Biometrics , 1–20

Precision Generalized Phase I-II Designs

Saijun Zhao[1], Peter F. Thall[2,*], Ying Yuan[2], Juhee Lee[3], Pavlos Msaouel[4], and Yong Zang[5,]**

[1]Department of Biostatistics and Health Data Science,

School of Medicine, Indiana University, Indianapolis, IN

[2]Department of Biostatistics, M.D. Anderson Cancer Center, Houston, TX

[3]Department of Statistics, University of California Santa Cruz, Santa Cruz, CA

[4]Departments of Genitourinary Medical Oncology and Translational Molecular Pathology,

M.D. Anderson Cancer Center, Houston, TX

[5]Department of Biostatistics and Health Data Science;

Center for Computational Biology and Bioinformatics,

Indiana University School of Medicine, Indianapolis, IN

*_email:_ rex@mdanderson.org

**_email:_ zangy@iu.edu

Summary: A new family of precision Bayesian dose optimization designs, PGen I-II, based on early efficacy, early toxicity, and long-term time to treatment failure is proposed. A PGen I-II design refines a Gen I-II design by accounting for patient heterogeneity characterized by subgroups that may be defined by prognostic levels, disease subtypes, or biomarker categories. The design makes subgroup-specific decisions, which may be to drop an unacceptably toxic or inefficacious dose, randomize patients among acceptable doses, or identify a best dose in terms of treatment success defined in terms of time to failure over long-term follow up. A piecewise exponential distribution for failure time is assumed, including subgroup-specific effects of dose, response, and toxicity. Latent variables are used to adaptively cluster subgroups found to have similar dose-outcome distributions, with the model simplified to borrow strength between subgroups in the same cluster. Guidelines and user-friendly computer software for implementing the design are provided. A simulation study is reported that shows the PGen I-II design is superior to similarly structured designs that either assume patient homogeneity or conduct separate trials within subgroups.

Key words: Bayesian Adaptive Design; Cell Therapy; Dose Optimization; Precision Medicine; Phase I-II Clinical Trials.

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

1. Introduction

We propose a new family of Bayesian dose-finding designs, based on early toxicity and efficacy and also a long-term outcome, that accounts for patient heterogeneity characterized by prespecified subgroups. The design, which we call precision generalized phase I-II (PGen I-II), extends the Bayesian generalized phase I-II (Gen I-II) design of Thall et al. (2023) by including dose-subgroup interactions in the underlying model and making subgroup-specific decisions. Subgroups may be defined by prognostic levels, disease subtypes, or biomarker categories. For each subgroup, a decision may be to drop an unacceptably toxic or ineffective dose, assign a patient to an acceptable dose, or identify a best dose. The design uses latent variables to adaptively cluster subgroups shown by interim data to have similar dose-outcome distributions. If subgroups are clustered, the model is simplified so that dose effect parameters are identical for all subgroups in each cluster, which improves the design's reliability by borrowing strength between subgroups. During the trial or at its conclusion, the design may identify different optimal doses for different subgroups.

Most dose-finding/optimization designs are based on clinical outcomes evaluated over a short follow up period, $[0,\ t_1]$, chosen to accommodate one or two courses of therapy or to evaluate biological variables, with t_1 often 30 to 60 days. Most phase I designs (Babb et al., 1998) use a short-term toxicity outcome, Y_T. Most phase I-II designs (Gooley, 1994; Thall and Russell, 1998; Thall and Cook, 2004), are based on both Y_T and short term efficacy, Y_E, which may be clinical response or a biological outcome. Depending on the setting, Y_T and Y_E may be a binary, ordinal, or a time-to-event variable, scored either during $[0,\ t_1]$ or at t_1. Using early $\boldsymbol{Y}_{ET} = (Y_E, Y_T)$ to evaluate doses is done to avoid unduly delaying accrual of future patients to evaluate recently treated patients' outcomes for adaptive decision making. Denoting dose by d, a phase I-II design uses $(d, \boldsymbol{Y}_{ET})$ data to screen out overly toxic or ineffective doses, and choose a best dose based on a criterion, $\phi_{ET}(d, \boldsymbol{\theta})$, defined in terms of

a bivariate probability model $p(\boldsymbol{y}_{ET} \mid d, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the parameter vector. Given a utility function $U(\boldsymbol{y}_{ET})$ elicited for all values $\boldsymbol{y}_{ET}$ of $\boldsymbol{Y}_{ET}$, the optimality criterion may be the mean utility $\phi_{ET}(d, \boldsymbol{\theta}) = E\{U(\boldsymbol{Y}_{ET}) \mid d, \boldsymbol{\theta}\}$ (Thall et al., 2012; Lin et al., 2020). For binary Y_T and Y_E, optimality criteria may be the response probability $\pi_E(d, \boldsymbol{\theta}) = \Pr(Y_E = 1 \mid d, \boldsymbol{\theta})$, a trade-off function of $\pi_E(d, \boldsymbol{\theta})$ and $\pi_T(d, \boldsymbol{\theta}) = \Pr(Y_T = 1 \mid d, \boldsymbol{\theta})$ used by the EffTox design (Thall and Cook, 2004), or a linear combination $\pi_E(d, \boldsymbol{\theta}) + \lambda\, \pi_T(d, \boldsymbol{\theta})$ with $\lambda < 0$.

A limitation of conventional phase I-II designs is that their decision criteria ignore long term outcomes, such as remission duration, progression-free survival (PFS) time or overall survival (OS) time. Let Y_S denote a time-to-event variable, evaluated over a longer follow up period $[0,\ t_2]$, where t_2 is a time, such as 6 or 12 months, substantially larger than t_1. Using $\boldsymbol{Y}_{ET}$ evaluated by short follow up time t_1 to characterize dose effects relies on the assumption that $\boldsymbol{Y}_{ET}$ is a surrogate for Y_S. In practice, while $\boldsymbol{Y}_{ET}$ and Y_S may be related, perfect surrogacy never holds. An undesirable consequence is that a dose optimizing criterion, $\phi_{ET}(d, \boldsymbol{\theta})$, defined in terms of $\boldsymbol{Y}_{ET}$ evaluated over $[0,\ t_1]$ may lead to choosing a dose that is suboptimal in terms of a longer term clinical success criterion, $\phi_S(d, \boldsymbol{\theta})$, based on Y_S evaluated up to time t_2. For example, an immunotherapy may achieve a high 30-day response rate at an optimal phase I-II dose, but responders may have a high six-month relapse rate.

To see how $\boldsymbol{Y}_{ET}$ and Y_S may disagree, consider doses d_1 and d_2, and assume that both doses are safe, so that Y_T may be ignored. Denote conditional 12-month PFS probabilities by $\overline{F}(d_j, y_E) = \Pr(Y_S > 12 \mid d_j, Y_E = y_E)$ for $j = 1, 2$ and $y_E = 0, 1$. Assume true probabilities $\pi_E^{true}(d_1) = .20$, $\pi_E^{true}(d_2) = .40$, $\overline{F}^{true}(d_1, 1) = .95$, $\overline{F}^{true}(d_1, 0) = .85$, $\overline{F}^{true}(d_2, 1) = .75$, and $\overline{F}^{true}(d_2, 0) = .65$. Averaging over $y_E = 0$ or 1 gives the unconditional 12-month PFS probabilities $\overline{F}^{true}(d_1) = .87$ and $\overline{F}^{true}(d_2) = .69$. Thus, while response is associated with a higher 12-month PFS probability for both doses, and d_2 has twice the response rate of d_1, in terms of unconditional 12-month PFS probabilities d_1 is greatly superior to d_2.

To address this problem, Thall et al. (2023) proposed the family of Bayesian Gen I-II designs, based on early Y_E and Y_T evaluated over $[0,\ t_1]$, and a time-to-failure outcome, Y_S, evaluated over $[0,\ t_2]$ for $t_2 > t_1$. For example, investigators may specify $t_1 = 30$ days and $t_2 = 180$ days for Y_S = PFS time. A Gen I-II design first uses a phase I-II design to choose doses for successive cohorts and screen out unacceptably toxic or inefficacious doses. The optimality criterion $\phi_{ET}(d, \boldsymbol{\theta})$, defined in terms of (d, Y_E, Y_T), is used to identify a set $\mathcal{C}$ of *candidate doses*, rather than selecting one dose. Additional patients are randomized among doses in $\mathcal{C}$, all patients are followed to time t_2 to obtain Y_S data, and a final dose is selected to maximize an optimality criterion $\phi_S(d, \boldsymbol{\theta})$ defined in terms of Y_S. Parametric models are assumed for $p(Y_S \mid Y_T, Y_E, d)$ and $p(Y_T, Y_E \mid d)$, including parameters based on $\boldsymbol{Y}_{ET}$ for the early phase I-II stages of the design, and effects of Y_E, Y_T, and d on Y_S for the final dose selection criterion $\phi_S(d, \boldsymbol{\theta})$. Our motivating example is the trial described by Thall et al. (2023) for the Gen I-II design, to optimize the dose of chimeric antigen receptor cord blood-derived CD70 natural killer (NK) cells given as immunotherapy for advanced hematologic malignancies. The four doses are $5.0 \times (10^6, 10^7, 10^8, 10^9)$ NK cells. We extend the Gen I-II design by constructing a PGen I-II design that accounts for the three disease subgroups given in Table S1 in the online supplementary materials, with the goal to choose optimal subgroup-specific NK cell doses.

Thall et al. (2023) assumed that the distributions of Y_E, Y_T, and Y_S as functions of NK cell dose were homogeneous across disease subgroups. In the present paper, using the CAR-NK cell trial, we extend the Gen I-II framework to a family of PGen I-II designs that make subgroup-specific decisions. For the PGen I-II design, the assumed models include dose-subgroup interactions to facilitate subgroup-specific decisions, including safety and efficacy monitoring, identifying a set of candidate doses, and selecting an optimal dose. A PGen I-II design has three stages, illustrated by Figure 1. In stage 1, an acceptable dose set is identified

for each subgroup based on early outcomes. Stage 2 identifies subgroup-specific candidate dose sets using both early and long-term outcomes. In stage 3, an optimal dose is determined for each subgroup based on the estimated long-term benefit criterion.

Thus, like the Gen I-II design, the PGen I-II design exploits data on the long term outcome Y_S to formulate a dose optimality criterion. Unlike the Gen I-II design, the PGen I-II design accounts for subgroups. Many precision phase I-II designs have been proposed that rely on early outcomes but ignore long term treatment success (Lee et al., 2021; McGovern et al., 2022; Park and Chang, 2024; Porter et al., 2024). Like precision phase I-II designs, the PGen I-II design makes subgroup-specific decisions, but unlike them it also uses data on the long term outcome Y_S. Consequently, the Gen I-II design and precision phase I-II designs may be considered competitors to PGen I-II, but for different reasons.

2. Dose-Outcome Model

Let $Y_j \in \{0, \cdots, L_j - 1\}$ denote ordinal efficacy ($j = E$) and toxicity ($j = T$) evaluated over follow interval $[0,\ t_1]$. For example, Y_E may take possible values 0 for progressive disease (PD), 1 for stable disease (SD), 2 for partial response (PR), or 3 for complete response (CR), with $L_E = 4$. Toxicity may be defined as the most severe grade up to time t_1, with possible values 0 for no toxicity, 1 for mild, 2 for moderate, 3 for severe, and 4 for life threatening or fatal, with $L_T = 5$. To stabilize computations, we replace raw dose values d_j^{raw} by the standardized values $d_j = \log(d_j^{raw})\ / \max_j \left\{\log(d_j^{raw})\right\}$ for dose levels $j = 1, \cdots, J$.

Let $g \in \{1, \cdots, G\}$ index pre-specified subgroups. Similarly to Lee et al. (2021), McGovern et al. (2022), and others, we adaptively cluster similar subgroups using latent cluster membership variables $\boldsymbol{z} = (z_1, \cdots, z_G)$. Fixing $z_1 = 1$, for each $g \geqslant 2$, if there is a subgroup $j \in \{1, \ldots, g-1\}$ in the same cluster as g, we set $z_g = z_j$, and otherwise $z_g = \max_{i \in \{1,\ldots,g-1\}} z_i + 1$. If $z_g = z_{g'}$ for $g \neq g'$, subgroups g and g' are combined in the same cluster, and the model is simplified so that dose effect parameters in subgroups g and g' are identical. This improves

the design's reliability. For example, if $G = 3$ and subgroups g=1 and g=2 are combined but $g = 3$ is distinct, the subgroup clusters {1,2} and {3} are identified by $\boldsymbol{z} = (1, 1, 2)$, and all subgroup-specific parameters indexed by $g = 1$ and $g = 2$ are identical.

We construct a parametric model for $p(Y_E, Y_T \mid d, g)$ by applying the widely used multivariate probit latent variable approach (Chib and Greenberg, 1998). Let $\boldsymbol{X} = (X_E, X_T)$ be real-valued latent variables following bivariate normal distribution $N_2((\mu_E(d,g),\ \mu_T(d,g)), \Sigma_X)$, where Σ_X is $2\times$ 2 with variances σ_{11}, σ_{22} and covariance σ_{12}. To ensure identifiability, we set $\sigma_{11} = \sigma_{22} = 1$ and constrain $-1 \leqslant \sigma_{12} \leqslant 1$. The latent variables $\boldsymbol{X}$ are used to define observed (Y_E, Y_T) as follows. Given real-valued cutpoints $-\infty = \eta_0 < \eta_1 = 0 < \eta_2 \cdots < \eta_{L_E-1} < \eta_{L_E} = +\infty$ and $-\infty = \zeta_0 < \zeta_1 = 0 < \zeta_2 < \cdots < \zeta_{L_T-1} < \zeta_{L_T} = +\infty$, we define

$$Y_E = \begin{cases} 0 & \text{if } \eta_0 < X_E < \eta_1 \\ 1 & \text{if } \eta_1 \leqslant X_E < \eta_2 \\ \cdots & \\ L_E - 1 & \text{if } \eta_{L_E-1} \leqslant X_E < \eta_{L_E} \end{cases}, \quad Y_T = \begin{cases} 0 & \text{if } \zeta_0 < X_T < \zeta_1 \\ 1 & \text{if } \zeta_1 \leqslant X_T < \zeta_2 \\ \cdots & \\ L_T - 1 & \text{if } \zeta_{L_T-1} \leqslant X_E < \zeta_{L_T} \end{cases}.$$

Denoting the pdf of $\boldsymbol{X}$ by $f_{\boldsymbol{X}}(x_E, x_T \mid d, g)$, the induced distribution of (Y_E, Y_T) is

$$\begin{aligned} \pi(y_E, y_T \mid d, g, \boldsymbol{\theta}_{ET}) &=_{def} \Pr(Y_E = y_E, Y_T = y_T \mid d, g, \boldsymbol{\theta}_{ET}) \\ &= \int_{\eta_{y_E}}^{\eta_{y_E+1}} \int_{\zeta_{y_T}}^{\zeta_{y_T+1}} f_{\boldsymbol{X}}(x_E, x_T \mid d, g) dx_E dx_T, \end{aligned} \tag{1}$$

for all outcomes (y_E, y_T), where $\boldsymbol{\theta}_{ET}$ denotes the model parameters.

