Supplementary Materials: Precision Bayesian Phase I-II Dose-Finding Based on Utilities Tailored to Prognostic Subgroups

1 Posterior Computation

Recall that $\boldsymbol{\theta}$ denotes the vector of all model parameters and $\tilde{\boldsymbol{\theta}}$ is the vector of all fixed hyperparameters, with

$\boldsymbol{\theta} = (h_{\text{T}0}, \boldsymbol{\beta}, \boldsymbol{\alpha}^\star, \boldsymbol{z}, \boldsymbol{\rho}, \Omega)$,

$\boldsymbol{\beta} = \{\beta_{j,\ell}, j = T, E, \ell = 1, 2, 3\}$,

$\boldsymbol{\alpha}^\star = \{\alpha^\star_{j,z}, j = T, E, z = 1, \ldots, R\}$,

$\boldsymbol{\rho} = \{\rho_{z,k}, z = 1, \ldots, R, k = 2, \ldots, K-1\}$,

$\tilde{\boldsymbol{\theta}} = (\bar{h}_{\text{T}0}, v_h^2, \bar{\boldsymbol{\beta}}, \boldsymbol{\tau}^2, \boldsymbol{\xi}, \bar{\boldsymbol{\alpha}}^\star, \boldsymbol{v}^2, \boldsymbol{\rho}_0, \boldsymbol{\kappa}, \Omega^0)$,

$\bar{\boldsymbol{\beta}} = \{\bar{\beta}_{j,\ell}, j = T, E, \ell = 1, 2, 3\}$,

$\boldsymbol{\tau}^2 = \{\tau^2_{j,\ell}, j = T, E, \ell = 1, 2, 3\}$,

$\boldsymbol{\xi} = (\xi_2, \ldots, \xi_G)$,

$\bar{\boldsymbol{\alpha}}^\star = \{\bar{\alpha}^\star_{j,z}, j = T, E, z = 1, \ldots, G\}$,

$\boldsymbol{v}^2 = \{v^2_{j,z}, j = T, E, z = 1, \ldots, G\}$,

$\boldsymbol{\rho}_0 = \{\rho_{0,k}, k = 2, \ldots, K-1\}$, and

$\boldsymbol{\kappa} = \{\kappa_{z,k}, z = 1, \ldots, G, k = 2, \ldots, K-1\}$.

We use the Markov chain Monte Carlo (MCMC) simulation to draw a sample of random parameters $\boldsymbol{\theta}$ and patient random effects $\boldsymbol{\gamma} = (\gamma_{\mathrm{T}}, \gamma_{\mathrm{E}})$ from the following posterior distribution.

$$\begin{aligned} p(\boldsymbol{\theta}, \boldsymbol{\gamma} \mid \mathcal{D}_{n(t)}, \tilde{\boldsymbol{\theta}}) &\propto p(\boldsymbol{\theta}, \boldsymbol{\gamma} \mid \tilde{\boldsymbol{\theta}}) \prod_{i=1}^{n(t)} p(y^o_{i,\mathrm{T}}, y_{i,\mathrm{E}}, \boldsymbol{\delta}_i \mid d_{[i]}, g_i, \boldsymbol{\gamma}_i, \boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}) \\ &= p(\boldsymbol{\theta}, \boldsymbol{\gamma} \mid \tilde{\boldsymbol{\theta}}) \prod_{i=1}^{n(t)} \Big[\left\{ h_{\mathrm{T}}(y^o_{i,\mathrm{T}} \mid g_i, d_{[i]}, \gamma_{i,\mathrm{T}}) \right\}^{\delta_{i,\mathrm{T}}} S_{\mathrm{T}}(y^o_{i,\mathrm{T}} \mid g_i, d_{[i]}, \gamma_{i,\mathrm{T}}) \Big] \\ &\qquad \times \left\{ \mathrm{P}(y_{i,\mathrm{E}} \mid g_i, d_{[i]}, \gamma_{i,\mathrm{E}}) \right\}^{\delta_{i,\mathrm{E}}}, \end{aligned} \tag{1}$$

where $\mathcal{D}_{n(t)} = \{\boldsymbol{y}^o_i, \boldsymbol{\delta}_i, g_i, d_{[i]}\}_{i=1}^{n(t)}$ denotes the data from $n(t)$ patients at trial time t. We assume a constant baseline hazard, $h_{0\mathrm{T}}(t) = h_{0\mathrm{T}}$, for all $t \in (0, C)$. The survival function is

$$\begin{aligned} S_{\mathrm{T}}(y_{i,\mathrm{T}} \mid g_i, d_{[i]}, \boldsymbol{\gamma}_i) &= \exp\left\{ - \int_0^{y_{i,\mathrm{T}}} h_{0\mathrm{T}} \exp\left(\eta_j(d_{[i]}) + \alpha_{\mathrm{T},g_i} + \gamma_{i,\mathrm{T}} \right) dt \right\} \\ &= \exp\left\{ -y_{i,\mathrm{T}} h_{0\mathrm{T}} \exp\left(\eta_j(d_{[i]}) + \alpha_{\mathrm{T},g_i} + \gamma_{i,\mathrm{T}} \right) \right\}. \end{aligned} \tag{2}$$

Recall that $\alpha_{j,g} = \alpha^{\star}_{j,z_g}$. We draw samples of most of the random parameters using the Metropolis-Hastings algorithm. The Markov chain parameters in the proposal distributions are automatically tuned by adaptive MCMC algorithms (Roberts and Rosenthal, 2009). We checked mixing and convergence of the Markov chain and did not find any evidence of converging to a wrong distribution. The full conditionals are given below.

1. Full conditionals of $\beta_{\mathrm{T},\ell}$

   Let $\boldsymbol{\theta}_{-\beta_{\mathrm{T},\ell}}$, $\ell = 1, 2, 3$ denote the vector of all random parameters, excluding $\beta_{\mathrm{T},\ell}$. For $\ell = 1$ and 3, we have

$$\begin{aligned} p(\beta_{\mathrm{T},\ell} \mid \mathcal{D}_{n(t)}, \boldsymbol{\gamma}, \boldsymbol{\theta}_{-\beta_{\mathrm{T},\ell}}) &\propto \exp\left\{ -\frac{(\beta_{\mathrm{T},\ell} - \bar{\beta}_{\mathrm{T},\ell})^2}{2\tau^2_{\mathrm{T},\ell}} \right\} \times \exp\Big[\sum_{i=1}^{n(t)} \Big\{ \delta_{\mathrm{T},i} \eta_{\mathrm{T}}(d_{[i]}) \\ &\qquad - y_{i,\mathrm{T}} h_{0\mathrm{T}} \exp\left((\eta_{\mathrm{T}}(d_{[i]}) + \alpha_{\mathrm{T},g_i} + \gamma_{i,\mathrm{T}}) \right) \Big\} \Big] \times 1(\beta_{\mathrm{T},\ell} > 0), \end{aligned}$$

where $\eta_{\mathrm{T}}(d_{[i]}) = \frac{\beta_{\mathrm{T},3}}{1+\exp(-\beta_{\mathrm{T},1}\times(10\times d_{[i]}-\beta_{\mathrm{T},2}))}$. For $\ell = 2$, we have

$$p(\beta_{\mathrm{T},2} \mid \mathcal{D}_{n(t)}, \boldsymbol{\gamma}, \boldsymbol{\theta}_{-\beta_{\mathrm{T},2}}) \propto \exp\left\{-\frac{(\beta_{\mathrm{T},2}-\bar{\beta}_{\mathrm{T},2})^2}{2\tau^2_{\mathrm{T},2}}\right\} \times \exp\Big[\sum_{i=1}^{n(t)}\Big\{\delta_{\mathrm{T},i}\eta_{\mathrm{T}}(d_{[i]}) - y_{i,\mathrm{T}}h_{0\mathrm{T}}\exp\big((\eta_{\mathrm{T}}(d_{[i]}) + \alpha_{\mathrm{T},g_i} + \gamma_{i,\mathrm{T}})\big)\Big\}\Big].$$

2. Full conditionals of $h_{0\mathrm{T}}$

   Let $\boldsymbol{\theta}_{-h_{0\mathrm{T}}}$ denote the vector of all random parameters, excluding $h_{0\mathrm{T}}$. We have

$$p(h_{0\mathrm{T}} \mid \mathcal{D}_{n(t)}, \boldsymbol{\gamma}, \boldsymbol{\theta}_{-h_{0\mathrm{T}}}) \propto \frac{1}{h_{0\mathrm{T}}}\exp\left\{-\frac{(\log(h_{0\mathrm{T}})-\bar{h}_{0\mathrm{T}})^2}{2v_h^2}\right\} \times \exp\Big[\sum_{i=1}^{n(t)}\Big\{\delta_{\mathrm{T},i}\log(h_{0\mathrm{T}}) - y_{i,\mathrm{T}}h_{0\mathrm{T}}\exp\big((\eta_{\mathrm{T}}(d_{[i]}) + \alpha_{\mathrm{T},g_i} + \gamma_{i,\mathrm{T}})\big)\Big\}\Big].$$

3. Full conditionals of $\alpha^{\star}_{\mathrm{T},z}$

   Let $\boldsymbol{\theta}_{-\alpha^{\star}_{\mathrm{T},z}}$, $z = 2,\ldots,R$ denote the vector of all random parameters, excluding $\alpha^{\star}_{\mathrm{T},z}$. For $z = 2,\ldots,R$, we have

$$p(\alpha^{\star}_{\mathrm{T},z} \mid \mathcal{D}_{n(t)}, \boldsymbol{\gamma}, \boldsymbol{\theta}_{-\alpha^{\star}_{\mathrm{T},z}}) \propto \exp\left\{-\frac{(\alpha^{\star}_{\mathrm{T},z}-\bar{\alpha}^{\star}_{\mathrm{T},z})^2}{2v^2_{\mathrm{T},z}}\right\} \times \exp\Big[\sum_{i=1|z_{g_i}=z}^{n(t)}\Big\{\delta_{\mathrm{T},i}\alpha^{\star}_{\mathrm{T},z} - y_{i,\mathrm{T}}h_{0\mathrm{T}}\exp\big((\eta_{\mathrm{T}}(d_{[i]}) + \alpha^{\star}_{\mathrm{T},z} + \gamma_{i,\mathrm{T}})\big)\Big\}\Big] \times \prod_{z'=2}^{R}\left\{1(\alpha^{\star}_{\mathrm{T},(z'-1)} < \alpha^{\star}_{\mathrm{T},z'})\frac{1}{\mathrm{P}(\alpha^{\star}_{\mathrm{T},(z'-1)} < \alpha^{\star}_{\mathrm{T},z'})}\right\},$$

   where $\alpha^{\star}_{\mathrm{T},1} = 0$,.

4. Full conditionals of $\beta_{\mathrm{E},\ell}$

   Let $\boldsymbol{\theta}_{-\beta_{\mathrm{E},\ell}}$, $\ell = 1,2,3$ denote the vector of all random parameters, excluding $\beta_{\mathrm{E},\ell}$. For

$\ell = 1$ and 3, we have

$$p(\beta_{\mathrm{E},\ell} \mid \mathcal{D}_{n(t)}, \boldsymbol{\gamma}, \boldsymbol{\theta}_{-\beta_{\mathrm{E},\ell}}) \quad \propto \quad \exp\left\{-\frac{(\beta_{\mathrm{E},\ell} - \bar{\beta}_{\mathrm{E},\ell})^2}{2\tau^2_{\mathrm{E},\ell}}\right\} \times \prod_{i=1|\delta_{i,\mathrm{E}}=1}^{n(t)} \pi_{y_{i,\mathrm{E}}}(g_i, d_{[i]}, \gamma_{i,\mathrm{E}}) \times 1(\beta_{\mathrm{E},\ell} > 0),$$

where

$$\pi_k(g_i, d_{[i]}, \gamma_{i,\mathrm{E}}) = \Phi(u_{g_i,k+1} \mid \eta_{\mathrm{E}}(d_{[i]}) + \alpha_{\mathrm{E},g_i} + \gamma_{i,\mathrm{E}}, \sigma^2_\pi) - \Phi(u_{g_i,k} \mid \eta_{\mathrm{E}}(d_{[i]}) + \alpha_{\mathrm{E},g_i} + \gamma_{i,\mathrm{E}}, \sigma^2_\pi),$$

and $\eta_{\mathrm{E}}(d_{[i]}) = \frac{\beta_{\mathrm{E},3}}{1+\exp(-\beta_{\mathrm{E},1} \times (10 \times d_{[i]} - \beta_{\mathrm{E},2}))}$. For $\ell = 2$, we have

$$p(\beta_{\mathrm{E},2} \mid \mathcal{D}_{n(t)}, \boldsymbol{\gamma}, \boldsymbol{\theta}_{-\beta_{\mathrm{E},2}}) \quad \propto \quad \exp\left\{-\frac{(\beta_{\mathrm{E},2} - \bar{\beta}_{\mathrm{E},2})^2}{2\tau^2_{\mathrm{E},2}}\right\} \times \prod_{i=1|\delta_{i,\mathrm{E}}=1}^{n(t)} \pi_{y_{i,\mathrm{E}}}(g_i, d_{[i]}, \gamma_{i,\mathrm{E}}).$$

5. Full conditionals of $\alpha^\star_{\mathrm{E},z}$

   Let $\boldsymbol{\theta}_{-\alpha^\star_{\mathrm{E},z}}$, $z = 1, \ldots, R$ denote the vector of all random parameters, excluding $\alpha^\star_{\mathrm{E},z}$. For $z = 1, \ldots, R$, we have

$$\begin{aligned} p(\alpha^\star_{\mathrm{E},z} \mid \mathcal{D}_{n(t)}, \boldsymbol{\gamma}, \boldsymbol{\theta}_{-\alpha^\star_{\mathrm{E},z}}) \quad \propto \quad & \exp\left\{-\frac{(\alpha^\star_{\mathrm{E},z} - \bar{\alpha}^\star_{\mathrm{E},z})^2}{2v^2_{\mathrm{E},z}}\right\} \prod_{i=1|\delta_{i,\mathrm{E}}=1}^{n(t)} \pi_{y_{i,\mathrm{E}}}(g_i, d_{[i]}, \gamma_{i,\mathrm{E}}) \\ & \times \prod_{z'=1}^{R} \left\{1(\alpha^\star_{\mathrm{E},(z'-1)} > \alpha^\star_{\mathrm{E},z'}) \frac{1}{\mathrm{P}(\alpha^\star_{\mathrm{E},(z'-1)} > \alpha^\star_{\mathrm{E},z'})}\right\}. \end{aligned}$$

6. Full conditionals of $\rho_{z,k}$

   Let $\boldsymbol{\theta}_{-\rho_{z,k}}$, $z = 1, \ldots, R$ and $k = 2, \ldots, K-1$ denote the vector of all random parameters, excluding $\rho_{z,k}$. For $z = 1, \ldots, R$ and $k = 2, \ldots, K-1$, we have

$$\begin{aligned} p(\rho_{z,k} \mid \mathcal{D}_{n(t)}, \boldsymbol{\gamma}, \boldsymbol{\theta}_{-\rho_{z,k}}) \quad \propto \quad & \frac{(\kappa_{z,k})^{(\rho_{z-1,k}\kappa_{z,k})}}{\Gamma(\rho_{z-1,k}\kappa_{z,k})} (\rho_{z,k})^{(\rho_{z-1,k}\kappa_{z,k})} \exp(-\kappa_{z,k}\rho_{z,k}) \\ & \times \prod_{i=1|\delta_{i,\mathrm{E}}=1}^{n(t)} \pi_{y_{i,\mathrm{E}}}(g_i, d_{[i]}, \gamma_{i,\mathrm{E}}). \end{aligned}$$

7. Full conditionals of $\boldsymbol{z}$

   Let $\boldsymbol{\theta}_{-\boldsymbol{z}}$ denote the vector of all random parameters, excluding $\boldsymbol{z}$. We have

$$\begin{aligned} p(\boldsymbol{z} \mid \mathcal{D}_{n(t)}, \boldsymbol{\gamma}, \boldsymbol{\theta}_{-\boldsymbol{z}}) \quad \propto \quad & \prod_{g=2}^{G} (\xi_x)^{I(z_x = z_{x-1})} (1-\xi_x)^{1-I(z_x = z_{x-1})} \\ & \times \exp\Big[\sum_{i=1}^{n(t)} \Big\{ \delta_{\mathrm{T},i} \alpha^{\star}_{\mathrm{T},z_{g_i}} - y_{i,\mathrm{T}} h_{0\mathrm{T}} \exp\Big((\eta_{\mathrm{T}}(d_{[i]}) + \alpha^{\star}_{\mathrm{T},z_{g_i}} + \gamma_{i,\mathrm{T}}) \Big) \Big\} \Big] \\ & \times \prod_{i=1 \mid \delta_{i,\mathrm{E}}=1}^{n(t)} \pi_{y_{i,\mathrm{E}}}(g_i, d_{[i]}, \gamma_{i,\mathrm{E}}). \end{aligned}$$

8. Full conditional of Ω

   Let $\boldsymbol{\theta}_{-\Omega}$ denote a vector of all random parameters, excluding Ω.

$$\begin{aligned} p(\Omega \mid \mathcal{D}_{n(t)}, \boldsymbol{\theta}_{-\Omega}, \boldsymbol{\gamma}) \quad & \propto \quad \prod_{i=1}^{n(t)} p(\boldsymbol{\gamma}_i \mid \Omega) p(\Omega \mid \nu, \Omega^0) \\ & \propto \quad \prod_{i=1}^{n(t)} |\Omega|^{-1/2} \exp\left(-\frac{1}{2} \boldsymbol{\gamma}_i' \Omega^{-1} \boldsymbol{\gamma}_i \right) |\Omega|^{\frac{\nu+4+1}{2}} \exp\left\{ -\frac{1}{2} \mathrm{tr}(\Omega^0 \Omega^{-1}) \right\}. \end{aligned}$$

   Thus, the full conditional distribution is $\Omega \mid \mathcal{D}_{n(t)}, \boldsymbol{\theta}, \boldsymbol{\gamma} \sim$ inv-Wishart$(\nu + n(t), \Omega^0 + \sum_{i=1}^{n(t)} \boldsymbol{\gamma}_i \boldsymbol{\gamma}_i')$.

9. Full conditional of $\boldsymbol{\gamma}_i$

$$\begin{aligned} p(\boldsymbol{\gamma}_i \mid \mathcal{D}_{n(t)}, \boldsymbol{\theta}_{-\boldsymbol{\gamma}_i}) \quad \propto \quad & \exp\left(-\frac{1}{2} \boldsymbol{\gamma}_i' \Omega^{-1} \boldsymbol{\gamma}_i \right) \exp\Big\{ \delta_{\mathrm{T},i} \gamma_{i,\mathrm{T}} - y_{i,\mathrm{T}} h_{0\mathrm{T}} \exp\big((\eta_{\mathrm{T}}(d_{[i]}) + \alpha_{\mathrm{T},g_i} + \gamma_{i,\mathrm{T}}) \big) \Big\} \\ & \times \{ \pi_{y_{i,\mathrm{E}}}(g_i, d_{[i]}, \gamma_{i,\mathrm{E}}) \}^{\delta_{i,\mathrm{E}}}. \end{aligned}$$

We generate a posterior sample of $\boldsymbol{\theta}$ values through posterior MCMC simulation and use it to evaluate quantities needed for subgroup-specific decisions, optimal dose selection, and dose acceptability monitoring during the trial. For selecting an optimal dose for subgroup

g, we proceed as follows: We generate a Mote Carlo sample according to $\boldsymbol{\theta} \sim p(\boldsymbol{\theta} \mid \mathcal{D}_{n(t)}, \tilde{\boldsymbol{\theta}})$, where $\mathcal{D}_{n(t)}$ is the data from the trial at time t and $\tilde{\boldsymbol{\theta}}$ fixed hyperparameters, and evaluate the posterior predictive mean utility of dose d_m given subgroup g,

$$\begin{aligned} u(d_m \mid x, \mathcal{D}_{n(t)}) &= \int_0^\infty \sum_{y_{\mathrm{E}}=0}^{K-1} U_g(\boldsymbol{y}) p(\boldsymbol{y} \mid g, d_m, \mathcal{D}_{n(t)}) dy_{\mathrm{T}} \\ &= \int_{\boldsymbol{\theta}} \int_0^\infty \sum_{y_{\mathrm{E}}=0}^{K-1} U_g(\boldsymbol{y}) p(\boldsymbol{y} \mid g, d_m, \boldsymbol{\gamma}, \boldsymbol{\theta}) p(\boldsymbol{\gamma}, \boldsymbol{\theta} \mid \mathcal{D}_{n(t)}) dy_{\mathrm{T}} d\boldsymbol{\gamma} d\boldsymbol{\theta}. \end{aligned}$$

in (8) of the main text using the sample. Evaluation of $u(d \mid g, \mathcal{D}_{n(t)})$ uses the posterior predictive distribution $p(\boldsymbol{Y} \mid g, d_m, \mathcal{D}_{n(t)})$ and requires marginalization of $p(\boldsymbol{Y}, \boldsymbol{\theta}, \boldsymbol{\gamma} \mid g, d, \mathcal{D}_{n(t)})$ over random parameters $\boldsymbol{\theta}$ and random frailty $\boldsymbol{\gamma}$. We utilize a Monte Carlo approximation for the integration using the posterior sample from the MCMC simulation by $\{\boldsymbol{\theta}^{(v_1)}$, $v_1 = 1, \ldots, V_1\}$ as follows; for each (g, d_m), the computation proceeds as follows:

1. For each $v_1 = 1, \ldots, V_1$ and $v_2 = 1, \ldots, V_2$
   (a) Simulate $\boldsymbol{\gamma}^{(v_1, v_2)}$ from $p(\boldsymbol{\gamma} \mid \boldsymbol{\theta}^{(v_1)})$.
   (b) Simulate $\boldsymbol{y}^{(v_1, v_2)}$ from $p(\boldsymbol{y} \mid g, d_m, \boldsymbol{\gamma}^{(v_1, v_2)}, \boldsymbol{\theta}^{(v_1)})$.
   (c) Evaluate $U_g(\boldsymbol{y}^{(v_1, v_2)})$.
2. Use the sample mean $\frac{1}{V_1 V_2} \sum_{v_1=1}^{V_1} \sum_{v_2=1}^{V_2} U_g(\boldsymbol{y}^{(v_1, v_2)})$ to estimate $u(d_m \mid g, \mathcal{D}_{n(t)})$.

The values $V_1 = 3000$ and $V_2 = 100$ are used for our simulation study in §5 of the main text. Similarly, for the acceptability criteria we use a Monte Carlo sample of $\boldsymbol{\theta}$ simulated from the posterior distribution. We evaluate $\mathrm{P}\{\zeta_j(g, d_m) > \bar{\zeta}_j, j = T, \text{ or } PD \mid \mathcal{D}_{n(t)}\}$ through a Monte Carlo approximation as follows: For each (g, d_m),

1. For each $v_1 = 1, \ldots, V_1$ and $v_2 = 1, \ldots, V_2$,
   (a) Simulate $\boldsymbol{\gamma}^{(v_1, v_2)}$ from $p(\boldsymbol{\gamma} \mid \boldsymbol{\theta}^{(v_1)})$.

(b) Simulate $\boldsymbol{y}^{(v_1,v_2)}$ from $p(\boldsymbol{y} \mid g, d_m, \boldsymbol{\gamma}^{(v_1,v_2)}, \boldsymbol{\theta}^{(v_1)})$.

(c)$'$ Compute $\tilde{\delta}_{\mathrm{T}}^{(v_1,v_2)} = \mathrm{I}(y_{\mathrm{T}}^{(v_1,v_2)} < C)$ and $\tilde{\delta}_{\mathrm{PD}}^{(v_1,v_2)} = \mathrm{I}(y_{\mathrm{E}}^{(v_1,v_2)} = 0)$.

2$'$ For each v_1, compute the sample mean $\hat{\zeta}_j(g, d_m, \boldsymbol{\theta}^{(v_1)}) = \sum_{v_2=1}^{V_2} \tilde{\delta}_j^{(v_1,v_2)}/V_2$, $j = T, PD$. We then evaluate

$$\frac{1}{V_1} \sum_{v_1=1}^{V_1} \mathrm{I}(\hat{\zeta}_j(g, d_m, \boldsymbol{\theta}^{(v_1)}) > \bar{\zeta}_j, \ j = T, \text{ or } PD),$$

and compare it to $p^\star$, to determine if dose d_m is unacceptable for subgroup g.

