BIOMETRICS 000, 1–14 DOI: 000

000 000

Supporting Information for Bayesian group sequential enrichment designs based on adaptive regression of response and survival time on baseline biomarkers by Yeonhee Park, Suyu Liu, Peter Thall, and Ying Yuan

Yeonhee Park[1], Suyu Liu[2], Peter Thall[2], and Ying Yuan[2,*]

[1]Department of Biostatistics and Medical Informatics, University of Wisconsin, Madison, WI, U.S.A.

[2]Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX, U.S.A.

**email:* yyuan@mdanderson.org

This paper has been submitted for consideration for publication in *Biometrics*

Web Appendix A: Accounting for Time to Evaluate Response

The following model elaboration accounts for settings where Y may be observed before Z can be evaluated. For example, if Z is scored at week 12, then it is possible that a patient may die before Z is evaluated. This is especially important for rapidly fatal diseases. Let T_0 be a latent survival time following the distribution F_0 with probability density function f_0, and let t_0 denote the time required to evaluate Z. To account for this possible complication the probability density function of Y may be formulated more generally as

$$\begin{aligned} f(y|G,\boldsymbol{x}) &= f(y|G,\boldsymbol{x},Y<T_0)\Pr(Y<T_0|G,\boldsymbol{x}) \\ &+ \sum_{z=0}^{1} f(y|Z=z,G,\boldsymbol{x},Y\geqslant T_0)\Pr(Z=z|G,\boldsymbol{x},Y\geqslant T_0)\Pr(Y\geqslant T_0|G,\boldsymbol{x}). \end{aligned}$$

Since Z is defined only if $Y \geqslant T_0$, we have $\Pr(Z=z|G,\boldsymbol{x},Y\geqslant T_0) = \Pr(Z=z|G,\boldsymbol{x})$. Let W be the indicator variable that Z is observed, and denote $\mathbf{W}_n = (W_1,\cdots,W_n)$ for n observations. This gives the generalized likelihood

$$\begin{aligned} &\mathcal{L}_n(\mathcal{O}_n,\mathbf{W}_n,\mathbf{G}_n,\mathbb{X}_n|\boldsymbol{\theta}_Z,\boldsymbol{\theta}_Y) \\ &= \prod_{i=1}^{n}\big[\big\{f_0(Y_i^o|G_i,\boldsymbol{x}_i,\boldsymbol{\theta}_Y)\big\}^{\delta_i}\big\{1-F_0(Y_i^o|G_i,\boldsymbol{x}_i,\boldsymbol{\theta}_Y)\big\}^{1-\delta_i}\big]^{1-W_i} \\ &\quad\times\big[\Phi\big(\tilde{\boldsymbol{x}}_i^\top\boldsymbol{\beta}_Z+G_i\tilde{\boldsymbol{x}}_i^\top\boldsymbol{\gamma}_Z\big)^{Z_i}\big\{1-\Phi\big(\tilde{\boldsymbol{x}}_i^\top\boldsymbol{\beta}_Z+G_i\tilde{\boldsymbol{x}}_i^\top\boldsymbol{\gamma}_Z\big)\big\}^{1-Z_i}\big\{f(Y_i^o-T_0|Z_i,G_i,\boldsymbol{x}_i,\boldsymbol{\theta}_Y)\big\}^{\delta_i} \\ &\quad\times\big\{1-F(Y_i^o-T_0|Z_i,G_i,\boldsymbol{x}_i,\boldsymbol{\theta}_Y)\big\}^{1-\delta_i}\big\{1-F_0(T_0|G_i,\boldsymbol{x}_i)\big\}\big]^{W_i}. \end{aligned}$$

Web Appendix B: Joint distribution of $\lambda_{Z,j}$ and $\lambda_{Y,j}$

The following describes the bivariate distribution of $\lambda_{Z,j}$ and $\lambda_{Y,j}$ for $j = 1,\ldots,2p+1$. For notational brevity, we suppress the index j. Let $p_Z = \Pr(\lambda_Z = 1)$, $p_Y = \Pr(\lambda_Y = 1)$ and let ρ denote the odds ratio. Let $\mathcal{B}(p_Z,p_Y,\rho)$ denote the joint Bernoulli distribution of binary variables

λ_Z and λ_Y, whose probability mass function is

$$p_{11}^{ab}\, p_{10}^{a(1-b)}\, p_{01}^{(1-a)b}\, p_{00}^{(1-a)(1-b)},$$

where $p_{ab} \equiv \Pr(\lambda_Z = a, \lambda_Y = b)$, $a, b = 0, 1$ are represented by three parameters, the marginal probabilities and the odds ratio, of the distribution: p_{10} is the solution of the quadratic equation

$$(1-\rho)p_{10}^2 - p_{10}\big\{1 - p_Y + p_Z + \rho(p_Y - p_Z)\big\} - (1-p_Y)p_Z = 0,$$

with $p_{00} = 1 - p_Y - p_{10}$, $p_{01} = p_Y - p_Z + p_{10}$ and $p_{11} = p_Z - p_{10}$, provided that $\max\{0, p_Z - p_Y\} \leqslant p_{10} \leqslant \min\{1 - p_Y, p_Z\}$.

Web Appendix C: Details regarding generation of the Markov Chain for variable selection

To identify the subset of promising predictor variables, the Stochastic Search Variable Selection (SSVS) uses the sampler from the full conditional distribution of the parameters. We generated a Markov Chain of length L on parameter $\boldsymbol{\lambda}_Z, \boldsymbol{\lambda}_Y, \boldsymbol{p}_Z, \boldsymbol{p}_Y, \boldsymbol{\rho}, \boldsymbol{\psi}_Z, \psi_{Z,0}, \boldsymbol{\psi}_Y, \alpha_Y, \boldsymbol{\phi}$ for fixed hyper-parameters. We arbitrarily choose initial values of parameters with $l = 1$. For each $l = 2, \ldots, L$, we iterated through Steps 1 - 5, described below.

Step 1 Generate indicator variables $\boldsymbol{\lambda}_Z^{(l)}$ and $\boldsymbol{\lambda}_Y^{(l)}$ by using the Metropolis-Hastings sampler on $\boldsymbol{\lambda}_Z^{(l)} = \big(\lambda_{Z,1}^{(l)}, \ldots, \lambda_{Z,2p+1}^{(l)}\big)$ and $\boldsymbol{\lambda}_Y^{(l)} = \big(\lambda_{Y,1}^{(l)}, \ldots, \lambda_{Y,2p+1}^{(l)}\big)$. For $j = 1, \ldots, p+1$, let $p_{Z,j}^{(l)} = \Pr(\lambda_{Z,j}^{(l)} = 1)$ and $p_{Y,j}^{(l)} = \Pr(\lambda_{Y,j}^{(l)} = 1)$. Then, each binary vector is obtained componentwise according to the Bernoulli distribution with probabilities