For efficacy, we assume a flexible Emax model,

$$E(X_E \mid d, g) = \mu_E(d, g) = \alpha_{0,g} + \frac{\alpha_{1,g} d^{\alpha_{3,g}}}{\alpha_{2,g}^{\alpha_{3,g}} + d^{\alpha_{3,g}}}$$

with all parameters real-valued, and denote $\boldsymbol{\alpha}_g = (\alpha_{0,g}, \alpha_{1,g}, \alpha_{2,g}, \alpha_{3,g})$ and $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_G)$. This model includes the case where $\mu_E(d, g)$ as a function of d increases to a plateau, which may arise if an administered dose reaches a maximum biological effectiveness due to saturation of the metabolized agent in the patient's body. For toxicity, we assume $E(X_T \mid d, g) =$

$\mu_T(d,g) = \beta_{0,g} + \beta_{1,g}d$, with $\beta_{0,g}$ real-valued and $\beta_{1,g} > 0$ to ensure that $\mu_T(d,g)$ increases with d, and denote $\boldsymbol{\beta}_g = (\beta_{0,g}, \beta_{1,g})$ and $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_G)$. Denote the i^{th} patient's dose by $d_{[i]}$ and subgroup by g_i. The dataset of early outcomes and doses for the first n patients is $\mathcal{D}_{ET}^n = \{(Y_{E,i}, Y_{T,i}, d_{[i]}, g_i) : \ i = 1, \cdots, n\}$, with likelihood $\mathcal{L}(\mathcal{D}_{ET}^n \mid \boldsymbol{\theta}_{ET}) = \prod_{i=1}^n \pi(y_{E,i}, y_{T,i} \mid d_{[i]}, g_i, \boldsymbol{\theta}_{ET})$, where $\boldsymbol{\theta}_{ET} = (\sigma_{12}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \eta_1, \cdots, \eta_{L_E-1}, \zeta_1, \cdots, \zeta_{L_T-1})$.

The time to failure, Y_S, over follow-up period $[0, t_2]$ may be PFS time, OS time, or remission duration among responders. For simplicity, henceforth we will use PFS time for Y_S. If $[Y_E = 0]$ = PD, then treatment has failed by time t_1. Since following a patient beyond t_1 to evaluate Y_S is only meaningful clinically if the patient is alive without PD at t_1, we define $Y_S' = (Y_S - t_1)\, I(Y_E > 0)$ as the treatment failure time starting from t_1. If $Y_E = 0$, we define $Y_S' = 0$. Given $0 < t_1 < t_2$, we define the long-term dose optimality criterion to be

$$\begin{aligned}\phi_S(d,g) = \Pr(Y_S > t_2 \mid d, g) &= \Pr(Y_S' > t_2 - t_1 \mid d, g) \\ &= \sum_{y=1}^{L_E-1} \Pr(Y_E = y \mid d, g) \Pr(Y_S' > t_2 - t_1 \mid d, g, Y_E = y).\end{aligned}$$

Given $Y_E > 0$, we denote the observed time to failure or right censoring, starting from t_1, by $Y_S'^o$, with $\delta = \mathrm{I}(Y_S' = Y_S'^o)$. For robustness, we assume that Y_S' follows a piecewise exponential (PE) distribution, defined as follows. For sample size n, denote $T_n = \max\{Y_{S,i}'^o, \ \ i = 1, \cdots, n\}$ and $m_n = \sum_{i=1}^n \delta_i$, the number of observed failures. We use the $\frac{1}{K}, \frac{2}{K}, \cdots, \frac{K-1}{K}$ quantiles of the m_n observed failure times as cutpoints to partition the observed failure time domain $[0, \ T_n]$ into K subintervals, $[0, \ c_1), [c_1, c_2), \cdots, [c_{K-1}, c_K]$, where $c_K = T_n$. Thus, the PE interval cutpoints $\boldsymbol{c} = (c_1, \cdots, c_K)$ are sample statistics that may change as the $(Y_{S,i}'^o, \delta_i)$ data accumulate during the trial. Writing Y_E^+ to represent values of $Y_E > 0$ (no PD) and $\boldsymbol{Y}_{ET}^+ = (Y_E^+, Y_T)$, the PE hazard function of Y_S' is

$$h_S(t \mid Y_E^+, Y_T, d, g, \boldsymbol{\theta}_S) = \Big\{\sum_{k=1}^K \lambda_k \, I(c_{k-1} \leqslant t < c_k)\Big\} \times \exp\big\{\gamma_{E,Y_E^+,g} + \gamma_{T,Y_T,g} + \sum_{j=1}^J \gamma_{D,j,g} \mathrm{I}(d = d_j)\big\}$$

for $t > 0$. Thus, λ_k is the baseline PE hazard on the k^{th} subinterval, and we denote $\boldsymbol{\lambda} =$

$(\lambda_1, \cdots, \lambda_K)$. Any fixed value λ_{K+1} may be assumed for $t > T_n$ since no Y_S' values are observed beyond T_n. Assuming $K = 3$, 4, or 5 intervals should work well for most applications. In the g^{th} subgroup, $\gamma_{E,r,g}$ is the effect of $[Y_E^+ = r]$, $\gamma_{T,k,g}$ is the effect of $[Y_T = k]$, and $\gamma_{D,j,g}$ is the effect of $[d = d_j]$ on the hazard of Y_S'. To ensure identifiability of the PE model, for each g, we set $\gamma_{E,1,g} = \gamma_{T,0,g} = \gamma_{D,1,g} = 0$. We denote $\boldsymbol{\gamma}_{E,g} = (\gamma_{E,2,g}, \cdots, \gamma_{E,L_E-1,g})$ with $\boldsymbol{\gamma}_E = (\boldsymbol{\gamma}_{E,1}, \cdots, \boldsymbol{\gamma}_{E,G})$, $\boldsymbol{\gamma}_{T,g} = (\gamma_{T,1,g}, \cdots, \gamma_{T,L_T-1,g})$ with $\boldsymbol{\gamma}_T = (\boldsymbol{\gamma}_{T,1}, \cdots, \boldsymbol{\gamma}_{T,G})$, and $\boldsymbol{\gamma}_{D,g} = (\gamma_{D,2,g}, \cdots, \gamma_{D,J,g})$, with $\boldsymbol{\gamma}_D = (\boldsymbol{\gamma}_{D,1}, \cdots, \boldsymbol{\gamma}_{D,G})$. The parameter vector for the PE hazard function h_S then can be written as $\boldsymbol{\theta}_S = (\boldsymbol{\lambda}, \boldsymbol{\gamma}_E, \boldsymbol{\gamma}_T, \boldsymbol{\gamma}_D)$.

Let $\mathcal{D}_S^n = \{(Y_{S,i}^{\prime o}, \delta_i, d_{[i]}, g_i) : i = 1, \cdots, n\}$ denote the time-to-event, dose, and subgroup data for the first n patients, so that the entire dataset is $\mathcal{D}^n = \mathcal{D}_{ET}^n \cup \mathcal{D}_S^n$. Denoting the conditional pdf and survival function of Y_S' by $f_S(\cdot \mid Y_E^+, Y_T, d, g, \boldsymbol{\theta}_S)$ and $\overline{F}_S(\cdot \mid Y_E^+, Y_T, d, g, \boldsymbol{\theta}_S)$, respectively, the likelihood is

$$\mathcal{L}(\mathcal{D}^n \mid \boldsymbol{\theta}) = \mathcal{L}(\mathcal{D}_{ET}^n \mid \boldsymbol{\theta}_{ET}) \prod_{i=1}^{n} \left\{ f_S(Y_{S,i}^{\prime o} \mid \boldsymbol{Y}_{ET,i}, d_{[i]}, g_i, \boldsymbol{\theta}_S) \right\}^{\delta_i} \left\{ \overline{F}_S(Y_{S,i}^{\prime o} \mid \boldsymbol{Y}_{ET,i}, d_{[i]}, g_i, \boldsymbol{\theta}_S) \right\}^{1-\delta_i}.$$

We use the latent cluster membership vector $\boldsymbol{z}$ to adaptively cluster similar subgroups by representing each possible configuration of $\boldsymbol{z}$ as a model, and index the models determined by $\boldsymbol{z}$ by $\nu = 1, \cdots, M$. For example, if $G = 3$ there are $M = 5$ possible models, indexed by $\nu = 1$ if $\boldsymbol{z} = (1, 1, 1)$, $\nu = 2$ if $\boldsymbol{z} = (1, 1, 2)$, $\nu = 3$ if $\boldsymbol{z} = (1, 2, 1)$, $\nu = 4$ if $\boldsymbol{z} = (1, 2, 2)$, $\nu = 5$ if $\boldsymbol{z} = (1, 2, 3)$. Assuming a discrete uniform prior on the set of models, in this example $\Pr(\nu = k)$ =1/5 for each $k = 1, \cdots, 5$. The model dimension may change with $\boldsymbol{z}$ and hence ν. For example, if $\nu = 1$ then all subgroup-specific parameters are identical. If $\nu = 2$ then all subgroup-specific parameters indexed by $g = 1$ and $g = 2$ are identical, while the parameters indexed by $g = 3$ are different. Specifically, if $z_g = z_{g'}$ for $g \neq g'$, then the following parameters for subgroups g and g' are identical: $\boldsymbol{\alpha}_g = \boldsymbol{\alpha}_{g'}$, $\boldsymbol{\beta}_g = \boldsymbol{\beta}_{g'}, \boldsymbol{\gamma}_{E,g} = \boldsymbol{\gamma}_{E,g'}$, $\boldsymbol{\gamma}_{T,g} = \boldsymbol{\gamma}_{T,g'}$, $\boldsymbol{\gamma}_{D,g} = \boldsymbol{\gamma}_{D,g'}$. Conditional on $\boldsymbol{z}$, for $k = 1, \cdots, \max(z_g)$ let $\left(\boldsymbol{\alpha}_k^\star, \boldsymbol{\beta}_k^\star, \boldsymbol{\gamma}_{E,k}^\star, \boldsymbol{\gamma}_{T,k}^\star, \boldsymbol{\gamma}_{D,k}^\star\right)$ denote the unique values of $(\boldsymbol{\alpha}_g, \boldsymbol{\beta}_g, \boldsymbol{\gamma}_{E,g}, \boldsymbol{\gamma}_{T,g}, \boldsymbol{\gamma}_{D,g})$ and if $z_g = k$ then $(\boldsymbol{\alpha}_g, \boldsymbol{\beta}_g, \boldsymbol{\gamma}_{E,g}, \boldsymbol{\gamma}_{T,g}, \boldsymbol{\gamma}_{D,g})$

$= (\boldsymbol{\alpha}_k^\star,\ \boldsymbol{\beta}_k^\star, \boldsymbol{\gamma}_{E,k}^\star, \boldsymbol{\gamma}_{T,k}^\star, \boldsymbol{\gamma}_{D,k}^\star)$. We denote $\boldsymbol{\alpha}^\star = (\boldsymbol{\alpha}_1^\star, \cdots, \boldsymbol{\alpha}_{\max(z_g)}^\star)$, $\boldsymbol{\beta}^\star = (\boldsymbol{\beta}_1^\star, \cdots, \boldsymbol{\beta}_{\max(z_g)}^\star)$, $\boldsymbol{\gamma}_E^\star = (\boldsymbol{\gamma}_{E,1}^\star, \cdots, \boldsymbol{\gamma}_{E,\max(z_g)}^\star)$, $\boldsymbol{\gamma}_T^\star = (\boldsymbol{\gamma}_{T,1}^\star, \cdots, \boldsymbol{\gamma}_{T,\max(z_g)}^\star)$, $\boldsymbol{\gamma}_D^\star = (\boldsymbol{\gamma}_{D,1}^\star, \cdots, \boldsymbol{\gamma}_{D,\max(z_g)}^\star)$, $\boldsymbol{\theta}_{ET}^\star = (\sigma_{12}, \boldsymbol{\alpha}^\star, \boldsymbol{\beta}^\star, \eta_1, \cdots, \eta_{L_E-1}, \zeta_1, \cdots, \zeta_{L_T-1})$ and $\boldsymbol{\theta}_S^\star = (\boldsymbol{\lambda}, \boldsymbol{\gamma}_E^\star, \boldsymbol{\gamma}_T^\star, \boldsymbol{\gamma}_D^\star)$. The overall model parameter vector is $\boldsymbol{\theta} = (\boldsymbol{z}, \boldsymbol{\theta}_{ET}^\star, \boldsymbol{\theta}_S^\star)$, abusing notation slightly since the dimensions of $\boldsymbol{\theta}_{ET}^\star$ and $\boldsymbol{\theta}_S^\star$ may change with $\boldsymbol{z}$. To accommodate different dimensions, we use reversible-jump Markov chain Monte Carlo (RJMCMC) to compute posteriors (Green, 1995).

For each subgroup g, the early and long-term dose optimality criteria are estimated, respectively, by their posterior means,

$$\widehat{\phi}_{ET}(d, g \mid \mathcal{D}_{ET}^n) = \int \phi_{ET}(d, g \mid \boldsymbol{\theta}_{ET}^\star, z_g)\ p(\boldsymbol{\theta}_{ET}^\star, z_g \mid \mathcal{D}_{ET}^n)\ d\boldsymbol{\theta}_{ET}^\star\ dz_g$$

and

$$\widehat{\phi}_S(d, g \mid \mathcal{D}^n) = \int \phi_S(d, g \mid \boldsymbol{\theta})\ p(\boldsymbol{\theta} \mid \mathcal{D}^n)\ d\boldsymbol{\theta},$$

where $p(\boldsymbol{\theta}_{ET}^\star, z_g \mid \mathcal{D}_{ET}^n)$ is the posterior based on the early outcome data and $p(\boldsymbol{\theta} \mid \mathcal{D}^n)$ is the posterior based on the data from all outcomes. Given a fixed lower bound $\underline{\phi}_S$ on $\phi_S(d_j, g)$, the posterior predictive probability that the long-term optimality criterion for dose d in subgroup g is greater than $\underline{\phi}_S$ is

$$\widehat{\Pr}\left\{\phi_S(d, g) > \underline{\phi}_S \mid \mathcal{D}^n\right\} = \int \Pr\left\{\phi_S(d, g) > \underline{\phi}_S \mid \boldsymbol{\theta}\right\} p(\boldsymbol{\theta} \mid \mathcal{D}^n)\ d\boldsymbol{\theta}.$$

If subgroup g belongs to a cluster with one or more other subgroups, then the above criteria for g are computed using the data from all subgroups in that cluster. These criteria will be used by the PGen I-II design to construct subgroup-specific futility rules.

To specify a prior on $\boldsymbol{\theta}$, we assume that the correlation $\sigma_{12} \sim \text{Unif}[-1, 1]$. For the interval cut-offs of the latent variables (X_E, X_T), we define $\kappa_{E,k} = \eta_k - \eta_{k-1}$ for $k = 2, \cdots, L_E - 1$ and $\kappa_{T,k} = \zeta_k - \zeta_{k-1}$ for $k = 2, \cdots, L_T - 1$, and assume all $\kappa_{j,k} \overset{\text{indep}}{\sim} \text{U}(0, \infty)$. To establish priors on the remaining parameters in $\boldsymbol{\theta}$, one may elicit probabilities of observable outcomes from the clinical investigators and apply the pseudo-sampling method of Thall et al. (2012). In the

simulations reported below, where elicited probabilities are not available, weakly-informative priors are used. Details are given in the Supplementary Materials.