2 Additional Details of Subgroup-Specific Utility Function Elicitation

In this section, we provide a detailed illustrative example of the procedure for eliciting subgroup-specific utility functions. Recall that we consider bivariate outcomes $\boldsymbol{Y} = (Y_{\mathrm{T}}, Y_{\mathrm{E}})$, where time to severe toxicity is $Y_{\mathrm{T}} \in \mathbb{R}^+$ and ordinal efficacy is $Y_{\mathrm{E}} \in \{0, 1, \ldots, K-1\}$. Consider a case with $K = 4$. Y_{T} is monitored for fixed follow-up time C, where $C = 84$ days in our example trial. Each patient is in one of G prognostic subgroups. In our motivating trial, $G = 3$ (Good prognosis, Intermediate prognosis, and Poor prognosis) and $C = 84$ days. We first elicit the marginal utility $U_{\mathrm{T},g}(Y_{\mathrm{T}})$ of Y_{T} using the following function,

$$U_{\mathrm{T},g}(Y_{\mathrm{T}}) \ = \ \begin{cases} U_{\mathrm{T,max}} \left(\frac{Y_{\mathrm{T}}}{C}\right)^{a_g} & \text{if} \quad Y_{\mathrm{T}} < C, \\ U_{\mathrm{T,max}} & \text{if} \quad Y_{\mathrm{T}} \geq C, \end{cases} \tag{3}$$

where $U_{\mathrm{T,max}}$ and a_g need to be specified. Eq (3) is continuously increasing in Y_{T} for $Y_{\mathrm{T}} \in [0, C]$ and remains constant at $U_{\mathrm{T,max}}$ for any $Y_{\mathrm{T}} > C$ to reward a later occurrence of severe toxicity. The power a_g in Eq (3) controls how fast $U_{\mathrm{T},g}$ increases in y_{T} for subgroup g. We

ask clinicians to provide numerical values using the following questions that are intuitive to practicing clinicians;

- Question A-1: What is $U_{\mathrm{T},\max}$ = the utility obtained if $Y_{\mathrm{T}} > C$ (i.e., no severe toxicity is observed during the follow-up period)?

- Question A-2: What is $y^{\star}_{\mathrm{T},g}$ = the time when $U_{\mathrm{T},g}$ is reduced by half for subgroup g, $g = 1, \ldots, G$? That is, we have $(y^{\star}_{\mathrm{T},g}/C)^{a_g} = 1/2$ at $Y_{\mathrm{T}} = y^{\star}_{\mathrm{T},g}$.

A small value of $y^{\star}_{\mathrm{T},g}$ implies that the occurrence of severe toxicity decreases the utility quickly, while the utility changes little as y_{T} gets close to C. A large value of $y^{\star}_{\mathrm{T},g}$ implies a slower decrease in the utility for smaller y_{T}, but a faster decrease for larger y_{T}. The elicited values for our motivating trial are (A-1): $U_{\mathrm{T},\max} = 140$, and (A-2): $y^{\star}_{\mathrm{T},g}$ =79, 42 and 28 for $g = 1$, 2, and 3, respectively. Using the elicited values of $y^{\star}_{\mathrm{T},g}$, we solve the equation $(y^{\star}_{\mathrm{T},g}/C)^{a_g} = 1/2$ for a_g with $C = 84$, and the resulting shape parameters are a_g=3.80, 1.00, and 0.63, respectively.

We next elicit and incorporate utilities of Y_{E}. Recall that $Y_{\mathrm{E}} = 0$, 1, 2 and 3 represent progressive disease (PD), stable disease (SD), partial response (PR) and complete response (CR), respectively. To derive the utility, $U_g(\boldsymbol{Y})$, we elicit the following information from clinical experts which may also include patients and patient advocates;

- Question B-1: How much do we penalize for the occurrence of PD?

- Question B-2: How much reward for occurrence of the events SD, PR and CR?

Here we considered PD separately from the other events since PD is a critically adverse outcome. The provided opinion was B-1: Reduce $U_{\mathrm{T},g}(Y_{\mathrm{T}})$ by half for all subgroups, i.e., $U_g(\boldsymbol{Y}) = U_{\mathrm{T},g}(Y_{\mathrm{T}})/2$ for $Y_{\mathrm{E}} = 0$, and B-2: Add rewards $U_{\mathrm{E},g}(Y_{\mathrm{E}})$ to $U_{\mathrm{T},g}(Y_{\mathrm{T}})$ to obtain the utility $U_g(\boldsymbol{Y}) = U_{\mathrm{T},g}(Y_{\mathrm{T}}) + U_{\mathrm{E},g}(Y_{\mathrm{E}})$ for $Y_{\mathrm{E}} > 0$. We next ask clinicians to the following;

- Question C-1: What utilities $U_{\text{E},g}(Y_{\text{T}})$ for $Y_{\text{E}} = 1, 2, 3$ and $g = 1, 2, 3$ are added for each $Y_{\text{E}} > 0$?

Considering that no occurrence of severe toxicity during the follow-up period has $U_{\text{T,max}}$, we fist specify the values of $U_{\text{E},g}(1)$ and $U_{\text{E},g}(3)$ for each g. The values provided by the clinicians were $U_{\text{E},g}(1) = 20$ and $U_{\text{E},g}(3) = 140$ for all g, i.e., the events of achieving CR and no severe toxicity give the same utilities, and there is a small amount of utility added for having SD. Considering the relative utility value of PR compared to SD and CR within a subgroup, $U_{\text{E},1}(2) = 60$, $U_{\text{E},2}(2) = 90$ and $U_{\text{E},3}(2) = 120$ are provided. In other words, the utility of PR relative to those of SD and CR is greater for less favorable prognostic subgroups. This reflects the practice that clinicians may accept a higher risk of toxicity to achieve PR for the poor prognostic subgroup than for the favorable prognostic subgroup.

The procedure of utility elicitation can be modified according to the specific needs of a trial. For example, a different functional form can be used for Eq (3). For certain diseases, such as some metastatic breast cancer subtypes, clinicians and patients may prefer treatments with lower toxicity risk for poorer prognostic subgroups

3 Computing Prior Hyperparameters

Recall that $\tilde{\boldsymbol{\theta}}$ denotes the vector of all fixed hyperparameters, with

$\tilde{\boldsymbol{\theta}} = (\bar{h}_{\text{T0}}, v_h^2, \bar{\boldsymbol{\beta}}, \boldsymbol{\tau}^2, \boldsymbol{\xi}, \bar{\boldsymbol{\alpha}}^\star, \boldsymbol{v}^2, \boldsymbol{\rho}_0, \boldsymbol{\kappa}, \Omega^0)$,

$\bar{\boldsymbol{\beta}} = \{\bar{\beta}_{j,\ell}, j = T, E, \ell = 1, 2, 3\}$,

$\boldsymbol{\tau}^2 = \{\tau^2_{j,\ell}, j = T, E, \ell = 1, 2, 3\}$,

$\boldsymbol{\xi} = (\xi_2, \ldots, \xi_G)$,

$\bar{\boldsymbol{\alpha}}^\star = \{\bar{\alpha}^\star_{j,z}, j = T, E, z = 1, \ldots, G\}$,

$\boldsymbol{v}^2 = \{v^2_{j,z}, j = T, E, z = 1, \ldots, G\}$,

$\boldsymbol{\rho}_0 = \{\rho_{0,k}, k = 2, \ldots, K-1\}$, and

$\boldsymbol{\kappa} = \{\kappa_{z,k}, z = 1, \ldots, G, k = 2, \ldots, K-1\}$.

To establish numerical values of the prior hyperparameters $\tilde{\boldsymbol{\theta}}$, we use historical data from previous studies and probabilities elicited from the clinicians and determine location parameters. Motzer *et al.* (2019) reports a trial, where 550 patients with renal cell carcinoma with a clear-cell component were treated by a combination of *nivolumab* and *ipilimumab*. Patients were grouped by International Metastatic Renal Cell Carcinoma Database Consortium (IMDC) risk status into three subgroups, 'favorable risk', 'intermediate risk' and 'poor risk'. They then provided efficacy data for the nivolumab plus ipilimumab immunonotherapy used as a backbone for our motivating trial which aims to determine the optimal doses of the drug sitravatinib added to the standard dose of nivolumab plus ipilimumab used by Motzer *et al.* (2019). Motzer *et al.* (2019) reported the proportion of patients achieving an objective response (either PD, SD, PR, CR, or "Unable to determine or not reported") to nivolumab plus ipilimumab separately for the 'favorable risk' IMDC subgroup and jointly for the 'intermediate risk' and 'poor risk' subgroups. Responses are often classified as "Unable to determine or not reported" in cases where due to rapid clinical deterioration (disease progression) the patients needed to be taken off trial or otherwise were unable to undergo restaging imaging. For this reason, we chose to be conservative and considered those patients classified as "Unable to determine or not reported" to be having "PD" as shown in Table 1. Table 2 has probabilities of event $Y_{\text{E}} = 0$ over doses and event $Y_{\text{T}} < 63$ over doses that are elicited by clinicians. Specifically, for $\bar{\boldsymbol{\alpha}}_{\text{E}}$ and $\boldsymbol{\rho}$, we simplify our model to fit the historical data and use the posterior distribution to specify their values. For the other location hyperparameters including $\bar{\boldsymbol{\alpha}}_{\text{T}}$, $\bar{\boldsymbol{\beta}}$ and $h_{\text{T}0}$, we take the approach of using pseudo samples in Thall and Nguyen (2012). We simulate datasets using the historical data or the elicited probabilities, fit the model after some simplification, and use the posterior means from the fitted model to specify the values of $\tilde{\boldsymbol{\theta}}$. We then calibrate dispersion parameters to obtain suitable prior uncertainty so that the data will dominate adaptive decisions. The details are

as follows:

1. Calibration of the priors of $\boldsymbol{\beta}_{\mathrm{T}}$, $\boldsymbol{\alpha}^{\star}_{\mathrm{T}}$ and $h_{\mathrm{T}0}$

   We used elicited probabilities of $\mathrm{P}(Y_{\mathrm{T}} < 63 \mid g, d_m)$, given in Table 2(b). The three patient subgroups were collapsed into two subgroups by combining the poor and intermediate subgroups. To simulate pseudo samples, we let $G = 2$ and $M = 5$ and fixed $\boldsymbol{z} = (1, 2)$ as given by the elicited probabilities. We assumed exponential distributions whose rate parameters are matched with $\mathrm{P}(Y_{\mathrm{T}} < 63 \mid g, d_m)$ in Table 2(b), and generated a set of $(y^{o}_{\mathrm{T}}, \delta_{\mathrm{T}})$ with $C = 63$ days. We simulated 9 observations for each (g, d_m), and in total, we have a pseudo sample of size 90. We simplified our model for the toxicity hazard in (1) by removing $\gamma_{i,\mathrm{T}}$, fitted the simplified model to the simulated data and obtained posterior samples of $\boldsymbol{\beta}_{\mathrm{T}}$, $\boldsymbol{\alpha}^{\star}_{\mathrm{T}}$ and $h_{\mathrm{T}0}$. To fit the model, we used the same prior for $\boldsymbol{\beta}_{\mathrm{T}}$, $\boldsymbol{\alpha}^{\star}_{\mathrm{T}}$ and $h_{\mathrm{T}0}$ that we use for our design, and fixed the hyperparameters to imply minimally informative priors; $\bar{\beta}_{\mathrm{T},\ell} = 0$ for all ℓ, $\bar{\alpha}^{\star}_{\mathrm{T},z} = 0$ and $\bar{h}_{0\mathrm{T}} = 0$ and their prior variances to be 25. We repeated this for 20 simulated datasets, and for each dataset, we computed the posterior means of $\boldsymbol{\beta}_{\mathrm{T}}$, $\boldsymbol{\alpha}^{\star}_{\mathrm{T}}$ and $h_{\mathrm{T}0}$. We finally set the means of the posterior means to be the values of $\bar{\beta}_{\mathrm{T},\ell}$, $\bar{\alpha}^{\star}_{\mathrm{T},z}$ and $\bar{h}_{0\mathrm{T}}$. The resulting numerical values are $\bar{\boldsymbol{\beta}}_{\mathrm{T}} = (1.348, -2.824, 1.616)$, $\bar{\lambda} = -6.996$, and $\bar{\alpha}^{\star}_{\mathrm{T},2} = 0.206$. Assuming that difference between latent clusters 2 and 3 is the same as that between clusters 1 and 2, we let $\bar{\alpha}^{\star}_{\mathrm{T},3} = 0.412$. We expressed prior uncertainty by setting $\tau^2_{\mathrm{T},\ell} = 9$ for all ℓ, $\sigma^2_{\mathrm{T}} = 9$ and $v^2_{\mathrm{T},z} = 9$ for all z.

2. Calibration of the priors of $\alpha^{\star}_{\mathrm{E},z}$ and $\rho_{z,k}$

   We used the historical data on $\mathrm{P}(Y_{\mathrm{E}} = k \mid g)$ in Table 1. In the historical data, no new drug was given to the patients, and we set $\eta_{\mathrm{E}} = 0$ under our model. The patients were classified into two subgroups, 'favorable' , 'poor/intermediate' . Similar to the calibration of $\boldsymbol{\alpha}^{\star}_{\mathrm{T}}$, we first set $G = 2$, and $K = 4$, and fixed $\eta_{\mathrm{E}} = 0$ and $\boldsymbol{z} = (1, 2)$

as given in the historical data. We modified our model and fit it to the historical data. Since marginal probabilities are given, we also dropped frailties $\gamma_{i,\mathrm{E}}$. The model simplified for the calibration is

$$\mathrm{P}(Y_{\mathrm{E}} = k \mid x) \;\; = \;\; \Phi(u^{\star}_{z,k+1} \mid \alpha^{\star}_{\mathrm{E},z_x}, \sigma^2_{\pi}) - \Phi(u^{\star}_{z_x,k} \mid \alpha^{\star}_{\mathrm{E},z_x}, \sigma^2_{\pi}),$$

where $\sigma^2_{\pi} = 4$. Similar to our model, we fixed $u^{\star}_{z,0} = -\infty$, $u^{\star}_{z,1} = 0$ and $u^{\star}_{z,4} = \infty$ for all $z = 1, 2$. Recall that $u^{\star}_{z,k} = u^{\star}_{z,(k-1)} + \rho_{z,k}$, $k = 2, 3$. We fit the simplified model to the data in Table 1, and we used their posterior means of $\alpha^{\star}_{\mathrm{E},z}$ and $\rho_{1,k}$ to specify $\bar{\alpha}^{\star}_{\mathrm{E},r}$ and $\rho_{0,k}$. We then placed a reasonable amount of prior uncertainty through dispersion hyperparameters $v^2_{\mathrm{E},z}$ and $\kappa_{z,k}$. Following this strategy, we let $\bar{\alpha}^{\star}_{\mathrm{E},1} = 1.853$, $\bar{\alpha}^{\star}_{\mathrm{E},2} = 0.929$, $\rho_{0,1} = 2.391$ and $\rho_{0,2} = 2.227$. Assuming that the distance of $\bar{\alpha}^{\star}_{\mathrm{E},3}$ from $\bar{\alpha}^{\star}_{\mathrm{E},2}$ is the same as that of $\bar{\alpha}^{\star}_{\mathrm{E},2}$ from $\bar{\alpha}^{\star}_{\mathrm{E},1}$, we let $\bar{\alpha}^{\star}_{\mathrm{E},3} = 0.005$. We reflect prior uncertainty by letting $v^2_{\mathrm{E},r} = 9$, $\kappa_{\rho,1} = 3$, and $\tau_{\kappa,2} = 50$.

3. Calibration of the priors of $\boldsymbol{\beta}_{\mathrm{E}}$

   Elicited probabilities in Table 2(a) is the probabilities of observing PD at dose d_m, and probabilities for the other disease statuses are not provided. We used this to calibrate the priors of $\boldsymbol{\beta}_{\mathrm{E}}$. Utilizing the pseudo sampling approach, we assumed $N = 90$ patients with each of 18 patients assigned to a dose level, and simulated binary $Y'_{i,\mathrm{E}} = 0$ if PD is observed, and $= 1$ otherwise. We simplified the model by letting $\gamma_{i,\mathrm{E}} = 0$, and fit it to the pseudo data,

$$\mathrm{P}(Y'_{\mathrm{E}} = 0 \mid d_m) = \Phi(0 \mid \eta_{\mathrm{E}}(d_m) + \alpha'_{\mathrm{E}}, \sigma^2_{\pi}), \tag{4}$$

   where $\sigma^2_{\pi} = 4$. We used the means of the resulting posterior distribution of $\boldsymbol{\beta}_{\mathrm{E}}$ to specify values of $\bar{\boldsymbol{\beta}}_{\mathrm{E}}$. From the simulation studies, we set $\bar{\boldsymbol{\beta}}_{\mathrm{E}} = (2.923, 1.142, 1.764)$.

We set $\tau^2_{\mathrm{E},\ell} = 25$ for all ℓ, and their priors are weakly informative. We tested the calibrated prior with pseudo data simulated under different scenarios and find that the model with the specified prior captures various patterns of $\mathrm{P}(Y'_{\mathrm{E}} = 0 \mid d_m)$ in d_m.

For clustering of the subgroups, we let $\xi_g = 0.1$, $g = 1, \ldots, G$, implying that a subgroup is collapsed with another subgroup with probability 0.1. In addition, we specified ν and Ω^0 for the prior distribution of the covariance matrix Ω of $\boldsymbol{\gamma}_i$. We let $\Omega^0_{j,j} = 0.10$, and $\Omega^0_{j,j'} = -0.5$ for $j' \neq j$. This specification of Ω^0 a priori implies the negative relationship between Y_{T} and Y_{E}. We let $\nu = 5$ to express weak prior information about Ω. The specified Ω^0 was calibrated through a preliminary simulation study and implies a priori correlation 0.5 or -0.5. In the preliminary study, we compared posterior predictive probabilities of outcomes occurring during the follow-up intervals with the corresponding empirical probabilities based on simulated data, and found that their differences are reasonably small. The resulting numerical values of $\tilde{\boldsymbol{\theta}}$ are summarized in Table 3. The simulation results in §5.2 of the main text show that the specified hyperparameter values in the table works reasonably well even when the simulation truth is very different from the historical data and elicited probabilities in Tables 1 and 2. This suggests that the model with our weakly informative priors adapts observed data well and produces reasonable posterior distributions.

4 Simulation Set-up

In § 5.1 of the main text, we considered eight different scenarios. For each scenario, three subgroups, five doses, and four categorical efficacy outcomes were assumed, i.e., $G = 3$, $M = 5$ and $K = 4$. We simulated the subgroup index of patient i g_i using a multinomial distribution with probability vector $\boldsymbol{p}_g = (0.23, 0.60, 0.17)$. The specification of $\boldsymbol{p}_g$ reflects the subgroup proportions of a similar trial. We set the maximum number of patients in a trial to $N_{\max} = 120$. For each scenario, we first specified the true latent clustering of the

three predefined subgroups, $\boldsymbol{z}^{\text{TR}} = (z_1^{\text{TR}}, z_2^{\text{TR}}, z_3^{\text{TR}})$. We specified the covariance matrix Ω^{TR} for the frailty vectors, and simulated frailties $\boldsymbol{\gamma}_i^{\text{TR}} \overset{iid}{\sim} \text{N}_2(\mathbf{0}, \Omega^{\text{TR}})$. We assumed $\Omega_{j,j}^{\text{TR}} = 0.01$, $j = 1, 2$, and $\Omega_{j,j'}^{\text{TR}} = -0.003$, $j \neq j'$ for Ω^{TR} for all scenarios. To simulate time to toxicity outcome $Y_{i,\text{T}}$, we specified marginal probabilities of toxicity occurring during the follow-up period $p_{\text{T}}^{\text{TR}}(g, d_m) = \text{P}(Y_{\text{T}} < C \mid g, d_m)$, for $g = 1$ and $d_m \in \{d_1, \ldots, d_5\}$. Assuming a constant hazard of toxicity, we find the hazard of toxicity of d_m for subgroup 1, $h_{0\text{T}}^{\text{TR}}(d_m) = -\log(1-p_{\text{T}}^{\text{TR}}(g, d_m))/C$, $g = 1$. To specify the hazard for subgroups 2 and 3, we set $\alpha_{\text{T},1}^{\star,\text{TR}} = 0$ and specify $\alpha_{\text{T},z}^{\star,\text{TR}} > \alpha_{\text{T},(z-1)}^{\star,\text{TR}}$, $z = 2, \ldots, R^{\text{TR}}$ if $R^{\text{TR}} > 1$. Given g_i, $d_{[i]}$ and $\boldsymbol{\gamma}_i^{\text{TR}}$, we generated $Y_{i,\text{T}}$ from the exponential distribution with specified hazard $h_{\text{T},i}^{\text{TR}} = h_{0\text{T}}^{\text{TR}}(d_m) + \alpha_{\text{T},z_{g_i}}^{\star,\text{TR}} + \gamma_{i,\text{T}}^{\text{TR}}$. We next generated $Y_{i,\text{E}}$ as follows; similar to $p_{\text{T}}^{\text{TR}}(g, d_m)$, we first specified the quantiles of the doses for subgroup 1, $\eta_{\text{E}}^{\text{TR}}(d_m)$, $m = 1, \ldots, M$ and $\alpha_{\text{E},z}^{\star,\text{TR}} < \alpha_{\text{E},(z-1)}^{\star,\text{TR}}$, $z = 2, \ldots, R^{\text{TR}}$ if $R^{\text{TR}} > 1$, with $\alpha_{\text{E},1}^{\star,\text{TR}} = 0$. We also specified $\sigma_\pi^{2,\text{TR}}$ and $u_{z,k}^{\star,\text{TR}}$, $z = 1, \ldots, R$ and $k = 2, 3$, while fixing $u_{z,0}^{\star,\text{TR}} = -\infty$, $u_{z,1}^{\star,\text{TR}} = 0$ and $u_{z,4}^{\star,\text{TR}} = \infty$ for all z. We then let

$$\begin{aligned}
\pi_{i,k}^{\text{TR}} &= \Phi\left(u_{z_{g_i}^{\text{TR}},k+1}^{\star,\text{TR}} \mid \eta_{\text{E}}^{\text{TR}}(d_{[i]}) + \alpha_{\text{E},z_{g_i}^{\text{TR}}}^{\star,\text{TR}} + \gamma_{i,\text{E}}^{\text{TR}},\ \sigma_\pi^{2,\text{TR}}\right) \\
&\quad -\Phi\left(u_{z_{g_i}^{\text{TR}},k}^{\star,\text{TR}} \mid \eta_{\text{E}}^{\text{TR}}(d_{[i]}) + \alpha_{\text{E},z_{g_i}^{\text{TR}}}^{\star,\text{TR}} + \gamma_{i,\text{E}}^{\text{TR}},\ \sigma_\pi^{2,\text{TR}}\right), \quad \text{for } k = 1, \ldots, K-1.
\end{aligned}$$

We finally drew $Y_{i,\text{E}}$ from $\{0, \ldots, K-1\}$ with probability $(\pi_{i,0}^{\text{TR}}, \ldots, \pi_{i,(K-1)}^{\text{TR}})$. We assumed $e_i - e_{i-1} \mid e_{i-1}, a_{\text{acc}} \overset{iid}{\sim} \text{Exp}(a_{\text{acc}}/30)$ with $e_0 = 0$ and $a_{\text{acc}} \overset{iid}{\sim} \text{Unif}(1, 4)$. This lets the number of new patients treated every 30 days range between 1 and 4, on average. Table 4 show the assumed values of $\boldsymbol{z}^{\text{TR}}$, $\alpha_{j,x}^{\text{TR}}$ and $u_{g,k}^{\text{TR}}$ for all scenarios. Table 6 gives the true probabilities of observing severe toxicity during the follow-up period and of observing PD given the frailty equals zero,

$$\begin{aligned}
p_{\text{T}}^{\text{TR}}(g, d_m) &= 1 - \exp\left[-C\left\{h_{0\text{T}}^{\text{TR}}(d_m) + \alpha_{\text{T},z_g}^{\star,\text{TR}}\right\}\right], \text{ and} \\
\pi_k^{\text{TR}}(g, d_m) &= \Phi(u_{z_g^{\text{TR}},k+1}^{\star,\text{TR}} \mid \eta_{\text{E}}^{\text{TR}}(d_m) + \alpha_{\text{E},z_g^{\text{TR}}}^{\star,\text{TR}},\ \sigma_\pi^{2,\text{TR}}) - \Phi(u_{z_g^{\text{TR}},k}^{\star,\text{TR}} \mid \eta_{\text{E}}^{\text{TR}}(d_m) + \alpha_{\text{E},z_g^{\text{TR}}}^{\star,\text{TR}},\ \sigma_\pi^{2,\text{TR}}),
\end{aligned}$$

with $k = 0$ for each scenario. Truly unacceptable doses and truly optimal doses are marked with red and blue, respectively. Recall the elicited thresholds $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all g, $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35, 0.35$ for $g = 1, 2, 3$. The expected utility under the truth $U^{\text{TR}}(g, d_m)$ are also given in the table. Table 5 gives the true probabilities of observing a disease status given the frailty equals zero, $\pi_k^{\text{TR}}(g, d_m)$ for all k, d_m and g.