$$\Pr\big(\lambda_{Z,j}^{(l)} = 1 | \text{else}\big) = \frac{p_{Z,j}^{(l-1)} f_N\big(\psi_{Z,j}^{(l-1)} | 0, u_{Z,j}^2 \tau_{Z,j}^2\big)}{p_{Z,j}^{(l-1)} f_N\big(\psi_{Z,j}^{(l-1)} | 0, u_{Z,j}^2 \tau_{Z,j}^2\big) + \big(1 - p_{Z,j}^{(l-1)}\big) f_N\big(\psi_{Z,j}^{(l-1)} | 0, \tau_{Z,j}^2\big)}$$

and

$$\Pr\big(\lambda_{Y,j}^{(l)} = 1 | \text{else}\big) = \frac{p_{Y,j}^{(l-1)} f_N\big(\psi_{Y,j}^{(l-1)} | 0, u_{Y,j}^2 \tau_{Y,j}^2\big)}{p_{Y,j}^{(l-1)} f_N\big(\psi_{Y,j}^{(l-1)} | 0, u_{Y,j}^2 \tau_{Y,j}^2\big) + \big(1 - p_{Y,j}^{(l-1)}\big) f_N\big(\psi_{Y,j}^{(l-1)} | 0, \tau_{Y,j}^2\big)},$$

where $f_N(\cdot|a,b)$ denotes the density of normal distribution with mean a and variance b. For $j = p+2,\ldots,2p+1$, $\lambda_{Z,j}^{(l)}$ and $\lambda_{Y,j}^{(l)}$ are generated from the Bernoulli distribution with probabilities

$$\begin{aligned}&\Pr\big(\lambda_{Z,j}^{(l)}=1|\text{else}\big)\\ &= \Pr\big(\lambda_{Z,j-p-1}^{(l)}=1|\text{else}\big)\Pr\big(\lambda_{Z,p+1}^{(l)}=1|\text{else}\big)\min\big\{\Pr\big(\lambda_{Z,j-p-1}^{(l)}=1|\text{else}\big),\Pr\big(\lambda_{Z,p+1}^{(l)}=1|\text{else}\big)\big\}\end{aligned}$$

and

$$\begin{aligned}&\Pr(\lambda_{Y,j}^{(l)}=1|\text{else})\\ &= \Pr(\lambda_{Y,j-p-1}^{(l)}=1|\text{else})\Pr(\lambda_{Y,p+1}^{(l)}=1|\text{else})\min\{\Pr(\lambda_{Y,j-p-1}^{(l)}=1|\text{else}),\Pr(\lambda_{Y,p+1}^{(l)}=1|\text{else})\}.\end{aligned}$$

Step 2 Generate $\boldsymbol{p}_Z^{(l)} = (p_{Z,1}^{(l)},\ldots,p_{Z,p+1}^{(l)})$, $\boldsymbol{p}_Y^{(l)} = (p_{Y,1}^{(l)},\ldots,p_{Y,p+1}^{(l)})$ and $\boldsymbol{\rho}^{(l)} = (\rho_1^{(l)},\ldots,\rho_{p+1}^{(l)})$ from

$$P(p_{Z,j}|\text{else}) \propto p_{00,j}^{(1-\lambda_{Z,j})(1-\lambda_{Y,j})}p_{01,j}^{(1-\lambda_{Z,j})\lambda_{Y,j}}p_{10,j}^{\lambda_{Z,j}(1-\lambda_{Y,j})}p_{11,j}^{\lambda_{Z,j}\lambda_{Y,j}}p_{Z,j}^{l_{Z1,j}-1}(1-p_{Z,j})^{l_{Z2,j}-1},$$

where $p_{00,j}, p_{01,j}, p_{10,j}$ and $p_{11,j}$ are obtained from $p_{Z,j}^{(l-1)}$, $p_{Y,j}^{(l-1)}$ and $\rho_j^{(l-1)}$,

$$P(p_{Y,j}|\text{else}) \propto p_{00,j}^{(1-\lambda_{Z,j})(1-\lambda_{Y,j})}p_{01,j}^{(1-\lambda_{Z,j})\lambda_{Y,j}}p_{10,j}^{\lambda_{Z,j}(1-\lambda_{Y,j})}p_{11,j}^{\lambda_{Z,j}\lambda_{Y,j}}p_{Y,j}^{l_{Y1,j}-1}(1-p_{Y,j})^{l_{Y2,j}-1},$$

where $p_{00,j}, p_{01,j}, p_{10,j}$ and $p_{11,j}$ are obtained from $p_{Z,j}^{(l)}$, $p_{Y,j}^{(l-1)}$ and $\rho_j^{(l-1)}$,

$$\begin{aligned}&P(\rho_j|\text{else})\\ &\propto p_{00,j}^{(1-\lambda_{Z,j})(1-\lambda_{Y,j})}p_{01,j}^{(1-\lambda_{Z,j})\lambda_{Y,j}}p_{10,j}^{\lambda_{Z,j}(1-\lambda_{Y,j})}p_{11,j}^{\lambda_{Z,j}\lambda_{Y,j}}\exp\big\{-(\log\rho_j - r_{1j})^2/(2r_{2j}^2)\big\}\rho_j^{-1},\end{aligned}$$

where $p_{00,j}, p_{01,j}, p_{10,j}$ and $p_{11,j}$ are obtained from $p_{Z,j}^{(l)}$, $p_{Y,j}^{(l)}$ and $\rho_j^{(l-1)}$, $j = 1,\ldots,p+1$.

Step 3 Generate $\boldsymbol{\psi}_Z^{(l)}$ from

$$
\begin{aligned}
& f(\boldsymbol{\psi}_Z|\text{else}) \\
& \propto \mathcal{L}_n(\mathcal{O}_n, \mathbf{G}_n, \mathbb{X}_n|\boldsymbol{\theta}_Z, \boldsymbol{\theta}_Y) f(\boldsymbol{\psi}_Z|\boldsymbol{\lambda}_Z^{(l)} = \mathbf{1}) \\
& \propto \prod_{i=1}^{n} \Phi\big(\psi_{Z,0} + \boldsymbol{x}_{Z,G,i}^{\top}\boldsymbol{\psi}_{Z,\lambda}\big)^{Z_i}\big\{1 - \Phi\big(\psi_{Z,0} + \boldsymbol{x}_{Z,G,i}^{\top}\boldsymbol{\psi}_{Z,\lambda}\big)\big\}^{1-Z_i} \exp(-\boldsymbol{\psi}_{Z,\lambda}^{\top}\boldsymbol{\Sigma}_{Z,\lambda}^{-1}\boldsymbol{\psi}_{Z,\lambda}/2),
\end{aligned}
$$

where $\boldsymbol{x}_{Z,G,i}$ denotes the predictor vector $(\boldsymbol{x}_i^{\top}, G\tilde{\boldsymbol{x}}_i^{\top})^{\top}$ corresponding to the selected subset of $\{1, \ldots, 2p+1\}$ based on short-term endpoint data for the i^{th} observation; $\boldsymbol{\psi}_{Z,\lambda}$ denotes the vector of coefficients for the subset and $\boldsymbol{\Sigma}_{Z,\lambda}$ denotes the restricted variance; and covariance matrix $\boldsymbol{\Sigma}_Z$, which is a diagonal matrix with $u_{Z,1}^2\tau_{Z,1}^2, \ldots, u_{Z,2p+1}^2\tau_{Z,2p+1}^2$, corresponding to the selected subset. Also, $\psi_{Z,0}^{(l)}$ is generated from