3. A Utility-Based PGen I-II Design

We illustrate the PGen I-II design for ordinal Y_E and Y_T using the following phase I-II design. The design uses the subgroup-specific mean utility as the early outcome optimality criterion, and the PFS probability at follow up time t_2 as the long-term optimality criterion. For $j = E, T$ and each (d, g), denote the L_j-vector of marginal early outcome probabilities

$$\boldsymbol{\pi}_j(d, g, \boldsymbol{\theta}_j) = (\pi_j(0 \mid d, g, \boldsymbol{\theta}_j), \pi_j(1 \mid d, g, \boldsymbol{\theta}_j), \cdots, \pi_j(L_j - 1 \mid d, g, \boldsymbol{\theta}_j))',$$

where $\boldsymbol{\theta}_E$ and $\boldsymbol{\theta}_T$ are the marginal model parameter vectors for efficacy and toxicity, respectively. Given current data $\mathcal{D}^n_{ET}$, *a dose d is acceptable for subgroup g* if

$$\Pr\{\boldsymbol{b}'_E\boldsymbol{\pi}_E(d, g, \boldsymbol{\theta}_E) > \underline{\pi}_E \mid \mathcal{D}^n_{ET}\} > p_E \quad \text{and} \quad \Pr\{\boldsymbol{b}'_T\boldsymbol{\pi}_T(d, g, \boldsymbol{\theta}_T) < \overline{\pi}_T \mid \mathcal{D}^n_{ET}\} > p_T, \qquad (2)$$

where $\boldsymbol{b}_E$ and $\boldsymbol{b}_T$ are design parameter vectors with all entries 0 or 1, and $\underline{\pi}_E$ and $\overline{\pi}_T$ are prespecified fixed limits. In our illustration, we use the simple criteria based on the best efficacy outcome $Y_E = L_E - 1$, and worst toxicity outcome $Y_T = L_T - 1$, so we set $\boldsymbol{b}_E = (0, \cdots, 0, 1)'_{L_E}$ and $\boldsymbol{b}_T = (0, \cdots, 0, 1)'_{L_T}$. Alternatively, for example, if the efficacy levels are PD, SD, PR, and CR, and the efficacy monitoring criterion is based on PR or CR, then $\boldsymbol{b}_E = (0, 0, 1, 1)'$. The decision cutoffs p_E and p_T should be calibrated by simulation, with values in the range .05 to .20 working well in most settings.

Given an elicited utility $U(y_E, y_T)$ for all possible pairs (y_E, y_T), the early outcome optimality criterion for dose d in subgroup g is the mean utility

$$\phi_{ET}(d, g, z_g, \boldsymbol{\theta}^\star_{ET}) = \sum_{y_E, y_T} \pi(y_E, y_T \mid d, g, z_g, \boldsymbol{\theta}^\star_{ET})\, U(y_E, y_T),$$

where $\boldsymbol{\theta}^\star_{ET} = (\sigma_{12}, \boldsymbol{\alpha}^\star, \boldsymbol{\beta}^\star, \eta_1, \cdots, \eta_{L_E-1}, \zeta_1, \cdots, \zeta_{L_T-1})$. Denoting $\boldsymbol{\theta} = (z_g, \boldsymbol{\theta}^\star_{ET}, \boldsymbol{\theta}^\star_S)$, the long-

term outcome optimality criterion is defined to be

$$\phi_S(d, g, \boldsymbol{\theta}) = \Pr(Y'_S > t_2 - t_1 \mid d, g, \boldsymbol{\theta}).$$

For example, if $t_1 = 1$ month and $t_2 = 6$ months, then $\phi_S(d, g, \boldsymbol{\theta})$ is the probability that PFS, starting at the time when $\boldsymbol{Y}_{ET}$ is evaluated, is at least 5 months.

A PGen I-II design has three stages, illustrated by Figure 1. For stage $s = 1, 2, 3$ the sample size is n_s, with overall sample size $N = n_1 + n_2 + n_3$. We denote the sample size of the g^{th} subgroup in stage s by $n_{s,g}$ and the sample size of dose d_j by $n_s(d_j)$. In stage 1, due to limited within-subgroup sample sizes, it is not feasible to identify an acceptable dose set for each subgroup. Consequently, the PGen I-II design temporarily ignores subgroups in stage 1. The version of the PGen I-II design that we consider here uses the BOIN12 design (Lin et al., 2020) to assign doses and determine an acceptable dose set for all subgroups combined in stage 1. To normalize the utilities, which range from 0 to 100, we divide them by 100, so the domain of $\phi_{ET}(d, g, z_g, \boldsymbol{\theta}^\star_{ET})/100$ is the interval $[0, 1]$. The BOIN12 design treats $\phi_{ET}(d, g, z_g, \boldsymbol{\theta}^\star_{ET})/100$ like a quasi-probability and uses a quasi-binomial model to compute its approximate posterior. The posterior means of $\phi_{ET}(d, g, z_g, \boldsymbol{\theta}^\star_{ET})/100$ are used to select successive doses in stage 1, temporarliy ignring the subgroup g. When the maximum overall stage 1 sample size n_1 has been reached, a stage 1 acceptable dose set $\mathcal{A}^{n_1}_{ET}$ satisfying the requirements in Equation (2) while ignoring subgroups is identified. Although subgroups are ignored by the decision rules in stage 1, subgroup information and the time-to-failure data are recorded for subsequent analyses.

Starting in stage 2, the PGen I-II design accounts for subgroups. For each subgroup g, $n_{2,g}$ additional patients are randomized fairly among the acceptable doses in $\mathcal{A}^{n_1}_{ET}$. Denote the combined sample size of subgroup g from Stages 1 and 2 by $n_{12,g} = n_{1,g} + n_{2,g}$, the per-dose combined sample size from Stages 1 and 2 by $n_{12}(d_j) = n_1(d_j) + n_2(d_j)$, and the overall combined sample size by $n_{12} = n_1 + n_2$. The data collected in Stage 2 includes both

early outcomes (Y_E, Y_T) up to time t_1, and observed $(Y_S'^{\,o},\ \delta)$ values up to a prespecified intermediate follow up time t_ℓ, where $t_1 < t_\ell \leqslant t_2$. At the end of Stage 2, for each g, a subgroup-specific acceptable dose set $\mathcal{A}_{ET,g}^{n_{12}}$ is determined using the criteria in (2), and the acceptable dose set in subgroup g is defined as

$$\mathcal{A}_{S,g}^{n_{12}} = \Big\{ d_j \in \mathcal{A}_{ET,g}^{n_{12}} : \Pr\left\{ \phi_S(d_j, g) > \underline{\phi}_S \mid \mathcal{D}^{n_{12}} \right\} > p_{S,L} \Big\}.$$

The posterior futility criterion, which uses a fixed decision cut-off such as $p_{S,L} = .10$ or possibly a larger value, is included to avoid choosing a dose that has a small probability of achieving at least the minimal PFS criterion $\underline{\phi}_S$ in any subgroup. For each subgroup g and dose $d \in \mathcal{A}_{ET,g}^{n_{12}}$, RJMCMC is used to approximate the posterior distribution and obtain an estimator $\widehat{\phi}_{ET}(d, g, z_g, \boldsymbol{\theta}_{ET}^{\star})$ of the phase I-II optimality criterion. The maximum over the set of acceptable doses for subgroup g is denoted by

$$\widehat{\phi}_{ET,g}^{max} = \max_{d_j \in \mathcal{A}_{ET,g}^{n_{12}}} \widehat{\phi}_{ET}(d_j, g \mid \mathcal{D}^{n_{12}}).$$

A *candidate dose set for subgroup* g then is defined as

$$\mathcal{C}_g^{n_{12}} = \Big\{ d_j \in \mathcal{A}_{S,g}^{n_{12}} : \widehat{\phi}_{ET}(d_j, g \mid \mathcal{D}^{n_{12}}) \geqslant \rho\ \widehat{\phi}_{ET,g}^{max} \Big\}.$$

The parameter $0 < \rho < 1$ quantifies how close $\widehat{\phi}_{ET}(d_j, g \mid \mathcal{D}^{n_{12}})$ must be to the empirically optimal dose to be included in the candidate set for subgroup g. To obtain a design with good operating characteristics (OCs), ρ may be chosen from the range .50 to .90, based on preliminary simulations.

In stage 3, for each subgroup g, $n_{3,g}$ additional patients are randomized among the doses in $\mathcal{C}_g^{n_{12}}$ and followed to time t_2. Denote $n_{123,g} = n_{1,g} + n_{2,g} + n_{3,g}$ and $N = \sum_{g=1}^{G} n_{123,g}$. Given final data $\mathcal{D}^N$ from all three stages, the candidate dose set for each subgroup g is updated from $\mathcal{C}_g^{n_{12}}$ to $\mathcal{C}_g^N$. Since all patients are followed to time t_2, the efficacy requirement in (2) is replaced by the minimal long-term success probability requirement. Thus, doses in the updated candidate dose set $\mathcal{C}_g^N$ must satisfy both the toxicity requirement in (2) and the

minimal long-term success probability requirement,

$$\mathcal{C}_g^N = \Big\{ d_j \in \mathcal{C}_g^{n_{12}} : \ \Pr\left\{ \boldsymbol{b}_T' \boldsymbol{\pi}_T(d_j, g, \boldsymbol{\theta}_T) < \overline{\pi}_T \mid \mathcal{D}^N \right\} > p_{\mathcal{C},T,L} \\ \textit{and} \ \Pr\left\{ \phi_S(d_j, g) > \underline{\phi}_S \mid \mathcal{D}^N \right\} > p_{\mathcal{C},S,L} \Big\}.$$

In our simulated design, described below, we used $p_{\mathcal{C},T,L} = .10$ and $p_{\mathcal{C},S,L} = .50$, based on preliminary simulations that showed these gave good design OCs. If the final candidate dose set $\mathcal{C}_g^N$ for subgroup g contains at least one dose, then the optimal dose for that subgroup is the acceptable dose that maximizes the estimated long-term benefit criterion,

$$d^{opt,g} = \underset{d_j \in \mathcal{C}_g^N}{argmax} \ \widehat{\phi}_S(d_j, g \mid \mathcal{D}^N). \tag{3}$$

4. Guidelines for Establishing Design Parameters

Given the large number of design parameters and prior hyperparameters that must be specified, in practice they may be determined as follows. First, G and N should be chosen so that the trial is feasible to conduct. Since specifying even $G = 2$ subgroups may be extremely useful compared to assuming homogeneity, whether a given G is too large may be decided by determining a value of N that is feasible, estimating the subgroup propensities, $\boldsymbol{p_g} = (p_1, \cdots, p_G)$, and computing the expected within-subgroup sample sizes, $p_1 N, \cdots, p_G N$. These may be used in preliminary simulations of the design. The ability of the PGen I-II design to cluster subgroups may mitigate difficulties presented by a large number of prespecified subgroups, such as $G = 5$ or 6.

The definitions of Y_E and Y_T determine t_1, and t_2 should provide a clinically meaningful long-term follow up time for evaluating how Y_S may depend on d in each subgroup. The fixed limits $\underline{\pi}_E$ for the early efficacy futility rule, $\overline{\pi}_T$ for the toxicity safety rule, and $\underline{\phi}_S$ for the long term success probability futility rule all are elicited from the clinical investigators. A numerical utility function may be elicited as described by Thall et al. (2023), Section 4,

by first specifying $U(y_E, y_T) = 0$ for the worst possible outcome and $U(y_E, y_T) = 100$ for the best possible outcome, and then eliciting utilities of all intermediate outcomes. In practice, this process is straightforward, since investigators readily understand utility functions, and provide numerical utilities based on their clinical experiences. The utility function used for the trial in the simulations reported here corresponds to binary Y_T and three-level ordinal Y_E, and is given in Table S.2 in the online supplementary materials.

Parameters that may be determined by preliminary simulations include the number of subintervals in the PE model, which usually should be 3 to 5, the fixed decision cutoffs p_L and p_E used in (5), the parameter ρ used to define the candidate dose set after 2 stages, and the decision cut-offs $p_{\mathcal{C},T,L}$ for the toxicity probability rule and $p_{\mathcal{C},S,L}$ for the long term success probability decision rule used to define the candidate dose sets $\mathcal{C}_g^N$ for $g = 1, \cdots, G$. Prior hyperparameters may be determined using either of the methods discussed above. Details are provided in the Supplementary Materials. Given G and N, the stage-specific subsample sizes n_1, n_2, n_3 must satisfy $N = n_1 + n_2 + n_3$. For a cohort size $c = 2$, 3, 4, or 5, each n_s should be a multiple of c. Given n_1, n_2, n_3, the subgroup-specific sample sizes $n_{1,g}, n_{2,g}, n_{3,g}$ for each g are determined based on estimates of the subgroup propensities $p_1, \cdots, p_G$.

5. Simulation Study

In this section, we describe a simulation study to evaluate the OCs of the PGen I-II design, and three competing designs. The simulations were designed to reflect the NK cell trial. We assumed three non-ordered subgroups ($G = 3$) corresponding to the disease groups in Table 1, four dose levels ($J = 4$), a binary toxicity indicator $Y_T = 0$ or 1, ($L_T = 2$) and a three-level ordinal efficacy outcome $Y_E = 0$ for PD, 1 for SD, or 2 for PR or CR ($L_E = 3$). The standardized doses were $(0.69, 0.79, 0.90, 1.00)$. We generated patient subgroups indexed by $g = 1$, ,2, 3 from a trinomial distribution with probability vector $\boldsymbol{p_g} = (0.3, 0.3, 0.4)$. Based on follow up times $t_1 = 1$ for (Y_E, Y_T), $t_\ell = 3$, and $t_2 = 6$ months for Y_S, we considered

an array of scenarios having different patterns of cluster membership variables $\boldsymbol{z}^{true} = (z_1, z_2, z_2)^{true}$, early outcome optimality criteria $\phi_{ET}^{true}(d, g)$, long term outcome optimality criteria $\phi_S^{true}(d, g) = \Pr(Y_S' > 5 \mid d, g)$, and true outcome distributions.

To specify simulation scenarios, we first fixed the marginal early outcome probabilities $\pi_T^{\text{true}}\ (0 \mid d, g)$, $\pi_E^{\text{true}}\ (1 \mid d, g)$, and $\pi_E^{\text{true}}\ (2 \mid d, g)$ for each (d, g). We used a latent bivariate normal distribution with $\sigma_{12}^{true} = 0.2$ to obtain joint probabilities $\pi^{true}(y_E, y_T \mid d, g)$ for $y_T = 0, 1$ and $y_E = 0, 1, 2$, and computed the resulting $\phi_{ET}^{true}(d, g)$. Given $Y_E > 0$, we generated failure time Y_S' following t_1 from a Weibull distribution with conditional hazard function

$$h_S\left(t \mid Y_E^+, Y_T, d, g\right) = \frac{\omega_g}{\psi_g}\left(\frac{t}{\psi_g}\right)^{\omega_g - 1} \exp\left\{\gamma'_{E,Y_E^+,g} + \gamma'_{T,Y_T,g} + \sum_{j=1}^{J} \gamma'_{D,j,g} \mathrm{I}(d = d_j)\right\} \qquad (4)$$

for $t > 0$. To ensure identifiability, we fixed $\gamma'_{E,1,g} = 0$, $\gamma'_{T,0,g} = 0$ and $\gamma'_{D,1,g} = 0$. We determined true values ω_g^{true}, ψ_g^{true}, $\gamma'^{\,true}_{E,r,g}$, $\gamma'^{\,true}_{T,k,g}$, $\gamma'^{\,true}_{D,j,g}$ to obtain specified values of $\phi_S^{true}(d, g)$. The dose acceptability criteria were determined by $\boldsymbol{b}_E = (0, 0, 1)'$, $\underline{\pi}_E = 0.50$, $\boldsymbol{b}_T = (0, 1)'$ and $\overline{\pi}_T = 0.30$ in Equation (2), with lower limit $\underline{\phi}_S = 0.40$ on $\phi_S^{true}(d, g)$, and $\rho = 0.7$. The simulation scenarios are illustrated in Figure S.1, showing that both $\pi_T^{\text{true}}\ (y_T \mid d, g)$ and $\pi_E^{\text{true}}\ (y_E \mid d, g)$ increased with dose, but we allowed $\phi_S^{\text{true}}\ (d_j, g)$ to be either increasing or non-monotone in dose, with this pattern differing between subgroups. Using cohort size 3 throughout, stage 1 had 10 cohorts, stage 2 had 30 cohorts, and stage 3 had 10 cohorts. The resulting sample size $N = 150$ is 2.5 times the maximum sample size of 60 considered by Thall et al. (2023), since the PGen I-II design accounts for three subgroups. An accrual rate of 3 patients per month was assumed. Simulation results for eight scenarios are presented in Tables 1-4, based on 1000 simulated trials for each design under each scenario.