5 Additional Simulation Results

Simulation results with $\boldsymbol{p}_x = (1/3, 1/3, 1/3)$ are summarized in Table 6 using $p^{\text{sel}}(g, d_m)$, $p^{\text{unacc}}(g, d_m)$ and $n^{\text{ptrt}}(g, d_m)$ under the three designs. It also includes the simulation truth to facilitate evaluation. Comparisons of D-Sub to the comparators, D-Comb and D-Sep, are also summarized in Figures 1 and 2.

As another comparator, we also evaluated a more conventional utility based design where $U_1(\boldsymbol{Y}) = U_2(\boldsymbol{Y}) = U_3(\boldsymbol{Y}) = U(\boldsymbol{Y})$, so that the same utility was used for all subgroups, and re-ran the simulations. This reflects what would be done by most utility-based designs, and here the truly optimal dose in each subgroup under each scenario may be different than the truly optimal dose using the subgroup-specific utilities. We let $U_2(\boldsymbol{Y})$ in Figure 1(c) of the main text be the common utility, while keeping the remaining simulation design the same. The results for all designs are summarized in Supplementary Table 7. Due to the change of the utility function for subgroups 1 and 3, the patterns of U^{TR} over d_m are changed for those subgroups and the true optimal doses are changed (e.g., subgroup 1 in Scenario 5). Since the true probability distributions of $\boldsymbol{Y}$ are subgroup-specific, subgroup-specific decision making is still more desirable, especially in Scenarios 1 and 6. While D-Sub outperforms D-Sep in most of the scenarios, D-Sub and D-Comb perform very similarly in many scenarios, especially in Scenarios 7 and 8 having $\boldsymbol{z}^{\text{TR}} = (1, 1, 1)$. However, in Scenarios 1 and 6, D-Sub performs the best by a wide margin under all criteria.

To compare D-Sub to a version of the design that does not induce clustering of subgroups, 'D-w/o clustering', we let $\boldsymbol{z} = (1, 2, 3)$ and assumed that each subgroup has its own α_{T} and α_{E}. The remaining model structure is the same as the model for D-Sub. The simulation summary, given in Table 11, indicates that the performance of D-w/o clustering is worse, especially for small subgroups, subgroups 1 and 3 in Scenarios 4 and 5. For the other scenarios, the performance of the two designs is similar.

To examine robustness of D-Sub, we assumed that the true hazard function of the time-to-toxicity outcome is not constant, while keeping the marginal true probabilities, $p_{\mathrm{T}}^{\mathrm{TR}}(g, d_m)$ and $\pi_k^{\mathrm{TR}}(g, d_m)$, the same. We evaluated the design for some selected scenarios, Scenarios 1, 3, 5, and 7, generating toxicity times from Weibull distributions with shape parameters $\zeta^{\mathrm{TR}} = 0.5$ and 2, where the hazard function $h_{0\mathrm{T}}^{\mathrm{TR}}(t \mid d_m) = h_{0\mathrm{T}}^{\mathrm{TR}}(d_m)\zeta^{\mathrm{TR}} t^{\zeta^{\mathrm{TR}}-1}$ for decreasing and increasing hazards, respectively. Recall that D-Sub assumes that the hazard function is constant. The simulation results are summarized in Table 8. Due to the changes in $h_{0\mathrm{T}}^{\mathrm{TR}}(t \mid d_m)$, $U^{\mathrm{TR}}(g, d_m)$ changes, and their patterns over d_m within g also change, as shown in the top two rows of the table. D-Sub's performance is slightly affected, but it still performs well.

We examined the performance of D-Sub with different maximum numbers of patients in a trial, $N_{\max} = 60$ and 180, under Scenarios 1, 3, 5, and 7. The results are summarized in Tables 9 and 10 for $N_{\max} = 60$ and 180, respectively. Comparing these tables to Table 1 of the main text, all three designs, D-Sub, D-Comb and D-Sep, all perform worse for the smaller $N_{\max} = 60$, and their performance improves as $N_{\max}$ increases. However, for each $N_{\max}$, D-Sub performs far better than the two comparators.

We compared the performance of D-Sub to that of a model that does not include patient-specific frailties $\boldsymbol{\gamma}_i$ ('D w/o frailty'). The results are summarized in Table 12. To facilitate comparison, the results from D-Sub are included in the table. While $p^{\mathrm{sel}}(g, d_m)$ are similar under both designs, $p^{\mathrm{unacc}}(g, d_m)$ and n_m^{ptrt} are greatly in favor of D-Sub. 'D w/o frailty'

is far less likely to identify truly unacceptable doses as unacceptable, and thus it assigns patients to truly unacceptable doses and selects truly unacceptable doses as optimal doses more often. In Scenario 1, where all doses are truly unacceptable for patients in subgroup 3, 'D w/o frailty' assigns patients in subgroup 3 to treatments and chooses dose 1 as optimal for subgroup 3 with high frequencies, while under D-Sub, patients in subgroup 3 do not receive a treatment and no dose is chosen as optimal most of the time.

To examine how a different adaptive randomization affects the performance of D-Sub, we considered a design that assigns a patient in subgroup g to dose $d_m \in \mathcal{A}(g,t)$ with probability proportional to $u_g(d_m \mid g, \mathcal{D}_{n(t)})$, where $\mathcal{A}(g,t)$ is the set of acceptable doses at time t for subgroup g and $u_g(d_m \mid g, \mathcal{D}_{n(t)})$ is the posterior expected utility of subgroup g at time t, and compared its performance to that of D-Sub. Recall that for D-Sub, a patient in subgroup g is assigned to dose $d_m \in \mathcal{A}(g,t)$ with probability proportional to $1/\{n_{g,m}(t)+1\}$, where $n_{g,m}(t)$ is the number of patients in subgroup g treated at dose d_m up to time t. The results in Table 13 show that any change in the performance of D-Sub with the different AR method is very small.

We considered a design that assumes a simpler linear regression model for the toxicity outcome, $\eta'_{\mathrm{T}}(d_m) = \beta'_{\mathrm{T}} d_m$ (D-Linear tox), and compared its performance to that of D-Sub. A normal distribution truncated below at 0 is used as a prior for β'_{T}. Recall that D-Sub uses the nonlinear function $\eta_{\mathrm{T}}(d_m) = \frac{\beta_{\mathrm{T3}}}{1+\exp\{-\beta_{\mathrm{T1}} \times (10 \times d_m - \beta_{\mathrm{T2}})\}}$. We compared D-Sub to D-Linear tox in Scenarios 1, 3, 5, and 7. The results are summarized in Table 14. D-Linear tox performs substantially worse than D-Sub for all criteria $p^{\mathrm{sel}}(g,d_m)$, $p^{\mathrm{unacc}}(g,d_m)$, and n_m^{ptrt}, in Scenarios 3, 5, and 7, with $p^{\mathrm{sel}}(g,d_m)$ much worse under D-Linear tox for the three scenarios. In Scenario 1, the performance of D-Linear tox is similar to that of D-Sub.

References

Motzer, R. J., Rini, B. I., McDermott, D. F., Frontera, O. A., Hammers, H. J., Carducci, M. A., Salman, P., Escudier, B., Beuselinck, B., Amin, A., *et al.* (2019). Nivolumab plus ipilimumab versus sunitinib in first-line treatment for advanced renal cell carcinoma: extended follow-up of efficacy and safety results from a randomised, controlled, phase 3 trial. *The Lancet Oncology* **20**, 10, 1370–1385.

Roberts, G. O. and Rosenthal, J. S. (2009). Examples of adaptive mcmc. *Journal of Computational and Graphical Statistics* **18**, 2, 349–367.

Thall, P. F. and Nguyen, H. Q. (2012). Adaptive randomization to improve utility-based dose-finding with bivariate ordinal outcomes. *Journal of biopharmaceutical statistics* **22**, 4, 785–801.

	PD	SD	PR	CR	total
$g = 1$	21	55	39	10	125
$g \geq 2$	137	110	130	48	425

Table 1: Historical Data of $\mathrm{P}(Y_{\mathrm{E}} = k \mid g)$ Obtained from Motzer *et al.* (2019)

g	Doses				
	d_1	d_2	d_3	d_4	d_5
$g \geq 2$	0.30	0.26	0.22	0.18	0.12

(a) Probabilities of $Y_{\mathrm{E}} = 0$ given $g \geq 2$ and d_m elicited by clinicians

x	Doses				
	d_1	d_2	d_3	d_4	d_5
$g = 1$	0.05	0.10	0.15	0.20	0.25
$g \geq 2$	0.10	0.15	0.20	0.25	0.30

(b) Probabilities of $Y_{\mathrm{T}} < 63$ given g and d_m elicited by clinicians

Table 2: Elicited probabilities of $Y_{\mathrm{E}} = 0$ and $Y_{\mathrm{T}} < 63$ over subgroup g and dose d_m

Hyperparameter	Elicited Numerical Values
$\bar{\beta}_{\mathrm{T},\ell}$	1.348, -2.824, and 1.616 for $\ell = 1, 2, 3$, and $\tau^2_{\mathrm{T},\ell} = 25$ for all ℓ
$\bar{\beta}_{\mathrm{E},\ell}$	2.923, 1.142, and 1.764 for $\ell = 1, 2, 3$, and $\tau^2_{\mathrm{E},\ell} = 25$ for all ℓ
$\bar{\alpha}^{\star}_{\mathrm{T},z}$	0.206, and 0.412, with $v^2_{\mathrm{T},z} = 9$, for $z = 2, 3$
$\bar{\alpha}^{\star}_{\mathrm{E},z}$	1.853, 0.929, and 0.005, with $v^2_{\mathrm{E},z} = 9$ for $z = 1, 2, 3$
$\rho_{0,k}$	2.391, and 2.227, for $k = 2$, 3
κ_k	30, and 50, for $k = 2$, 3
Ω	$\Omega_{j,j} = 0.1$ and $\Omega_{j,j'} = -0.05$ with $j \neq j'$
σ^2_{π}	4
ξ_g	0.1 for all g

Table 3: [Prior Elicitation] Elicited hyperparameter values.

$\boldsymbol{z}^{\mathrm{TR}}$		$\alpha_{\mathrm{T},g}^{\mathrm{TR}}$	$\alpha_{\mathrm{E},g}^{\mathrm{TR}}$	$u_{g,2}^{\mathrm{TR}}$	$u_{g,3}^{\mathrm{TR}}$	$\alpha_{\mathrm{T},g}^{\mathrm{TR}}$	$\alpha_{\mathrm{E},g}^{\mathrm{TR}}$	$u_{g,2}^{\mathrm{TR}}$	$u_{g,3}^{\mathrm{TR}}$
		Scenario 1				Scenario 2			
$(1,2,3)$	$g=1$	0.0	0.0	4.0	8.0	0.0	0.0	5.0	7.0
	$g=2$	0.4	1.0	3.0	6.0	0.3	0.5	3.0	7.0
	$g=3$	0.8	3.5	4.0	7.0	0.5	1.0	3.0	7.0
		Scenario 3				Scenario 4			
$(1,1,2)$	$g=1$	0.0	0.0	3.0	7.0	0.0	0.0	5.5	8.5
	$g=2$	0.0	0.0	3.0	7.0	0.0	0.0	5.5	8.5
	$g=3$	1.0	1.5	3.0	6.0	0.0	2.3	3.0	6.0
		Scenario 5				Scenario 6			
$(1,2,2)$	$g=1$	0.0	0.0	5.5	8.0	0.0	0.0	4.0	8.0
	$g=2$	0.0	0.5	3.0	7.0	1.5	0.5	5.0	9.0
	$g=3$	0.0	0.5	3.0	7.0	1.5	0.5	5.0	9.0
		Scenario 7				Scenario 8			
$(1,1,1)$	$g=1$	0.0	0.0	5.0	7.0	0.0	0.0	6.0	8.0
	$g=2$	0.0	0.0	5.0	7.0	0.0	0.0	6.0	8.0
	$g=3$	0.0	0.0	5.0	7.0	0.0	0.0	6.0	8.0

Table 4: [Simulation Truth] Specified values of $\alpha_{j,g}^{\mathrm{TR}}$, $j = E$ or T, $g = 1, 2, 3$, $u_{g,k}^{\mathrm{TR}}$, $k = 2, 3$ for the eight simulation scenarios.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 1					Scenario 2				
$g=1$	PD	0.16	0.16	0.12	0.09	0.07	0.50	0.37	0.25	0.20	0.16
	SD	0.47	0.47	0.44	0.41	0.37	0.45	0.54	0.59	0.60	0.59
	PR	0.32	0.32	0.37	0.41	0.44	0.04	0.07	0.11	0.14	0.16
	CR	0.05	0.05	0.07	0.09	0.12	0.01	0.02	0.05	0.07	0.09
$g=2$	PD	0.25	0.25	0.20	0.16	0.12	0.57	0.43	0.31	0.25	0.20
	SD	0.38	0.38	0.36	0.34	0.31	0.31	0.36	0.38	0.38	0.36
	PR	0.28	0.28	0.31	0.34	0.36	0.12	0.19	0.28	0.32	0.37
	CR	0.09	0.09	0.12	0.16	0.20	0.01	0.02	0.03	0.05	0.07
$g=3$	PD	0.57	0.57	0.50	0.43	0.37	0.63	0.50	0.37	0.31	0.25
	SD	0.37	0.37	0.41	0.44	0.47	0.28	0.34	0.38	0.38	0.38
	PR	0.06	0.06	0.08	0.11	0.14	0.09	0.15	0.23	0.28	0.32
	CR	0.01	0.01	0.01	0.02	0.02	0.00	0.01	0.02	0.03	0.05
		Scenario 3					Scenario 4				
$g=1$	PD	0.16	0.09	0.03	0.02	0.02	0.16	0.12	0.09	0.09	0.09
	SD	0.34	0.28	0.17	0.14	0.14	0.64	0.63	0.60	0.60	0.60
	PR	0.41	0.47	0.49	0.47	0.47	0.17	0.20	0.24	0.24	0.24
	CR	0.09	0.16	0.31	0.37	0.37	0.03	0.05	0.07	0.07	0.07
$g=2$	PD	0.16	0.09	0.03	0.02	0.02	0.16	0.12	0.09	0.09	0.09
	SD	0.34	0.28	0.17	0.14	0.14	0.64	0.63	0.60	0.60	0.60
	PR	0.41	0.47	0.49	0.47	0.47	0.17	0.20	0.24	0.24	0.24
	CR	0.09	0.16	0.31	0.37	0.37	0.03	0.05	0.07	0.07	0.07
$g=3$	PD	0.31	0.20	0.09	0.07	0.07	0.41	0.34	0.29	0.29	0.29
	SD	0.38	0.36	0.28	0.24	0.24	0.37	0.38	0.38	0.38	0.38
	PR	0.24	0.31	0.38	0.38	0.38	0.18	0.22	0.26	0.26	0.26
	CR	0.07	0.12	0.25	0.31	0.31	0.04	0.05	0.08	0.08	0.08
		Scenario 5					Scenario 6				
$g=1$	PD	0.12	0.12	0.12	0.07	0.04	0.16	0.16	0.16	0.05	0.05
	SD	0.63	0.63	0.63	0.56	0.49	0.47	0.47	0.47	0.32	0.32
	PR	0.19	0.19	0.19	0.25	0.29	0.32	0.32	0.32	0.47	0.47
	CR	0.07	0.07	0.07	0.12	0.18	0.05	0.05	0.05	0.16	0.16
$g=2$	PD	0.16	0.16	0.16	0.09	0.05	0.20	0.20	0.20	0.07	0.07
	SD	0.34	0.34	0.34	0.28	0.22	0.60	0.60	0.60	0.50	0.50
	PR	0.41	0.41	0.41	0.47	0.49	0.19	0.19	0.19	0.37	0.37
	CR	0.09	0.09	0.09	0.16	0.23	0.02	0.02	0.02	0.07	0.07
$g=3$	PD	0.16	0.16	0.16	0.09	0.05	0.20	0.20	0.20	0.07	0.07
	SD	0.34	0.34	0.34	0.28	0.22	0.60	0.60	0.60	0.50	0.50
	PR	0.41	0.41	0.41	0.47	0.49	0.19	0.19	0.19	0.37	0.37
	CR	0.09	0.09	0.09	0.16	0.23	0.02	0.02	0.02	0.07	0.07

Table 5: [Simulation Truth] The true probabilities of ordinal disease status outcomes with the frailty being zero, $\pi_k^{\text{TR}}(g, d_m)$ with $\gamma_{\text{E}} = 0$, $k = 0, 1, 2, 3, 4$ for PD, SD, PR and CR, respectively.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 7					Scenario 8				
$g=1$	PD	0.54	0.46	0.31	0.16	0.09	0.31	0.16	0.12	0.12	0.12
	SD	0.42	0.48	0.57	0.59	0.54	0.62	0.68	0.68	0.68	0.68
	PR	0.03	0.05	0.09	0.16	0.21	0.05	0.11	0.14	0.14	0.14
	CR	0.01	0.01	0.03	0.09	0.16	0.02	0.05	0.07	0.07	0.07
$g=2$	PD	0.54	0.46	0.31	0.16	0.09	0.31	0.16	0.12	0.12	0.12
	SD	0.42	0.48	0.57	0.59	0.54	0.62	0.68	0.68	0.68	0.68
	PR	0.03	0.05	0.09	0.16	0.21	0.05	0.11	0.14	0.14	0.14
	CR	0.01	0.01	0.03	0.09	0.16	0.02	0.05	0.07	0.07	0.07
$g=3$	PD	0.54	0.46	0.31	0.16	0.09	0.31	0.16	0.12	0.12	0.12
	SD	0.42	0.48	0.57	0.59	0.54	0.62	0.68	0.68	0.68	0.68
	PR	0.03	0.05	0.09	0.16	0.21	0.05	0.11	0.14	0.14	0.14
	CR	0.01	0.01	0.03	0.09	0.16	0.02	0.05	0.07	0.07	0.07

Table 5 continued: The true probabilities of ordinal disease status outcomes with the frailty being zero, $\pi_k^{\text{TR}}(g, m)$ with $\gamma_{\text{E}} = 0$, $k = 0, 1, 2, 3, 4$ for PD, SD, PR and CR, respectively.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 1					Scenario 2				
u^{TR}	$g=1$	**56.89**	51.38	47.69	43.81	39.46	34.34	35.57	34.21	32.38	32.45
	$g=2$	**58.17**	53.25	51.91	49.83	47.67	35.85	38.94	40.73	40.27	41.70
	$g=3$	39.85	35.41	33.63	32.13	30.50	34.59	38.40	41.17	41.94	43.86
p_{T}^{TR}	$g=1$	0.05	0.20	0.35	*0.50*	*0.65*	0.25	0.35	*0.50*	*0.60*	*0.65*
	$g=2$	0.07	0.28	*0.47*	*0.64*	*0.79*	0.32	*0.44*	*0.61*	*0.71*	*0.76*
	$g=3$	0.11	0.39	*0.61*	*0.78*	*0.90*	0.38	*0.51*	*0.68*	*0.78*	*0.82*
π_0^{TR}	$g=1$	0.16	0.16	0.12	0.09	0.07	*0.50*	*0.37*	*0.25*	*0.20*	0.16
	$g=2$	0.25	0.25	0.20	0.16	0.12	*0.57*	*0.43*	0.31	0.25	0.20
	$g=3$	*0.57*	*0.56*	*0.50*	*0.43*	*0.37*	*0.63*	*0.50*	*0.37*	0.31	0.25

- $p_m^{\text{sel}}(x)$

D-Sub	$g=1$	**0.80**	0.09	0.05	0.02	0.02	0.01	0.01	0.02	0.00	0.01
	$g=2$	**0.76**	0.11	0.07	0.02	0.02	0.03	0.03	0.02	0.00	0.00
	$g=3$	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D-Comb	all g	*0.52*	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D-Sep	$g=1$	**0.66**	0.13	0.09	0.02	0.06	0.02	0.01	0.03	0.01	0.02
	$g=2$	**0.72**	0.09	0.09	0.03	0.05	0.06	0.06	0.04	0.01	0.01
	$g=3$	0.03	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00

- $p_m^{\text{Accp}}(x)$

D-Sub	$g=1$	0.06	0.06	0.28	*0.59*	*0.67*	*0.99*	*0.98*	*0.96*	*0.98*	*0.99*
	$g=2$	0.03	0.06	*0.44*	*0.86*	*0.91*	*0.94*	*0.93*	*0.97*	*1.00*	*1.00*
	$g=3$	*0.94*	*0.97*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*
D-Comb	all g	*0.46*	*0.50*	*0.87*	*0.99*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*
D-Sep	$g=1$	0.07	0.08	0.23	*0.48*	*0.59*	*0.98*	*0.96*	*0.94*	*0.96*	*0.97*
	$g=2$	0.02	0.08	*0.46*	*0.83*	*0.90*	*0.89*	*0.87*	*0.93*	*0.98*	*0.99*
	$g=3$	*0.97*	*0.98*	*0.99*	*1.00*	*1.00*	*0.99*	*0.98*	*0.99*	*1.00*	*1.00*

- $n_m^{\text{ptrt}}(x)$

D-Sub	$g=1$	10.48	10.48	7.98	*5.32*	*4.67*	*3.02*	*3.80*	*3.83*	*2.66*	*2.39*
	$g=2$	13.44	12.69	*7.66*	*3.10*	*2.23*	*5.20*	*5.12*	*2.98*	*1.20*	*0.73*
	$g=3$	*5.85*	*4.47*	*2.05*	*0.66*	*0.31*	*3.50*	*2.93*	*1.26*	*0.42*	*0.21*
D-Comb	$g=1$	11.72	10.10	4.42	*1.56*	*1.19*	*3.33*	*2.96*	*1.87*	*0.87*	*0.69*
	$g=2$	11.63	9.86	*4.39*	*1.56*	*1.25*	*3.37*	*3.02*	*1.80*	*0.88*	*0.68*
	$g=3$	11.83	9.83	*4.47*	*1.50*	*1.28*	*3.28*	*3.00*	*1.83*	*0.90*	*0.67*
D-Sep	$g=1$	10.26	10.05	8.26	*5.88*	*5.12*	*8.02*	*7.50*	*5.70*	*3.81*	*3.30*
	$g=2$	13.13	11.89	*7.70*	*3.93*	*3.03*	*10.36*	*8.80*	*5.01*	*2.48*	*1.83*
	$g=3$	*10.54*	*8.72*	*4.70*	*1.95*	*1.38*	*9.77*	*7.56*	*4.00*	*1.69*	*1.18*

Table 6: [**Simulation Results with $\boldsymbol{p}_g = (1/3, 1/3, 1/3)$**]. $p^{\text{unacc}}(g,m)$ = P(declare dose d_m unacceptable for subgroup g), $p^{\text{sel}}(g,m)$ = P(select dose d_m as optimal for subgroup g), and n_m^{ptrt} = mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g- = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 3					Scenario 4				
u^{TR}	$g=1$	59.95	64.84	**72.19**	64.57	58.12	**54.07**	48.75	48.63	40.37	36.04
	$g=2$	64.96	70.83	**79.05**	74.63	70.18	**56.47**	54.15	55.52	49.64	46.55
	$g=3$	56.53	63.35	**73.50**	65.89	60.18	51.47	52.05	**54.98**	50.88	48.50
p_{T}^{TR}	$g=1$	0.05	0.08	0.10	0.35	*0.50*	0.05	0.25	0.30	*0.50*	*0.60*
	$g=2$	0.05	0.08	0.10	0.35	*0.50*	0.05	0.25	0.30	*0.50*	*0.60*
	$g=3$	0.13	0.20	0.25	*0.69*	*0.85*	0.05	0.25	0.30	*0.50*	*0.60*
π_0^{TR}	$g=1$	0.16	0.09	0.03	0.02	0.02	0.16	0.12	0.09	0.09	0.09
	$g=2$	0.16	0.09	0.03	0.02	0.02	0.16	0.12	0.09	0.09	0.09
	$g=3$	0.31	0.20	0.09	0.07	0.07	*0.41*	0.34	0.29	0.29	0.29