$$
\begin{aligned}
& f(\psi_{Z,0}|\text{else}) \\
& \propto \mathcal{L}_n(\mathcal{O}_n, \mathbf{G}_n, \mathbb{X}_n|\boldsymbol{\theta}_Z, \boldsymbol{\theta}_Y) f(\psi_{Z,0}) \\
& \propto \prod_{i=1}^{n} \Phi\big(\psi_{Z,0} + \boldsymbol{x}_{Z,G,i}^{\top}\boldsymbol{\psi}_{Z,\lambda}\big)^{Z_i}\big\{1 - \Phi\big(\psi_{Z,0} + \boldsymbol{x}_{Z,G,i}^{\top}\boldsymbol{\psi}_{Z,\lambda}\big)\big\}^{1-Z_i} \exp\{-(\psi_{Z,0} - u_0)^2/(2\tau_0^2)\}.
\end{aligned}
$$

Step 4 Generate $\boldsymbol{\psi}_Y^{(l)}$ from

$$
\begin{aligned}
& f(\boldsymbol{\psi}_Y|\text{else}) \\
& \propto \mathcal{L}_n(\mathcal{O}_n, \mathbf{G}_n, \mathbb{X}_n|\boldsymbol{\theta}_Z, \boldsymbol{\theta}_Y) f(\boldsymbol{\psi}_Y|\boldsymbol{\lambda}_Y^{(l)} = \mathbf{1}) \\
& \propto \prod_{i=1}^{n} f(Y_i|Z_i, G_i, \boldsymbol{x}_{Y,i}, \boldsymbol{\theta}_Y)^{\delta_i}\big\{1 - F(Y_i|Z_i, G_i, \boldsymbol{x}_{Y,i}, \boldsymbol{\theta}_Y)\big\}^{1-\delta_i} \exp\big(-\boldsymbol{\psi}_{Y,\lambda}^{\top}\boldsymbol{\Sigma}_{Y,\lambda}^{-1}\boldsymbol{\psi}_{Y,\lambda}/2\big),
\end{aligned}
$$

where $\boldsymbol{x}_{Y,i}$ denotes the selected predictors for the i^{th} observation based on long-term endpoint data, $\boldsymbol{\psi}_{Y,\lambda}$ denotes the vector of coefficients for the selected subset of $\{1, \ldots, 2p+1\}$ and $\boldsymbol{\Sigma}_{Y,\lambda}$ denotes the restricted variance and covariance matrix $\boldsymbol{\Sigma}_Y$, which is a diagonal matrix with $u_{Y,1}^2\tau_{Y,1}^2, \ldots, u_{Y,2p+1}^2\tau_{Y,2p+1}^2$, corresponding to the selected subset. Generate $\alpha_Y^{(l)}$

from

$$f(\alpha_Y|\text{else}) \propto \prod_{i=1}^{n} f(Y_i|Z_i, G_i, \boldsymbol{x}_{Y,i}, \boldsymbol{\theta}_Y)^{\delta_i} \big\{1 - F(Y_i|Z_i, G_i, \boldsymbol{x}_{Y,i}, \boldsymbol{\theta}_Y)\big\}^{1-\delta_i} \exp\big\{-(\alpha_Y - u_a)^2/(2\tau_a^2)\big\}.$$

Step 5 Generate $\boldsymbol{\phi}^{(l)} = (\phi_1^{(l)}, \ldots, \phi_M^{(l)})$ from

$$f(\phi_m|\text{else}) \propto \prod_{i=1}^{n} f(Y_i|Z_i, G_i, \boldsymbol{x}_{Y,i}, \boldsymbol{\theta}_Y)^{\delta_i} \big\{1 - F(Y_i|Z_i, G_i, \boldsymbol{x}_{Y,i}, \boldsymbol{\theta}_Y)\big\}^{1-\delta_i} \phi_m^{\tilde{c}\tilde{\phi}_m^* - 1} \exp(-\tilde{c}\phi_m),$$

for $m = 1, \ldots, M$.

For implementation, the hyperparameters are specified as follows. We choose large $u_{Z,j}$ and small $\tau_{Z,j}$ in the spike-and-slab prior so that $\lambda_{Z,j} = 1$ implies that a nonzero estimate of $\psi_{Z,j}$ is included, whereas $\lambda_{Z,j} = 0$ implies that the covariate corresponding to $\psi_{Z,j}$ has a negligible effect. In our simulations, $u_{Z,j} = 100$ and $\tau_{Z,j} = 0.1$ were chosen for $j = 1, \ldots, 2p + 1$. Similar choices were applied to the regression coefficient for survival in order to obtain sparse vectors of coefficient estimates for the long-term endpoint Y. In our simulations, $u_{Y,j} = 100$ and $\tau_{Y,j} = 0.1$ were chosen for $j = 1, \ldots, 2p + 1$. A sensitivity analysis for $u_{\cdot,j}$ and $\tau_{\cdot,j}$ is presented in Web Appendix F. The results of the sensitivity analysis show that the operating characteristics of the design are similar to the case $u_{Z,j} = u_{Y,j} = 100$ and $\tau_{Z,j} = \tau_{Y,j} = 0.1$. In practice, it is useful to perform preliminary simulations to calibrate these values to obtain the appropriate sparsity in terms of the interpretation and performance of the proposed method or design. Previous research on the target disease might provide a sense of sparsity among a large number of possible clinical covariates. Assuming that a few markers are predictive to characterize the treatment-sensitive patients, a parsimonious model is appropriate. In general, when $\tau_{\cdot,j}$ is small, the prior concentrates its mass on parsimonious models,

whereas when $\tau_{\cdot,j}$ is large, the prior encourages models with many variables and thus is more likely to include noninformative covariates. Also, $u_{\cdot,j}$ needs to be large enough to include informative covariates. In our experience, except for the computational burden and interpretation of the models with many covariates, the performance of the proposed design is good if informative covariates are included for enrichment of the trial population and monitoring of treatment effects.

For the remaining parameters used for variable selection, we considered noninformative normal priors on $\psi_{Z,0}$, α_Y and $\log \rho_j$, with the hyperparameters $u_0 = 0, \tau_0 = 10, u_a = 0, \tau_a = 10$, $r_{1j} = 0$ and $r_{2j} = 100$. We considered $\phi_m \sim$ Gamma$(\tilde{c}\tilde{\phi}_m, \tilde{c})$, $m = 1, \ldots, M$, which has mean $\tilde{\phi}_m$ and variance $\tilde{\phi}_m/\tilde{c}$. The parameter $\tilde{c}$ controls the amount of smoothness. The small value of $\tilde{c}$ provides less information in the smoothing of ϕ_m. In the simulation study, we used $\tilde{\phi}_m = 1$ and $\tilde{c} = 0.01$ so that mean of ϕ_m is 1 and variance of ϕ_m is 100. For beta priors on $p_{Z,j}$ and $p_{Y,j}$, $l_{Z1,j} = l_{Z2,j} = l_{Y1,j} = l_{Y2,j} = 1$, $j = 1, \ldots, p+1$, are prespecified to make them noninformative.