The simulation scenarios included three configurations of the subgroup clustering variables $\boldsymbol{z}^{true}$. Scenarios 1 and 4 had $\boldsymbol{z}^{true} = (1, 1, 1)$, corresponding to complete homogeneity, Scenarios 2, 3, 5, 6, and 7 had $\boldsymbol{z}^{true} = (1, 2, 1)$, corresponding to clusters {1, 3} and {2}, and Scenario 8 had $\boldsymbol{z} = (1, 2, 3)$, corresponding to complete heterogeneity. We compared the

PGen I-II design to three competing designs. PGen I-II-Comb behaves like the original Gen I-II design by ignoring patient subgroups. PGen I-II-ET is a utility-based phase I-II design that ignores Y_S and selects an optimal dose in each subgroup based on $U(Y_E, Y_T)$. PGen I-II-Sep conducts a separate trial within each subgroup. The OCs include dose selection percentages and the mean number of patients treated at each dose for each subgroup.

In Scenario 1, no dose is acceptable for any subgroup due to excessive toxicity probabilities or unacceptably low values of $\phi_S^{\text{true}}(d, g)$. The PGen I-II design terminates the trial early 93.2%, 92.5%, and 94.3% of the time for the three subgroups, which are much larger than the early stopping probabilities for the PGen I-II-ET and PGen I-II-Sep designs, and slightly smaller than the value 97.0 for the PGen I-II-Comb design. In Scenario 2, no dose is acceptable for subgroups 1 and 3, while d_3 is optimal for subgroup 2. Table 1 shows that the PGen I-II design performs far better than the PGen I-II-Comb and PGen I-II-ET designs. Compared to the PGen I-II-Sep design, the PGen I-II design performs better for subgroups 1 and 3, but worse for subgroup 2. This is not unexpected, since subgroups 1 and 3 form a cluster, and the PGen I-II design shares information between them by identifying the cluster. In Scenario 3, d_4 is optimal for subgroups 1 and 3, while no dose is acceptable for subgroup 2. Similarly to Scenario 2, the PGen I-II design performs significantly better than both the PGen I-II-Comb and PGen I-II-ET designs. Compared to the PGen I-II-Sep design, the PGen I-II design performs better for subgroups 1 and 3, but worse for subgroup 2.

In Scenarios 4 and 7, d_3 is optimal for all three subgroups. In scenario 4, the three subgroups are homogeneous and form a single cluster. In Scenario 7, the subgroups form two clusters, $\{1, 3\}$ and $\{2\}$. In both scenarios 4 and 7, the PGen I-II design has significantly superior performance compared to the PGen I-II-ET and PGen I-II-Sep designs. Since the optimal dose for all three subgroups is the same, the PGen I-II design shows slightly less effective performance compared to the PGen I-II-Comb design, which makes the correct assumption.

In Scenarios 5 and 6, there are two clusters, $\{1, 3\}$ and $\{2\}$, but with different optimal doses. In scenarios 5 and 6, the PGen I-II design performs significantly better than both the PGen I-II-Comb and PGen I-II-ET designs. Compared to the PGen I-II-Sep design, the PGen I-II design performs better for the cluster $(1, 3)$ and slightly worse for $\{2\}$, In Scenario 8, each subgroup is a singleton cluster, with different optimal doses for each subgroup. In this scenario, the PGen I-II-Sep design performs best, while the PGen I-II design shows slightly less effective performance. However, PGen I-II has greatly superior performance compared to the PGen I-II-Comb and PGen I-II-ET designs.

In summary, across the eight scenarios, the PGen I-II design outperforms each of the other three designs, often by a wide margin. Compared to the PGen I-II-Sep design, due to adaptive clustering the PGen I-II design has much better OCs when some subgroups are truly identical. If there are no clusters, PGen I-II has slightly worse OCs than PGen I-II-Sep, as expected. In general, PGen I-II is far superior to both the PGen I-II-Comb design, which ignores subgroups, and the PGen I-II-ET design, which ignores Y_S'.

Five additional sets of simulations were conducted to assess the PGen I-II design's sensitivity to (1) values of the interim follow up time t_ℓ, (2) the parameter ρ used to determine the candidate dose set, (3) the data-generating distribution for Y_S', (4) the maximum sample size N, and (5) the assumed distribution of Y_S'. The results are summarized in Tables S.3 to S.7 in the online supplementary materials. Table S.3 shows that both selection percentages and numbers of patients treated were insensitive to the different values of t_ℓ. Table S.4 shows that values of the closeness parameter ρ had no substantive effect on any of the OCs in nearly all scenarios, with a drop of about 4% seen in correct dose selection percentage for subgroups $g = 1$ and $g = 3$ for $\rho = 0.9$ in scenario 6. Table S.5 shows that, while in general the PGen II design had high probabilities of correctly selecting no dose in Scenarios 1 and 2 where no doses are acceptable, these probabilities were roughly 6% to 7% smaller under a

Weibull distribution compared to a lognormal. For Scenarios 3 - 8, where there were some acceptable doses for each subgroup, the correct subgroup-specific dose selection probabilities were similar for a Weibull distribution compared to a log-normal or gamma.

Overall, the simulation results in Table S.6 for the case studied, with three subgroups and four doses, are quite striking. They show that decreasing the maximum sample size N from 150 to 120 or 90 substantively decreases subgroup-specific probabilities of correctly selecting no dose in scenarios 1 and 2, and of correctly selecting optimal doses in scenarios 3 – 8. Thus, running the trial with the smaller values $N = 120$ or 90 gives a design that is far less reliable. For example, when N is decreased from 150 to 90, in Scenario 6 correct selection probabilities drop by as much as 17% for some subgroups, and in Scenario 4 by as much as 25%. These results show that, if one wishes to use the PGen I-II design to select personalized doses, it is extremely important to enroll an adequate overall sample size. The results in Table S.7 show that, when the data are generated from a Weibull distribution, the PE distribution-based design performs comparably to the Weibull distribution-based design. However, in scenarios with greater heterogeneity, such as Scenario 8, the PE distribution-based design outperforms the Weibull-based design, which shows the advantage of using a PE distribution.

6. Discussion

Both a clinically and scientifically, it is highly desirable for a clinical trial design to make personalized treatment decisions that account for treatment-subgroup interactions. Our simulations show that, across a wide array of scenarios, the PGen I-II design is greatly superior to similarly structured designs that either ignore subgroups or conduct separate trials within subgroups.

The PGen I-II design is complex. This is because, in addition to incorporating the outcomes and decision making structure of the Gen I-II design, PGen I-II accounts for subgroup effects and does adaptive subgroup clustering. To implement the design, practical requirements

include eliciting many prior parameters, specifying a nontrivial number of design parameters, specifying a set of fairly complex simulation scenarios, and doing computer simulations to calibrate prior hyperparameters and establish the design's OCs. Given the great advantages provided by the PGen I-II design, however, this effort seems worthwhile.

The PGen I-II design provides an attractive an alternative to conducting G separate trials, one in each subgroup. Given that the Gen I-II design presented by Thall et al. (2023), which assumed homogeneity, had $N = 60$, extending it to a precision design accommodating three subgroups with $N = 150$ rather than 180 seems worthwhile. Moreover, the PGen I-II paradigm is not limited to dose optimization, and may be applied in trials with the goal to optimize subgroup-specific schedules, dose-schedule combinations, or different treatments. In such cases, however, it would be necessary to modify the underlying model accordingly.

While we have used PFS time as the long-term endpoint, the PGen I-II design can accommodate other endpoints, such as remission duration or Y_{OS} = overall survival time. For example, since $Y_E = 0$ due to PD occurring prior to t_1 does not imply that $Y_{OS} < t_2$, a model for Y_{OS} can be specified without the need to define Y'_{OS}, as done for PFS.

Acknowledgments

The authors thank the Associate Editor and Referee for their valuable comments. Yong Zang's research was partially supported by the NIH grants R01 GM150808, P30 CA082709 and the Showalter Scholar program funded by Ralph W. and Grace M. Showalter Research Trust Fund. Peter Thall's research was supported by NIH/NCI grants 1R01CA261978 and 5 P30 CA016672 45.

Supplementary materials

Web Appendices, Tables, and Figures are available with this paper at the Biometrics website on Oxford Academic. R code for implementing the PGen I-II design is available from https://github.com/yongzang2020/PGEN12.

Data Availability

Data sharing is not applicable as no new data were created or analyzed in this paper.

References

Babb, J., Rogatko, A., and Zacks, S. (1998). Cancer phase I clinical trials: Efficient dose escalation with overdose control. *Statistics in Medicine*, **17**, 1103-1120.

Chib, S., and Greenberg, E. (1998). Analysis of multivariate probit models. *Biometrika*, **85**, 347-361.

Gooley, T.A., Martin, P.J., Fisher, L.D. and Pettinger M. (1994) Simulating as a design tool for phase I/II clinical trials: an example from bone marrow transplantation. *Controlled Clinical Trials*, **15**, 450–462.

Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, **82(4)**, 711-732.

Guo, B., and Yuan, Y. (2017). Bayesian phase I/II biomarker-based dose finding for precision medicine with molecularly targeted agents. *Journal of the American Statistical Association*, **112**, 508-520.

Lee, J., Thall, P.F., and Msaouel, P. (2021). Precision Bayesian phase I-II dose-finding based on utilities tailored to prognostic subgroups. *Statistics in Medicine*, **40**, 5199-5217.

Lin, R., Zhou, Y., Yan, F., Li, D., and Yuan, Y. (2020). BOIN12: Bayesian optimal interval phase I/II trial design for utility-based dose finding in immunotherapy and targeted therapies. *JCO Precision Oncology*, **4**, 1393-1402.

McGovern, A., Chapple, A.G., and Ma, C. (2022). Two-stage subgroup-specific time-to-event (2S-Sub-TITE): An adaptive two-stage time-to-toxicity design for subgroup-specific dose finding in phase I oncology trials. *Pharmaceutical Statistics*, **21**(6), 1138-1148.

Park, Y. and Chang, W. (2024). A personalized dose-finding algorithm based on adaptive Gaussian process regression. *Pharmaceutical Statistics*. Wiley Online Library.

Porter, S., Murray, T. A., and Eaton, A. (2024). Phase I/II design for selecting subgroup-specific optimal biological doses for prespecified subgroups. *Statistics in Medicine*, Wiley Online Library.

Thall, P.F., and Cook, J.D. (2004). Dose-finding based on efficacy-toxicity trade-offs. *Biometrics*, **60**, 684-693.

Thall, P.F., and Russell, K. (1998). A strategy for dose-finding and safety monitoring based on efficacy and adverse outcomes in phase I/II clinical trials. *Biometrics*, **54**, 251-264.

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**, 785-801.

Thall, P.F., Zang, Y., and Yuan, Y. (2023). Generalized phase I-II designs to increase long term therapeutic success rate. *Pharmaceutical Statistics*, **22**, 692-706.

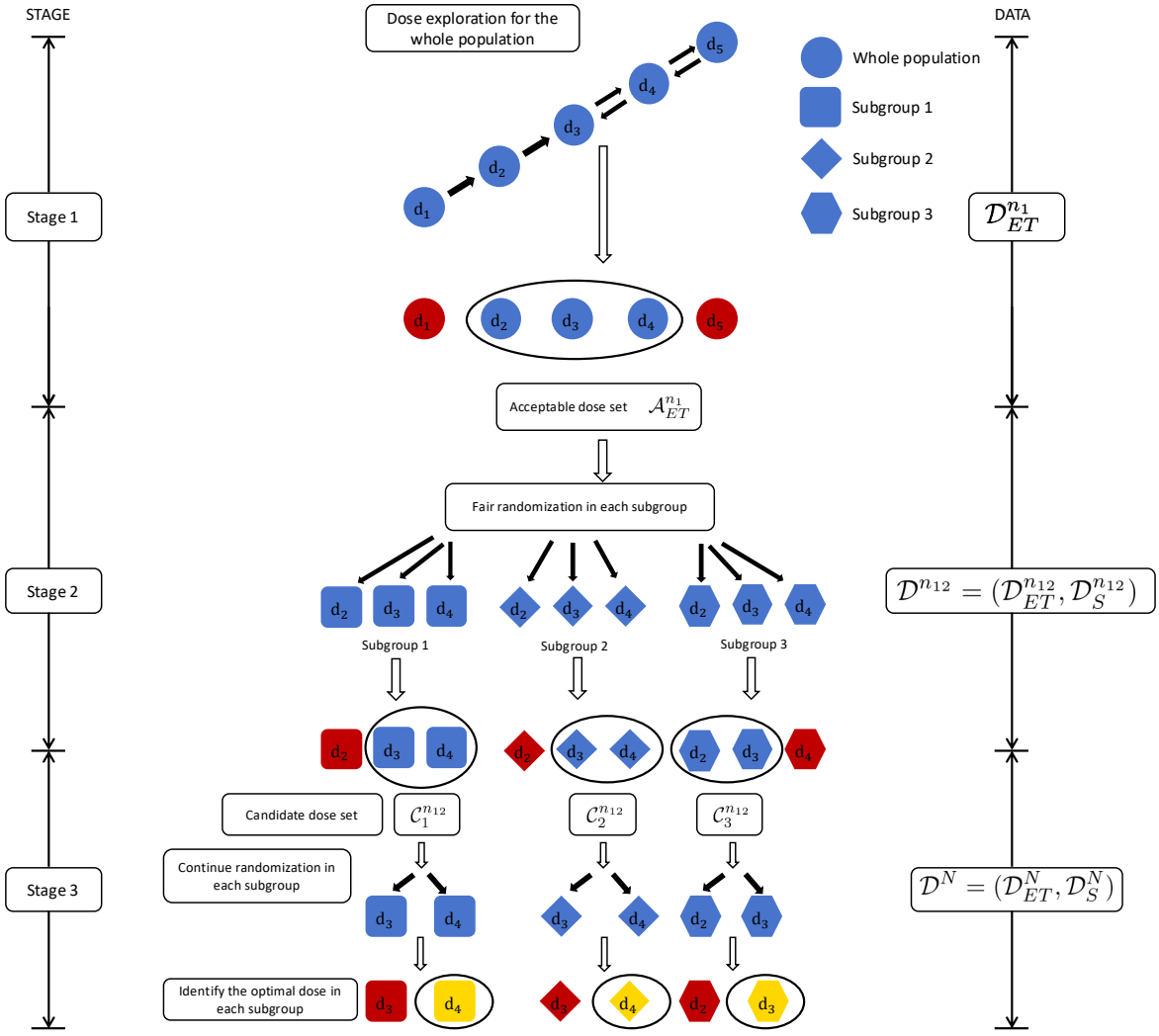

Figure 1: Schematic for the PGen I-II design with three subgroups. $N = n_1 + n_2 + n_3$ is the overall sample size for the whole trial.

Table 1: Selection % and mean number of patients treated at each dose level under the PGen I-II, PGen I-II-Comb, PGen I-II-ET and PGen I-II-Sep designs. Boldface indicates results for the true optimal decision under each subgroup. The number under d_0 is the percentage of trials terminated early for the specific subgroup with no dose is selected.