- $p_m^{\text{sel}}(x)$

D-Sub	$g=1$	0.01	0.03	**0.90**	0.03	0.03	**0.73**	0.15	0.11	0.00	0.00
	$g=2$	0.01	0.02	**0.86**	0.06	0.05	**0.57**	0.16	0.21	0.02	0.04
	$g=3$	0.01	0.04	**0.91**	0.02	0.01	0.39	0.16	**0.23**	0.03	0.06
D-Comb	all g	0.01	0.04	**0.85**	0.07	0.03	**0.65**	0.13	**0.15**	0.02	0.04
D-Sep	$g=1$	0.06	0.12	**0.54**	0.14	0.15	**0.68**	0.13	0.13	0.02	0.02
	$g=2$	0.01	0.06	**0.53**	0.16	0.24	**0.46**	0.17	0.19	0.03	0.15
	$g=3$	0.08	0.11	**0.73**	0.05	0.03	0.16	0.12	**0.20**	0.07	0.31

- $p_m^{\text{Accp}}(x)$

D-Sub	$g=1$	0.01	0.00	0.00	0.11	*0.24*	0.03	0.02	0.07	*0.29*	*0.44*
	$g=2$	0.01	0.00	0.00	0.17	*0.35*	0.01	0.01	0.07	*0.34*	*0.50*
	$g=3$	0.05	0.03	0.04	*0.85*	*0.96*	*0.19*	0.20	0.30	*0.71*	*0.80*
D-Comb	all g	0.01	0.00	0.01	*0.49*	*0.79*	*0.04*	0.05	0.13	*0.44*	*0.58*
D-Sep	$g=1$	0.01	0.00	0.01	0.13	*0.23*	0.05	0.06	0.17	0.39	*0.49*
	$g=2$	0.00	0.00	0.01	0.12	*0.22*	0.00	0.02	0.10	0.34	*0.44*
	$g=3$	0.03	0.01	0.07	*0.81*	*0.95*	*0.29*	0.24	0.26	0.48	*0.55*

- $n_m^{\text{ptrt}}(x)$

D-Sub	$g=1$	8.40	8.60	8.51	7.74	*6.81*	9.17	9.18	8.60	*6.84*	*5.95*
	$g=2$	8.79	8.88	8.84	7.27	*6.11*	9.75	9.70	8.85	*6.27*	*5.24*
	$g=3$	10.81	11.19	11.07	*3.80*	*2.07*	*10.48*	10.00	8.40	*4.16*	*3.20*
D-Comb	$g=1$	10.19	10.02	9.93	5.84	*4.11*	10.45	9.80	8.48	*5.78*	*4.92*
	$g=2$	10.14	10.16	9.79	6.05	*3.98*	10.30	9.71	8.72	*5.87*	*4.89*
	$g=3$	10.08	9.98	9.63	*6.07*	*3.92*	10.31	9.68	8.36	*5.88*	*4.98*
D-Sep	$g=1$	8.59	8.77	8.47	7.34	*6.84*	10.19	9.69	8.17	*6.12*	*5.39*
	$g=2$	8.52	8.58	8.39	7.57	*6.91*	10.11	9.53	8.45	*6.34*	*5.58*
	$g=3$	11.28	11.22	10.25	*4.43*	*2.77*	*9.23*	9.13	8.31	*6.26*	*5.64*

Table 6 continued:[**Simulation Results with $\boldsymbol{p}_g = (1/3, 1/3, 1/3)$**]. $p^{\text{unacc}}(g,m)$ = P(declare dose d_m unacceptable for subgroup g), $p^{\text{sel}}(g,m)$ = P(select dose d_m as optimal for subgroup g), and n_m^{ptrt} = mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g- = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 5					Scenario 6				
u^{TR}	$g=1$	**57.64**	48.14	47.08	49.66	52.30	57.62	55.12	53.76	**63.17**	56.94
	$g=2$	65.44	59.25	58.92	63.74	**67.67**	**52.93**	46.64	44.45	52.85	42.68
	$g=3$	70.05	65.27	64.67	70.81	**75.03**	**55.65**	50.80	48.84	59.80	51.27
p_{T}^{TR}	$g=1$	0.03	0.28	0.30	0.35	0.38	0.03	0.10	0.13	0.15	0.30
	$g=2$	0.03	0.28	0.30	0.35	0.38	0.13	0.38	*0.46*	*0.52*	*0.80*
	$g=3$	0.03	0.28	0.30	0.35	0.38	0.13	0.37	*0.46*	*0.52*	*0.80*
π_0^{TR}	$g=1$	0.12	0.12	0.12	0.07	0.04	0.16	0.16	0.16	0.05	0.05
	$g=2$	0.16	0.16	0.16	0.09	0.06	0.20	0.20	0.20	0.07	0.07
	$g=3$	0.16	0.16	0.16	0.09	0.05	0.20	0.20	0.20	0.07	0.07

- $p_m^{\text{sel}}(x)$

D-Sub	$g=1$	**0.62**	0.01	0.01	0.01	0.36	0.20	0.01	0.03	**0.32**	0.45
	$g=2$	0.26	0.01	0.01	0.02	**0.70**	**0.73**	0.02	0.02	0.18	0.04
	$g=3$	0.16	0.01	0.01	0.02	**0.81**	**0.68**	0.03	0.03	0.16	0.03
D-Comb	all g	**0.27**	0.01	0.01	0.02	**0.70**	**0.43**	0.02	0.03	**0.32**	0.19
D-Sep	$g=1$	**0.56**	0.04	0.03	0.03	0.34	0.09	0.05	0.05	**0.24**	0.58
	$g=2$	0.20	0.03	0.03	0.04	**0.70**	**0.69**	0.05	0.05	0.13	0.07
	$g=3$	0.11	0.02	0.03	0.04	**0.81**	**0.59**	0.06	0.07	0.15	0.11

- $p_m^{\text{Accp}}(x)$

D-Sub	$g=1$	0.00	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.01	0.02
	$g=2$	0.00	0.01	0.03	0.04	0.05	0.01	0.11	*0.42*	*0.69*	*0.87*
	$g=3$	0.01	0.03	0.06	0.07	0.08	0.08	0.20	*0.53*	*0.79*	*0.94*
D-Comb	all g	0.00	0.01	0.04	0.05	0.06	0.02	0.04	*0.15*	*0.33*	*0.57*
D-Sep	$g=1$	0.02	0.04	0.08	0.12	0.13	0.04	0.03	0.02	0.03	0.05
	$g=2$	0.00	0.03	0.06	0.09	0.11	0.01	0.10	*0.40*	*0.68*	*0.85*
	$g=3$	0.00	0.02	0.05	0.09	0.11	0.03	0.13	*0.40*	*0.69*	*0.85*

- $n_m^{\text{ptrt}}(x)$

D-Sub	$g=1$	8.22	8.19	8.02	7.83	7.79	7.96	8.12	8.00	7.98	7.84
	$g=2$	8.50	8.22	7.87	7.61	7.59	12.69	10.86	*7.63*	*5.05*	*3.51*
	$g=3$	9.26	8.43	7.59	7.12	7.05	13.60	10.86	*6.60*	*3.47*	*1.85*
D-Comb	$g=1$	8.75	8.55	7.73	7.53	7.38	10.58	9.91	8.03	6.19	4.72
	$g=2$	8.89	8.47	7.88	7.51	7.44	10.70	9.84	*8.33*	*6.38*	*4.71*
	$g=3$	8.95	8.29	7.68	7.56	7.40	10.71	9.74	*7.88*	*6.19*	*4.76*
D-Sep	$g=1$	9.17	8.56	7.80	7.29	7.10	8.13	8.10	8.01	7.96	7.79
	$g=2$	9.18	8.49	7.91	7.29	7.11	12.99	11.22	*7.56*	*4.60*	*3.26*
	$g=3$	8.98	8.43	7.97	7.42	7.21	12.84	11.00	*7.58*	*4.71*	*3.31*

Table 6 continued:[**Simulation Results with $\boldsymbol{p}_g = (1/3, 1/3, 1/3)$**]. $p^{\text{unacc}}(g,m)$ = P(declare dose d_m unacceptable for subgroup g), $p^{\text{sel}}(g,m)$ = P(select dose d_m as optimal for subgroup g), and n_m^{ptrt} = mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g- = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 7					Scenario 8				
u^{TR}	$g=1$	39.68	42.13	46.41	50.73	**54.11**	46.93	**53.95**	46.36	40.27	36.10
	$g=2$	40.33	43.08	48.81	55.18	**60.06**	48.04	**55.63**	52.02	47.72	44.74
	$g=3$	40.83	43.79	50.06	57.99	**63.55**	48.90	**57.16**	55.02	51.73	49.26
p_{T}^{TR}	$g=1$	0.03	0.05	0.10	0.20	0.25	0.05	0.05	0.30	*0.45*	*0.55*
	$g=2$	0.03	0.05	0.10	0.20	0.25	0.05	0.05	0.30	*0.45*	*0.55*
	$g=3$	0.03	0.05	0.10	0.20	0.25	0.05	0.05	0.30	*0.45*	*0.55*
π_0^{TR}	$g=1$	*0.54*	*0.46*	*0.31*	0.16	0.09	0.31	0.16	0.12	0.12	0.12
	$g=2$	*0.54*	*0.46*	0.31	0.16	0.09	0.31	0.16	0.12	0.12	0.12
	$g=3$	*0.54*	*0.46*	0.31	0.16	0.09	0.31	0.16	0.12	0.12	0.12

- $p_m^{\text{sel}}(x)$

D-Sub	$g=1$	0.00	0.00	0.06	0.15	**0.79**	0.12	**0.78**	0.09	0.00	0.01
	$g=2$	0.00	0.00	0.03	0.10	**0.86**	0.07	**0.76**	0.14	0.01	0.03
	$g=3$	0.00	0.00	0.02	0.09	**0.87**	0.04	**0.67**	0.19	0.02	0.08
D-Comb	all g	0.00	0.00	0.03	0.11	**0.84**	0.06	**0.75**	0.14	0.01	0.04
D-Sep	$g=1$	0.00	0.00	0.08	0.15	**0.71**	0.21	**0.47**	0.14	0.02	0.05
	$g=2$	0.00	0.00	0.06	0.12	**0.81**	0.11	**0.50**	0.19	0.04	0.16
	$g=3$	0.00	0.00	0.04	0.10	**0.85**	0.04	**0.42**	0.21	0.06	0.26

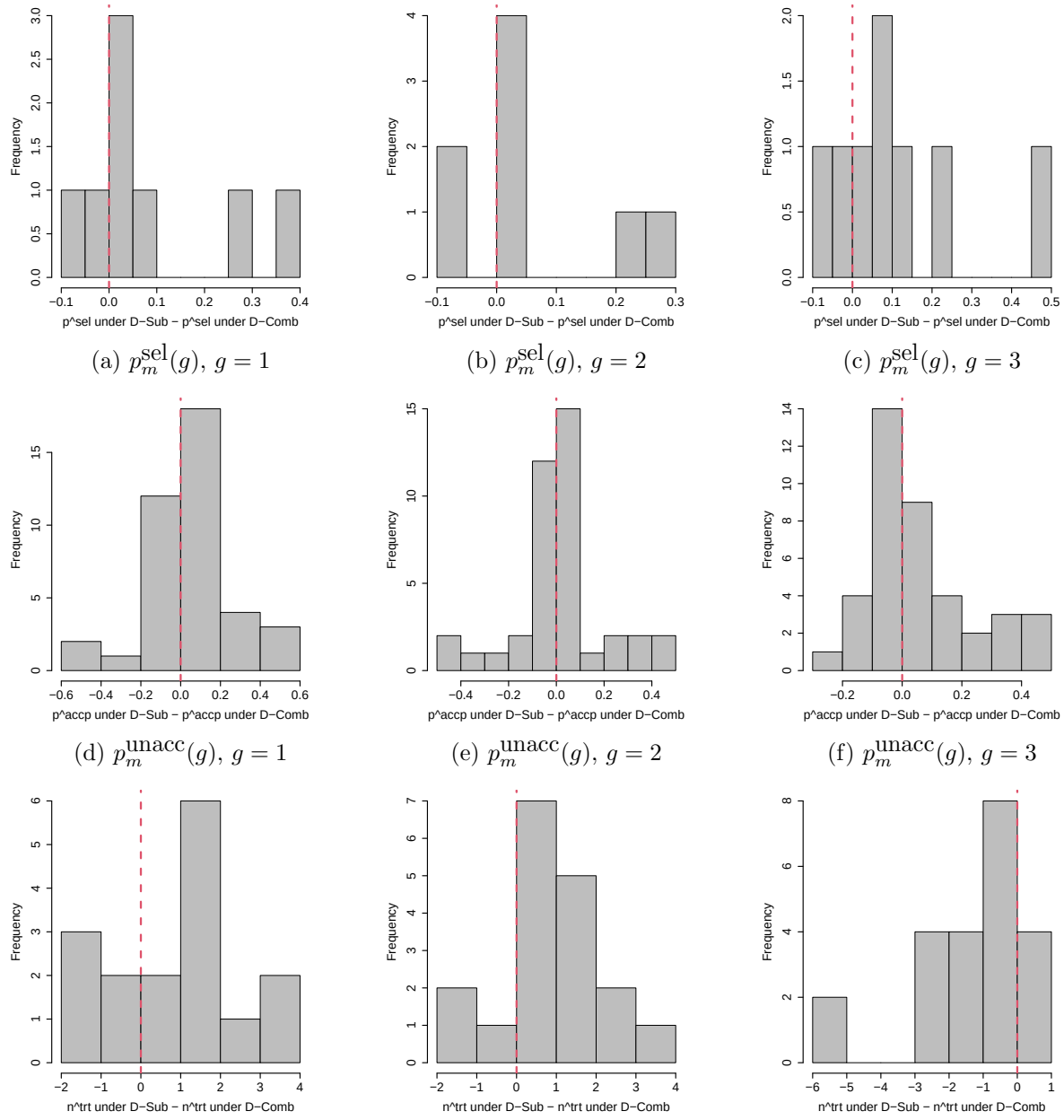
- $p_m^{\text{Accp}}(x)$

D-Sub	$g=1$	*0.99*	*0.98*	*0.64*	0.03	0.01	0.27	0.03	0.08	*0.31*	*0.39*
	$g=2$	*0.64*	*0.55*	0.16	0.01	0.01	0.04	0.00	0.10	*0.38*	*0.48*
	$g=3$	*0.72*	*0.63*	0.21	0.03	0.03	0.05	0.01	0.14	*0.47*	*0.58*
D-Comb	all g	*0.91*	*0.85*	*0.33*	0.04	0.03	0.11	0.02	0.10	*0.41*	*0.50*
D-Sep	$g=1$	*0.96*	*0.91*	*0.55*	0.12	0.07	0.31	0.20	0.23	*0.40*	*0.46*
	$g=2$	*0.51*	*0.42*	0.11	0.03	0.03	0.03	0.01	0.08	*0.30*	*0.37*
	$g=3$	*0.49*	*0.41*	0.13	0.04	0.04	0.03	0.01	0.12	*0.32*	*0.41*

- $n_m^{\text{ptrt}}(x)$

D-Sub	$g=1$	*2.11*	*2.90*	*7.32*	13.42	14.01	8.23	9.73	8.79	*6.79*	*6.40*
	$g=2$	*4.88*	*5.69*	8.68	10.21	10.35	9.73	10.16	8.47	*6.13*	*5.42*
	$g=3$	*4.10*	*4.96*	8.49	10.89	10.88	10.25	10.71	8.54	*5.27*	*4.54*
D-Comb	$g=1$	*3.25*	*3.97*	*8.32*	11.75	12.15	9.50	10.05	8.84	*6.08*	*5.32*
	$g=2$	*3.40*	*4.00*	8.31	11.72	11.81	9.41	10.10	8.73	*5.95*	*5.31*
	$g=3$	*3.26*	*3.99*	8.26	11.61	11.86	9.54	10.22	8.69	*6.03*	*5.26*
D-Sep	$g=1$	*5.64*	*5.87*	*7.52*	9.76	10.49	8.54	9.01	8.41	*6.58*	*6.25*
	$g=2$	*6.89*	*7.22*	8.32	8.82	8.59	9.39	9.37	8.39	*6.63*	*6.13*
	$g=3$	*6.82*	*7.21*	8.49	8.77	8.59	9.50	9.62	8.35	*6.52*	*5.94*

Table 6 continued:[**Simulation Results with $\boldsymbol{p}_g = (1/3, 1/3, 1/3)$**].$p^{\text{unacc}}(g,m) = \text{P}$(declare dose d_m unacceptable for subgroup g), $p^{\text{sel}}(g,m) = \text{P}$(select dose d_m as optimal for subgroup g), and $n_m^{\text{ptrt}} =$ mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g- = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

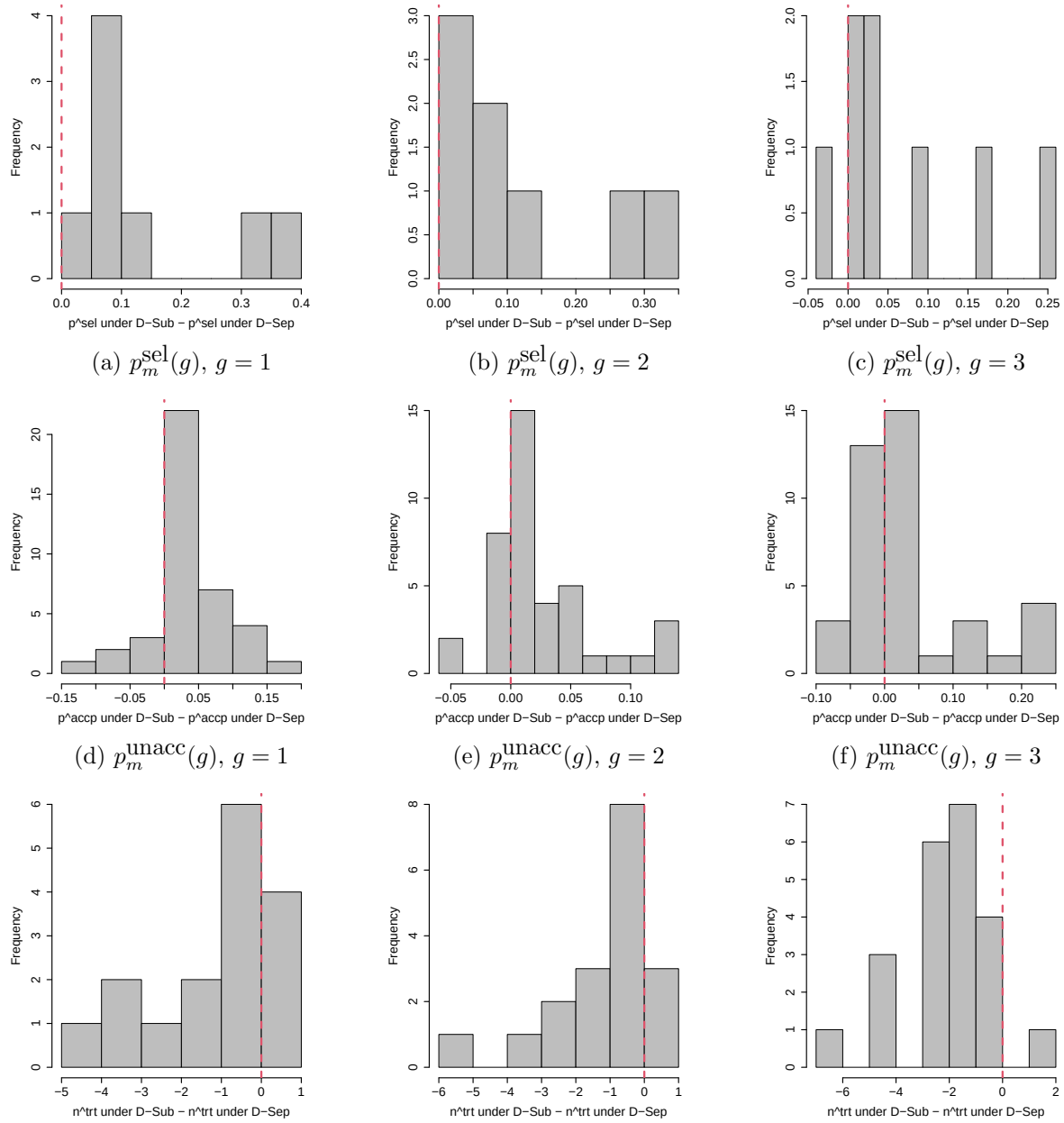


(a) $p_m^{\text{sel}}(g)$, $g = 1$ (b) $p_m^{\text{sel}}(g)$, $g = 2$ (c) $p_m^{\text{sel}}(g)$, $g = 3$

(d) $p_m^{\text{unacc}}(g)$, $g = 1$ (e) $p_m^{\text{unacc}}(g)$, $g = 2$ (f) $p_m^{\text{unacc}}(g)$, $g = 3$

(g) $n_m^{\text{ptrt}}(g)$, $g = 1$ (h) $n_m^{\text{ptrt}}(g)$, $g = 2$ (i) $n_m^{\text{ptrt}}(g)$, $g = 3$

Figure 1: [Comparison between D-Sub and D-Comb when $\boldsymbol{p}_g = (1/3, 1/3, 1/3)$] Histograms of differences in $p_m^{\text{sel}}(g)$ between D-Sub and D-Comb are in panels (a)-(c). Histograms of differences in $p_m^{\text{unacc}}(g)$ between D-Sub and D-Comb for the truly unacceptable doses and those between D-Comb and D-Sub for the truly acceptable doses are in panels (d)-(f). Histograms of differences in $n_m^{\text{ptrt}}(g)$ between D-Sub and D-Comb for the truly unacceptable doses are in panels (g)-(i). The left, middle and right columns are for subgroups $g = 1$(favorable), $g = 2$ (intermediate), and $g = 3$ (poor). A positive value indicates better performance of D-Sub than D-Comb in panels (a)-(f), and a worse performance in panels (g)-(i).



(a) $p_m^{\text{sel}}(g)$, $g = 1$ (b) $p_m^{\text{sel}}(g)$, $g = 2$ (c) $p_m^{\text{sel}}(g)$, $g = 3$

(d) $p_m^{\text{unacc}}(g)$, $g = 1$ (e) $p_m^{\text{unacc}}(g)$, $g = 2$ (f) $p_m^{\text{unacc}}(g)$, $g = 3$

(g) $n_m^{\text{ptrt}}(g)$, $g = 1$ (h) $n_m^{\text{ptrt}}(g)$, $g = 2$ (i) $n_m^{\text{ptrt}}(g)$, $g = 3$

Figure 2: [Comparison between D-Sub and D-Sep when $\boldsymbol{p}_g = (1/3, 1/3, 1/3)$] Histograms of differences in $p_m^{\text{sel}}(g)$ between the D-Sub and D-Sep designs, in panels (a)-(c). Histograms of differences in $p_m^{\text{unacc}}(g)$ between D-Sub and D-Sep for the truly unacceptable doses and those between D-Sep and D-Sub for the truly acceptable doses are in panels (d)-(f). Histograms of differences in $n_m^{\text{ptrt}}(g)$ between D-Sub and D-Sep for the truly unacceptable doses are in panels (g)-(i). The left, middle and right columns are for subgroups $g = 1$(favorable), $g = 2$ (intermediate), and $g = 3$ (poor). A positive value indicates better performance of D-Sub than D-Sep in panels (a)-(f), and a worse performance in panels (g)-(i).