Web Appendix D: Calibration of design parameters in the enrichment criteria

The enrichment criteria involve the two designs parameters v and g. At the beginning of the trial at the first interim decision, there are few observed events for the long-term endpoint Y. As a result, ω_1 is small and thus the enrichment rule can be approximated as

$$\Omega(\boldsymbol{x} \,|\, \mathcal{D}_1) \approx \Pr\big\{\Delta_Z(\boldsymbol{x}_Z^{(k)}, \boldsymbol{\theta}_Z) > \epsilon_1 \,|\, \mathcal{D}_1\big\} > v(n_1/N)^g.$$

At the end of the trial, many events are observed, thus ω_K is close to 1 and the enrichment rule can be approximated as

$$\Omega(\boldsymbol{x} \,|\, \mathcal{D}_K) \approx \Pr\big\{\Delta_Y(\boldsymbol{x}_Y^{(k)}, \boldsymbol{\theta}_Y) < \epsilon_2 \,|\, \mathcal{D}_K\big\} > v.$$

These approximations may not be very accurate, but they are sufficient here because the purpose of the approximations is to facilitate determination of the tuning parameters v and g. The above equations motivate the following procedure for calibrating the values of v and g :

Step 1 Elicit from the clinicians two probability cutoffs p_1 and p_2, where p_1 represents the event that there is more than a $p_1 \times 100\%$ chance that a patient is expected to benefit from E in terms of the short-term endpoint, and thus it is desirable to enroll that patient into the trial. p_2 represents the event that there is more than a $p_2 \times 100\%$ chance that a patient is expected to benefit from E in terms of the long-term endpoint, and thus it is desirable to enroll that patient into the trial. Typically, we require that $p_1 \leqslant p_2$, so a relatively low cutoff is used at the beginning.

Step 2 Calculate v and g from

$$v = p_2 \qquad \text{and} \qquad g = \frac{\log(p_1/p_2)}{\log(n_1/N)}.$$

The explicit forms of v and g are obtained from the solution of two equations by setting the cutoff at the first interim value $v(n_1/N)^g$ to be p_1, and setting the cutoff at the end of the trial v to be p_2 from Step 1.

Step 3 Run preliminary simulations and present the operating characteristics of the design to the clinicians. If needed, adjust the values of p_1 and p_2, and repeat Step 1 until desirable operating characteristics are obtained.

As a example, consider a trial with two interim tests, i.e., $K = 3$, and cohort sizes $c_1 = 200$ and $c_2 = c_3 = 100$. Suppose $p_1 = 60\%$ and $p_2 = 76.6\%$ are elicited from clinicians. Then, $v = 0.766$ and $g = \log(0.6/0.766)/\log(200/400) = 0.352$.

Web Appendix E: Elicitation rule and calibration of cutoffs for the Bayesian sequential monitoring rule

The Bayesian sequential monitoring rule involves two design parameters, B_1 and B_2, where B_1 controls the type I error rate and B_2 controls the type II error rate. To calibrate their values, first specify decision rule cutoffs of b_1 and b_2 to denote the minimal improvement in survival and elicit targeted type I and II error rates, say α and β, from the clinicians. For complicated adaptive designs such as the proposed adaptive enrichment design, it is not possible to calculate the type I and II error rates analytically. Thus, simulation is used to determine the empirical type I and II errors. In our setting, we assign initial values for B_1 and B_2, and perform simulations to calculate the type I and II error rates for the given values of b_1 and b_2. If the empirical type I error rate is lower/higher than the specified level, we decrease/increase the value of B_1, and if the calculated type II error rate is lower/higher than the desirable level, we decrease/increase the values of B_2. We repeat this calibration process until the specified type I and II error rates are obtained. Based on our experience, reasonable initial values of (B_1, B_2) are $B_1 = 1 - \alpha$, $B_2 = 1 - \beta$, and it takes several rounds of calibration to achieve the target type I and II error rates.

Web Appendix F: True model parameters used in the simulation study

Each scenario in Simulation study is generated from the regression models (8) and (9) using Web Table 1 and $\alpha_Y = -0.5$.

[Web Table 1 about here.]

We assume that ten covariates are generated from Bernoulli distributions with response probabilities $0.5, 0.5, 0.5, 0.1, 0.2, 0.2, 0.4, 0.6, 0.8, 0.8$. Under the null distribution, we consider $\pi(\boldsymbol{x}, 1, \boldsymbol{\theta}_Z) = \pi(\boldsymbol{x}, 0, \boldsymbol{\theta}_Z) = 0.5$ (i.e., $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z)$ = 0) and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 1$ for all $\boldsymbol{x}$. However, under the alter-

native distribution, the sample consists of E-sensitive patients and E-insensitive patients, and the response rates and the survival times for E and C differ according to the biomarker profile pattern. For example, in scenario 2, we generated responses for E-sensitive patients with $x_1 = 1$ from $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.65 - 0.46 = 0.19$ and for the E-insensitive patients with $x_1 = 0$ from $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.46 - 0.5 = -0.04$. Survival time was generated from $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 0.6/1.22 = 0.49$ for E-sensitive patients and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 1.22/1 = 1.22$ for E-insensitive patients. In scenario 3, there are four different biomarker profiles according to the values of x_1 and x_2. When $x_1 = 1$ and $x_2 = 1$ representing the E-sensitive patients, we considered $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.65 - 0.42 = 0.23$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 0.49/0.82 = 0.6$; When either $x_1 = 1$ and $x_2 = 0$ or $x_1 = 0$ and $x_2 = 1$, $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.46 - 0.46 = 0$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 0.95/0.91 = 1.04$; When $x_1 = 0$ and $x_2 = 0$, $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.28 - 0.5 = -0.22$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 1.82/1 = 1.82$. In scenario 4, responses and survival times for E-sensitive patients with $x_1 = 1$ and $x_2 = 0$ were generated from $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.65 - 0.46 = 0.19$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 0.54/0.91 = 0.59$. For E-insensitive patients, when $x_1 = 1$ and $x_2 = 1$, $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.42 - 0.42 = 0$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 0.84/0.82 = 1.02$; When $x_1 = 0$ and $x_2 = 1$, $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.26 - 0.46 = -0.2$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 2.83/0.91 = 3.11$; When $x_1 = 0$ and $x_2 = 0$, $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.48 - 0.5 = -0.02$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 1.82/1 = 1.82$. In scenario 5, there are eight biomarker profiles determined by the values of x_1, x_2 and x_3. When $(x_1, x_2, x_3) = (1, 1, 1)$, which characterizes the marker profile of E-sensitive patients, we considered $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.65 - 0.44 = 0.21$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 0.45/0.74 = 0.61$; When $(x_1, x_2, x_3) = (1, 1, 0), (1, 0, 1), (0, 1, 1)$, $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.46 - 0.46 = 0$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 0.83/0.82 = 1.01$; When $(x_1, x_2, x_3) = (1, 0, 0), (0, 1, 0), (0, 0, 1)$, $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.27 - 0.48 = -0.21$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 1.54/0.91 = 1.69$; When $(x_1, x_2, x_3) = (0, 0, 0)$, $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.14 - 0.5 = -0.36$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 2.86/1 = 2.86$. In scenario 6, when $(x_1, x_2, x_3) = (1, 1, 0)$, which