Scenario 1																
	Subgroup	**g=1**					**g=2**					**g=3**				
	z_g	**1**					**1**					**1**				
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.12	0.22	0.36	0.50		0.12	0.22	0.36	0.50		0.12	0.22	0.36	0.50
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.35	0.30	0.20	0.15		0.35	0.30	0.20	0.15		0.35	0.30	0.20	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.45	0.60	0.75	0.80		0.45	0.60	0.75	0.80		0.45	0.60	0.75	0.80
	$\phi_{ET}^{true}(d, g)$		62.7	69.5	72.9	70.1		62.7	69.5	72.9	70.1		62.7	69.5	72.9	70.1
	$\phi_S^{true}(d, g)$		0.10	0.20	0.30	0.45		0.10	0.20	0.30	0.45		0.10	0.20	0.30	0.45
PGen I-II	Selection %	**93.2**	0.1	0.0	1.6	5.1	**92.5**	0.1	0.0	1.5	5.9	**94.3**	0.0	0.0	1.4	4.3
	Patients		12.5	11.8	10.5	3.7		12.7	11.9	10.6	3.8		17.2	15.9	13.6	4.3
PGen I-II-Comb	Selection %	**97.0**	0.0	0.0	1.5	1.5	**97.0**	0.0	0.0	1.5	1.5	**97.0**	0.0	0.0	1.5	1.5
	Patients		12.4	11.6	10.0	2.7		12.7	11.7	10.0	2.8		17.3	15.8	13.0	3.6
PGen I-II-ET	Selection %	**7.2**	26.9	39.4	22.9	3.6	**6.9**	26.2	40.2	23.5	3.2	**6.8**	28.2	39.0	23.3	2.7
	Patients		15.1	15.1	10.2	2.7		15.2	15.1	10.3	2.8		20.7	20.6	13.3	3.5
PGen I-II-Sep	Selection %	**84.8**	2.3	1.2	6.0	5.7	**86.5**	1.6	0.9	4.4	6.6	**88.2**	0.7	0.4	4.6	6.1
	Patients		13.8	14.4	11.6	3.0		13.9	14.2	11.5	3.1		18.2	17.9	14.7	3.8
Scenario 2																
	Subgroup	**g=1**					**g=2**					**g=3**				
	z_g	**1**					**2**					**1**				
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.14	0.18	0.24		0.10	0.12	**0.14**	0.16		0.10	0.14	0.18	0.24
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.35	0.35	0.35	0.30		0.30	0.30	**0.20**	0.15		0.35	0.35	0.35	0.30
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.45	0.50	0.50	0.60		0.35	0.50	**0.70**	0.80		0.45	0.50	0.50	0.60
	$\phi_{ET}^{true}(d, g)$		63.3	65.9	64.7	68.8		54.0	65.1	**76.9**	82.5		63.3	65.9	64.7	68.8
	$\phi_S^{true}(d, g)$		0.10	0.20	0.25	0.30		0.20	0.30	**0.70**	0.40		0.10	0.20	0.25	0.30
PGen I-II	Selection %	**83.9**	1.8	0.6	6.1	7.6	22.3	2.1	2.3	**70.5**	2.8	**86.1**	1.1	0.8	4.6	7.4
	Patients		8.9	10.4	11.7	10.1		10.6	12.8	**14.7**	10.4		12.2	13.7	15.6	13.7
PGen I-II-Comb	Selection %	**64.4**	0.1	0.5	31.0	4.0	64.4	0.1	0.5	**31.0**	4.0	**64.4**	0.1	0.5	31.0	4.0
	Patients		8.6	10.1	13.5	9.8		8.8	9.9	**13.6**	9.8		11.9	13.4	18.4	13.1
PGen I-II-ET	Selection %	**2.6**	18.5	19.4	19.5	40.0	3.2	13.9	20.9	**19.5**	42.5	**2.8**	18.3	17.9	19.3	41.7
	Patients		10.5	12.0	11.6	9.7		10.1	12.1	**12.0**	9.8		14.4	16.0	15.6	13.1
PGen I-II-Sep	Selection %	**72.3**	4.1	3.5	6.5	13.6	15.8	2.7	3.0	**76.9**	1.6	**78.5**	3.0	2.5	4.6	11.4
	Patients		9.5	11.4	11.7	10.5		10.6	12.6	**13.3**	9.8		12.7	14.7	15.6	14.0

Table 2: (Continued).

Scenario 3																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**2**					**1**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}(1 \mid d,g)$		0.10	0.12	0.14	**0.16**		0.10	0.14	0.18	0.24		0.10	0.12	0.14	**0.16**
	$\pi_E^{\text{true}}(1 \mid d,g)$		0.30	0.30	0.20	**0.15**		0.35	0.35	0.35	0.30		0.30	0.30	0.20	**0.15**
	$\pi_E^{\text{true}}(2 \mid d,g)$		0.35	0.50	0.70	**0.80**		0.45	0.50	0.50	0.60		0.35	0.50	0.70	**0.80**
	$\phi_{ET}^{true}(d,g)$		54.0	65.1	76.9	**82.5**		63.3	65.9	64.7	68.8		54.0	65.1	76.9	**82.5**
	$\phi_S^{true}(d,g)$		0.20	0.30	0.40	**0.60**		0.05	0.10	0.15	0.20		0.20	0.30	0.40	**0.60**
PGen I-II	Selection %	22.2	1.8	2.7	5.5	**67.8**	**91.3**	1.3	0.4	0.6	6.4	22.9	1.0	1.9	5.5	**68.7**
	Patients		8.9	11.1	12.7	**12.3**		8.6	9.4	11.3	11.4		12.1	14.4	17.3	**16.2**
PGen I-II-Comb	Selection %	34.4	0.2	0.6	2.1	**62.7**	**34.4**	0.2	0.6	2.1	62.7	34.4	0.2	0.6	2.1	**62.7**
	Patients		8.1	9.8	11.9	**13.0**		8.4	9.7	12.3	13.2		11.3	12.7	16.1	**17.1**
PGen I-II-ET	Selection %	2.7	3.2	7.0	18.0	**69.1**	**2.5**	5.6	8.3	20.3	63.3	2.5	3.6	6.7	17.9	**69.3**
	Patients		9.1	11.2	12.6	**11.0**		10.1	11.1	12.6	11.0		12.6	14.8	17.0	**14.5**
PGen I-II-Sep	Selection %	22.2	6.5	8.6	8.2	**54.5**	**96.1**	0.6	0.7	0.4	2.2	21.8	3.1	6.3	7.2	**61.6**
	Patients		9.5	11.6	12.9	**11.8**		8.6	9.4	11.2	11.0		12.7	15.1	17.6	**15.7**
Scenario 4																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**1**					**1**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}(1 \mid d,g)$		0.10	0.12	**0.14**	0.16		0.10	0.12	**0.14**	0.16		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}(1 \mid d,g)$		0.30	0.30	**0.20**	0.15		0.30	0.30	**0.20**	0.15		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}(2 \mid d,g)$		0.35	0.50	**0.70**	0.80		0.35	0.50	**0.70**	0.80		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d,g)$		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5
	$\phi_S^{true}(d,g)$		0.40	0.50	**0.70**	0.60		0.40	0.50	**0.70**	0.60		0.40	0.50	**0.70**	0.60
PGen I-II	Selection %	2.3	4.1	6.4	**78.8**	8.4	2.1	3.6	4.7	**81.1**	8.5	2.2	4.5	4.4	**81.4**	7.5
	Patients		8.1	11.1	**13.6**	11.5		8.5	11.1	**13.5**	11.5		11.0	14.5	**18.1**	15.2
PGen I-II-Comb	Selection %	2.2	2.2	5.5	**81.2**	8.9	2.2	2.2	5.5	**81.2**	8.9	2.2	2.2	5.5	**81.2**	8.9
	Patients		7.7	11.3	**13.7**	11.6		8.0	11.2	**13.8**	11.5		10.6	14.7	**18.3**	15.3
PGen I-II-ET	Selection %	2.0	1.2	4.5	**14.1**	78.2	2.0	1.0	4.5	**14.2**	78.3	1.9	1.6	4.2	**13.7**	78.6
	Patients		8.1	11.2	**13.5**	11.6		8.4	11.1	**13.5**	11.5		11.0	14.4	**17.9**	15.4
PGen I-II-Sep	Selection %	2.3	11.7	17.3	**58.8**	9.9	2.5	14.4	17.6	**56.3**	9.2	2.1	11.4	15.3	**61.9**	9.3
	Patients		8.3	11.1	**13.4**	11.5		8.6	11.0	**13.5**	11.3		11.2	14.6	**17.9**	15.2

Table 3: (Continued).

Scenario 5																
	Subgroup			**g=1**					**g=2**					**g=3**		
	z_g			**1**					**2**					**1**		
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.08	0.08	0.10	**0.10**		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	0.30	**0.20**	0.15		0.40	0.40	0.40	**0.35**		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.35	0.50	**0.70**	0.80		0.55	0.55	0.55	**0.60**		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	**76.9**	82.5		73.3	73.3	72.6	**75.0**		54.0	65.1	**76.9**	82.5
	$\phi_S^{true}(d, g)$		0.40	0.50	**0.70**	0.60		0.20	0.35	0.50	**0.75**		0.40	0.50	**0.70**	0.60
PGen I-II	Selection %	2.4	7.7	10.3	**67.3**	12.3	5.3	2.2	4.7	24.0	**63.8**	2.3	7.7	10.1	**67.1**	12.8
	Patients		9.5	11.6	**12.8**	10.6		9.5	11.4	13.1	**10.8**		12.7	15.2	**17.3**	14.1
PGen I-II-Comb	Selection %	4.0	2.0	4.0	**64.0**	26.0	4.0	2.0	4.0	64.0	**26.0**	4.0	2.0	4.0	**64.0**	26.0
	Patients		9.1	12.4	**12.2**	10.6		9.5	12.1	12.8	**10.9**		11.6	15.5	**16.7**	14.6
PGen I-II-ET	Selection %	1.7	4.3	5.4	**18.8**	69.8	1.0	10.9	8.6	18.4	**61.1**	1.5	3.8	5.8	**18.5**	70.4
	Patients		9.5	11.6	**12.8**	10.6		10.3	11.5	12.6	**10.6**		12.7	15.1	**17.3**	14.1
PGen I-II-Sep	Selection %	2.1	13.5	14.9	**59.5**	10.0	6.1	3.0	4.1	18.7	**68.1**	2.4	11.6	15.9	**61.8**	8.3
	Patients		9.6	11.6	**12.7**	10.6		9.4	11.4	13.1	**10.9**		12.8	15.1	**17.4**	14.1
Scenario 6																
	Subgroup			**g=1**					**g=2**					**g=3**		
	z_g			**1**					**2**					**1**		
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	**0.12**	0.14	0.16		0.08	0.08	0.10	**0.10**		0.10	**0.12**	0.14	0.16
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	**0.20**	0.15	0.10		0.40	0.40	0.40	**0.35**		0.30	**0.20**	0.15	0.10
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.50	**0.70**	0.80	0.90		0.55	0.55	0.55	**0.60**		0.50	**0.70**	0.80	0.90
	$\phi_{ET}^{true}(d, g)$		65.7	**77.7**	83.2	88.8		73.3	73.3	72.6	**75.0**		65.7	**77.7**	83.2	88.8
	$\phi_S^{true}(d, g)$		0.40	**0.70**	0.60	0.50		0.20	0.35	0.50	**0.75**		0.40	**0.70**	0.60	0.50
PGen I-II	Selection %	0.5	3.9	**80.8**	13.2	1.6	8.2	4.5	12.4	29.0	**45.9**	0.4	3.0	**83.0**	12.6	1.0
	Patients		11.7	**14.1**	12.1	7.2		10.9	13.9	12.3	**7.4**		15.8	**18.7**	16.0	9.5
PGen I-II-Comb	Selection %	0.7	2.4	**59.3**	25.4	12.2	0.7	2.4	59.3	25.4	**12.2**	0.7	2.4	**59.3**	25.4	12.2
	Patients		11.2	**14.4**	12.1	7.2		11.2	14.1	12.3	**7.3**		15.3	**18.8**	16.2	9.6
PGen I-II-ET	Selection %	0.2	10.8	**14.0**	28.8	46.2	0.4	16.7	14.1	27.6	**41.2**	0.2	10.9	**14.0**	28.7	46.2
	Patients		11.7	**14.1**	11.9	7.2		11.9	13.7	12.0	**7.1**		16.0	**18.6**	15.8	9.6
PGen I-II-Sep	Selection %	0.4	12.2	**70.6**	15.7	1.1	9.2	5.5	8.7	27.8	**48.8**	0.5	7.2	**79.2**	12.2	0.9
	Patients		11.8	**14.2**	12.0	7.2		10.9	13.9	12.4	**7.5**		16.0	**18.7**	15.8	9.43

Table 4: (Continued).

Scenario 7																
	Subgroup			**g=1**					**g=2**					**g=3**		
	z_g			**1**					**2**					**1**		
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}(1 \mid d,g)$		0.10	0.12	**0.14**	0.16		0.08	0.08	**0.10**	0.10		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}(1 \mid d,g)$		0.30	0.30	**0.20**	0.15		0.40	0.40	**0.40**	0.35		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}(2 \mid d,g)$		0.35	0.50	**0.70**	0.80		0.55	0.55	**0.55**	0.60		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d,g)$		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5
	$\phi_S^{true}(d,g)$		0.40	0.50	**0.70**	0.60		0.20	0.35	**0.60**	0.40		0.40	0.50	**0.70**	0.60
PGen I-II	Selection %	2.3	7.3	9.8	**70.5**	10.1	6.9	4.4	8.2	**69.3**	11.2	2.1	6.7	11.5	**71.2**	8.5
	Patients		9.5	11.7	**12.8**	10.6		9.5	11.5	**13.0**	10.7		12.6	15.2	**17.4**	14.1
PGen I-II-Comb	Selection %	2.4	2.6	6.5	**80.2**	8.3	2.4	2.6	6.5	**80.2**	8.3	2.4	2.6	6.5	**80.2**	8.3
	Patients		9.2	11.6	**12.9**	10.7		9.3	11.5	**13.1**	10.8		12.5	15.4	**17.4**	14.1
PGen I-II-ET	Selection %	1.7	4.3	5.4	**18.8**	69.8	1.0	10.9	8.6	**18.4**	61.1	1.5	3.8	5.8	**18.5**	70.4
	Patients		9.5	11.6	**12.8**	10.6		10.3	11.5	**12.6**	10.6		12.7	15.1	**17.3**	14.1
PGen I-II-Sep	Selection %	2.2	12.3	16.3	**58.4**	10.8	7.3	4.4	9.0	**68.2**	11.1	2.6	10.7	15.8	**61.6**	9.3
	Patients		9.5	11.5	**12.8**	10.7		9.4	11.4	**13.2**	10.8		12.8	15.1	**17.3**	14.1
Scenario 8																
	Subgroup			**g=1**					**g=2**					**g=3**		
	z_g			**1**					**2**					**3**		
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}(1 \mid d,g)$		0.10	0.12	**0.14**	0.16		0.16	**0.20**	0.26	0.30		0.08	0.08	0.10	**0.10**
	$\pi_E^{\text{true}}(1 \mid d,g)$		0.20	0.20	**0.20**	0.15		0.35	**0.30**	0.30	0.30		0.30	0.30	0.25	**0.20**
	$\pi_E^{\text{true}}(2 \mid d,g)$		0.40	0.45	**0.55**	0.65		0.55	**0.65**	0.65	0.65		0.30	0.45	0.60	**0.70**
	$\phi_{ET}^{true}(d,g)$		54.9	58.2	**65.3**	70.9		69.2	**74.0**	72.0	70.6		50.7	62.4	72.0	**78.4**
	$\phi_S^{true}(d,g)$		0.30	0.40	**0.70**	0.40		0.30	**0.60**	0.40	0.35		0.20	0.30	0.35	**0.75**
PGen I-II	Selection %	12.0	2.5	9.6	**70.0**	5.9	14.4	5.0	**64.2**	11.1	5.3	18.3	0.8	2.7	5.7	**72.5**
	Patients		9.4	11.3	**12.3**	10.7		9.6	**11.8**	12.0	10.8		11.5	14.2	16.3	**16.6**
PGen I-II-Comb	Selection %	8.7	0.8	14.8	**41.8**	33.9	8.7	0.8	**14.8**	41.8	33.9	8.7	0.8	14.8	41.8	**33.9**
	Patients		8.5	11.3	**12.6**	11.1		8.6	**11.1**	12.8	11.3		11.7	15.1	16.8	**15.2**
PGen I-II-ET	Selection %	2.8	10.1	11.8	**14.3**	61.0	3.1	21.4	**15.9**	16.4	43.2	4.3	5.3	14.0	13.8	**62.6**
	Patients		9.7	11.1	**11.9**	11.0		10.0	**11.1**	12.0	10.9		12.2	14.9	16.4	**15.3**
PGen I-II-Sep	Selection %	10.5	3.2	8.8	**76.1**	1.4	12.3	4.8	**69.5**	11.1	2.3	18.4	0.9	1.5	3.7	**75.5**
	Patients		9.6	11.3	**12.3**	10.7		9.7	**11.9**	12.1	10.6		11.5	14.0	16.2	**16.7**

Supplementary Material for "Precision Generalized Phase I-II Designs" by S. Zhao, P.F. Thall, Y. Yuan, J. Lee, P. Msaouel and Y. Zang.