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
$\boldsymbol{z}^{\text{TR}} = (1,2,3)$		Scenario 1					Scenario 2				
U^{TR}	$g=1$	**61.00**	57.47	56.30	54.83	52.67	37.45	40.37	41.46	41.18	42.19
	$g=2$	**58.17**	53.25	51.91	49.83	47.67	35.85	38.94	40.73	40.27	41.70
	$g=3$	38.75	33.13	29.97	27.18	24.41	32.14	34.53	35.35	34.91	35.96
p_{T}^{TR}	$g=1$	0.05	0.20	0.35	*0.50*	*0.65*	0.25	0.35	*0.50*	*0.60*	*0.65*
	$g=2$	0.07	0.28	*0.47*	*0.64*	*0.79*	0.32	*0.44*	*0.61*	*0.71*	*0.76*
	$g=3$	0.11	0.39	*0.61*	*0.78*	*0.90*	0.38	*0.51*	*0.68*	*0.78*	*0.82*
π_0^{TR}	$g=1$	0.16	0.16	0.12	0.09	0.07	*0.50*	*0.37*	*0.25*	*0.20*	0.16
	$g=2$	0.25	0.25	0.20	0.16	0.12	*0.57*	*0.43*	0.31	0.25	0.20
	$g=3$	*0.57*	*0.56*	*0.50*	*0.43*	*0.37*	*0.63*	*0.50*	*0.37*	0.31	0.25

- $p_m^{\text{sel}}(g)$

D-Sub	$g=1$	**0.67**	0.08	0.07	0.02	0.13	0.00	0.01	0.03	0.01	0.02
	$g=2$	**0.83**	0.08	0.05	0.00	0.02	0.02	0.02	0.01	0.00	0.00
	$g=3$	0.14	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D-Comb	all g	*0.75*	0.04	0.02	0.00	0.01	0.00	0.00	0.00	0.00	0.00
D-Sep	$g=1$	**0.38**	0.19	0.09	0.06	0.25	0.02	0.03	0.07	0.03	0.10
	$g=2$	**0.83**	0.06	0.06	0.01	0.02	0.02	0.02	0.01	0.00	0.00
	$g=3$	0.10	0.03	0.03	0.01	0.01	0.07	0.04	0.05	0.01	0.01

- $p_m^{\text{unacc}}(g)$

D-Sub	$g=1$	0.05	0.06	0.31	*0.57*	*0.64*	*0.99*	*0.98*	*0.95*	*0.97*	*0.97*
	$g=2$	0.03	0.04	*0.54*	*0.93*	*0.95*	*0.98*	*0.97*	*0.99*	*1.00*	*1.00*
	$g=3$	*0.86*	*0.91*	*0.99*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*
D-Comb	all g	*0.18*	*0.23*	*0.74*	*0.96*	*0.97*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*
D-Sep	$g=1$	0.05	0.06	0.21	*0.44*	*0.53*	*0.94*	*0.91*	*0.85*	*0.88*	*0.88*
	$g=2$	0.03	0.07	*0.57*	*0.91*	*0.94*	*0.97*	*0.96*	*0.99*	*1.00*	*1.00*
	$g=3$	*0.87*	*0.89*	*0.94*	*0.97*	*0.98*	*0.89*	*0.90*	*0.93*	*0.98*	*0.99*

- $n_m^{\text{ptrt}}(g)$

D-Sub	$g=1$	7.28	7.22	5.43	*3.69*	*3.41*	*1.90*	*2.75*	*2.97*	*2.14*	*1.99*
	$g=2$	25.50	24.53	*12.87*	*4.57*	*3.24*	*8.97*	*8.69*	*4.60*	*1.80*	*1.21*
	$g=3$	*3.86*	*3.21*	*1.22*	*0.28*	*0.08*	*1.59*	*1.79*	*0.68*	*0.16*	*0.05*
D-Comb	$g=1$	9.68	8.43	3.90	*1.41*	*1.07*	*2.44*	*2.08*	*1.27*	*0.54*	*0.45*
	$g=2$	24.76	22.25	*10.07*	*3.69*	*2.72*	*6.21*	*5.63*	*3.35*	*1.50*	*1.16*
	$g=3$	*7.06*	*6.25*	*2.80*	*1.05*	*0.78*	*1.66*	*1.52*	*0.99*	*0.44*	*0.34*
D-Sep	$g=1$	6.92	6.89	5.76	*4.46*	*3.94*	*7.75*	*7.16*	*5.18*	*3.25*	*2.84*
	$g=2$	25.85	23.47	*12.53*	*5.22*	*3.81*	*11.74*	*10.01*	*5.66*	*2.49*	*1.86*
	$g=3$	*6.97*	*6.14*	*3.56*	*1.82*	*1.32*	*6.93*	*5.46*	*2.96*	*1.56*	*1.16*

Table 7: [**Simulation Results when $U_g(\boldsymbol{Y})$ is the same for all g**]]. $p^{\text{unacc}}(g, d_m)$ = P(declare dose d_m unacceptable for subgroup g), $p^{\text{sel}}(g, d_m)$ = P(select dose d_m as optimal for subgroup g), and n_m^{ptrt} = mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
$\boldsymbol{z}^{\text{TR}} = (1,1,2)$		Scenario 3					Scenario 4				
U^{TR}	$g=1$	65.01	71.02	**78.92**	74.56	70.02	**56.54**	54.30	55.31	49.60	46.40
	$g=2$	64.96	70.83	**79.05**	74.63	70.18	**56.47**	54.15	55.52	49.64	46.55
	$g=3$	53.30	58.95	**68.12**	57.76	50.98	49.29	48.50	**50.74**	45.58	42.66
p_{T}^{TR}	$g=1$	0.05	0.08	0.10	0.35	*0.50*	0.05	0.25	0.30	*0.50*	*0.60*
	$g=2$	0.05	0.08	0.10	0.35	*0.50*	0.05	0.25	0.30	*0.50*	*0.60*
	$g=3$	0.13	0.20	0.25	*0.69*	*0.85*	0.05	0.25	0.30	*0.50*	*0.60*
π_0^{TR}	$g=1$	0.16	0.09	0.03	0.02	0.02	0.16	0.12	0.09	0.09	0.09
	$g=2$	0.16	0.09	0.03	0.02	0.02	0.16	0.12	0.09	0.09	0.09
	$g=3$	0.31	0.20	0.09	0.07	0.07	*0.41*	0.34	0.29	0.29	0.29

- $p_m^{\text{sel}}(g)$

D-Sub	$g=1$	0.00	0.02	**0.84**	0.09	0.06	**0.58**	0.15	0.21	0.02	0.04
	$g=2$	0.00	0.02	**0.86**	0.07	0.05	**0.60**	0.15	0.20	0.01	0.03
	$g=3$	0.01	0.06	**0.86**	0.02	0.01	0.51	0.14	**0.16**	0.01	0.02
D-Comb	all g	0.01	0.03	**0.86**	0.07	0.04	**0.65**	0.13	**0.17**	0.02	0.03
D-Sep	$g=1$	0.02	0.06	**0.44**	0.15	0.33	**0.40**	0.15	0.17	0.05	0.21
	$g=2$	0.01	0.04	**0.70**	0.14	0.12	**0.64**	0.12	0.16	0.03	0.05
	$g=3$	0.13	0.17	**0.55**	0.08	0.07	0.20	0.13	**0.18**	0.07	0.36

- $p_m^{\text{unacc}}(g)$

D-Sub	$g=1$	0.00	0.00	0.00	0.09	*0.25*	0.01	0.01	0.08	*0.32*	*0.48*
	$g=2$	0.00	0.00	0.00	0.14	*0.38*	0.00	0.00	0.09	*0.41*	*0.59*
	$g=3$	0.05	0.04	0.05	*0.68*	*0.89*	*0.20*	0.21	0.36	*0.75*	*0.84*
D-Comb	all g	0.00	0.00	0.01	*0.31*	*0.60*	*0.01*	0.02	0.11	*0.39*	*0.55*
D-Sep	$g=1$	0.01	0.01	0.02	0.14	*0.22*	0.05	0.06	0.15	*0.37*	*0.45*
	$g=2$	0.00	0.00	0.01	0.14	*0.28*	0.00	0.03	0.11	*0.37*	*0.52*
	$g=3$	0.03	0.02	0.08	*0.64*	*0.83*	*0.23*	0.21	0.21	*0.38*	*0.43*

- $n_m^{\text{ptrt}}(g)$

D-Sub	$g=1$	5.69	6.03	5.94	5.26	*4.81*	6.49	6.67	5.81	*4.54*	*4.04*
	$g=2$	15.81	15.84	15.91	13.38	*10.93*	18.50	18.36	15.77	*10.73*	*8.58*
	$g=3$	5.11	5.34	5.22	*2.53*	*1.46*	*5.56*	5.26	3.84	*1.83*	*1.30*
D-Comb	$g=1$	6.66	6.56	6.27	4.59	*3.60*	7.17	6.74	5.98	*4.11*	*3.51*
	$g=2$	16.93	16.97	16.61	12.33	*9.10*	18.88	17.62	15.17	*11.15*	*9.16*
	$g=3$	4.74	4.80	4.60	*3.52*	*2.62*	*5.28*	5.04	4.20	*3.03*	*2.64*
D-Sep	$g=1$	5.98	6.06	5.89	5.20	*4.87*	6.92	6.65	5.81	*4.55*	*4.04*
	$g=2$	15.97	15.63	15.28	13.26	*11.79*	18.55	17.23	15.28	*11.19*	*9.63*
	$g=3$	5.12	5.24	4.69	*2.80*	*2.11*	*4.70*	4.68	4.21	*3.37*	*3.00*

Table 7 continued:[**Simulation Results when $U_g(\boldsymbol{Y})$ is the same for all g**]]. $p^{\text{unacc}}(g, d_m)$ = P(declare dose d_m unacceptable for subgroup g), $p^{\text{sel}}(g, d_m)$ = P(select dose d_m as optimal for subgroup g), and n_m^{ptrt} = mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
$\boldsymbol{z}^{\text{TR}} = (1,2,2)$		Scenario 5					Scenario 6				
U^{TR}	$g=1$	60.06	53.86	53.10	57.16	**60.73**	61.48	59.91	58.96	**70.38**	66.25
	$g=2$	65.44	59.25	58.92	63.74	**67.67**	**52.93**	46.64	44.45	52.85	42.68
	$g=3$	65.49	59.39	58.71	63.80	**67.59**	**52.99**	46.81	44.36	52.90	42.58
p_{T}^{TR}	$g=1$	0.03	0.28	0.30	0.35	0.38	0.03	0.10	0.13	0.15	0.30
	$g=2$	0.03	0.28	0.30	0.35	0.38	0.13	0.38	*0.46*	*0.52*	*0.80*
	$g=3$	0.03	0.28	0.30	0.35	0.38	0.13	0.37	*0.46*	*0.52*	*0.80*
π_0^{TR}	$g=1$	0.12	0.12	0.12	0.07	0.04	0.16	0.16	0.16	0.05	0.05
	$g=2$	0.16	0.16	0.16	0.09	0.06	0.20	0.20	0.20	0.07	0.07
	$g=3$	0.16	0.16	0.16	0.09	0.05	0.20	0.20	0.20	0.07	0.07

- $p_m^{\text{sel}}(g)$

D-Sub	$g=1$	0.22	0.01	0.01	0.01	**0.75**	0.06	0.01	0.02	**0.27**	0.65
	$g=2$	0.26	0.01	0.01	0.01	**0.71**	**0.76**	0.02	0.03	0.15	0.03
	$g=3$	0.30	0.01	0.01	0.02	**0.65**	**0.74**	0.02	0.02	0.11	0.01
D-Comb	all g	0.27	0.01	0.01	0.02	**0.70**	**0.60**	0.01	0.03	**0.22**	0.11
D-Sep	$g=1$	0.24	0.04	0.04	0.03	**0.65**	0.02	0.02	0.04	**0.16**	0.76
	$g=2$	0.24	0.02	0.01	0.03	**0.70**	**0.77**	0.03	0.03	0.12	0.03
	$g=3$	0.15	0.05	0.05	0.06	**0.69**	**0.52**	0.08	0.10	0.14	0.14

- $p_m^{\text{unacc}}(g)$

D-Sub	$g=1$	0.00	0.00	0.01	0.02	0.02	0.00	0.00	0.01	0.01	0.03
	$g=2$	0.00	0.00	0.02	0.04	0.05	0.01	0.09	*0.47*	*0.73*	*0.91*
	$g=3$	0.01	0.03	0.06	0.08	0.09	0.10	0.20	*0.59*	*0.82*	*0.95*
D-Comb	all g	0.00	0.02	0.03	0.04	0.05	0.03	0.08	*0.28*	*0.54*	*0.77*
D-Sep	$g=1$	0.02	0.04	0.09	0.13	0.15	0.05	0.04	0.03	0.03	0.04
	$g=2$	0.00	0.02	0.04	0.06	0.07	0.03	0.14	*0.49*	*0.77*	*0.91*
	$g=3$	0.01	0.02	0.07	0.13	0.16	0.03	0.10	*0.32*	*0.57*	*0.74*

- $n_m^{\text{ptrt}}(g)$

D-Sub	$g=1$	5.58	5.68	5.42	5.42	5.33	5.51	5.73	5.59	5.56	5.41
	$g=2$	15.46	15.13	14.27	13.56	13.65	23.89	21.48	*13.24*	*7.81*	*4.75*
	$g=3$	4.48	4.40	3.89	3.59	3.51	6.96	5.80	*3.01*	*1.48*	*0.64*
D-Comb	$g=1$	6.15	5.66	5.33	5.23	5.22	8.31	7.16	5.32	3.69	2.54
	$g=2$	15.97	15.02	13.89	13.59	13.50	21.86	18.73	*13.79*	*9.32*	*6.48*
	$g=3$	4.41	4.27	4.06	3.96	3.75	6.29	5.34	*3.87*	*2.72*	*1.86*
D-Sep	$g=1$	6.32	6.14	5.46	5.12	4.95	5.68	5.79	5.61	5.46	5.46
	$g=2$	16.10	15.14	14.17	13.37	13.22	25.84	20.59	*12.58*	*7.08*	*4.46*
	$g=3$	4.51	4.35	4.00	3.61	3.52	5.53	5.24	*4.02*	*2.75*	*2.27*

Table 7 continued:[**Simulation Results when $U_g(\boldsymbol{Y})$ is the same for all g**]]. $p^{\text{unacc}}(g, d_m)$ = P(declare dose d_m unacceptable for subgroup g), $p^{\text{sel}}(g, d_m)$ = P(select dose d_m as optimal for subgroup g), and n_m^{ptrt} = mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
$\boldsymbol{z}^{\text{TR}} = (1,1,1)$		Scenario 7					Scenario 8				
U^{TR}	$g=1$	40.33	43.08	48.81	55.18	**60.06**	48.04	**55.63**	52.02	47.72	44.74
	$g=2$	40.33	43.08	48.81	55.18	**60.06**	48.04	**55.63**	52.02	47.72	44.74
	$g=3$	40.33	43.08	48.81	55.18	**60.06**	48.04	**55.63**	52.02	47.72	44.74
p_{T}^{TR}	$g=1$	0.03	0.05	0.10	0.20	0.25	0.05	0.05	0.30	*0.45*	*0.55*
	$g=2$	0.03	0.05	0.10	0.20	0.25	0.05	0.05	0.30	*0.45*	*0.55*
	$g=3$	0.03	0.05	0.10	0.20	0.25	0.05	0.05	0.30	*0.45*	*0.55*
π_0^{TR}	$g=1$	*0.54*	*0.46*	*0.31*	0.16	0.09	0.31	0.16	0.12	0.12	0.12
	$g=2$	*0.54*	*0.46*	0.31	0.16	0.09	0.31	0.16	0.12	0.12	0.12
	$g=3$	*0.54*	*0.46*	0.31	0.16	0.09	0.31	0.16	0.12	0.12	0.12

- $p_m^{\text{sel}}(g)$

D-Sub	$g=1$	0.00	0.00	0.03	0.11	**0.85**	0.04	**0.74**	0.17	0.01	0.05
	$g=2$	0.00	0.00	0.04	0.12	**0.84**	0.04	**0.77**	0.15	0.01	0.03
	$g=3$	0.00	0.00	0.04	0.13	**0.77**	0.05	**0.78**	0.14	0.00	0.02
D-Comb	all g	0.00	0.00	0.03	0.12	**0.82**	0.06	**0.74**	0.15	0.00	0.03
D-Sep	$g=1$	0.00	0.00	0.03	0.10	**0.83**	0.09	**0.37**	0.18	0.05	0.24
	$g=2$	0.00	0.00	0.05	0.12	**0.82**	0.08	**0.64**	0.17	0.02	0.09
	$g=3$	0.00	0.01	0.06	0.10	**0.82**	0.10	**0.34**	0.19	0.07	0.30

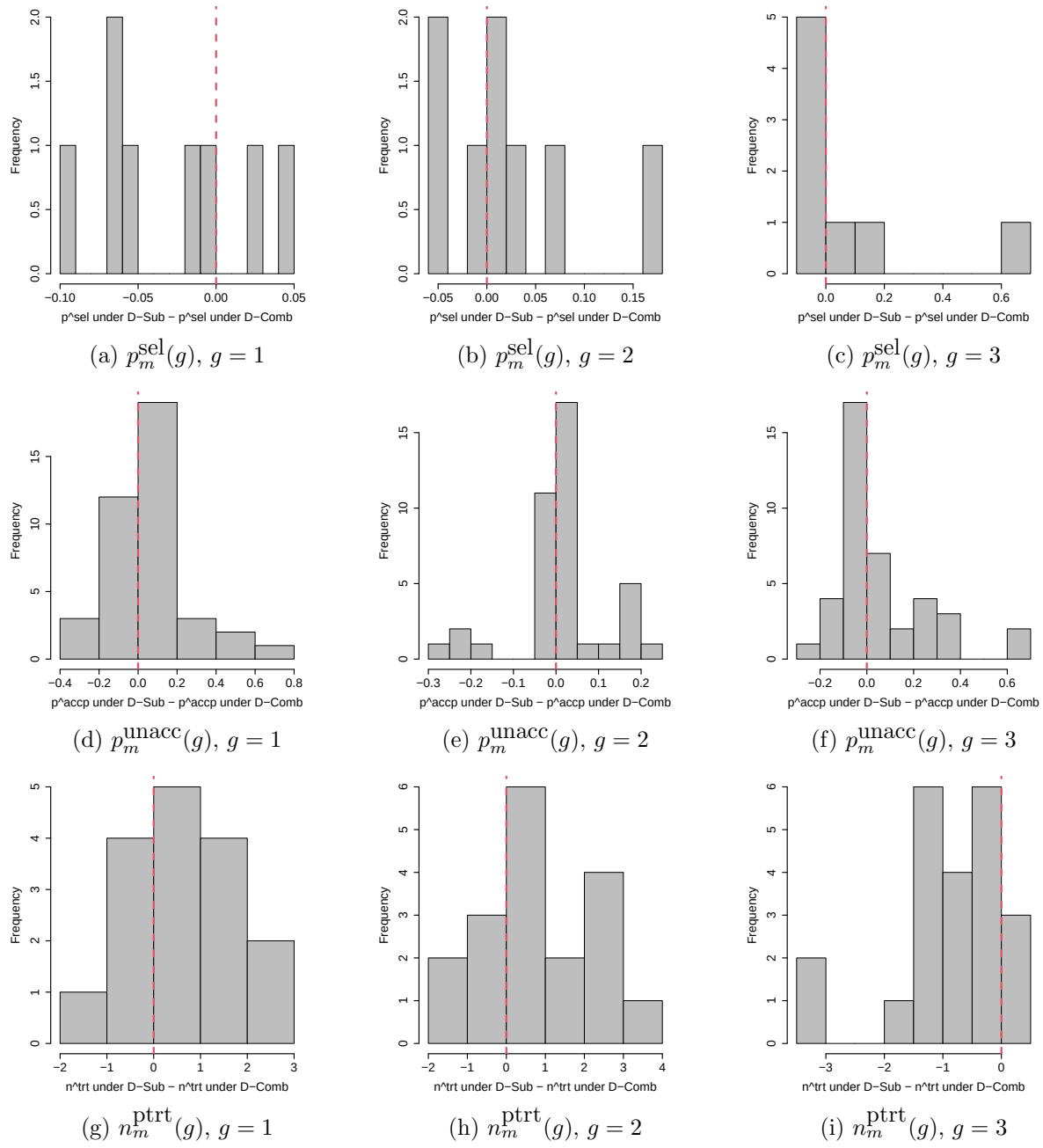
- $p_m^{\text{unacc}}(g)$

D-Sub	$g=1$	*0.98*	*0.96*	*0.61*	0.02	0.01	0.27	0.02	0.06	*0.27*	*0.36*
	$g=2$	*0.65*	*0.56*	0.16	0.01	0.01	0.05	0.00	0.08	*0.37*	*0.48*
	$g=3$	*0.77*	*0.69*	0.26	0.06	0.06	0.08	0.01	0.13	*0.50*	*0.61*
D-Comb	all g	*0.90*	*0.85*	*0.34*	0.03	0.03	0.10	0.02	0.12	*0.41*	*0.51*
D-Sep	$g=1$	*0.92*	*0.86*	*0.51*	0.12	0.06	0.27	0.17	0.20	0.34	0.39
	$g=2$	*0.62*	*0.54*	0.17	0.02	0.03	0.04	0.01	0.10	*0.32*	*0.42*
	$g=3$	*0.46*	*0.38*	0.12	0.04	0.04	0.04	0.02	0.09	*0.26*	*0.33*

- $n_m^{\text{ptrt}}(g)$

D-Sub	$g=1$	*1.40*	*2.26*	*5.21*	8.98	9.52	5.30	6.47	6.13	*4.81*	*4.59*
	$g=2$	*8.88*	*10.23*	*15.39*	18.69	18.63	17.43	17.99	15.90	*11.23*	*9.80*
	$g=3$	*1.67*	*2.44*	4.23	5.39	5.45	5.04	5.58	4.52	*2.69*	*2.18*
D-Comb	$g=1$	*2.32*	*2.64*	*5.72*	8.36	8.19	6.46	7.05	5.78	*4.23*	*3.62*
	$g=2$	*6.03*	*7.24*	14.84	21.34	21.28	17.20	18.49	15.36	*10.72*	*9.57*
	$g=3$	*1.70*	*2.05*	4.23	5.95	6.04	4.95	5.09	4.32	*3.10*	*2.69*
D-Sep	$g=1$	*5.36*	*5.39*	*5.55*	5.81	5.79	6.40	6.48	5.82	*4.77*	*4.43*
	$g=2$	*9.97*	*10.90*	15.42	17.65	17.53	17.23	17.40	15.29	*11.42*	*10.39*
	$g=3$	*3.97*	*4.18*	4.07	3.98	3.81	4.43	4.59	4.20	*3.52*	*3.27*

Table 7 continued:[**Simulation Results when $U_g(\boldsymbol{Y})$ is the same for all g**]]. $p^{\text{unacc}}(g, d_m)$ = P(declare dose d_m unacceptable for subgroup g), $p^{\text{sel}}(g, d_m)$ = P(select dose d_m as optimal for subgroup g), and n_m^{ptrt} = mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

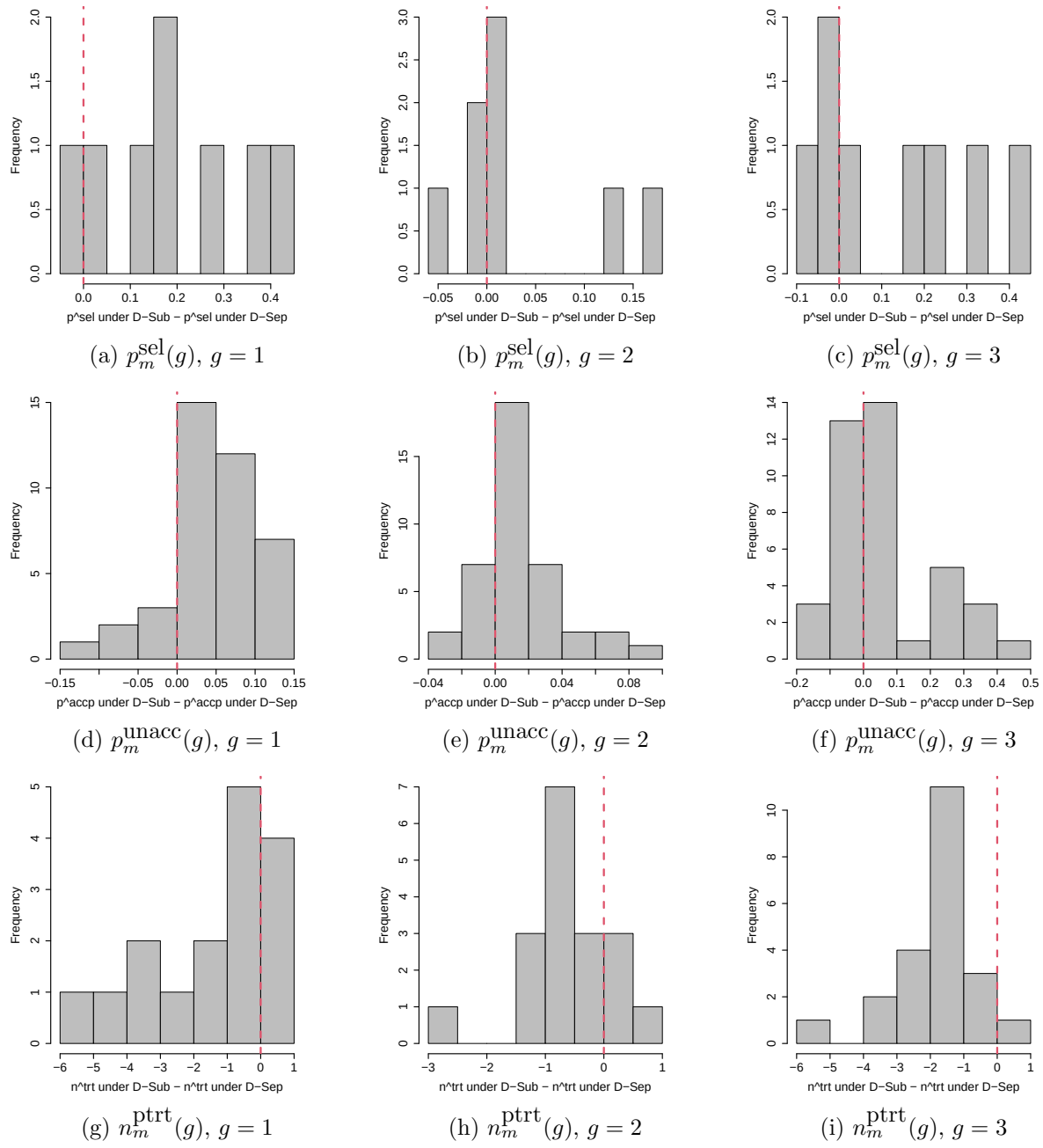


(a) $p_m^{\text{sel}}(g)$, $g = 1$ (b) $p_m^{\text{sel}}(g)$, $g = 2$ (c) $p_m^{\text{sel}}(g)$, $g = 3$

(d) $p_m^{\text{unacc}}(g)$, $g = 1$ (e) $p_m^{\text{unacc}}(g)$, $g = 2$ (f) $p_m^{\text{unacc}}(g)$, $g = 3$

(g) $n_m^{\text{ptrt}}(g)$, $g = 1$ (h) $n_m^{\text{ptrt}}(g)$, $g = 2$ (i) $n_m^{\text{ptrt}}(g)$, $g = 3$

Figure 3: [Comparison between D-Sub and D-Comb when $U_g(\boldsymbol{Y})$ is the same for all g] Histograms of differences in $p_m^{\text{sel}}(g)$ between D-Sub and D-Comb are in panels (a)-(c). Histograms of differences in $p_m^{\text{unacc}}(g)$ between D-Sub and D-Comb for the truly unacceptable doses and those between D-Comb and D-Sub for the truly acceptable doses are in panels (d)-(f). Histograms of differences in $n_m^{\text{ptrt}}(g)$ between D-Sub and D-Comb for the truly unacceptable doses are in panels (g)-(i). The left, middle and right columns are for subgroups $g = 1$(favorable), $g = 2$ (intermediate), and $g = 3$ (poor). A positive value indicates better performance of D-Sub than D-Comb in panels (a)-(f), and a worse performance in panels (g)-(i).



