecified sample space.

ACKNOWLEDGEMENT

We thank the referees for their comments and the Medical Research Council for funding.

APPENDIX

Construction of sufficient statistics for \( \theta_1, \ldots, \theta_{k+1} \)

Let \( \theta_{k+1} \) be orthogonal to \( \theta_i \), that is, the information matrix term \( i_0^T \theta_{k+1} = 0 \) (Cox & Reid, 1987), and write \( \mu_0 = \sum_{i=0}^{k+1} d_i \theta_i \) for some \( d_i \) \((i = 1, \ldots, k + 1)\). Since we may scale \( \theta_{k+1} \) arbitrarily and retain orthogonality, set \( d_{k+1} = 1 \) to get \( \mu_0 = \sum_{i=1}^{k+1} d_i \theta_i + \theta_{k+1} \) and

\[
\mu_i = \sum_{j=1}^{k+1} d_j \theta_j + \theta_{k+1} + \theta_i \quad (i = 1, \ldots, k). \tag{A1}
\]

Let \( X = (X_0^T, \ldots, X_k^T)^T; \) then \( E(X) = \Lambda \theta \) where \( \Lambda \) is a \((\sum_{i=0}^{k} r_i) \times (k + 1)\) matrix with blocks of \( r_i \) rows equal to \((d_1, \ldots, d_{k}, 1) + 1, (i = 0, \ldots, k)\); here \( 1_i \) denotes the \((k + 1)\) row vector with element \( i \) equal to \( 1 \) \((i = 1, \ldots, k)\) and other elements equal to 0, and \( 1_0 = (0, \ldots, 0)^T \). We have \( \text{var}(X) = T \), diagonal with elements \((\tau_0^1, \ldots, \tau_{r_0}^1, \tau_1^2, \ldots, \tau_{r_1}^2, \ldots, \tau_{r_k}^k)\). The likelihood is proportional to \( \exp \{ (X - \Lambda \theta)^T T^{-1} (X - \Lambda \theta) \} = g(X)h(\theta) \exp \{ -2X^T T^{-1} \Lambda \theta \} \), for some \( g(X) \) and \( h(\theta) \). Thus \( X^T T^{-1} \Lambda \) is complete and sufficient for \( \theta \) (Lehmann & Romano, 2005, §4.3). Denoting \( X^T T^{-1} \Lambda \) by \((W_1, \ldots, W_{k+1})^T \), we have \( W_{k+1} = \sum_{i=0}^k X_i^T \tau_i = \sum_{i=0}^k Z_i \tau_i \) and

\[
W_i = d_i \sum_{i'=0}^k X_i^T \tau_{i'} + X_i^T = d_i \sum_{i'=0}^k Z_i \tau_{i'} + Z_i \tau_i \quad (i = 1, \ldots, k). \tag{A2}
\]
For orthogonality $$(\partial \Lambda \theta / \partial \theta_i) T^{-1} (\partial \Lambda \theta / \partial \theta_{i+1}) = d_i \sum_{i'=0}^k \tau_{i'} + \tau_i = 0$$, so $$d_i = -\tau_i / \sum_{i'=0}^k \tau_{i'}$$ and, from (A1) and (A2), $$\theta_{i+1} = \sum_{i'=0}^k \tau_{i'} / \sum_{i'=0}^k \tau_{i'} W_{k+1} = \sum_{i'=0}^k Z_{i'} \tau_{i'}$$ and $$W_i = \tau_i Z_i - \tau_i \sum_{i'=0}^k Z_{i'} \tau_{i'} / \sum_{i'=0}^k \tau_{i'} (i = 1, \ldots, k)$$, so $$Z$$ is complete and sufficient.

With no selection, i.e., $$s_r = k$$, $$Z_i$$ and $$Z_i - Z_0$$ are the usual uniformly minimum variance unbiased estimators for $$\mu_i$$ and $$\theta_i$$ ($$i = 1, \ldots, k$$). By the factorization theorem, sufficient statistics are the same for the conditional and unconditional likelihoods, so $$Z$$ is also sufficient in this case.

**REFERENCES**


[Received on 13 September 2017. Editorial decision on 7 January 2018]