1 Prior Specification and Pseudo Sampling

To specify a prior on $\boldsymbol{\theta}_S^\star$, we assume λ_k normal priors truncated below at 0 independently:

$$p\left(\lambda_k \mid \mu_{\lambda_k}, \varepsilon^2_{\lambda_k}\right) \propto \exp\left\{-\left(\lambda_k - \mu_{\lambda_k}\right)^2 / \left(2\varepsilon^2_{\lambda_k}\right)\right\}, \quad \lambda_k > 0,$$

and denote $\boldsymbol{\mu}_\lambda = (\mu_{\lambda,1}, \cdots, \mu_{\lambda,K})$ and $\boldsymbol{\varepsilon}^2_\lambda = (\varepsilon^2_{\lambda,1}, \cdots, \boldsymbol{\varepsilon}^2_{\lambda,K})$. For each cluster denoted by $k = 1, \cdots, \max(z_g)$, we assume the following independent normal priors:

$$\gamma^\star_{T,j,k} \overset{\text{indep}}{\sim} \mathrm{N}\left(\mu_{\gamma^\star_{T,j,k}}, \varepsilon^2_{\gamma^\star_{T,j,k}}\right), \quad j = 1, \cdots, L_T - 1$$

$$\gamma^\star_{E,j,k} \overset{\text{indep}}{\sim} \mathrm{N}\left(\mu_{\gamma^\star_{E,j,k}}, \varepsilon^2_{\gamma^\star_{E,j,k}}\right), \quad j = 2, \cdots, L_E - 1$$

and

$$\gamma^\star_{D,j,k} \overset{\text{indep}}{\sim} \mathrm{N}\left(\mu_{\gamma^\star_{D,j,k}}, \varepsilon^2_{\gamma^\star_{D,j,k}}\right), \quad j = 2, \cdots, J.$$

We denote

$$\boldsymbol{\mu}_{\gamma^\star_{T,k}} = (\mu_{\gamma^\star_{T,1,k}}, \cdots, \mu_{\gamma_{T,L_T-1,k}}), \quad \boldsymbol{\mu}_{\gamma^\star_T} = (\boldsymbol{\mu}_{\gamma^\star_{T,1}}, \cdots, \boldsymbol{\mu}_{\gamma^\star_{T,\max(z_g)}}),$$

$$\boldsymbol{\mu}_{\gamma^\star_{E,k}} = (\mu_{\gamma^\star_{E,2,k}}, \cdots, \mu_{\gamma_{E,L_E-1,k}}), \quad \boldsymbol{\mu}_{\gamma^\star_E} = (\boldsymbol{\mu}_{\gamma^\star_{E,1}}, \cdots, \boldsymbol{\mu}_{\gamma^\star_{E,\max(z_g)}}),$$

$$\boldsymbol{\mu}_{\gamma^\star_{D,k}} = (\mu_{\gamma^\star_{D,2,k}}, \cdots, \mu_{\gamma_{D,J,k}}), \quad \boldsymbol{\mu}_{\gamma^\star_D} = (\boldsymbol{\mu}_{\gamma^\star_{D,1}}, \cdots, \boldsymbol{\mu}_{\gamma^\star_{D,\max(z_g)}}),$$

$$\boldsymbol{\varepsilon}^2_{\gamma^\star_{T,k}} = (\varepsilon^2_{\gamma^\star_{T,1,k}}, \cdots, \varepsilon^2_{\gamma^\star_{T,L_T-1,k}}), \quad \boldsymbol{\varepsilon}^2_{\gamma^\star_T} = (\boldsymbol{\varepsilon}^2_{\gamma^\star_{T,1}}, \cdots, \boldsymbol{\varepsilon}^2_{\gamma^\star_{T,\max(z_g)}}),$$

$$\boldsymbol{\varepsilon}^2_{\gamma^\star_{E,k}} = (\varepsilon^2_{\gamma^\star_{E,2,k}}, \cdots, \varepsilon^2_{\gamma^\star_{E,L_E-1,k}}), \quad \boldsymbol{\varepsilon}^2_{\gamma^\star_E} = (\boldsymbol{\varepsilon}^2_{\gamma^\star_{E,1}}, \cdots, \boldsymbol{\varepsilon}^2_{\gamma^\star_{E,\max(z_g)}}),$$

$$\boldsymbol{\varepsilon}^2_{\gamma^\star_{D,k}} = (\varepsilon^2_{\gamma^\star_{D,2,k}}, \cdots, \varepsilon^2_{\gamma^\star_{D,J,k}}), \quad \boldsymbol{\varepsilon}^2_{\gamma^\star_d} = (\boldsymbol{\varepsilon}^2_{\gamma^\star_{D,1}}, \cdots, \boldsymbol{\varepsilon}^2_{\gamma^\star_{D,\max(z_g)}}).$$

The vector of all hyperparameters for $\boldsymbol{\theta}^\star_S$ is $\widetilde{\boldsymbol{\theta}}_S = \left(\boldsymbol{\mu}_\lambda, \boldsymbol{\varepsilon}^2_\lambda, \boldsymbol{\mu}_{\gamma^\star_T}, \boldsymbol{\mu}_{\gamma^\star_E}, \boldsymbol{\mu}_{\gamma^\star_D}, \boldsymbol{\varepsilon}^2_{\gamma^\star_T}, \boldsymbol{\varepsilon}^2_{\gamma^\star_E}, \boldsymbol{\varepsilon}^2_{\gamma^\star_d}\right)$.

To apply the pseudo sampling method in the present setting, parametric priors first are defined as follows. For the early outcome dose effect parameters $\boldsymbol{\alpha}^\star_k$ and $\boldsymbol{\beta}^\star_k$ in cluster $k = 1, \cdots, \max(z_g)$, we assume $\alpha^\star_{r,k} \overset{\text{indep}}{\sim} \mathrm{N}\left(\mu_{\alpha^\star_{r,k}}, \varepsilon^2_{\alpha^\star_{r,k}}\right), r = 0, \cdots, 3$ and $\beta^\star_{0,k} \overset{\text{indep}}{\sim} \mathrm{N}\left(\mu_{\beta^\star_{0,k}}, \varepsilon^2_{\beta^\star_{0,k}}\right)$. To ensure $\beta^\star_{1,k} > 0$, we assume a normal prior truncated below at 0:

$$p\left(\beta^\star_{1,k} \mid \mu_{\beta^\star_{1,k}}, \varepsilon^2_{\beta^\star_{1,k}}\right) \propto \exp\left\{-\left(\beta^\star_{1,k} - \mu_{\beta^\star_{1,k}}\right)^2 / \left(2\varepsilon^2_{\beta^\star_{1,k}}\right)\right\}, \quad \beta^\star_{1,k} > 0.$$

Denoting $\boldsymbol{\mu}_{\alpha^\star,k} = (\mu_{\alpha^\star_{0,k}}, \mu_{\alpha^\star_{1,k}}, \mu_{\alpha^\star_{2,k}}, \mu_{\alpha^\star_{3,k}})$, $\boldsymbol{\mu}_{\alpha^\star} = (\boldsymbol{\mu}_{\alpha^\star,1}, \cdots, \boldsymbol{\mu}_{\alpha^\star,\max(z_g)})$; $\boldsymbol{\varepsilon}^2_{\alpha^\star,k} = (\varepsilon^2_{\alpha^\star_{0,k}}, \varepsilon^2_{\alpha^\star_{1,k}}, \varepsilon^2_{\alpha^\star_{2,k}}, \varepsilon^2_{\alpha^\star_{3,k}})$, $\boldsymbol{\varepsilon}^2_{\alpha^\star} = (\boldsymbol{\varepsilon}^2_{\alpha^\star,1}, \cdots, \boldsymbol{\varepsilon}^2_{\alpha^\star,\max(z_g)})$; $\boldsymbol{\mu}_{\beta^\star,k} = (\mu_{\beta^\star_{0,k}}, \mu_{\beta^\star_{1,k}})$, $\boldsymbol{\mu}_{\beta^\star} = (\boldsymbol{\mu}_{\beta^\star,1}, \cdots, \boldsymbol{\mu}_{\beta^\star,\max(z_g)})$; $\boldsymbol{\varepsilon}^2_{\beta^\star,k} = (\varepsilon^2_{\beta^\star_{0,k}}, \varepsilon^2_{\beta^\star_{1,k}})$ and $\boldsymbol{\varepsilon}^2_{\beta^\star} = (\boldsymbol{\varepsilon}^2_{\beta^\star,1}, \cdots, \boldsymbol{\varepsilon}^2_{\beta^\star,\max(z_g)})$, and collecting terms, the vector of all hyperparameters for $\boldsymbol{\theta}^\star_{ET}$ is $\widetilde{\boldsymbol{\theta}}_{ET} = \left(\boldsymbol{\mu}_{\alpha^\star}, \boldsymbol{\varepsilon}^2_{\alpha^\star}, \boldsymbol{\mu}_{\beta^\star}, \boldsymbol{\varepsilon}^2_{\beta^\star}\right)$. The entire hyperparameter vector is $\widetilde{\boldsymbol{\theta}} = (\widetilde{\boldsymbol{\theta}}_{ET}, \widetilde{\boldsymbol{\theta}}_S)$.

In Step 1 of the pseudo sampling method, a vector of outcome probabilities are elicited from the clinical investigators. In Step 2, the elicited probabilities are used to simulate a sample of pseudo outcomes, $\widetilde{\boldsymbol{Y}}_1, \cdots, \widetilde{\boldsymbol{Y}}_N$. In Step 3, starting with a non-informative prior on $\boldsymbol{\theta}'$, a pseudo-posterior $p(\boldsymbol{\theta}' \mid \widetilde{\boldsymbol{Y}}_1, \cdots, \widetilde{\boldsymbol{Y}}_N)$ is derived, and the vector $E(\boldsymbol{\theta} - \sigma_{12} \mid \widetilde{\boldsymbol{Y}}_1, \cdots, \widetilde{\boldsymbol{Y}}_N)$ of pseudo-posterior means is computed. Steps 2 and 3 are repeated a large number of times, and the overall mean of the pseudo-posterior sample means is used as the prior mean of $\boldsymbol{\theta}'$. Prior variances are calibrated by computer simulation to obtain a suitable prior effective sample size (ESS), in the range .10 to 1.00. Finally, the pseudo-posterior sample means and variances are used to solve for the hyperparameters $\widetilde{\boldsymbol{\theta}}$. While a variety of other methods may be used to establish a prior, pseudo sampling is a practical way to incorporate expert

opinion about expected behavior of $\boldsymbol{Y}$ for given doses and patient subgroups.

To apply pseudo-sampling in the present setting, one may begin by eliciting $\widetilde{\pi}_j(y_j \mid d, g)$ for all combinations of (y_j, d, g) and $j = E, T$, and use these probabilities to simulate pseudo-samples $\{\widetilde{Y}_{E,i}(d,g), \widetilde{Y}_{T,i}(d,g), i = 1, \cdots, 100\}$ for each (d, g). Additionally, for distinct fixed time points t_1 and t_2, which may be specified by the clinical investigators, $\widetilde{S}(t_1 \mid d, g)$ and $\widetilde{S}(t_2 \mid d, g)$ may be elicited for each (d, g) pair. Ignoring dependence of Y_S on (Y_E, Y_T), and approximating $p(Y_S \mid d, g)$ by a Weibull distribution with survival function

$$S(t \mid d, g, \alpha_{W,d,g}, \beta_{W,d,g}) = \exp\left\{-(t/\beta_{W,d,g})^{\alpha_{W,d,g}}\right\},$$

one may solve the two equations $\widetilde{S}(t_r \mid d, g) = S(t_r \mid d, g, \alpha_{W,d,g}, \beta_{W,d,g})$, $r = 1, 2$, for $\widehat{\alpha}_{W,d,g}$ and $\widehat{\beta}_{W,d,g}$. This gives estimated survival functions $S(t \mid d, g, \widehat{\alpha}_{W,d,g}, \widehat{\beta}_{W,d,g})$, which may be used to simulate 100 pseudo failure times $\{\widetilde{Y}_{S,i}(d,g), i = 1, \cdots, 100\}$ for each (d, g). The resulting simulated pseudo dataset is

$$\widetilde{\mathcal{D}}(d,g) = \left\{\left(\widetilde{Y}_{E,i}(d,g), \widetilde{Y}_{T,i}(d,g), \widetilde{Y}_{S,i}(d,g)\right), i = 1, \ldots, 100\right\},$$

and the combined pseudo-sample is $\widetilde{\mathcal{D}} = \cup_{d,g}\widetilde{\mathcal{D}}(d,g)$. We then assume a very non-informative pseudo-prior on $\boldsymbol{\theta'}$ and compute a pseudo posterior of $[\boldsymbol{\theta'} \mid \widetilde{\mathcal{D}}]$. Repeating this simulation-to-posterior computation process 1000 times, one may use the pseudo posterior means as the prior means of $\boldsymbol{\theta'}$. To ensure appropriate prior uncertainty, the prior dispersion parameters may be calibrated to obtain prior ESS = 0.10. The prior means and dispersion parameters then determine the hyperparameters $\widetilde{\boldsymbol{\theta}} = (\widetilde{\boldsymbol{\theta}}_{ET}, \widetilde{\boldsymbol{\theta}}_S)$.

2 Weakly-informative priors used in the simulations

$$\Pr(\nu = k) = 1/5,\ k = 1, \cdots, 5,$$

$$\sigma_{12} \sim \text{Unif}[-1, 1],$$

$$\gamma^{\star}_{E,j,k} \overset{\text{indep}}{\sim} \text{N}(0, 10),$$

$$\gamma^{\star}_{T,j,k} \overset{\text{indep}}{\sim} \text{N}(0, 10),$$

$$\gamma^{\star}_{D,j,k} \overset{\text{indep}}{\sim} \text{N}(0, 10),$$

$$\kappa_{j,k} \overset{\text{indep}}{\sim} \text{U}(0, \infty),$$

$$\alpha^{\star}_{r,k} \overset{\text{indep}}{\sim} \text{N}(0, 10),$$

$$\beta^{\star}_{0,k} \overset{\text{indep}}{\sim} \text{N}(0, 10),$$

$$p\left(\beta^{\star}_{1,k}\right) \propto \exp\left\{-\left(\beta^{\star}_{1,k}\right)^2 / (20)\right\}, \quad \beta^{\star}_{1,k} > 0,$$

$$p(\lambda_k) \propto \exp\left\{-(\lambda_k)^2 / (20)\right\}, \quad \lambda_k > 0.$$

Table S.1: Three Disease Subgroups in the CD70 CAR NK Cell Trial. AML = acute myelogeneous leukemia, ALL = acute lymphocytic leukemia, CMML = chronic myelomonocytic leukemia, MDS = myelodysplastic syndrome, and CML = chronic myelogeneous leukemia

Disease Group	Description
1	AML, ALL, CMML, MDS, and blastic CML
2	Non-T-cell Hodgkin's and non-Hodgkin's lymphoma
3	T-cell non-Hodgkin's lymphoma

Table S.2: Utility Function for the Early Outcomes

	$Y_E = 0$	$Y_E = 1$	$Y_E = 2$
$Y_T = 0$	20	50	100
$Y_T = 1$	0	30	60

Table S.3: Simulation results to assess sensitivity of the PGen I-II design to t_l = 1, 3, or 6 months.