(a) $p_m^{\text{sel}}(g)$, $g = 1$ (b) $p_m^{\text{sel}}(g)$, $g = 2$ (c) $p_m^{\text{sel}}(g)$, $g = 3$

(d) $p_m^{\text{unacc}}(g)$, $g = 1$ (e) $p_m^{\text{unacc}}(g)$, $g = 2$ (f) $p_m^{\text{unacc}}(g)$, $g = 3$

(g) $n_m^{\text{ptrt}}(g)$, $g = 1$ (h) $n_m^{\text{ptrt}}(g)$, $g = 2$ (i) $n_m^{\text{ptrt}}(g)$, $g = 3$

Figure 4: [Comparison between D-Sub and D-Sep when $U_g(\boldsymbol{Y})$ is the same for all g] Histograms of differences in $p_m^{\text{sel}}(g)$ between the D-Sub and D-Sep designs, in panels (a)-(c). Histograms of differences in $p_m^{\text{unacc}}(g)$ between D-Sub and D-Sep for the truly unacceptable doses and those between D-Sep and D-Sub for the truly acceptable doses are in panels (d)-(f). Histograms of differences in $n_m^{\text{ptrt}}(g)$ between D-Sub and D-Sep for the truly unacceptable doses are in panels (g)-(i). The left, middle and right columns are for subgroups $g = 1$(favorable), $g = 2$ (intermediate), and $g = 3$ (poor). A positive value indicates better performance of D-Sub than D-Sep in panels (a)-(f), and a worse performance in panels (g)-(i).

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 1 ($\boldsymbol{z}^{\text{TR}} = (1,2,3)$)					Scenario 3 ($\boldsymbol{z}^{\text{TR}} = (1,1,2)$)				
Dec. U^{TR}	$g=1$	**56.68**	50.61	46.37	42.04	37.37	59.75	64.51	**71.76**	63.18	56.30
	$g=2$	**57.63**	51.19	48.40	44.98	41.73	64.58	70.20	**78.23**	71.75	66.10
	$g=3$	39.16	32.89	29.38	26.33	23.47	55.56	61.75	**71.40**	59.69	52.40
Inc. U^{TR}	$g=1$	**57.05**	52.02	48.80	45.35	41.35	60.12	65.12	**72.55**	65.75	59.71
	$g=2$	**58.50**	54.52	54.13	53.02	51.74	65.20	71.22	**79.55**	76.43	72.77
	$g=3$	40.21	36.78	36.01	35.50	34.79	57.04	64.20	**74.62**	69.42	64.81
p_{T}^{TR}	$g=1$	0.05	0.20	0.35	*0.50*	*0.65*	0.05	0.08	0.10	0.35	*0.50*
	$g=2$	0.07	0.28	*0.47*	*0.64*	*0.79*	0.05	0.08	0.10	0.35	*0.50*
	$g=3$	0.11	0.39	*0.62*	*0.79*	*0.90*	0.13	0.20	0.25	*0.69*	*0.85*
π_0^{TR}	$g=1$	0.16	0.16	0.12	0.09	0.07	0.16	0.09	0.03	0.02	0.02
	$g=2$	0.25	0.25	0.20	0.16	0.12	0.16	0.09	0.03	0.02	0.02
	$g=3$	*0.57*	*0.57*	*0.50*	*0.43*	*0.37*	0.31	0.20	0.09	0.07	0.07

- $p_m^{\text{sel}}(g)$

Dec.	$g=1$	**0.79**	0.08	0.05	0.01	0.02	0.01	0.06	**0.86**	0.04	0.03
	$g=2$	**0.82**	0.08	0.05	0.00	0.02	0.01	0.03	**0.87**	0.06	0.04
	$g=3$	0.12	0.01	0.00	0.00	0.00	0.01	0.05	**0.80**	0.02	0.01
Inc.	$g=1$	**0.81**	0.06	0.04	0.01	0.03	0.01	0.04	**0.86**	0.05	0.04
	$g=2$	**0.81**	0.08	0.07	0.01	0.03	0.00	0.02	**0.83**	0.08	0.06
	$g=3$	0.14	0.01	0.00	0.00	0.00	0.01	0.04	**0.86**	0.04	0.02

- $p_m^{\text{unacc}}(g)$

Dec.	$g=1$	0.08	0.10	0.43	*0.69*	*0.75*	0.01	0.00	0.01	0.22	*0.46*
	$g=2$	0.03	0.07	*0.65*	*0.94*	*0.96*	0.00	0.00	0.01	0.32	*0.61*
	$g=3$	*0.87*	*0.93*	*1.00*	*1.00*	*1.00*	0.14	0.11	0.13	*0.80*	*0.96*
Inc.	$g=1$	0.07	0.06	0.26	*0.50*	*0.56*	0.01	0.00	0.00	0.05	*0.13*
	$g=2$	0.01	0.02	*0.44*	*0.88*	*0.92*	0.00	0.00	0.00	0.08	*0.22*
	$g=3$	*0.86*	*0.90*	*0.99*	*1.00*	*1.00*	0.05	0.03	0.03	*0.65*	*0.86*

- $n_m^{\text{ptrt}}(g)$

Dec.	$g=1$	8.16	7.88	4.99	*2.91*	*2.40*	6.22	6.46	6.17	4.95	*3.88*
	$g=2$	28.15	25.76	*11.17*	*3.11*	*1.84*	17.60	17.58	17.17	11.53	*8.08*
	$g=3$	*3.49*	*2.98*	*0.99*	*0.18*	*0.04*	5.03	5.42	5.15	*1.80*	*0.78*
Inc.	$g=1$	6.75	6.74	5.31	*4.20*	*3.92*	5.48	5.75	5.72	5.57	*5.15*
	$g=2$	23.61	22.91	*14.18*	*6.14*	*4.67*	15.17	15.34	15.14	13.93	*12.39*
	$g=3$	*3.72*	*3.18*	*1.36*	*0.40*	*0.19*	4.87	5.31	5.20	*2.72*	*1.81*

Table 8: [**Simulation Results - Nonconstant Hazard**]. $p^{\text{unacc}}(g, d_m) = \text{P}(\text{declare dose } d_m \text{ unacceptable for subgroup } g)$, $p^{\text{sel}}(g, d_m) = \text{P}(\text{select dose } d_m \text{ as optimal for subgroup } g)$, and $n_m^{\text{ptrt}} =$ mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g- = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 5 ($\boldsymbol{z}^{\text{TR}} = (1,2,2)$)					Scenario 7 ($\boldsymbol{z}^{\text{TR}} = (1,1,1)$)				
Dec. U^{TR}	$g=1$	**57.51**	47.06	45.92	48.31	50.83	39.58	41.97	46.04	49.94	**53.10**
	$g=2$	65.20	57.11	56.64	60.96	**64.60**	40.14	42.76	48.11	53.64	**58.09**
	$g=3$	69.82	62.99	62.21	67.80	**71.70**	40.64	43.47	49.32	56.35	**61.43**
Dec. U^{TR}	$g=1$	**57.74**	49.04	48.05	50.82	53.56	39.76	42.27	46.72	51.39	**54.96**
	$g=2$	65.58	60.57	60.33	65.48	**69.58**	40.44	43.27	49.24	56.13	**61.27**
	$g=3$	70.18	66.50	65.99	72.43	**76.85**	40.93	43.97	50.45	58.87	**64.70**
p_{T}^{TR}	$g=1$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
	$g=2$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
	$g=3$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
π_0^{TR}	$g=1$	0.12	0.12	0.12	0.07	0.04	*0.54*	*0.46*	*0.31*	0.16	0.09
	$g=2$	0.16	0.16	0.16	0.09	0.05	*0.54*	*0.46*	0.31	0.16	0.09
	$g=3$	0.16	0.16	0.16	0.09	0.05	*0.54*	*0.46*	0.31	0.16	0.09

- $p_m^{\text{sel}}(g)$

Dec.	$g=1$	**0.70**	0.01	0.01	0.02	0.26	0.00	0.00	0.06	0.18	**0.72**
	$g=2$	0.39	0.01	0.02	0.02	**0.55**	0.00	0.00	0.05	0.15	**0.76**
	$g=3$	0.28	0.01	0.02	0.03	**0.62**	0.00	0.00	0.04	0.11	**0.73**
Inc.	$g=1$	**0.57**	0.01	0.01	0.02	0.39	0.00	0.00	0.04	0.15	**0.80**
	$g=2$	0.21	0.01	0.01	0.02	**0.75**	0.00	0.00	0.02	0.10	**0.88**
	$g=3$	0.12	0.00	0.01	0.02	**0.84**	0.00	0.00	0.02	0.08	**0.87**

- $p_m^{\text{unacc}}(g)$

Dec.	$g=1$	0.01	0.02	0.08	0.11	0.13	*0.98*	*0.97*	*0.57*	0.07	0.07
	$g=2$	0.01	0.04	0.12	0.16	0.18	*0.67*	*0.58*	0.16	0.06	0.08
	$g=3$	0.06	0.11	0.20	0.23	0.26	*0.78*	*0.71*	0.26	0.14	0.15
Inc.	$g=1$	0.00	0.00	0.00	0.01	0.01	*0.98*	*0.96*	*0.58*	0.02	0.01
	$g=2$	0.00	0.00	0.01	0.01	0.02	*0.66*	*0.56*	0.15	0.01	0.00
	$g=3$	0.01	0.03	0.04	0.05	0.05	*0.77*	*0.70*	0.22	0.03	0.03

- $n_m^{\text{ptrt}}(g)$

Dec.	$g=1$	6.35	6.24	5.30	4.83	4.73	*1.46*	*2.27*	*5.29*	8.81	8.66
	$g=2$	18.20	16.60	13.46	11.83	11.29	*9.00*	*10.33*	16.10	17.73	17.10
	$g=3$	5.29	4.49	3.40	2.87	2.64	*1.67*	*2.44*	4.30	4.95	4.76
Inc.	$g=1$	5.37	5.63	5.45	5.47	5.59	*1.40*	*2.22*	*5.27*	9.17	9.46
	$g=2$	14.71	14.67	14.49	14.15	13.95	*8.63*	*9.97*	15.82	18.60	18.90
	$g=3$	4.25	4.21	3.98	3.88	3.76	*1.58*	*2.42*	4.33	5.68	5.64

Table 8 continued:[**Simulation Results - Nonconstant Hazards**]. $p^{\text{unacc}}(g, d_m)$ = P(declare dose d_m unacceptable for subgroup g), $p^{\text{sel}}(g, d_m)$ = P(select dose d_m as optimal for subgroup g), and n_m^{ptrt} = mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 1 ($\boldsymbol{z}^{\text{TR}} = (1,2,3)$)					Scenario 3 ($\boldsymbol{z}^{\text{TR}} = (1,1,2)$)				
U^{TR}	$g=1$	**56.89**	51.38	47.69	43.81	39.46	59.95	64.84	**72.19**	64.57	58.12
	$g=2$	**58.17**	53.25	51.91	49.83	47.67	64.96	70.83	**79.05**	74.63	70.18
	$g=3$	39.85	35.41	33.63	32.13	30.50	56.53	63.35	**73.50**	65.89	60.18
p_{T}^{TR}	$g=1$	0.05	0.20	0.35	*0.50*	*0.65*	0.05	0.08	0.10	0.35	*0.50*
	$g=2$	0.07	0.28	*0.47*	*0.64*	*0.79*	0.05	0.08	0.10	0.35	*0.50*
	$g=3$	0.11	0.39	*0.62*	*0.79*	*0.90*	0.13	0.20	0.25	*0.69*	*0.85*
π_0^{TR}	$g=1$	0.16	0.16	0.12	0.09	0.07	0.16	0.09	0.03	0.02	0.02
	$g=2$	0.25	0.25	0.20	0.16	0.12	0.16	0.09	0.03	0.02	0.02
	$g=3$	*0.57*	*0.57*	*0.50*	*0.43*	*0.37*	0.31	0.20	0.09	0.07	0.07

- $p_m^{\text{sel}}(g)$

D-Sub	$g=1$	**0.67**	0.11	0.07	0.02	0.10	0.02	0.09	**0.69**	0.09	0.11
	$g=2$	**0.69**	0.13	0.10	0.02	0.05	0.01	0.06	**0.67**	0.12	0.14
	$g=3$	0.22	0.03	0.01	0.00	0.00	0.03	0.09	**0.68**	0.08	0.07
D-Comb	all g	*0.67*	0.08	0.06	0.01	0.02	0.01	0.06	**0.68**	0.13	0.12
D-Sep	$g=1$	**0.47**	0.14	0.12	0.05	0.19	0.07	0.13	**0.35**	0.12	0.34
	$g=2$	**0.64**	0.13	0.10	0.04	0.07	0.02	0.06	**0.51**	0.16	0.26
	$g=3$	0.12	0.06	0.07	0.03	0.05	0.10	0.12	**0.46**	0.15	0.17

- $p_m^{\text{unacc}}(g)$

D-Sub	$g=1$	0.06	0.06	0.22	*0.43*	*0.49*	0.01	0.00	0.00	0.06	*0.17*
	$g=2$	0.02	0.05	*0.43*	*0.82*	*0.90*	0.01	0.00	0.00	0.14	*0.30*
	$g=3$	*0.75*	*0.81*	*0.96*	*1.00*	*1.00*	0.10	0.08	0.09	*0.56*	*0.77*
D-Comb	all g	*0.18*	*0.23*	*0.60*	*0.91*	*0.94*	0.01	0.01	0.01	*0.23*	*0.43*
D-Sub	$g=1$	0.10	0.10	0.18	*0.36*	*0.43*	0.02	0.02	0.02	0.12	*0.19*
	$g=2$	0.03	0.06	*0.40*	*0.76*	*0.87*	0.00	0.00	0.01	0.12	*0.21*
	$g=3$	*0.78*	*0.80*	*0.86*	*0.93*	*0.94*	0.06	0.05	0.12	*0.52*	*0.68*

- $n_m^{\text{ptrt}}(g)$

D-Sub	$g=1$	3.12	3.32	2.83	*2.25*	*2.10*	2.62	2.95	2.89	2.85	*2.58*
	$g=2$	11.51	11.09	*7.09*	*3.36*	*2.46*	7.64	7.78	7.76	6.79	*5.84*
	$g=3$	*2.48*	*2.40*	*1.12*	*0.31*	*0.11*	2.20	2.65	2.52	*1.49*	*1.02*
D-Comb	$g=1$	4.55	3.93	2.34	*1.11*	*0.85*	3.17	3.08	3.04	2.40	*2.10*
	$g=2$	11.72	10.51	*6.16*	*2.71*	*2.11*	8.09	8.13	7.91	6.47	*5.39*
	$g=3$	*3.29*	*2.97*	*1.75*	*0.80*	*0.61*	2.34	2.30	2.16	*1.78*	*1.54*
D-Sep	$g=1$	3.20	3.34	2.90	*2.32*	*2.22*	2.74	3.08	3.01	2.68	*2.49*
	$g=2$	11.25	10.54	*7.27*	*3.81*	*2.84*	7.70	7.74	7.57	6.68	*6.28*
	$g=3$	*2.78*	*2.80*	*1.96*	*1.33*	*1.05*	2.16	2.45	2.24	*1.70*	*1.44*

Table 9: [**Simulation Results with** $N_{\max} = 60$]. $p^{\text{unacc}}(g, d_m) = \text{P}(\text{declare dose } d_m \text{ unacceptable for subgroup } g)$, $p^{\text{sel}}(g, d_m) = \text{P}(\text{select dose } d_m \text{ as optimal for subgroup } g)$, and $n_m^{\text{ptrt}} =$ mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g- = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 5 ($\boldsymbol{z}^{\text{TR}} = (1,2,2)$)					Scenario 7 ($\boldsymbol{z}^{\text{TR}} = (1,1,1)$)				
U^{TR}	$g=1$	**57.64**	48.14	47.08	49.66	52.30	39.68	42.13	46.41	50.73	**54.11**
	$g=2$	65.44	59.25	58.92	63.74	**67.67**	40.33	43.08	48.81	55.18	**60.06**
	$g=3$	70.05	65.27	64.67	70.81	**75.03**	40.83	43.79	50.06	57.99	**63.55**
p_{T}^{TR}	$g=1$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
	$g=2$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
	$g=3$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
π_0^{TR}	$g=1$	0.12	0.12	0.12	0.07	0.04	*0.54*	*0.46*	*0.31*	0.16	0.09
	$g=2$	0.16	0.16	0.16	0.09	0.05	*0.54*	*0.46*	0.31	0.16	0.09
	$g=3$	0.16	0.16	0.16	0.09	0.05	*0.54*	*0.46*	0.31	0.16	0.09

- $p_m^{\text{sel}}(g)$

D-Sub	$g=1$	**0.45**	0.03	0.02	0.04	0.47	0.00	0.00	0.10	0.14	**0.75**
	$g=2$	0.22	0.02	0.03	0.03	**0.69**	0.00	0.00	0.06	0.13	**0.79**
	$g=3$	0.19	0.02	0.03	0.04	**0.70**	0.00	0.00	0.06	0.11	**0.75**
D-Comb	all g	**0.24**	0.03	0.02	0.03	**0.68**	0.00	0.00	0.05	0.12	**0.81**
D-Sep	$g=1$	**0.39**	0.06	0.05	0.06	0.44	0.01	0.01	0.08	0.13	**0.72**
	$g=2$	0.20	0.02	0.03	0.04	**0.71**	0.00	0.00	0.05	0.13	**0.80**
	$g=3$	0.08	0.04	0.05	0.07	**0.75**	0.00	0.00	0.03	0.08	**0.87**

- $p_m^{\text{unacc}}(g)$

D-Sub	$g=1$	0.01	0.01	0.02	0.04	0.05	*0.89*	*0.83*	*0.38*	0.05	0.02
	$g=2$	0.00	0.01	0.04	0.07	0.09	*0.52*	*0.43*	0.13	0.03	0.03
	$g=3$	0.03	0.06	0.12	0.16	0.18	*0.71*	*0.64*	0.29	0.11	0.09
D-Comb	all g	0.00	0.02	0.06	0.07	0.08	*0.77*	*0.68*	*0.27*	0.04	0.03
D-Sep	$g=1$	0.06	0.07	0.12	0.17	0.19	*0.80*	*0.75*	*0.46*	0.15	0.08
	$g=2$	0.00	0.02	0.06	0.10	0.12	*0.46*	*0.38*	0.13	0.03	0.04
	$g=3$	0.02	0.04	0.08	0.13	0.16	*0.37*	*0.32*	0.11	0.04	0.05

- $n_m^{\text{ptrt}}(g)$

D-Sub	$g=1$	2.59	2.85	2.83	2.72	2.70	*1.26*	*1.91*	*2.89*	3.73	3.93
	$g=2$	7.88	7.68	7.17	6.56	6.45	*5.51*	*6.00*	7.61	8.39	8.39
	$g=3$	2.35	2.39	1.98	1.74	1.65	*0.86*	*1.66*	2.13	2.36	2.43
D-Comb	$g=1$	3.03	2.91	2.69	2.56	2.50	*1.73*	*1.93*	*2.85*	3.52	3.60
	$g=2$	8.20	7.70	7.04	6.67	6.63	*4.62*	*5.11*	7.54	9.25	9.20
	$g=3$	2.28	2.12	1.95	1.86	1.82	*1.31*	*1.48*	2.11	2.49	2.57
D-Sep	$g=1$	3.01	3.09	2.82	2.61	2.46	*2.68*	*2.93*	*2.90*	2.79	2.71
	$g=2$	8.25	7.75	7.08	6.61	6.29	*6.36*	*6.75*	7.40	7.65	7.71
	$g=3$	2.00	2.24	2.06	1.91	1.77	*1.75*	*2.13*	2.11	2.04	1.97

Table 9 continued:[**Simulation Results with** $N_{\max} = 60$]. $p^{\text{unacc}}(g, d_m)$ = P(declare dose d_m unacceptable for subgroup g), $p^{\text{sel}}(g, d_m)$ = P(select dose d_m as optimal for subgroup g), and n_m^{ptrt} = mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 1 ($\boldsymbol{z}^{\text{TR}} = (1,2,3)$)					Scenario 3 ($\boldsymbol{z}^{\text{TR}} = (1,1,2)$)				
U^{TR}	$g=1$	**56.89**	51.38	47.69	43.81	39.46	59.95	64.84	**72.19**	64.57	58.12
	$g=2$	**58.17**	53.25	51.91	49.83	47.67	64.96	70.83	**79.05**	74.63	70.18
	$g=3$	39.85	35.41	33.63	32.13	30.50	56.53	63.35	**73.50**	65.89	60.18
p_{T}^{TR}	$g=1$	0.05	0.20	0.35	*0.50*	*0.65*	0.05	0.08	0.10	0.35	*0.50*
	$g=2$	0.07	0.28	*0.47*	*0.64*	*0.79*	0.05	0.08	0.10	0.35	*0.50*
	$g=3$	0.11	0.39	*0.62*	*0.79*	*0.90*	0.13	0.20	0.25	*0.69*	*0.85*
π_0^{TR}	$g=1$	0.16	0.16	0.12	0.09	0.07	0.16	0.09	0.03	0.02	0.02
	$g=2$	0.25	0.25	0.20	0.16	0.12	0.16	0.09	0.03	0.02	0.02
	$g=3$	*0.57*	*0.57*	*0.50*	*0.43*	*0.37*	0.31	0.20	0.09	0.07	0.07

- $p_m^{\text{sel}}(g)$

D-Sub	$g=1$	**0.88**	0.04	0.03	0.01	0.02	0.01	0.02	**0.94**	0.03	0.01
	$g=2$	**0.90**	0.04	0.03	0.01	0.01	0.00	0.01	**0.91**	0.06	0.02
	$g=3$	0.08	0.00	0.00	0.00	0.00	0.01	0.03	**0.91**	0.02	0.00
D-Comb	all g	*0.80*	0.02	0.01	0.00	0.01	0.01	0.02	0.90	0.05	0.02
D-Sep	$g=1$	**0.70**	0.12	0.06	0.02	0.05	0.06	0.12	**0.57**	0.12	0.13
	$g=2$	**0.87**	0.05	0.04	0.01	0.01	0.01	0.03	**0.80**	0.09	0.07
	$g=3$	0.05	0.01	0.01	0.00	0.00	0.07	0.13	**0.68**	0.08	0.04