indicates the marker profile of E-sensitive patients, we considered $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.65 - 0.46 = 0.19$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 0.49/0.82 = 0.60$, and for E-insensitive patients, when $(x_1, x_2, x_3) = (1, 0, 0)$ or $(0, 1, 0)$, we considered $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.42 - 0.48 = -0.06$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 1.18/0.91 = 1.30$; when $(x_1, x_2, x_3) = (0, 0, 1)$, we considered $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.09 - 0.48 = -0.39$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 4.39/0.91 = 4.82$; when $(x_1, x_2, x_3) = (1, 1, 1)$, we considered $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.43 - 0.44 = -0.01$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 0.75/0.74 = 1.01$; when $(x_1, x_2, x_3) = (1, 0, 1)$ or $(0, 1, 1)$, we considered $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.26 - 0.46 = -0.2$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 1.82/0.82 = 2.22$; when $(x_1, x_2, x_3) = (0, 0, 0)$, we considered $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.22 - 0.5 = -0.28$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 2.86/1 = 2.86$. In scenario 7, when $(x_1, x_2, x_3) = (1, 0, 0)$, which indicates the marker profile of E-sensitive patients, we considered $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.65 - 0.48 = 0.17$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 0.54/0.91 = 0.59$, and for E-insensitive patients, when $(x_1, x_2, x_3) = (1, 1, 1)$, we considered $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.27 - 0.48 = -0.21$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 0.83/0.82 = 1.01$; when $(x_1, x_2, x_3) = (0, 1, 0)$ or $(0, 0, 1)$, we considered $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.10 - 0.48 = -0.38$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 4.35/0.91 = 4.78$; when $(x_1, x_2, x_3) = (1, 1, 0)$, we considered $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.45 - 0.46 = -0.01$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 1.26/0.74 = 1.70$; when $(x_1, x_2, x_3) = (1, 0, 1)$, we considered $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.45 - 0.46 = -0.01$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 0.83/0.82 = 1.01$; when $(x_1, x_2, x_3) = (0, 1, 1)$, we considered $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.04 - 0.46 = -0.42$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 6.62/0.82 = 8.07$; when $(x_1, x_2, x_3) = (0, 0, 0)$, we considered $\Delta_Z(\boldsymbol{x}, \boldsymbol{\theta}_Z) = 0.22 - 0.46 = -0.24$ and $\Delta_Y(\boldsymbol{x}, \boldsymbol{\theta}_Y) = 2.86/1 = 2.86$.

Web Appendix G: Comparison results of simulation study

We added comparison results of simulation study (Section 3) with the four designs. Web Table 2 shows the results of probability that stop early due to superiority and futility obtained from 1000

simulated datasets when survival time follows a Weibull distribution with deceasing hazard. AED is more likely than CGS to correctly conclude that E is more effective than C in the identified sensitive subgroups and stop the trial early for superiority, and AED also is less likely to incorrectly stop the trial for futility when E actually is effective for the sensitive subgroup.

[Web Table 2 about here.]

Web Appendix H: Sensitivity to different hazard functions

As sensitivity analyses, Web Table 3 shows the results for AED when Y is generated from a Weibull distribution with scale parameter 1 and shape parameter 2 to obtain an increasing hazard, and Web Table 4 shows the results for AED when Y is generated from a log-logistic distribution with a $\cap$-shaped hazard. The results are generally similar to those reported when Y is generated from a Weibull distribution with a decreasing hazard.

[Web Table 3 about here.]

[Web Table 4 about here.]

Web Appendix I: Further investigation of the performance of AED

We evaluated the AED with larger maximum sample size, more covariates, consideration of additional main effect of covariate, covariates effect associated with Y but not with Z, no treatment-covariate interaction effects, design parameters for enrichment, and sparsity parameters. Assume that survival time follows a Weibull distribution with a decreasing hazard, which is the same simulation setting as Tables 2-3.

Simulations with a larger sample of size 800 We generated the data under the same settings in the simulation study (Section 3), but increased the maximum sample size to 800. Two interim analyses

were performed when the total accrual reached 400 and 600 patients. The results are provided in Web Table 5. The (generalized) power is larger compared to one with $n = 400$, while the other results are generally similar to each other.

[Web Table 5 about here.]

Simulations with a larger number of covariates We used the same settings as the simulation study (Section 3) except that we added 40 more covariates, generated from a Bernoulli distribution with response probability 0.5. The true model parameters for the first ten covariates are the same as in Web Table 1 and the model parameters of the last 40 covariates in the true model are set to 0 for all scenarios. The results were provided in Web Table 6, which is similar to those reported when only the first ten biomarkers only were used.

[Web Table 6 about here.]

Simulations for consideration of additional main effect of covariate We chose scenario 2 of the simulation study (Section 3), which has E-sensitive patients with $x_1 = 1$. In our simulation study, main and interaction effects of x_1 were considered. To consider additional main effect of covariate, we used the same simulation setting as the simulation study, but replaced $\beta_{Z,2} = \beta_{Y,2}$ = 0 with $\beta_{Z,2} = 0.5$ and $\beta_{Y,2} = 0.8$. Thus, main/interaction effects of x_1 and main effect of x_2 are included. The results were provided in Web Table 7, which is similar to those reported when no additional main effect of x_2 is considered.

[Web Table 7 about here.]

Simulations for consideration of effects associated with Y but not associated with Z We chose scenario 2 of the simulation study (Section 3), which has E-sensitive patients with $x_1 = 1$.

In our simulation study, the covariate x_1 was associated with both Y and Z. Specifically, we used $\beta_{Y,1} = 0.2$, $\gamma_{Y,1} = -0.911$, $\beta_{Z,1} = -0.1$, and $\gamma_{Z,1} = 0.585$. To investigate the performance of the proposed design when covariate effects are associated with Y but not associated with Z, we forced $\beta_{Z,1}$ and $\gamma_{Z,1}$ to be zero, but still had nonzero $\beta_{Y,1}$ and $\gamma_{Y,1}$ (i.e., $\beta_{Y,1} = 0.2$, $\gamma_{Y,1} = -0.911$). Our methods are likely to detect estimates of main and interaction effects $\beta_{Y,1}$, $\beta_{Z,1}$, $\gamma_{Y,1}$ and $\gamma_{Z,1}$ when the covariate effects are associated with both Y and Z (i.e., (Y, Z) case). However, they rarely detect the effects, $\beta_{Z,1}$ and $\gamma_{Z,1}$, when the covariate effects are associated with Y but not associated with Z (i.e., Y only case). The results were provided in Web Table 8. Compared with (Y, Z) case, Y only case did not borrow information from the binary response at early of trial, which makes less enrich the E-sensitive patients and less reject the null hypothesis.