Scenario 1

	Subgroup	**g=1**					**g=2**					**g=3**				
	z_g	**1**					**1**					**1**				
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.12	0.22	0.36	0.50		0.12	0.22	0.36	0.50		0.12	0.22	0.36	0.50
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.35	0.30	0.20	0.15		0.35	0.30	0.20	0.15		0.35	0.30	0.20	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.45	0.60	0.75	0.80		0.45	0.60	0.75	0.80		0.45	0.60	0.75	0.80
	$\phi_{ET}^{true}(d, g)$		62.7	69.5	72.9	70.1		62.7	69.5	72.9	70.1		62.7	69.5	72.9	70.1
t_l	$\phi_S^{true}(d, g)$		0.10	0.20	0.30	0.45		0.10	0.20	0.30	0.45		0.10	0.20	0.30	0.45
1	Selection %	**93.2**	0.2	0.1	0.8	5.7	**93.2**	0.2	0.0	0.7	5.9	**93.9**	0.2	0.0	1.1	4.8
	Patients		12.4	11.8	10.4	3.7		12.7	11.9	10.6	3.7		17.2	16.0	13.8	4.3
3	Selection %	**93.2**	0.1	0.0	1.6	5.1	**92.5**	0.1	0.0	1.5	5.9	**94.3**	0.0	0.0	1.4	4.3
	Patients		12.5	11.8	10.5	3.7		12.7	11.9	10.6	3.8		17.2	15.9	13.6	4.3
6	Selection %	**93.1**	0.2	0.0	1.0	5.7	**93.4**	0.1	0.0	0.7	5.8	**94.2**	0.0	0.0	0.6	5.2
	Patients		12.4	11.7	10.3	3.8		12.7	11.9	10.6	3.8		17.3	15.8	13.5	4.4

Scenario 2

	Subgroup	**g=1**					**g=2**					**g=3**				
	z_g	**1**					**2**					**1**				
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.14	0.18	0.24		0.10	0.12	**0.14**	0.16		0.10	0.14	0.18	0.24
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.35	0.35	0.35	0.30		0.30	0.30	**0.20**	0.15		0.35	0.35	0.35	0.30
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.45	0.50	0.50	0.60		0.35	0.50	**0.70**	0.80		0.45	0.50	0.50	0.60
	$\phi_{ET}^{true}(d, g)$		63.3	65.9	64.7	68.8		54.0	65.1	**76.9**	82.5		63.3	65.9	64.7	68.8
t_l	$\phi_S^{true}(d, g)$		0.10	0.20	0.25	0.30		0.20	0.30	**0.70**	0.40		0.10	0.20	0.25	0.30
1	Selection %	**85.1**	1.4	0.8	5.1	7.6	21.4	1.8	2.4	**71.3**	3.1	**87.2**	1.2	0.6	3.7	7.3
	Patients		8.9	10.4	11.6	10.2		10.5	12.9	**14.7**	10.2		12.3	13.8	15.5	13.6
3	Selection %	**83.9**	1.8	0.6	6.1	7.6	22.3	2.1	2.3	**70.5**	2.8	**86.1**	1.1	0.8	4.6	7.4
	Patients		8.9	10.4	11.7	10.1		10.6	12.8	**14.7**	10.4		12.2	13.7	15.6	13.7
6	Selection %	**85.3**	1.6	0.5	5.9	6.7	23.6	2.4	1.5	**70.2**	2.3	**87.2**	1.3	0.3	4.5	6.7
	Patients		8.9	10.5	11.6	10.2		10.7	12.9	**14.7**	10.4		12.2	13.8	15.4	13.5

Table S.3: (Continued).

Scenario 3																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**2**					**1**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	0.14	**0.16**		0.10	0.14	0.18	0.24		0.10	0.12	0.14	**0.16**
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	0.30	0.20	**0.15**		0.35	0.35	0.35	0.30		0.30	0.30	0.20	**0.15**
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.35	0.50	0.70	**0.80**		0.45	0.50	0.50	0.60		0.35	0.50	0.70	**0.80**
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	76.9	**82.5**		63.3	65.9	64.7	68.8		54.0	65.1	76.9	**82.5**
t_l	$\phi_S^{true}(d, g)$		0.20	0.30	0.40	**0.60**		0.05	0.10	0.15	0.20		0.20	0.30	0.40	**0.60**
1	Selection %	22.5	1.2	2.2	4.6	**69.5**	**91.5**	1.2	0.3	0.9	6.1	22.1	1.1	2.4	4.5	**69.9**
	Patients		8.9	11.1	12.8	**12.3**		8.7	9.5	11.4	11.4		12.1	14.4	17.1	**16.1**
3	Selection %	22.2	1.8	2.7	5.5	**67.8**	**91.3**	1.3	0.4	0.6	6.4	22.9	1.0	1.9	5.5	**68.7**
	Patients		8.9	11.1	12.7	**12.3**		8.6	9.4	11.3	11.4		12.1	14.4	17.3	**16.2**
6	Selection %	21.8	1.9	1.7	4.9	**69.7**	**92.0**	1.1	0.4	0.6	5.9	21.4	1.2	1.8	5.5	**70.1**
	Patients		8.9	11.1	12.8	**12.3**		8.6	9.5	11.3	11.4		12.2	14.4	17.2	**16.2**
Scenario 4																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**1**					**1**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.10	0.12	**0.14**	0.16		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	0.30	**0.20**	0.15		0.30	0.30	**0.20**	0.15		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.35	0.50	**0.70**	0.80		0.35	0.50	**0.70**	0.80		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5
t_l	$\phi_S^{true}(d, g)$		0.40	0.50	**0.70**	0.60		0.40	0.50	**0.70**	0.60		0.40	0.50	**0.70**	0.60
1	Selection %	2.2	4.5	7.0	**77.8**	8.5	2.2	4.6	6.3	**78.5**	8.4	2.2	4.5	5.6	**79.8**	7.9
	Patients		8.2	11.1	**13.7**	11.6		8.5	11.0	**13.5**	11.5		11.0	14.4	**18.0**	15.3
3	Selection %	2.3	4.1	6.4	**78.8**	8.4	2.1	3.6	4.7	**81.1**	8.5	2.2	4.5	4.4	**81.4**	7.5
	Patients		8.1	11.1	**13.6**	11.5		8.5	11.1	**13.5**	11.5		11.0	14.5	**18.1**	15.2
6	Selection %	2.3	3.7	7.5	**78.8**	7.7	2.3	3.5	7.6	**79.2**	7.4	2.1	4.0	6.3	**80.6**	7.0
	Patients		8.1	11.2	**13.6**	11.6		8.4	11.0	**13.6**	11.4		11.1	14.5	**18.0**	15.2

Table S.3: (Continued).

Scenario 5

	Subgroup	**g=1**					**g=2**					**g=3**				
	z_g	**1**					**2**					**1**				
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}(1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.08	0.08	0.10	**0.10**		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}(1 \mid d, g)$		0.30	0.30	**0.20**	0.15		0.40	0.40	0.40	**0.35**		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}(2 \mid d, g)$		0.35	0.50	**0.70**	0.80		0.55	0.55	0.55	**0.60**		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	**76.9**	82.5		73.3	73.3	72.6	**75.0**		54.0	65.1	**76.9**	82.5
t_l	$\phi_S^{true}(d, g)$		0.40	0.50	**0.70**	0.60		0.20	0.35	0.50	**0.75**		0.40	0.50	**0.70**	0.60
1	Selection %	2.3	6.8	8.8	**68.5**	13.6	6.1	2.5	4.1	22.5	**64.8**	2.3	6.8	9.5	**67.7**	13.7
	Patients		9.4	11.7	**12.7**	10.7		9.5	11.4	13.0	**10.8**		12.7	15.2	**17.3**	14.2
3	Selection %	2.4	7.7	10.3	**67.3**	12.3	5.3	2.2	4.7	24.0	**63.8**	2.3	7.7	10.1	**67.1**	12.8
	Patients		9.5	11.6	**12.8**	10.6		9.5	11.4	13.1	**10.8**		12.7	15.2	**17.3**	14.1
6	Selection %	2.2	8.7	10.0	**67.8**	11.3	5.8	2.4	5.3	24.2	**62.3**	2.1	7.4	10.1	**68.3**	12.1
	Patients		9.5	11.7	**12.8**	10.7		9.4	11.4	13.0	**10.8**		12.6	15.2	**17.4**	14.1

Scenario 6

	Subgroup	**g=1**					**g=2**					**g=3**				
	z_g	**1**					**2**					**1**				
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}(1 \mid d, g)$		0.10	**0.12**	0.14	0.16		0.08	0.08	0.10	**0.10**		0.10	**0.12**	0.14	0.16
	$\pi_E^{\text{true}}(1 \mid d, g)$		0.30	**0.20**	0.15	0.10		0.40	0.40	0.40	**0.35**		0.30	**0.20**	0.15	0.10
	$\pi_E^{\text{true}}(2 \mid d, g)$		0.50	**0.70**	0.80	0.90		0.55	0.55	0.55	**0.60**		0.50	**0.70**	0.80	0.90
	$\phi_{ET}^{true}(d, g)$		65.7	**77.7**	83.2	88.8		73.3	73.3	72.6	**75.0**		65.7	**77.7**	83.2	88.8
t_l	$\phi_S^{true}(d, g)$		0.40	**0.70**	0.60	0.50		0.20	0.35	0.50	**0.75**		0.40	**0.70**	0.60	0.50
1	Selection %	0.4	3.9	**80.2**	13.4	2.1	8.7	4.0	12.3	28.6	**46.4**	0.3	3.4	**82.8**	12.0	1.5
	Patients		11.5	**14.2**	12.1	7.2		11.0	13.9	12.3	**7.5**		15.9	**18.7**	15.9	9.5
3	Selection %	0.5	3.9	**80.8**	13.2	1.6	8.2	4.5	12.4	29.0	**45.9**	0.4	3.0	**83.0**	12.6	1.0
	Patients		11.7	**14.1**	12.1	7.2		10.9	13.9	12.3	**7.4**		15.8	**18.7**	16.0	9.5
6	Selection %	0.5	4.4	**80.1**	13.7	1.3	7.8	4.9	11.9	29.6	**45.8**	0.5	3.6	**81.9**	12.6	1.4
	Patients		11.6	**14.2**	12.1	7.2		11.0	13.9	12.3	**7.5**		15.8	**18.7**	16.0	9.6

Table S.3: (Continued).

Scenario 7

	Subgroup	**g=1**					**g=2**					**g=3**				
	z_g	**1**					**2**					**1**				
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.08	0.08	**0.10**	0.10		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	0.30	**0.20**	0.15		0.40	0.40	**0.40**	0.35		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.35	0.50	**0.70**	0.80		0.55	0.55	**0.55**	0.60		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5
t_l	$\phi_S^{true}(d, g)$		0.40	0.50	**0.70**	0.60		0.20	0.35	**0.60**	0.40		0.40	0.50	**0.70**	0.60
1	Selection %	2.2	7.8	10.2	**70.6**	9.2	6.8	4.8	6.7	**69.1**	12.6	2.1	6.6	10.0	**72.6**	8.7
	Patients		9.4	11.6	**12.8**	10.7		9.5	11.4	**13.0**	10.8		12.8	15.2	**17.3**	14.1
3	Selection %	2.3	7.3	9.8	**70.5**	10.1	6.9	4.4	8.2	**69.3**	11.2	2.1	6.7	11.5	**71.2**	8.5
	Patients		9.5	11.7	**12.8**	10.6		9.5	11.5	**13.0**	10.7		12.6	15.2	**17.4**	14.1
6	Selection %	2.2	7.8	9.6	**71.7**	8.7	6.4	4.5	8.1	**70.1**	10.9	2.2	6.8	10.1	**72.8**	8.1
	Patients		9.4	11.6	**12.8**	10.7		9.5	11.5	**13.1**	10.7		12.7	15.2	**17.4**	14.1

Scenario 8

	Subgroup	**g=1**					**g=2**					**g=3**				
	z_g	**1**					**2**					**3**				
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.16	**0.20**	0.26	0.30		0.08	0.08	0.10	**0.10**
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.20	0.20	**0.20**	0.15		0.35	**0.30**	0.30	0.30		0.30	0.30	0.25	**0.20**
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.40	0.45	**0.55**	0.65		0.55	**0.65**	0.65	0.65		0.30	0.45	0.60	**0.70**
	$\phi_{ET}^{true}(d, g)$		54.9	58.2	**65.3**	70.9		69.2	**74.0**	72.0	70.6		50.7	62.4	72.0	**78.4**
t_l	$\phi_S^{true}(d, g)$		0.30	0.40	**0.70**	0.40		0.30	**0.60**	0.40	0.35		0.20	0.30	0.35	**0.75**
1	Selection %	10.8	2.5	9.4	**71.6**	5.7	12.7	4.9	**66.2**	11.3	4.9	18.7	0.9	2.5	5.2	**72.7**
	Patients		9.5	11.3	**12.4**	10.7		9.5	**11.8**	12.0	10.7		11.5	14.3	16.2	**16.4**
3	Selection %	12.0	2.5	9.6	**70.0**	5.9	14.4	5.0	**64.2**	11.1	5.3	18.3	0.8	2.7	5.7	**72.5**
	Patients		9.4	11.3	**12.3**	10.7		9.6	**11.8**	12.0	10.8		11.5	14.2	16.3	**16.6**
6	Selection %	11.7	3.0	8.5	**71.1**	5.7	13.9	4.7	**65.4**	12.4	3.6	17.4	1.0	2.4	6.4	**72.8**
	Patients		9.5	11.3	**12.5**	10.6		9.6	**11.8**	11.9	10.7		11.5	14.2	16.2	**16.6**

Table S.4: Simulation results for the sensitivity of the PGen I-II design to $\rho = 0.5$, 0.7, or 0.9.