- $p_m^{\text{unacc}}(g)$

D-Sub	$g=1$	0.05	0.05	0.33	*0.61*	*0.70*	0.00	0.00	0.00	0.11	*0.32*
	$g=2$	0.02	0.03	*0.60*	*0.96*	*0.97*	0.00	0.00	0.00	0.15	*0.40*
	$g=3$	*0.91*	*0.95*	*1.00*	*1.00*	*1.00*	0.05	0.04	0.05	*0.83*	*0.97*
D-Comb	all g	*0.18*	*0.24*	*0.79*	*0.99*	*0.99*	0.00	0.01	0.01	*0.30*	*0.69*
D-Sub	$g=1$	0.06	0.07	0.22	*0.49*	*0.59*	0.01	0.01	0.02	0.12	*0.22*
	$g=2$	0.03	0.08	*0.63*	*0.95*	*0.97*	0.00	0.00	0.01	*0.14*	*0.36*
	$g=3$	*0.94*	*0.96*	*0.99*	*1.00*	*1.00*	0.03	0.02	0.07	*0.74*	*0.91*

- $n_m^{\text{ptrt}}(g)$

D-Sub	$g=1$	11.39	11.42	7.82	*5.05*	*4.52*	8.73	9.02	9.01	7.92	*6.80*
	$g=2$	40.10	38.52	*18.74*	*5.31*	*3.81*	24.11	24.29	23.75	*19.80*	*15.73*
	$g=3$	*4.50*	*3.54*	*1.26*	*0.30*	*0.10*	8.25	8.52	8.34	*2.97*	*1.59*
D-Comb	$g=1$	15.29	12.80	5.32	*1.57*	*1.28*	10.13	10.05	9.62	7.00	*4.61*
	$g=2$	39.67	33.72	*13.64*	*3.94*	*3.21*	26.04	25.57	25.28	*18.43*	*12.48*
	$g=3$	*11.13*	*9.77*	*3.70*	*1.13*	*0.85*	7.41	7.24	7.31	*5.05*	*3.45*
D-Sep	$g=1$	10.44	10.29	8.42	*6.06*	*5.27*	8.88	8.79	8.66	7.67	*7.01*
	$g=2$	40.85	36.59	*17.66*	*5.99*	*4.33*	23.89	23.57	23.24	*20.05*	*17.05*
	$g=3$	*9.88*	*8.39*	*4.58*	*1.97*	*1.46*	8.44	8.43	7.73	*3.77*	*2.54*

Table 10: [**Simulation Results with** $N_{\max} = 180$]. $p^{\text{unacc}}(g, d_m) = \text{P}(\text{declare dose } d_m \text{ unacceptable for subgroup } g)$, $p^{\text{sel}}(g, d_m) = \text{P}(\text{select dose } d_m \text{ as optimal for subgroup } g)$, and $n_m^{\text{ptrt}} =$ mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g- = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 5 ($\boldsymbol{z}^{\text{TR}} = (1,2,2)$)					Scenario 7 ($\boldsymbol{z}^{\text{TR}} = (1,1,1)$)				
U^{TR}	$g=1$	**57.64**	48.14	47.08	49.66	52.30	39.68	42.13	46.41	50.73	**54.11**
	$g=2$	65.44	59.25	58.92	63.74	**67.67**	40.33	43.08	48.81	55.18	**60.06**
	$g=3$	70.05	65.27	64.67	70.81	**75.03**	40.83	43.79	50.06	57.99	**63.55**
p_{T}^{TR}	$g=1$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
	$g=2$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
	$g=3$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
π_0^{TR}	$g=1$	0.12	0.12	0.12	0.07	0.04	*0.54*	*0.46*	*0.31*	0.16	0.09
	$g=2$	0.16	0.16	0.16	0.09	0.05	*0.54*	*0.46*	0.31	0.16	0.09
	$g=3$	0.16	0.16	0.16	0.09	0.05	*0.54*	*0.46*	0.31	0.16	0.09

- $p_m^{\text{sel}}(g)$

D-Sub	$g=1$	**0.69**	0.00	0.00	0.01	0.30	0.00	0.00	0.03	0.15	**0.81**
	$g=2$	0.23	0.00	0.00	0.01	**0.75**	0.00	0.00	0.02	0.10	**0.87**
	$g=3$	0.11	0.00	0.00	0.02	**0.85**	0.00	0.00	0.01	0.08	**0.87**
D-Comb	all g	**0.27**	0.00	0.00	0.01	**0.72**	0.00	0.00	0.01	0.11	**0.86**
D-Sep	$g=1$	**0.57**	0.04	0.02	0.04	0.33	0.01	0.00	0.07	0.15	**0.71**
	$g=2$	0.24	0.01	0.01	0.02	**0.73**	0.00	0.00	0.05	0.12	**0.81**
	$g=3$	0.11	0.02	0.04	0.05	**0.79**	0.00	0.00	0.03	0.10	**0.85**

- $p_m^{\text{unacc}}(g)$

D-Sub	$g=1$	0.00	0.00	0.01	0.02	0.02	*1.00*	*0.99*	*0.70*	0.02	0.02
	$g=2$	0.00	0.01	0.03	0.03	0.04	*0.75*	*0.63*	0.15	0.02	0.02
	$g=3$	0.02	0.04	0.06	0.07	0.08	*0.82*	*0.73*	0.21	0.04	0.04
D-Comb	all g	0.00	0.02	0.03	0.03	0.03	*0.95*	*0.90*	*0.33*	0.03	0.03
D-Sep	$g=1$	0.03	0.04	0.08	0.10	0.11	0.95	0.92	0.58	0.12	0.07
	$g=2$	0.00	0.02	0.04	0.04	0.05	0.72	0.62	0.17	0.02	0.02
	$g=3$	0.00	0.02	0.06	0.10	0.12	0.52	0.41	0.12	0.03	0.03

- $n_m^{\text{ptrt}}(g)$

D-Sub	$g=1$	8.47	8.37	8.25	8.17	8.07	*1.43*	*2.25*	*7.12*	14.84	15.16
	$g=2$	23.26	22.22	21.33	20.68	20.45	*11.28*	*13.52*	24.27	29.03	28.99
	$g=3$	6.78	6.26	5.70	5.41	5.44	*2.10*	*3.15*	6.64	8.43	8.54
D-Comb	$g=1$	8.98	8.48	8.10	8.12	7.74	*2.49*	*3.22*	*8.59*	13.06	13.32
	$g=2$	23.63	22.09	21.29	20.49	20.48	*6.59*	*8.34*	22.68	34.06	34.43
	$g=3$	6.67	6.23	5.88	5.97	5.86	*1.90*	*2.33*	6.44	9.62	9.65
D-Sep	$g=1$	9.36	8.67	7.96	7.57	7.32	*5.53*	*5.73*	*7.62*	10.25	11.08
	$g=2$	23.79	22.22	21.09	20.58	20.33	*12.86*	*14.68*	23.36	27.84	27.90
	$g=3$	6.91	6.67	6.24	5.68	5.50	*5.76*	*5.99*	6.37	6.48	6.35

Table 10 continued:[**Simulation Results with** $N_{\max} = 180$]. $p^{\text{unacc}}(g, d_m) = \text{P}(\text{declare dose } d_m \text{ unacceptable for subgroup } g)$, $p^{\text{sel}}(g, d_m) = \text{P}(\text{select dose } d_m \text{ as optimal for subgroup } g)$, and n_m^{ptrt} = mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
$\boldsymbol{z}^{\text{TR}} = (1,2,3)$		Scenario 1					Scenario 2				
U^{TR}	$g=1$	**56.89**	51.38	47.69	43.81	39.46	34.34	35.57	34.21	32.38	32.45
	$g=2$	**58.17**	53.25	51.91	49.83	47.67	35.85	38.94	40.73	40.27	41.70
	$g=3$	39.85	35.41	33.63	32.13	30.50	34.59	38.40	41.17	41.94	43.86
p_{T}^{TR}	$g=1$	0.05	0.20	0.35	*0.50*	*0.65*	0.25	0.35	*0.50*	*0.60*	*0.65*
	$g=2$	0.07	0.28	*0.47*	*0.64*	*0.79*	0.32	*0.44*	*0.61*	*0.71*	*0.76*
	$g=3$	0.11	0.39	*0.61*	*0.78*	*0.90*	0.38	*0.51*	*0.68*	*0.78*	*0.82*
π_0^{TR}	$g=1$	0.16	0.16	0.12	0.09	0.07	*0.50*	*0.37*	*0.25*	*0.20*	0.16
	$g=2$	0.25	0.25	0.20	0.16	0.12	*0.57*	*0.43*	0.31	0.25	0.20
	$g=3$	*0.57*	*0.56*	*0.50*	*0.43*	*0.37*	*0.63*	*0.50*	*0.37*	0.31	0.25

- $p_m^{\text{sel}}(g)$

D-Sub	$g=1$	**0.84**	0.06	0.03	0.01	0.03	0.00	0.02	0.03	0.00	0.02
	$g=2$	**0.85**	0.07	0.04	0.01	0.02	0.01	0.02	0.01	0.00	0.00
	$g=3$	0.13	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
D-w/o clustering	$g=1$	**0.80**	0.09	0.05	0.01	0.04	0.01	0.03	0.05	0.02	0.03
	$g=2$	**0.81**	0.10	0.06	0.01	0.01	0.01	0.01	0.01	0.00	0.00
	$g=3$	0.06	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

- $p_m^{\text{unacc}}(g)$

D-Sub	$g=1$	0.06	0.07	0.32	*0.55*	*0.62*	*0.99*	*0.97*	*0.95*	*0.97*	*0.97*
	$g=2$	0.02	0.05	*0.54*	*0.92*	*0.95*	*0.98*	*0.96*	*0.99*	*1.00*	*1.00*
	$g=3$	*0.86*	*0.91*	*0.99*	*1.00*	*1.00*	*1.00*	*0.99*	*1.00*	*1.00*	*1.00*
D-w/o clustering	$g=1$	0.02	0.02	0.14	*0.37*	*0.45*	*0.97*	*0.94*	*0.90*	*0.92*	*0.93*
	$g=2$	0.02	0.04	*0.57*	*0.96*	*0.98*	*0.99*	*0.98*	*0.98*	*1.00*	*1.00*
	$g=3$	*0.94*	*0.97*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*	*1.00*

- $n_m^{\text{ptrt}}(g)$

D-Sub	$g=1$	7.35	7.19	5.31	*3.80*	*3.43*	*1.87*	*2.66*	*2.90*	*2.16*	*2.04*
	$g=2$	25.71	23.94	*13.13*	*4.55*	*3.31*	*8.90*	*8.53*	*4.60*	*1.76*	*1.16*
	$g=3$	*3.88*	*3.06*	*1.30*	*0.30*	*0.08*	*1.67*	*1.81*	*0.66*	*0.16*	*0.05*
D-w/o clustering	$g=1$	6.46	6.58	5.75	*4.56*	*4.09*	*2.31*	*3.27*	*3.59*	*2.84*	*2.60*
	$g=2$	26.24	24.96	*13.11*	*4.09*	*2.63*	*8.08*	*8.33*	*4.73*	*1.65*	*1.09*
	$g=3$	*2.44*	*2.23*	*0.91*	*0.21*	*0.04*	*1.38*	*1.57*	*0.45*	*0.08*	*0.03*

Table 11: [**Simulation Results - Comparison to the design without clustering**]. $p^{\text{unacc}}(g, d_m) = \text{P}(\text{declare dose } d_m \text{ unacceptable for subgroup } g)$, $p^{\text{sel}}(g, d_m) = \text{P}(\text{select dose } d_m \text{ as optimal for subgroup } g)$, and $n_m^{\text{ptrt}} =$ mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g- = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
$\boldsymbol{z}^{\text{TR}} = (1,1,2)$		Scenario 3					Scenario 4				
U^{TR}	$g=1$	59.95	64.84	**72.19**	64.57	58.12	**54.07**	48.75	48.63	40.37	36.04
	$g=2$	64.96	70.83	**79.05**	74.63	70.18	**56.47**	54.15	55.52	49.64	46.55
	$g=3$	56.53	63.35	**73.50**	65.89	60.18	51.47	52.05	**54.98**	50.88	48.50
p_{T}^{TR}	$g=1$	0.05	0.08	0.10	0.35	*0.50*	0.05	0.25	0.30	*0.50*	*0.60*
	$g=2$	0.05	0.08	0.10	0.35	*0.50*	0.05	0.25	0.30	*0.50*	*0.60*
	$g=3$	0.13	0.20	0.25	*0.69*	*0.85*	0.05	0.25	0.30	*0.50*	*0.60*
π_0^{TR}	$g=1$	0.16	0.09	0.03	0.02	0.02	0.16	0.12	0.09	0.09	0.09
	$g=2$	0.16	0.09	0.03	0.02	0.02	0.16	0.12	0.09	0.09	0.09
	$g=3$	0.31	0.20	0.09	0.07	0.07	*0.41*	0.34	0.29	0.29	0.29

- $p_m^{\text{sel}}(g)$

D-Sub	$g=1$	0.01	0.04	**0.87**	0.05	0.03	**0.74**	0.13	0.13	0.00	0.00
	$g=2$	0.00	0.02	**0.85**	0.08	0.05	**0.60**	0.16	0.19	0.01	0.04
	$g=3$	0.02	0.04	**0.85**	0.04	0.01	0.44	0.14	**0.22**	0.01	0.05
D-w/o clustering	$g=1$	0.01	0.03	**0.87**	0.05	0.05	**0.61**	0.22	0.16	0.01	0.01
	$g=2$	0.00	0.02	**0.87**	0.06	0.04	**0.52**	0.21	0.23	0.01	0.03
	$g=3$	0.01	0.04	**0.86**	0.02	0.01	0.34	0.15	**0.20**	0.01	0.02

- $p_m^{\text{unacc}}(g)$

D-Sub	$g=1$	0.01	0.00	0.00	0.10	*0.26*	0.01	0.01	0.07	*0.32*	*0.47*
	$g=2$	0.00	0.00	0.00	0.15	*0.37*	0.01	0.01	0.08	*0.40*	*0.57*
	$g=3$	0.08	0.05	0.06	*0.74*	*0.92*	*0.17*	0.20	0.34	*0.73*	*0.82*
D-w/o clustering	$g=1$	0.00	0.00	0.00	0.05	*0.11*	0.01	0.00	0.04	*0.17*	*0.30*
	$g=2$	0.00	0.00	0.01	0.22	*0.46*	0.01	0.01	0.11	*0.54*	*0.67*
	$g=3$	0.10	0.07	0.08	*0.86*	*0.96*	*0.33*	0.36	0.54	*0.93*	*0.96*

- $n_m^{\text{ptrt}}(g)$

D-Sub	$g=1$	5.83	6.02	5.86	5.22	*4.73*	6.49	6.49	5.91	*4.64*	*4.10*
	$g=2$	15.86	15.87	15.86	13.17	*11.08*	18.29	17.82	15.72	*10.87*	*8.97*
	$g=3$	5.14	5.45	5.40	*2.33*	*1.32*	*5.49*	5.20	4.02	*1.95*	*1.43*
D-w/o clustering	$g=1$	5.55	5.74	5.61	5.52	*5.20*	6.02	6.18	5.83	*5.08*	*4.62*
	$g=2$	16.47	16.45	16.20	12.72	*10.06*	19.37	18.93	16.47	*9.44*	*7.38*
	$g=3$	5.27	5.67	5.62	*1.78*	*0.73*	*5.35*	5.13	3.51	*0.97*	*0.50*

Table 11 continued:[**Simulation Results - Comparison to the design without clustering**]. $p^{\text{unacc}}(g, d_m) = \text{P}(\text{declare dose } d_m \text{ unacceptable for subgroup } g)$, $p^{\text{sel}}(g, d_m) = \text{P}(\text{select dose } d_m \text{ as optimal for subgroup } g)$, and $n_m^{\text{ptrt}} = \text{mean number of patients in subgroup } g$ treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
$\boldsymbol{z}^{\text{TR}} = (1,2,2)$		Scenario 5					Scenario 6				
U^{TR}	$g=1$	**57.64**	48.14	47.08	49.66	52.30	57.62	55.12	53.76	**63.17**	56.94
	$g=2$	65.44	59.25	58.92	63.74	**67.67**	**52.93**	46.64	44.45	52.85	42.68
	$g=3$	70.05	65.27	64.67	70.81	**75.03**	**55.65**	50.80	48.84	59.80	51.27
p_{T}^{TR}	$g=1$	0.03	0.28	0.30	0.35	0.38	0.03	0.10	0.13	0.15	0.30
	$g=2$	0.03	0.28	0.30	0.35	0.38	0.13	0.38	*0.46*	*0.52*	*0.80*
	$g=3$	0.03	0.28	0.30	0.35	0.38	0.13	0.37	*0.46*	*0.52*	*0.80*
π_0^{TR}	$g=1$	0.12	0.12	0.12	0.07	0.04	0.16	0.16	0.16	0.05	0.05
	$g=2$	0.16	0.16	0.16	0.09	0.06	0.20	0.20	0.20	0.07	0.07
	$g=3$	0.16	0.16	0.16	0.09	0.05	0.20	0.20	0.20	0.07	0.07
• $p_m^{\text{sel}}(g)$											
D-Sub	$g=1$	**0.64**	0.01	0.00	0.01	0.34	0.21	0.02	0.03	**0.26**	0.49
	$g=2$	0.28	0.01	0.00	0.02	**0.68**	**0.77**	0.03	0.03	0.14	0.03
	$g=3$	0.15	0.01	0.00	0.02	**0.81**	**0.69**	0.02	0.04	0.12	0.02
D-w/o clustering	$g=1$	**0.52**	0.02	0.02	0.02	0.43	0.19	0.02	0.02	**0.26**	0.51
	$g=2$	0.25	0.02	0.02	0.03	**0.69**	**0.74**	0.03	0.03	0.16	0.03
	$g=3$	0.34	0.02	0.02	0.02	**0.54**	**0.65**	0.02	0.02	0.05	0.00
• $p_m^{\text{unacc}}(g)$											
D-Sub	$g=1$	0.00	0.01	0.02	0.02	0.03	0.00	0.00	0.01	0.02	0.03
	$g=2$	0.00	0.01	0.03	0.04	0.05	0.01	0.10	*0.46*	*0.75*	*0.92*
	$g=3$	0.02	0.04	0.06	0.08	0.09	0.11	0.23	*0.59*	*0.84*	*0.96*
D-w/o clustering	$g=1$	0.00	0.00	0.01	0.02	0.03	0.01	0.01	0.01	0.01	0.02
	$g=2$	0.00	0.01	0.04	0.06	0.07	0.01	0.09	*0.43*	*0.72*	*0.90*
	$g=3$	0.07	0.18	0.30	0.36	0.37	0.26	0.47	*0.81*	*0.93*	*1.00*
• $n_m^{\text{ptrt}}(g)$											
D-Sub	$g=1$	5.63	5.64	5.48	5.40	5.41	5.59	5.58	5.53	5.40	5.29
	$g=2$	15.55	15.00	14.28	13.50	13.44	24.17	20.77	*13.62*	*7.98*	*4.87*
	$g=3$	4.58	4.31	3.93	3.63	3.56	6.91	5.73	*3.22*	*1.49*	*0.65*
D-w/o clustering	$g=1$	5.48	5.67	5.54	5.55	5.30	5.42	5.50	5.48	5.54	5.48
	$g=2$	15.93	15.50	14.24	13.37	12.97	23.78	21.05	*13.68*	*8.05*	*4.91*
	$g=3$	5.84	4.57	3.24	2.61	2.47	6.91	4.74	*2.03*	*0.83*	*0.21*

Table 11 continued:[**Simulation Results - Comparison to the design without clustering**]. $p^{\text{unacc}}(g, d_m) = \text{P}(\text{declare dose } d_m \text{ unacceptable for subgroup } g)$, $p^{\text{sel}}(g, d_m) = \text{P}(\text{select dose } d_m \text{ as optimal for subgroup } g)$, and $n_m^{\text{ptrt}} = \text{mean number of patients in subgroup } g$ treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35,$ and 0.35 for $g = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
$\boldsymbol{z}^{\text{TR}} = (1,1,1)$		Scenario 7					Scenario 8				
U^{TR}	$g=1$	39.68	42.13	46.41	50.73	**54.11**	46.93	**53.95**	46.36	40.27	36.10
	$g=2$	40.33	43.08	48.81	55.18	**60.06**	48.04	**55.63**	52.02	47.72	44.74
	$g=3$	40.83	43.79	50.06	57.99	**63.55**	48.90	**57.16**	55.02	51.73	49.26
p_{T}^{TR}	$g=1$	0.03	0.05	0.10	0.20	0.25	0.05	0.05	0.30	*0.45*	*0.55*
	$g=2$	0.03	0.05	0.10	0.20	0.25	0.05	0.05	0.30	*0.45*	*0.55*
	$g=3$	0.03	0.05	0.10	0.20	0.25	0.05	0.05	0.30	*0.45*	*0.55*
π_0^{TR}	$g=1$	*0.54*	*0.46*	*0.31*	0.16	0.09	0.31	0.16	0.12	0.12	0.12
	$g=2$	*0.54*	*0.46*	0.31	0.16	0.09	0.31	0.16	0.12	0.12	0.12
	$g=3$	*0.54*	*0.46*	0.31	0.16	0.09	0.31	0.16	0.12	0.12	0.12
• $p_m^{\text{sel}}(g)$											
D-Sub	$g=1$	0.00	0.00	0.05	0.17	**0.76**	0.08	**0.82**	0.09	0.00	0.00
	$g=2$	0.00	0.00	0.04	0.11	**0.84**	0.04	**0.76**	0.17	0.00	0.03
	$g=3$	0.00	0.00	0.02	0.10	**0.82**	0.03	**0.68**	0.21	0.01	0.05
D-w/o clustering	$g=1$	0.00	0.00	0.05	0.13	**0.81**	0.05	**0.81**	0.12	0.00	0.01
	$g=2$	0.00	0.00	0.04	0.12	**0.83**	0.03	**0.77**	0.16	0.01	0.03
	$g=3$	0.00	0.00	0.04	0.11	**0.72**	0.03	**0.72**	0.16	0.01	0.02
• $p_m^{\text{unacc}}(g)$											
D-Sub	$g=1$	*0.98*	*0.97*	*0.58*	0.03	0.02	0.27	0.02	0.05	*0.29*	*0.38*
	$g=2$	*0.66*	*0.54*	0.15	0.02	0.02	0.04	0.00	0.07	*0.38*	*0.49*
	$g=3$	*0.77*	*0.68*	0.23	0.06	0.05	0.08	0.02	0.12	*0.51*	*0.60*
D-w/o clustering	$g=1$	*0.95*	*0.90*	*0.38*	0.02	0.01	0.18	0.00	0.03	*0.14*	*0.19*
	$g=2$	*0.76*	*0.65*	0.17	0.03	0.03	0.06	0.01	0.11	*0.48*	*0.58*
	$g=3$	*0.97*	*0.94*	0.46	0.15	0.14	0.23	0.06	0.37	*0.86*	*0.90*
• $n_m^{\text{ptrt}}(g)$											
D-Sub	$g=1$	*1.46*	*2.27*	*5.18*	8.93	9.25	5.43	6.53	6.20	*4.92*	*4.49*
	$g=2$	*8.96*	*10.19*	15.49	18.37	18.43	17.46	18.00	15.96	*11.21*	*9.69*
	$g=3$	*1.65*	*2.47*	4.21	5.49	5.49	4.87	5.50	4.49	*2.55*	*2.07*
D-w/o clustering	$g=1$	*2.00*	*2.90*	*5.98*	8.21	8.41	5.45	6.25	5.83	*5.25*	*5.02*
	$g=2$	*7.64*	*9.32*	16.31	18.98	19.08	18.31	19.25	16.11	*9.88*	*8.40*
	$g=3$	*0.67*	*1.67*	3.92	5.79	5.81	5.26	6.89	4.25	*1.29*	*0.84*

Table 11 continued:[**Simulation Results - Comparison to the design without clustering**]. $p^{\text{unacc}}(g, d_m) = \text{P}(\text{declare dose } d_m \text{ unacceptable for subgroup } g)$, $p^{\text{sel}}(g, d_m) = \text{P}(\text{select dose } d_m \text{ as optimal for subgroup } g)$, and $n_m^{\text{ptrt}} = \text{mean number of patients in subgroup } g$ treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35,$ and 0.35 for $g = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 1 ($\boldsymbol{z}^{\text{TR}} = (1,2,3)$)					Scenario 3 ($\boldsymbol{z}^{\text{TR}} = (1,1,2)$)				
U^{TR}	$g=1$	**56.89**	51.38	47.69	43.81	39.46	59.95	64.84	**72.19**	64.57	58.12
	$g=2$	**58.17**	53.25	51.91	49.83	47.67	64.96	70.83	**79.05**	74.63	70.18
	$g=3$	39.85	35.41	33.63	32.13	30.50	56.53	63.35	**73.50**	65.89	60.18
p_{T}^{TR}	$g=1$	0.05	0.20	0.35	*0.50*	*0.65*	0.05	0.08	0.10	0.35	*0.50*
	$g=2$	0.07	0.28	*0.47*	*0.64*	*0.79*	0.05	0.08	0.10	0.35	*0.50*
	$g=3$	0.11	0.39	*0.62*	*0.79*	*0.90*	0.13	0.20	0.25	*0.69*	*0.85*
π_0^{TR}	$g=1$	0.16	0.16	0.12	0.09	0.07	0.16	0.09	0.03	0.02	0.02
	$g=2$	0.25	0.25	0.20	0.16	0.12	0.16	0.09	0.03	0.02	0.02
	$g=3$	*0.57*	*0.57*	*0.50*	*0.43*	*0.37*	0.31	0.20	0.09	0.07	0.07