[Web Table 8 about here.]

Simulations for consideration of no treatment-covariate interaction effect We performed additional simulations to evaluate the performance of AED and compare it to the all comer CGS design when there are no treatment-covariate interactions, since CGS is appropriate in this case. We generated the data from regression models (8) and (9) with regression coefficients $\beta_{Z,1} = -0.1$, $\gamma_{Z,0} = 0.4$, $\beta_{Y,1} = 0.2$ and $\gamma_{Y,0} = -0.8$, with all other regression coefficients set to 0. We considered three prevalence of $x_1 = 1$: 65%, 50%, and 35%, and generated the other covariates, $x_2, \ldots, x_{10}$, from Bernoulli distributions with response probabilities 0.5, 0.5, 0.1, 0.2, 0.2, 0.4, 0.6, 0.8, 0.8. Simulation results are shown in Web Table 9. In this case, the GP and power are essentially identical, since there is no E-sensitive subgroup. The results show that, when there are no treatment-covariate interactions, AED performs well and has properties very similar to those of CGS.

[Web Table 9 about here.]

Simulations with different design parameters Scenario 2 of the simulation study (Section 3) was considered for the sensitivity analysis of the design parameters v and g regarding enrichment. In the simulation study, we used $v = 0.766$ and $g = 0.352$, which correspond to a $\geqslant 60\%$ chance that a patient is expected to benefit from E in terms of the short-term endpoint and $\geqslant 70\%$ chance that a patient is expected to benefit from E in terms of the long-term endpoint. Here, we considered $v = 0.85$ and $g = 0.503$, so that there is a $\geqslant 60\%$ chance that a patient is expected to benefit from E in terms of the short-term endpoint and $\geqslant 74\%$ chance that a patient is expected to benefit from E in terms of the long-term endpoint. The results were provided in Web Table 10. This change in tuning parameters allows us to enroll more E-sensitive patients in 2nd and 3rd cohorts, and thus increases the power up to 6%.

[Web Table 10 about here.]

Simulations with different sparsity parameters We again considered scenario 2 of the simulation study (Section 3) for this sensitivity analysis. In the simulation study, we used $u_{Z,j} = u_{Y,j} = 100$ and $\tau_{Z,j} = \tau_{Y,j} = 0.1$ for $j = 1, \ldots, 2p + 1$. Here, sensitivity analyses for the parameters $u_{Z,j}$, $u_{Y,j}$, $\tau_{Z,j}$ and $\tau_{Y,j}$ were performed with relatively smaller or bigger values to $u_{Z,j} = u_{Y,j} = 100$ and $\tau_{Z,j} = \tau_{Y,j} = 0.1$. The results were provided in Web Table 11 and are similar to those reported when $u_{Z,j} = u_{Y,j} = 100$ and $\tau_{Z,j} = \tau_{Y,j} = 0.1$.

[Web Table 11 about here.]

Received 000 000. *Revised 000* 000. *Accepted 000* 000.

Web Table 1: True model parameters used in the simulation study for Tables 2-3. In all scenarios, $\beta_{Z,j} = \gamma_{Z,j} = \beta_{Y,j} = \gamma_{Y,j} = 0$ for $j = 4, \ldots, 10$.

	Scenarios					
	2	3	4	5	6	7
$\beta_{Z,0}$	0	0	0	0	0	0
$\beta_{Z,1}$	-0.1	-0.1	-0.1	-0.05	-0.05	-0.05
$\beta_{Z,2}$	0	-0.1	-0.1	-0.05	-0.05	-0.05
$\beta_{Z,3}$	0	0	0	-0.05	-0.05	-0.05
$\gamma_{Z,0}$	-0.1	-0.59	-0.050	-1.10	-0.78	-0.78
$\gamma_{Z,1}$	0.585	0.588	0.535	0.545	0.633	1.215
$\gamma_{Z,2}$	0	0.588	-0.490	0.545	0.633	-0.45
$\gamma_{Z,3}$	0	0	0	0.545	-0.5	-0.45
$\beta_{Y,1}$	0.2	-0.1	-0.1	-0.1	-0.1	-0.1
$\beta_{Y,2}$	0	-0.1	-0.1	-0.1	-0.1	-0.1
$\beta_{Y,3}$	0	0	0	-0.1	-0.1	-0.1
$\gamma_{Y,0}$	0.2	0.601	0.601	1.050	1.05	1.05
$\gamma_{Y,1}$	-0.911	-0.556	-1.112	-0.520	-0.781	-1.561
$\gamma_{Y,2}$	0	-0.556	0.540	-0.520	-0.781	0.52
$\gamma_{Y,3}$	0	0	0	-0.520	0.53	0.52

Web Table 2: Results of probability that stop early due to superiority and futility obtained from 1000 simulated datasets when survival time follows a Weibull distribution with deceasing hazard.

Scen	% sensitive patients in 1st cohort	Prob that Stop Early superiority (futility)				
		AED	GSED	InterAdapt	Simon	CGS
1	0	0.05(0.72)	0.05(0.69)	0.06(0.94)	0.05(0.69)	0.05(0.06)
2	0.65	0.67(0.12)	0.70(0.08)	0.65(0.35)	0.64(0.10)	0.65(0.25)
	0.50	0.64(0.17)	0.46(0.26)	0.65(0.35)	0.39(0.27)	0.51(0.46)
	0.35	0.62(0.18)	0.35(0.45)	0.62(0.38)	0.18(0.46)	0.10(0.65)
3	0.65	0.63(0.25)	0.35(0.27)	0.69(0.31)	0.51(0.20)	0.58(0.34)
	0.50	0.59(0.27)	0.23(0.43)	0.58(0.42)	0.26(0.39)	0.53(0.38)
	0.35	0.54(0.30)	0.23(0.46)	0.42(0.58)	0.08(0.55)	0.46(0.41)
4	0.65	0.68(0.20)	0.36(0.26)	0.68(0.32)	0.48(0.24)	0.58(0.42)
	0.50	0.63(0.24)	0.21(0.48)	0.66(0.34)	0.30(0.33)	0.44(0.54)
	0.35	0.57(0.32)	0.17(0.59)	0.63(0.37)	0.11(0.50)	0.02(0.75)
5	0.65	0.79(0.07)	0.31(0.47)	0.40(0.60)	0.20(0.43)	0.24(0.14)
	0.50	0.78(0.08)	0.19(0.66)	0.37(0.63)	0.15(0.48)	0.21(0.26)
	0.35	0.73(0.14)	0.07(0.86)	0.23(0.77)	0.09(0.54)	0.21(0.40)
6	0.65	0.83(0.04)	0.29(0.39)	0.36(0.64)	0.37(0.26)	0.09(0.41)
	0.50	0.83(0.09)	0.13(0.61)	0.28(0.72)	0.23(0.44)	0.09(0.49)
	0.35	0.77(0.10)	0.04(0.83)	0.17(0.84)	0.07(0.60)	0.07(0.50)
7	0.65	0.91(0.05)	0.23(0.46)	0.36(0.64)	0.61(0.12)	0.04(0.53)
	0.50	0.80(0.14)	0.16(0.56)	0.25(0.75)	0.32(0.32)	0.02(0.63)
	0.35	0.80(0.17)	0.07(0.74)	0.16(0.84)	0.20(0.46)	0.01(0.67)

Web Table 3: Simulation results obtained from 1000 simulated datasets for the proposed AED when survival time follows a Weibull distribution with increasing hazard.