- $p_m^{\text{sel}}(g)$

D-Sub	$g=1$	**0.84**	0.06	0.03	0.01	0.03	0.01	0.04	**0.87**	0.05	0.03
	$g=2$	**0.85**	0.07	0.04	0.01	0.02	0.00	0.02	**0.85**	0.08	0.05
	$g=3$	0.13	0.01	0.00	0.00	0.00	0.02	0.04	**0.85**	0.04	0.01
D-w/o frailty	$g=1$	**0.82**	0.11	0.05	0.01	0.01	0.01	0.05	**0.85**	0.06	0.03
	$g=2$	**0.71**	0.13	0.09	0.02	0.05	0.00	0.03	**0.81**	0.10	0.06
	$g=3$	0.66	0.09	0.06	0.01	0.02	0.01	0.04	**0.84**	0.07	0.03

- $p_m^{\text{unacc}}(g)$

D-Sub	$g=1$	0.06	0.07	0.32	*0.55*	*0.62*	0.01	0.00	0.00	0.10	*0.26*
	$g=2$	0.02	0.05	*0.54*	*0.92*	*0.95*	0.00	0.00	0.00	0.15	*0.37*
	$g=3$	*0.86*	*0.91*	*0.99*	*1.00*	*1.00*	0.08	0.05	0.06	*0.74*	*0.92*
D-w/o frailty	$g=1$	0.00	0.00	0.01	*0.02*	*0.08*	0.00	0.00	0.00	0.01	*0.01*
	$g=2$	0.00	0.00	*0.02*	*0.05*	*0.19*	0.00	0.00	0.00	0.02	*0.02*
	$g=3$	*0.16*	*0.23*	*0.58*	*0.84*	*0.90*	0.01	0.01	0.02	*0.21*	*0.35*

- $n_m^{\text{ptrt}}(g)$

D-Sub	$g=1$	7.35	7.19	5.31	*3.80*	*3.43*	5.83	6.02	5.86	5.22	*4.73*
	$g=2$	25.71	23.94	*13.13*	*4.55*	*3.31*	15.86	15.87	15.86	13.17	*11.08*
	$g=3$	*3.88*	*3.06*	*1.30*	*0.30*	*0.08*	5.14	5.45	5.40	*2.33*	*1.32*
D-w/o frailty	$g=1$	5.63	5.64	5.62	*5.47*	*5.23*	5.38	5.61	5.56	5.50	*5.54*
	$g=2$	15.65	15.74	*14.71*	*13.57*	*12.33*	14.51	14.55	14.64	14.15	*14.12*
	$g=3$	*6.39*	*5.49*	*3.08*	*1.47*	*1.09*	4.37	4.54	4.53	*3.61*	*3.08*

Table 12: [**Simulation Results - Comparison to the model without frailties**]. $p^{\text{unacc}}(g, d_m) = \text{P}(\text{declare dose } d_m \text{ unacceptable for subgroup } g)$, $p^{\text{sel}}(g, d_m) = \text{P}(\text{select dose } d_m \text{ as optimal for subgroup } g)$, and $n_m^{\text{ptrt}} =$ mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g- = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 5 ($\boldsymbol{z}^{\text{TR}} = (1,2,2)$)					Scenario 7 ($\boldsymbol{z}^{\text{TR}} = (1,1,1)$)				
U^{TR}	$g=1$	**57.64**	48.14	47.08	49.66	52.30	39.68	42.13	46.41	50.73	**54.11**
	$g=2$	65.44	59.25	58.92	63.74	**67.67**	40.33	43.08	48.81	55.18	**60.06**
	$g=3$	70.05	65.27	64.67	70.81	**75.03**	40.83	43.79	50.06	57.99	**63.55**
p_{T}^{TR}	$g=1$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
	$g=2$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
	$g=3$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
π_0^{TR}	$g=1$	0.12	0.12	0.12	0.07	0.04	*0.54*	*0.46*	*0.31*	0.16	0.09
	$g=2$	0.16	0.16	0.16	0.09	0.05	*0.54*	*0.46*	0.31	0.16	0.09
	$g=3$	0.16	0.16	0.16	0.09	0.05	*0.54*	*0.46*	0.31	0.16	0.09

- $p_m^{\text{sel}}(g)$

D-Sub	$g=1$	**0.64**	0.01	0.00	0.01	0.34	0.00	0.00	0.05	0.17	**0.76**
	$g=2$	0.28	0.01	0.00	0.02	**0.68**	0.00	0.00	0.04	0.11	**0.84**
	$g=3$	0.15	0.01	0.00	0.02	**0.81**	0.00	0.00	0.02	0.10	**0.82**
D-w/o frailty	$g=1$	**0.58**	0.01	0.01	0.02	0.38	0.00	0.00	0.06	0.19	**0.74**
	$g=2$	0.22	0.01	0.00	0.04	**0.73**	0.00	0.00	0.04	0.15	**0.81**
	$g=3$	0.09	0.00	0.00	0.04	**0.86**	0.00	0.00	0.03	0.13	**0.82**

- $p_m^{\text{unacc}}(g)$

D-Sub	$g=1$	0.00	0.01	0.02	0.02	0.03	*0.98*	*0.97*	*0.58*	0.03	0.02
	$g=2$	0.00	0.01	0.03	0.04	0.05	*0.66*	*0.54*	0.15	0.02	0.02
	$g=3$	0.02	0.04	0.06	0.08	0.09	*0.77*	*0.68*	0.23	0.06	0.05
D-w/o frailty	$g=1$	0.00	0.00	0.00	0.00	0.00	*0.00*	*0.00*	*0.00*	0.00	0.00
	$g=2$	0.00	0.00	0.01	0.01	0.01	*0.00*	*0.00*	0.00	0.00	0.01
	$g=3$	0.00	0.00	0.01	0.01	0.01	*0.01*	*0.01*	0.02	0.02	0.02

- $n_m^{\text{ptrt}}(g)$

D-Sub	$g=1$	5.63	5.64	5.48	5.40	5.41	*1.46*	*2.27*	*5.18*	8.93	9.25
	$g=2$	15.55	15.00	14.28	13.50	13.44	*8.96*	*10.19*	15.49	18.37	18.43
	$g=3$	4.58	4.31	3.93	3.63	3.56	*1.65*	*2.47*	4.21	5.49	5.49
D-w/o frailty	$g=1$	5.49	5.61	5.60	5.51	5.46	*5.40*	*5.59*	*5.55*	5.53	5.51
	$g=2$	14.61	14.41	14.32	14.17	14.18	*14.40*	*14.47*	14.43	14.43	14.30
	$g=3$	4.06	4.26	4.08	4.10	4.06	*3.92*	*4.05*	4.05	4.07	4.06

Table 12 continued:[**Simulation Results - Comparison to the model without frailties**]. $p^{\text{unacc}}(g, d_m) = \text{P}(\text{declare dose } d_m \text{ unacceptable for subgroup } g)$, $p^{\text{sel}}(g, d_m) = \text{P}(\text{select dose } d_m \text{ as optimal for subgroup } g)$, and n_m^{ptrt} = mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 1 ($\boldsymbol{z}^{\text{TR}} = (1,2,3)$)					Scenario 3 ($\boldsymbol{z}^{\text{TR}} = (1,1,2)$)				
U^{TR}	$g=1$	**56.89**	51.38	47.69	43.81	39.46	59.95	64.84	**72.19**	64.57	58.12
	$g=2$	**58.17**	53.25	51.91	49.83	47.67	64.96	70.83	**79.05**	74.63	70.18
	$g=3$	39.85	35.41	33.63	32.13	30.50	56.53	63.35	**73.50**	65.89	60.18
p_{T}^{TR}	$g=1$	0.05	0.20	0.35	*0.50*	*0.65*	0.05	0.08	0.10	0.35	*0.50*
	$g=2$	0.07	0.28	*0.47*	*0.64*	*0.79*	0.05	0.08	0.10	0.35	*0.50*
	$g=3$	0.11	0.39	*0.62*	*0.79*	*0.90*	0.13	0.20	0.25	*0.69*	*0.85*
π_0^{TR}	$g=1$	0.16	0.16	0.12	0.09	0.07	0.16	0.09	0.03	0.02	0.02
	$g=2$	0.25	0.25	0.20	0.16	0.12	0.16	0.09	0.03	0.02	0.02
	$g=3$	*0.57*	*0.57*	*0.50*	*0.43*	*0.37*	0.31	0.20	0.09	0.07	0.07

- $p_m^{\text{sel}}(g)$

D-Sub	$g=1$	**0.84**	0.06	0.03	0.01	0.03	0.01	0.04	**0.87**	0.05	0.03
	$g=2$	**0.85**	0.07	0.04	0.01	0.02	0.00	0.02	**0.85**	0.08	0.05
	$g=3$	0.13	0.01	0.00	0.00	0.00	0.02	0.04	**0.85**	0.04	0.01
D-Diff AR	$g=1$	**0.84**	0.05	0.03	0.01	0.03	0.01	0.05	**0.87**	0.04	0.03
	$g=2$	**0.83**	0.08	0.05	0.01	0.03	0.00	0.02	**0.86**	0.06	0.05
	$g=3$	0.14	0.01	0.00	0.00	0.00	0.01	0.05	**0.87**	0.03	0.01

- $p_m^{\text{unacc}}(g)$

D-Sub	$g=1$	0.06	0.07	0.32	*0.55*	*0.62*	0.01	0.00	0.00	0.10	*0.26*
	$g=2$	0.02	0.05	*0.54*	*0.92*	*0.95*	0.00	0.00	0.00	0.15	*0.37*
	$g=3$	*0.86*	*0.91*	*0.99*	*1.00*	*1.00*	0.08	0.05	0.06	*0.74*	*0.92*
D-Diff AR	$g=1$	0.06	0.06	0.30	*0.56*	*0.62*	0.00	0.00	0.00	0.11	*0.27*
	$g=2$	0.02	0.05	*0.52*	*0.88*	*0.92*	0.00	0.00	0.00	0.17	*0.38*
	$g=3$	*0.85*	*0.89*	*0.98*	*1.00*	*1.00*	0.06	0.04	0.05	*0.72*	*0.92*

- $n_m^{\text{ptrt}}(g)$

D-Sub	$g=1$	7.35	7.19	5.31	*3.80*	*3.43*	5.83	6.02	5.86	5.22	*4.73*
	$g=2$	25.71	23.94	*13.13*	*4.55*	*3.31*	15.86	15.87	15.86	13.17	*11.08*
	$g=3$	*3.88*	*3.06*	*1.30*	*0.30*	*0.08*	5.14	5.45	5.40	*2.33*	*1.32*
D-Diff AR	$g=1$	7.18	7.53	5.32	*3.58*	*3.44*	4.95	6.17	6.52	5.37	*4.77*
	$g=2$	25.71	24.52	*13.01*	*4.39*	*3.18*	14.45	15.80	17.24	13.41	*10.78*
	$g=3$	*3.67*	*3.22*	*1.23*	*0.28*	*0.09*	4.32	5.64	6.09	*2.42*	*1.32*

Table 13: [**Simulation Results - Comparison to the design with different AR**]. $p^{\text{unacc}}(g, d_m) = \text{P}(\text{declare dose } d_m \text{ unacceptable for subgroup } g)$, $p^{\text{sel}}(g, d_m) = \text{P}(\text{select dose } d_m \text{ as optimal for subgroup } g)$, and $n_m^{\text{ptrt}} =$ mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g- = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 5 ($\boldsymbol{z}^{\text{TR}} = (1,2,2)$)					Scenario 7 ($\boldsymbol{z}^{\text{TR}} = (1,1,1)$)				
U^{TR}	$g=1$	**57.64**	48.14	47.08	49.66	52.30	39.68	42.13	46.41	50.73	**54.11**
	$g=2$	65.44	59.25	58.92	63.74	**67.67**	40.33	43.08	48.81	55.18	**60.06**
	$g=3$	70.05	65.27	64.67	70.81	**75.03**	40.83	43.79	50.06	57.99	**63.55**
p_{T}^{TR}	$g=1$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
	$g=2$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
	$g=3$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
π_0^{TR}	$g=1$	0.12	0.12	0.12	0.07	0.04	*0.54*	*0.46*	*0.31*	0.16	0.09
	$g=2$	0.16	0.16	0.16	0.09	0.05	*0.54*	*0.46*	0.31	0.16	0.09
	$g=3$	0.16	0.16	0.16	0.09	0.05	*0.54*	*0.46*	0.31	0.16	0.09

- $p_m^{\text{sel}}(g)$

D-Sub	$g=1$	**0.64**	0.01	0.00	0.01	0.34	0.00	0.00	0.05	0.17	**0.76**
	$g=2$	0.28	0.01	0.00	0.02	**0.68**	0.00	0.00	0.04	0.11	**0.84**
	$g=3$	0.15	0.01	0.00	0.02	**0.81**	0.00	0.00	0.02	0.10	**0.82**
D-Diff AR	$g=1$	**0.59**	0.01	0.01	0.02	0.37	0.00	0.00	0.06	0.16	**0.77**
	$g=2$	0.26	0.01	0.01	0.03	**0.70**	0.00	0.00	0.04	0.12	**0.84**
	$g=3$	0.16	0.00	0.01	0.03	**0.79**	0.00	0.00	0.02	0.11	**0.81**

- $p_m^{\text{unacc}}(g)$

D-Sub	$g=1$	0.00	0.01	0.02	0.02	0.03	*0.98*	*0.97*	*0.58*	0.03	0.02
	$g=2$	0.00	0.01	0.03	0.04	0.05	*0.66*	*0.54*	0.15	0.02	0.02
	$g=3$	0.02	0.04	0.06	0.08	0.09	*0.77*	*0.68*	0.23	0.06	0.05
D-Diff AR	$g=1$	0.01	0.01	0.02	0.03	0.03	*0.98*	*0.96*	*0.57*	0.03	0.02
	$g=2$	0.00	0.01	0.03	0.05	0.05	*0.62*	*0.52*	0.16	0.02	0.02
	$g=3$	0.02	0.04	0.07	0.09	0.09	*0.74*	*0.65*	0.25	0.06	0.05

- $n_m^{\text{ptrt}}(g)$

D-Sub	$g=1$	5.63	5.64	5.48	5.40	5.41	*1.46*	*2.27*	*5.18*	8.93	9.25
	$g=2$	15.55	15.00	14.28	13.50	13.44	*8.96*	*10.19*	15.49	18.37	18.43
	$g=3$	4.58	4.31	3.93	3.63	3.56	*1.65*	*2.47*	4.21	5.49	5.49
D-Diff AR	$g=1$	4.97	5.64	5.40	5.53	5.72	*1.04*	*2.33*	*4.94*	9.20	10.01
	$g=2$	14.95	14.92	14.12	13.77	14.39	*7.25*	*9.41*	14.89	19.19	20.59
	$g=3$	4.21	4.41	3.78	3.68	3.89	*1.16*	*2.34*	4.04	5.59	5.87

Table 13 continued:[**Simulation Results - Comparison to the design with different AR**]. $p^{\text{unacc}}(g, d_m) = \text{P}(\text{declare dose } d_m \text{ unacceptable for subgroup } g)$, $p^{\text{sel}}(g, d_m) = \text{P}(\text{select dose } d_m \text{ as optimal for subgroup } g)$, and n_m^{ptrt} = mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 1 ($\boldsymbol{z}^{\text{TR}} = (1,2,3)$)					Scenario 3 ($\boldsymbol{z}^{\text{TR}} = (1,1,2)$)				
U^{TR}	$g=1$	**56.89**	51.38	47.69	43.81	39.46	59.95	64.84	**72.19**	64.57	58.12
	$g=2$	**58.17**	53.25	51.91	49.83	47.67	64.96	70.83	**79.05**	74.63	70.18
	$g=3$	39.85	35.41	33.63	32.13	30.50	56.53	63.35	**73.50**	65.89	60.18
p_{T}^{TR}	$g=1$	0.05	0.20	0.35	*0.50*	*0.65*	0.05	0.08	0.10	0.35	*0.50*
	$g=2$	0.07	0.28	*0.47*	*0.64*	*0.79*	0.05	0.08	0.10	0.35	*0.50*
	$g=3$	0.11	0.39	*0.62*	*0.79*	*0.90*	0.13	0.20	0.25	*0.69*	*0.85*
π_0^{TR}	$g=1$	0.16	0.16	0.12	0.09	0.07	0.16	0.09	0.03	0.02	0.02
	$g=2$	0.25	0.25	0.20	0.16	0.12	0.16	0.09	0.03	0.02	0.02
	$g=3$	*0.57*	*0.57*	*0.50*	*0.43*	*0.37*	0.31	0.20	0.09	0.07	0.07

- $p_m^{\text{sel}}(g)$

D-Sub	$g=1$	**0.84**	0.06	0.03	0.01	0.03	0.01	0.04	**0.87**	0.05	0.03
	$g=2$	**0.85**	0.07	0.04	0.01	0.02	0.00	0.02	**0.85**	0.08	0.05
	$g=3$	0.13	0.01	0.00	0.00	0.00	0.02	0.04	**0.85**	0.04	0.01
D-Linear Tox	$g=1$	**0.87**	0.02	0.05	0.04	0.00	0.03	0.04	**0.68**	0.25	0.00
	$g=2$	**0.88**	0.03	0.06	0.01	0.00	0.01	0.03	**0.65**	0.32	0.00
	$g=3$	0.10	0.00	0.01	0.00	0.00	0.06	0.05	**0.69**	0.17	0.00

- $p_m^{\text{unacc}}(g)$

D-Sub	$g=1$	0.06	0.07	0.32	*0.55*	*0.62*	0.01	0.00	0.00	0.10	*0.26*
	$g=2$	0.02	0.05	*0.54*	*0.92*	*0.95*	0.00	0.00	0.00	0.15	*0.37*
	$g=3$	*0.86*	*0.91*	*0.99*	*1.00*	*1.00*	0.08	0.05	0.06	*0.74*	*0.92*
D-Linear Tox	$g=1$	0.04	0.03	0.13	*0.53*	*0.96*	0.00	0.00	0.00	0.01	*0.56*
	$g=2$	0.02	0.02	*0.37*	*0.93*	*1.00*	0.00	0.00	0.00	0.02	*0.75*
	$g=3$	*0.90*	*0.93*	*0.99*	*1.00*	*1.00*	0.05	0.04	0.08	*0.43*	*0.98*

- $n_m^{\text{ptrt}}(g)$

D-Sub	$g=1$	7.35	7.19	5.31	*3.80*	*3.43*	5.83	6.02	5.86	5.22	*4.73*
	$g=2$	25.71	23.94	*13.13*	*4.55*	*3.31*	15.86	15.87	15.86	13.17	*11.08*
	$g=3$	*3.88*	*3.06*	*1.30*	*0.30*	*0.08*	5.14	5.45	5.40	*2.33*	*1.32*
D-Linear Tox	$g=1$	7.47	7.49	6.75	*4.25*	*1.25*	6.10	6.28	6.10	5.59	*3.42*
	$g=2$	25.43	25.21	*15.73*	*3.98*	*0.68*	16.91	16.95	16.87	14.98	*6.45*
	$g=3$	*3.91*	*3.23*	*1.07*	*0.17*	*0.02*	5.26	5.37	5.10	*3.36*	*0.63*

Table 14: [**Simulation Results - Comparison to the model with linear regression for toxicity outcome**]. $p^{\text{unacc}}(g, d_m)$ = P(declare dose d_m unacceptable for subgroup g), $p^{\text{sel}}(g, d_m)$ = P(select dose d_m as optimal for subgroup g), and n_m^{ptrt} = mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g- = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.

		d_1	d_2	d_3	d_4	d_5	d_1	d_2	d_3	d_4	d_5
		Scenario 5 ($\boldsymbol{z}^{\text{TR}} = (1,2,2)$)					Scenario 7 ($\boldsymbol{z}^{\text{TR}} = (1,1,1)$)				
U^{TR}	$g=1$	**57.64**	48.14	47.08	49.66	52.30	39.68	42.13	46.41	50.73	**54.11**
	$g=2$	65.44	59.25	58.92	63.74	**67.67**	40.33	43.08	48.81	55.18	**60.06**
	$g=3$	70.05	65.27	64.67	70.81	**75.03**	40.83	43.79	50.06	57.99	**63.55**
p_{T}^{TR}	$g=1$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
	$g=2$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
	$g=3$	0.03	0.28	0.30	0.35	0.38	0.03	0.05	0.10	0.20	0.25
π_0^{TR}	$g=1$	0.12	0.12	0.12	0.07	0.04	*0.54*	*0.46*	*0.31*	0.16	0.09
	$g=2$	0.16	0.16	0.16	0.09	0.05	*0.54*	*0.46*	0.31	0.16	0.09
	$g=3$	0.16	0.16	0.16	0.09	0.05	*0.54*	*0.46*	0.31	0.16	0.09

- $p_m^{\text{sel}}(g)$

D-Sub	$g=1$	**0.64**	0.01	0.00	0.01	0.34	0.00	0.00	0.05	0.17	**0.76**
	$g=2$	0.28	0.01	0.00	0.02	**0.68**	0.00	0.00	0.04	0.11	**0.84**
	$g=3$	0.15	0.01	0.00	0.02	**0.81**	0.00	0.00	0.02	0.10	**0.82**
D-Linear Tox	$g=1$	**0.58**	0.02	0.03	0.25	0.12	0.00	0.00	0.08	0.46	**0.22**
	$g=2$	0.43	0.01	0.03	0.33	**0.19**	0.00	0.01	0.12	0.42	**0.27**
	$g=3$	0.38	0.01	0.03	0.36	**0.20**	0.00	0.01	0.09	0.35	**0.30**

- $p_m^{\text{unacc}}(g)$

D-Sub	$g=1$	0.00	0.01	0.02	0.02	0.03	*0.98*	*0.97*	*0.58*	0.03	0.02
	$g=2$	0.00	0.01	0.03	0.04	0.05	*0.66*	*0.54*	0.15	0.02	0.02
	$g=3$	0.02	0.04	0.06	0.08	0.09	*0.77*	*0.68*	0.23	0.06	0.05
D-Linear Tox	$g=1$	0.00	0.00	0.05	0.13	0.39	*0.98*	*0.97*	*0.58*	0.30	0.33
	$g=2$	0.00	0.00	0.09	0.20	0.53	*0.72*	*0.63*	0.25	0.30	0.34
	$g=3$	0.03	0.03	0.13	0.25	0.60	*0.83*	*0.77*	0.36	0.33	0.38

- $n_m^{\text{ptrt}}(g)$

D-Sub	$g=1$	5.63	5.64	5.48	5.40	5.41	*1.46*	*2.27*	*5.18*	8.93	9.25
	$g=2$	15.55	15.00	14.28	13.50	13.44	*8.96*	*10.19*	15.49	18.37	18.43
	$g=3$	4.58	4.31	3.93	3.63	3.56	*1.65*	*2.47*	4.21	5.49	5.49
D-Linear Tox	$g=1$	6.27	6.50	5.86	5.05	3.87	*2.10*	*2.83*	*6.71*	5.72	5.51
	$g=2$	18.26	18.36	15.15	12.30	8.13	*9.34*	*10.84*	19.06	11.84	11.10
	$g=3$	5.11	5.33	4.06	3.22	1.86	*1.57*	*2.47*	5.06	3.50	3.09

Table 14 continued:[**Simulation Results - Comparison to the model with linear regression for toxicity outcome**]. $p^{\text{unacc}}(g, d_m)$ = P(declare dose d_m unacceptable for subgroup g), $p^{\text{sel}}(g, d_m)$ = P(select dose d_m as optimal for subgroup g), and n_m^{ptrt} = mean number of patients in subgroup g treated at dose d_m. Values for *True unacceptable* and **true optimal** doses are given in *red italics* and **blue bold**, respectively. $\bar{\zeta}_{\text{PD}}(g) = 0.20, 0.35$, and 0.35 for $g = 1, 2, 3$ and $\bar{\zeta}_{\text{T}}(g) = 0.40$ for all subgroups.