	% sensitive patients in cohort			Prob identify	Power		Pr(Stop Early)		Mean
Scen	1st	2nd	3rd	target popn	traditional	generalized	Superiority	Futility	sample size
1	0	0	0	NA	0.05	NA	0.05	0.87	213
2	0.65	0.89	0.90	0.94	0.87	0.84	0.52	0.13	274
	0.50	0.80	0.82	0.88	0.86	0.81	0.48	0.14	279
	0.35	0.71	0.73	0.80	0.82	0.74	0.38	0.16	291
3	0.65	0.82	0.85	0.88	0.77	0.69	0.58	0.23	251
	0.50	0.64	0.71	0.73	0.70	0.57	0.49	0.28	251
	0.35	0.61	0.62	0.71	0.66	0.56	0.49	0.31	251
4	0.65	0.79	0.79	0.81	0.85	0.77	0.52	0.14	272
	0.50	0.71	0.71	0.74	0.82	0.73	0.52	0.18	271
	0.35	0.53	0.55	0.61	0.80	0.71	0.49	0.20	271
5	0.65	0.75	0.83	0.89	0.83	0.49	0.81	0.12	214
	0.50	0.72	0.78	0.81	0.82	0.47	0.79	0.14	225
	0.35	0.63	0.67	0.69	0.77	0.47	0.75	0.20	230
6	0.65	0.87	0.94	0.98	0.93	0.72	0.81	0.07	235
	0.50	0.86	0.90	0.94	0.89	0.68	0.71	0.11	254
	0.35	0.80	0.84	0.90	0.73	0.49	0.52	0.27	255
7	0.65	0.90	0.94	0.97	0.95	0.81	0.89	0.02	232
	0.50	0.87	0.96	0.96	0.85	0.72	0.85	0.15	227
	0.35	0.82	0.86	0.94	0.83	0.68	0.80	0.17	220

Web Table 4: Simulation results obtained from 1000 simulated datasets for the proposed AED when survival time is generated from a log-logistic distribution.

Scen	% sensitive patients in cohort			Prob identify	Power		Pr(Stop Early)		Mean
	1st	2nd	3rd	target popn	traditional	generalized	Superiority	Futility	sample size
1	0	0	0	NA	0.05	NA	0.05	0.92	205
2	0.65	0.83	0.84	0.91	0.77	0.68	0.42	0.23	273
	0.50	0.76	0.76	0.85	0.75	0.64	0.38	0.25	279
	0.35	0.67	0.68	0.79	0.69	0.61	0.33	0.30	276
3	0.65	0.80	0.81	0.86	0.66	0.60	0.35	0.31	274
	0.50	0.76	0.77	0.81	0.62	0.54	0.35	0.39	264
	0.35	0.66	0.67	0.73	0.59	0.50	0.31	0.39	265
4	0.65	0.75	0.77	0.80	0.76	0.67	0.50	0.23	261
	0.50	0.64	0.67	0.72	0.73	0.64	0.48	0.26	259
	0.35	0.51	0.52	0.59	0.71	0.62	0.39	0.28	273
5	0.65	0.88	0.88	0.89	0.77	0.57	0.70	0.18	241
	0.50	0.81	0.83	0.87	0.74	0.49	0.66	0.19	243
	0.35	0.72	0.72	0.78	0.70	0.47	0.58	0.20	256
6	0.65	0.91	0.91	0.94	0.89	0.60	0.74	0.11	244
	0.50	0.86	0.88	0.94	0.84	0.57	0.62	0.16	257
	0.35	0.76	0.77	0.80	0.82	0.57	0.62	0.18	261
7	0.65	0.92	0.92	0.94	0.91	0.79	0.85	0.06	222
	0.50	0.82	0.86	0.91	0.86	0.72	0.79	0.14	221
	0.35	0.76	0.72	0.78	0.71	0.58	0.71	0.26	215

Web Table 5: Simulation results obtained from 1000 simulated datasets for the proposed AED when survival time follows a Weibull distribution with deceasing hazard when $n = 800$.

	% sensitive patients in cohort			Prob identify	Power		Pr(Stop Early)		Mean
Scen	1st	2nd	3rd	target popn	traditional	generalized	Superiority	Futility	sample size
1	0	0	0	NA	0.05	NA	0.05	0.66	500
2	0.65	0.91	0.92	0.94	0.90	0.89	0.73	0.09	489
	0.50	0.85	0.86	0.90	0.90	0.89	0.73	0.10	485
	0.35	0.75	0.78	0.83	0.89	0.87	0.68	0.10	502
3	0.65	0.83	0.83	0.85	0.76	0.72	0.67	0.24	457
	0.50	0.75	0.75	0.79	0.74	0.71	0.60	0.26	485
	0.35	0.59	0.60	0.63	0.73	0.66	0.53	0.27	502
4	0.65	0.82	0.83	0.83	0.84	0.82	0.68	0.16	477
	0.50	0.73	0.73	0.74	0.81	0.80	0.64	0.18	483
	0.35	0.59	0.61	0.61	0.80	0.79	0.63	0.20	483
5	0.65	0.77	0.89	0.96	0.90	0.62	0.83	0.11	438
	0.50	0.76	0.84	0.91	0.85	0.61	0.73	0.19	443
	0.35	0.70	0.76	0.83	0.79	0.56	0.67	0.21	454
6	0.65	0.93	0.94	0.97	0.96	0.70	0.77	0.14	469
	0.50	0.88	0.90	0.94	0.91	0.66	0.76	0.19	468
	0.35	0.80	0.84	0.92	0.90	0.64	0.75	0.19	467
7	0.65	0.94	0.95	0.97	0.97	0.96	0.91	0.03	460
	0.50	0.89	0.92	0.93	0.91	0.91	0.91	0.09	419
	0.35	0.81	0.88	0.89	0.88	0.88	0.88	0.12	419

Web Table 6: Simulation results obtained from 1000 simulated datasets for the proposed AED when survival time follows a Weibull distribution with deceasing hazard when $p = 50$.

	% sensitive patients in cohort			Prob identify	Power		Pr(Stop Early)		Mean
Scen	1st	2nd	3rd	target popn	traditional	generalized	Superiority	Futility	sample size
1	0	0	0	NA	0.05	NA	0.05	0.70	233
2	0.65	0.82	0.85	0.88	0.94	0.68	0.84	0.06	235
	0.50	0.74	0.76	0.80	0.91	0.65	0.76	0.09	248
	0.35	0.63	0.69	0.71	0.89	0.64	0.76	0.11	248
3	0.65	0.76	0.77	0.80	0.86	0.65	0.80	0.14	229
	0.50	0.67	0.67	0.71	0.77	0.64	0.66	0.23	242
	0.35	0.53	0.54	0.58	0.72	0.53	0.59	0.27	249
4	0.65	0.83	0.85	0.86	0.83	0.56	0.76	0.17	232
	0.50	0.74	0.75	0.77	0.78	0.51	0.67	0.22	244
	0.35	0.62	0.63	0.66	0.74	0.43	0.64	0.25	245
5	0.65	0.77	0.78	0.87	0.92	0.70	0.91	0.04	219
	0.50	0.77	0.78	0.85	0.82	0.55	0.81	0.13	231
	0.35	0.64	0.69	0.74	0.76	0.46	0.68	0.14	252
6	0.65	0.83	0.89	0.92	0.94	0.79	0.87	0.06	233
	0.50	0.83	0.85	0.87	0.93	0.66	0.83	0.06	230
	0.35	0.70	0.75	0.87	0.86	0.53	0.80	0.14	227
7	0.65	0.90	0.93	0.97	0.95	0.82	0.90	0.03	232
	0.50	0.86	0.87	0.92	0.92	0.80	0.90	0.08	206
	0.35	0.85	0.87	0.87	0.90	0.80	0.90	0.11	208

Web Table 7: Simulation results obtained from 1000 simulated datasets for the proposed AED when survival time follows a Weibull distribution with deceasing hazard when a main effect of x_2 is added to scenario 2 of Simulation study (i.e., the case where there are main/interaction effects of x_1 only).

	% sensitive patients in cohort			Prob identify	Power		Pr(Stop Early)		Mean
Scen	1st	2nd	3rd	target popn	traditional	generalized	Superiority	Futility	sample size
Main/Interaction	0.65	0.88	0.89	0.91	0.87	0.79	0.67	0.12	250
effect of x_1	0.50	0.81	0.81	0.85	0.84	0.73	0.64	0.17	244
	0.35	0.70	0.72	0.79	0.82	0.73	0.62	0.18	248
Main/Interaction	0.65	0.90	0.90	0.91	0.83	0.77	0.68	0.17	245
effect of x_1 +	0.50	0.84	0.85	0.85	0.80	0.77	0.65	0.20	239
Main effect of x_2	0.35	0.75	0.77	0.81	0.80	0.69	0.65	0.20	247

Web Table 8: Simulation results obtained from 1000 simulated datasets for the proposed AED when survival time follows a Weibull distribution with deceasing hazard when covariate x_1 is associated with Y and Z or associated with Y only in scenario 2 of Simulation study.

Effect of x_1	% sensitive patients in cohort			Prob identify	Power		Pr(Stop Early)		Mean
associated with	1st	2nd	3rd	target popn	traditional	generalized	Superiority	Futility	sample size
(Y, Z)	0.65	0.88	0.89	0.91	0.87	0.79	0.67	0.12	250
	0.50	0.81	0.81	0.85	0.84	0.73	0.64	0.17	244
	0.35	0.70	0.72	0.79	0.82	0.73	0.62	0.18	248
Y only	0.65	0.83	0.87	0.89	0.82	0.61	0.63	0.17	248
	0.50	0.75	0.79	0.82	0.82	0.59	0.62	0.18	247
	0.35	0.60	0.63	0.66	0.76	0.58	0.56	0.25	247

Web Table 9: Simulation results for cases with no treatment-covariate interactions.

			Pr(Stop early)				Mean	
	Power		Superiority		Futility		Sample Size	
Pr($x_1 = 1$)	AED	CGS	AED	CGS	AED	CGS	AED	CGS
0.65	0.80	0.78	0.24	0.23	0.19	0.21	308	319
0.50	0.79	0.79	0.25	0.22	0.21	0.22	302	314
0.35	0.80	0.80	0.23	0.20	0.19	0.18	307	321

Web Table 10: Simulation results obtained from 1000 simulated datasets for the proposed AED when survival time follows a Weibull distribution with deceasing hazard when different tuning parameters are used in scenario 2 of Simulation study.

	% sensitive patients in cohort			Prob identify	Power		Pr(Stop Early)		Mean
(v, g)	1st	2nd	3rd	target popn	traditional	generalized	Superiority	Futility	sample size
$(0.77, 0.35)$	0.65	0.88	0.89	0.91	0.87	0.79	0.67	0.12	250
	0.50	0.81	0.81	0.85	0.84	0.73	0.64	0.17	244
	0.35	0.70	0.72	0.79	0.82	0.73	0.62	0.18	248
$(0.85, 0.50)$	0.65	0.88	0.90	0.94	0.93	0.83	0.74	0.07	246
	0.50	0.82	0.85	0.91	0.89	0.75	0.68	0.12	250
	0.35	0.75	0.77	0.83	0.88	0.75	0.65	0.12	257

Web Table 11: Simulation results obtained from 1000 simulated datasets for the proposed AED when survival time follows a Weibull distribution with deceasing hazard when different sparsity parameters are used in scenario 2 of Simulation study.

	% sensitive patients in cohort			Prob identify	Power		Pr(Stop Early)		Mean sample
$(u_{\cdot,j}, \tau_{\cdot,j})$	1st	2nd	3rd	target popn	traditional	generalized	Superiority	Futility	size
$(100, 0.1)$	0.65	0.88	0.89	0.91	0.87	0.79	0.67	0.12	250
	0.50	0.81	0.81	0.85	0.84	0.73	0.64	0.17	244
	0.35	0.70	0.72	0.79	0.82	0.73	0.62	0.18	248
$(100, 0.05)$	0.65	0.84	0.86	0.89	0.88	0.77	0.69	0.12	262
	0.50	0.79	0.80	0.88	0.82	0.75	0.62	0.18	271
	0.35	0.67	0.68	0.73	0.81	0.70	0.47	0.19	275
$(100, 0.5)$	0.65	0.88	0.88	0.91	0.86	0.77	0.64	0.14	249
	0.50	0.81	0.82	0.86	0.83	0.76	0.53	0.15	271
	0.35	0.70	0.71	0.77	0.80	0.72	0.51	0.19	267
$(100, 1)$	0.65	0.88	0.89	0.92	0.86	0.76	0.61	0.14	256
	0.50	0.82	0.83	0.85	0.86	0.75	0.48	0.14	280
	0.35	0.71	0.71	0.77	0.82	0.75	0.40	0.17	287
$(10, 0.1)$	0.65	0.84	0.87	0.88	0.89	0.85	0.40	0.11	302
	0.50	0.73	0.82	0.88	0.86	0.78	0.37	0.14	294
	0.35	0.67	0.69	0.72	0.83	0.76	0.34	0.16	304
$(50, 0.1)$	0.65	0.88	0.90	0.93	0.89	0.85	0.67	0.12	253
	0.50	0.80	0.81	0.86	0.85	0.75	0.60	0.13	267
	0.35	0.71	0.74	0.79	0.82	0.74	0.59	0.15	247