Scenario 1																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**1**					**1**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.12	0.22	0.36	0.50		0.12	0.22	0.36	0.50		0.12	0.22	0.36	0.50
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.35	0.30	0.20	0.15		0.35	0.30	0.20	0.15		0.35	0.30	0.20	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.45	0.60	0.75	0.80		0.45	0.60	0.75	0.80		0.45	0.60	0.75	0.80
	$\phi_{ET}^{true}(d, g)$		62.7	69.5	72.9	70.1		62.7	69.5	72.9	70.1		62.7	69.5	72.9	70.1
ρ	$\phi_S^{true}(d, g)$		0.10	0.20	0.30	0.45		0.10	0.20	0.30	0.45		0.10	0.20	0.30	0.45
0.5	Selection %	**93.2**	0.1	0.0	1.6	5.1	**92.5**	0.1	0.0	1.5	5.9	**94.3**	0.0	0.0	1.4	4.3
	Patients		12.5	11.8	10.5	3.7		12.7	11.9	10.6	3.8		17.2	15.9	13.6	4.3
0.7	Selection %	**93.2**	0.1	0.0	1.6	5.1	**92.5**	0.1	0.0	1.5	5.9	**94.3**	0.0	0.0	1.4	4.3
	Patients		12.5	11.8	10.5	3.7		12.7	11.9	10.6	3.8		17.2	15.9	13.6	4.3
0.9	Selection %	**93.4**	0.1	0.0	1.6	4.9	**92.8**	0.1	0.0	1.5	5.6	**94.3**	0.0	0.0	1.4	4.3
	Patients		12.5	11.8	10.5	3.6		12.7	11.9	10.6	3.7		17.3	15.9	13.6	4.3
Scenario 2																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**2**					**1**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.14	0.18	0.24		0.10	0.12	**0.14**	0.16		0.10	0.14	0.18	0.24
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.35	0.35	0.35	0.30		0.30	0.30	**0.20**	0.15		0.35	0.35	0.35	0.30
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.45	0.50	0.50	0.60		0.35	0.50	**0.70**	0.80		0.45	0.50	0.50	0.60
	$\phi_{ET}^{true}(d, g)$		63.3	65.9	64.7	68.8		54.0	65.1	**76.9**	82.5		63.3	65.9	64.7	68.8
ρ	$\phi_S^{true}(d, g)$		0.10	0.20	0.25	0.30		0.20	0.30	**0.70**	0.40		0.10	0.20	0.25	0.30
0.5	Selection %	**83.9**	1.8	0.6	6.1	7.6	22.3	2.1	2.3	**70.5**	2.8	**86.1**	1.1	0.8	4.6	7.4
	Patients		8.9	10.4	11.7	10.1		10.6	12.8	**14.7**	10.4		12.2	13.7	15.6	13.7
0.7	Selection %	**83.9**	1.8	0.6	6.1	7.6	22.3	2.1	2.3	**70.5**	2.8	**86.1**	1.1	0.8	4.6	7.4
	Patients		8.9	10.4	11.7	10.1		10.6	12.8	**14.7**	10.4		12.2	13.7	15.6	13.7
0.9	Selection %	**84.2**	1.8	0.6	5.8	7.6	22.9	1.8	2.1	**70.4**	2.8	**86.4**	1.1	0.8	4.5	7.2
	Patients		8.9	10.4	11.8	10.1		10.5	12.8	**14.8**	10.5		12.2	13.7	15.7	13.6

Table S.4: (Continued).

Scenario 3																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**2**					**1**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	0.14	**0.16**		0.10	0.14	0.18	0.24		0.10	0.12	0.14	**0.16**
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	0.30	0.20	**0.15**		0.35	0.35	0.35	0.30		0.30	0.30	0.20	**0.15**
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.35	0.50	0.70	**0.80**		0.45	0.50	0.50	0.60		0.35	0.50	0.70	**0.80**
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	76.9	**82.5**		63.3	65.9	64.7	68.8		54.0	65.1	76.9	**82.5**
ρ	$\phi_S^{true}(d, g)$		0.20	0.30	0.40	**0.60**		0.05	0.10	0.15	0.20		0.20	0.30	0.40	**0.60**
0.5	Selection %	22.1	1.5	3.8	4.7	**67.9**	**92.8**	0.9	0.2	0.5	5.6	22.1	1.0	2.8	4.6	**69.5**
	Patients		8.9	11.1	12.8	**12.4**		8.6	9.4	11.4	11.4		12.2	14.4	17.4	**16.1**
0.7	Selection %	22.2	1.8	2.7	5.5	**67.8**	**91.3**	1.3	0.4	0.6	6.4	22.9	1.0	1.9	5.5	**68.7**
	Patients		8.9	11.1	12.7	**12.3**		8.6	9.4	11.3	11.4		12.1	14.4	17.3	**16.2**
0.9	Selection %	21.7	1.5	3.8	4.9	**68.1**	**93.2**	0.5	0.2	0.5	5.6	22.0	0.9	2.6	5.1	**69.4**
	Patients		8.6	10.9	13.0	**12.7**		8.6	9.4	11.4	11.3		11.9	14.0	17.6	**16.7**
Scenario 4																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**1**					**1**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.10	0.12	**0.14**	0.16		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	0.30	**0.20**	0.15		0.30	0.30	**0.20**	0.15		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.35	0.50	**0.70**	0.80		0.35	0.50	**0.70**	0.80		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5
ρ	$\phi_S^{true}(d, g)$		0.40	0.50	**0.70**	0.60		0.40	0.50	**0.70**	0.60		0.40	0.50	**0.70**	0.60
0.5	Selection %	2.4	4.2	5.7	**78.6**	9.1	2.4	4.0	6.4	**77.7**	9.5	2.2	5.0	5.5	**79.0**	8.3
	Patients		8.3	11.1	**13.4**	11.5		8.5	11.1	**13.5**	11.4		11.3	14.5	**17.8**	15.3
0.7	Selection %	2.3	4.1	6.4	**78.8**	8.4	2.1	3.6	4.7	**81.1**	8.5	2.2	4.5	4.4	**81.4**	7.5
	Patients		8.1	11.1	**13.6**	11.5		8.5	11.1	**13.5**	11.5		11.0	14.5	**18.1**	15.2
0.9	Selection %	2.4	3.4	6.0	**79.9**	8.3	2.4	3.7	5.8	**79.8**	8.3	2.2	3.8	5.0	**81.5**	7.5
	Patients		7.6	10.5	**14.1**	12.2		7.8	10.4	**14.1**	12.1		10.3	13.6	**18.8**	16.1

Table S.4: (Continued).

Scenario 5

	Subgroup	**g=1**					**g=2**					**g=3**				
	z_g	**1**					**2**					**1**				
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.08	0.08	0.10	**0.10**		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	0.30	**0.20**	0.15		0.40	0.40	0.40	**0.35**		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.35	0.50	**0.70**	0.80		0.55	0.55	0.55	**0.60**		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	**76.9**	82.5		73.3	73.3	72.6	**75.0**		54.0	65.1	**76.9**	82.5
ρ	$\phi_S^{true}(d, g)$		0.40	0.50	**0.70**	0.60		0.20	0.35	0.50	**0.75**		0.40	0.50	**0.70**	0.60
0.5	Selection %	2.3	8.0	10.0	**68.3**	11.4	6.3	3.0	4.9	23.4	**62.4**	2.3	7.3	9.7	**69.2**	11.5
	Patients		9.5	11.6	**12.8**	10.7		9.4	11.4	13.0	**10.8**		12.7	15.1	**17.4**	14.1
0.7	Selection %	2.4	7.7	10.3	**67.3**	12.3	5.3	2.2	4.7	24.0	**63.8**	2.3	7.7	10.1	**67.1**	12.8
	Patients		9.5	11.6	**12.8**	10.6		9.5	11.4	13.1	**10.8**		12.7	15.2	**17.3**	14.1
0.9	Selection %	2.3	6.1	9.3	**69.7**	12.6	6.2	2.5	5.2	23.6	**62.5**	2.4	6.1	9.4	**68.9**	13.2
	Patients		9.0	11.3	**13.1**	11.0		9.3	11.3	13.1	**10.8**		12.2	14.7	**17.9**	14.6

Scenario 6

	Subgroup	**g=1**					**g=2**					**g=3**				
	z_g	**1**					**2**					**1**				
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	**0.12**	0.14	0.16		0.08	0.08	0.10	**0.10**		0.10	**0.12**	0.14	0.16
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	**0.20**	0.15	0.10		0.40	0.40	0.40	**0.35**		0.30	**0.20**	0.15	0.10
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.50	**0.70**	0.80	0.90		0.55	0.55	0.55	**0.60**		0.50	**0.70**	0.80	0.90
	$\phi_{ET}^{true}(d, g)$		65.7	**77.7**	83.2	88.8		73.3	73.3	72.6	**75.0**		65.7	**77.7**	83.2	88.8
ρ	$\phi_S^{true}(d, g)$		0.40	**0.70**	0.60	0.50		0.20	0.35	0.50	**0.75**		0.40	**0.70**	0.60	0.50
0.5	Selection %	0.5	3.9	**80.8**	13.2	1.6	8.2	4.5	12.4	29.0	**45.9**	0.4	3.0	**83.0**	12.6	1.0
	Patients		11.7	**14.1**	12.1	7.2		10.9	13.9	12.3	**7.4**		15.8	**18.7**	16.0	9.5
0.7	Selection %	0.5	3.9	**80.8**	13.2	1.6	8.2	4.5	12.4	29.0	**45.9**	0.4	3.0	**83.0**	12.6	1.0
	Patients		11.7	**14.1**	12.1	7.2		10.9	13.9	12.3	**7.4**		15.8	**18.7**	16.0	9.5
0.9	Selection %	0.5	3.6	**76.8**	16.5	2.6	8.2	3.8	12.0	29.1	**46.9**	0.4	2.7	**79.2**	15.8	1.9
	Patients		11.3	**14.1**	12.3	7.4		10.9	13.9	12.3	**7.5**		15.4	**18.7**	16.3	9.8

Table S.4: (Continued).

Scenario 7																
	Subgroup				**g=1**					**g=2**					**g=3**	
	z_g				**1**					**2**					**1**	
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.08	0.08	**0.10**	0.10		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	0.30	**0.20**	0.15		0.40	0.40	**0.40**	0.35		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.35	0.50	**0.70**	0.80		0.55	0.55	**0.55**	0.60		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5
ρ	$\phi_S^{true}(d, g)$		0.40	0.50	**0.70**	0.60		0.20	0.35	**0.60**	0.40		0.40	0.50	**0.70**	0.60
0.5	Selection %	2.3	7.7	10.3	**71.1**	8.6	6.1	4.4	7.6	**69.1**	12.8	2.2	6.9	9.9	**72.5**	8.5
	Patients		9.5	11.6	**12.8**	10.7		9.5	11.4	**13.2**	10.7		12.7	15.2	**17.3**	14.0
0.7	Selection %	2.3	7.3	9.8	**70.5**	10.1	6.9	4.4	8.2	**69.3**	11.2	2.1	6.7	11.5	**71.2**	8.5
	Patients		9.5	11.7	**12.8**	10.6		9.5	11.5	**13.0**	10.7		12.6	15.2	**17.4**	14.1
0.9	Selection %	2.3	6.0	10.3	**73.3**	8.1	6.6	3.8	8.3	**68.5**	12.8	2.3	6.0	10.0	**74.2**	7.5
	Patients		9.1	11.3	**13.1**	11.0		9.3	11.4	**13.3**	10.8		12.2	14.7	**17.8**	14.6
Scenario 8																
	Subgroup				**g=1**					**g=2**					**g=3**	
	z_g				**1**					**2**					**3**	
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.16	**0.20**	0.26	0.30		0.08	0.08	0.10	**0.10**
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.20	0.20	**0.20**	0.15		0.35	**0.30**	0.30	0.30		0.30	0.30	0.25	**0.20**
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.40	0.45	**0.55**	0.65		0.55	**0.65**	0.65	0.65		0.30	0.45	0.60	**0.70**
	$\phi_{ET}^{true}(d, g)$		54.9	58.2	**65.3**	70.9		69.2	**74.0**	72.0	70.6		50.7	62.4	72.0	**78.4**
ρ	$\phi_S^{true}(d, g)$		0.30	0.40	**0.70**	0.40		0.30	**0.60**	0.40	0.35		0.20	0.30	0.35	**0.75**
0.5	Selection %	11.1	2.6	10.1	**71.0**	5.2	13.5	4.5	**64.2**	13.8	4.0	18.2	0.4	3.2	5.9	**72.3**
	Patients		9.4	11.4	**12.5**	10.6		9.5	**11.8**	11.9	10.8		11.5	14.2	16.34	**16.5**
0.7	Selection %	12.0	2.5	9.6	**70.0**	5.9	14.4	5.0	**64.2**	11.1	5.3	18.3	0.8	2.7	5.7	**72.5**
	Patients		9.4	11.3	**12.3**	10.7		9.6	**11.8**	12.0	10.8		11.5	14.2	16.3	**16.6**
0.9	Selection %	11.0	2.5	10.1	**71.7**	4.7	13.7	4.5	**63.6**	14.0	4.2	18.3	0.4	3.4	5.7	**72.2**
	Patients		9.4	11.4	**12.5**	10.7		9.5	**11.8**	12.0	10.7		11.4	14.0	16.5	**16.7**

Table S.5: Simulation results for sensitivity of the PGen I-II design to the data-generating distribution for Y_S'.

Scenario 1																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**1**					**1**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.12	0.22	0.36	0.50		0.12	0.22	0.36	0.50		0.12	0.22	0.36	0.50
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.35	0.30	0.20	0.15		0.35	0.30	0.20	0.15		0.35	0.30	0.20	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.45	0.60	0.75	0.80		0.45	0.60	0.75	0.80		0.45	0.60	0.75	0.80
	$\phi_{ET}^{true}(d, g)$		62.7	69.5	72.9	70.1		62.7	69.5	72.9	70.1		62.7	69.5	72.9	70.1
	$\phi_S^{true}(d, g)$		0.10	0.20	0.30	0.45		0.10	0.20	0.30	0.45		0.10	0.20	0.30	0.45
Weibull	Selection %	**93.2**	0.1	0.0	1.6	5.1	**92.5**	0.1	0.0	1.5	5.9	**94.3**	0.0	0.0	1.4	4.3
	Patients		12.5	11.8	10.5	3.7		12.7	11.9	10.6	3.8		17.2	15.9	13.6	4.3
Log-normal	Selection %	**99.9**	0.0	0.0	0.0	0.1	**99.7**	0.0	0.0	0.2	0.1	**100.0**	0.0	0.0	0.0	0.0
	Patients		12.8	12.2	9.1	3.0		13.1	12.2	9.1	3.3		16.8	15.8	11.0	3.7
Gamma	Selection %	**99.2**	0.0	0.0	0.0	0.8	**99.9**	0.0	0.0	0.0	0.1	**99.4**	0.0	0.0	0.0	0.6
	Patients		13.2	11.9	8.6	3.4		13.3	11.9	8.9	3.4		17.2	15.6	11.0	4.3
Scenario 2																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**2**					**1**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.14	0.18	0.24		0.10	0.12	**0.14**	0.16		0.10	0.14	0.18	0.24
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.35	0.35	0.35	0.30		0.30	0.30	**0.20**	0.15		0.35	0.35	0.35	0.30
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.45	0.50	0.50	0.60		0.35	0.50	**0.70**	0.80		0.45	0.50	0.50	0.60
	$\phi_{ET}^{true}(d, g)$		63.3	65.9	64.7	68.8		54.0	65.1	**76.9**	82.5		63.3	65.9	64.7	68.8
	$\phi_S^{true}(d, g)$		0.10	0.20	0.25	0.30		0.20	0.30	**0.70**	0.40		0.10	0.20	0.25	0.30
Weibull	Selection %	**83.9**	1.8	0.6	6.1	7.6	22.3	2.1	2.3	**70.5**	2.8	**86.1**	1.1	0.8	4.6	7.4
	Patients		8.9	10.4	11.7	10.1		10.6	12.8	**14.7**	10.4		12.2	13.7	15.6	13.7
Log-normal	Selection %	**89.2**	0.4	0.4	7.5	2.5	23.8	0.6	1.1	**68.7**	5.8	**89.9**	0.3	0.4	6.8	2.6
	Patients		9.0	10.3	11.8	9.7		10.0	12.3	**15.6**	11.6		12.4	13.7	15.4	13.3
Gamma	Selection %	**87.8**	1.0	0.2	7.6	3.4	27.4	0.7	0.5	**65.6**	5.8	**88.5**	0.5	0.2	6.8	4.0
	Patients		9.3	10.0	12.0	10.2		10.4	12.2	**15.4**	11.1		12.5	13.1	15.6	13.2

Table S.5: (Continued).

Scenario 3																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**2**					**1**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	0.14	**0.16**		0.10	0.14	0.18	0.24		0.10	0.12	0.14	**0.16**
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	0.30	0.20	**0.15**		0.35	0.35	0.35	0.30		0.30	0.30	0.20	**0.15**
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.35	0.50	0.70	**0.80**		0.45	0.50	0.50	0.60		0.35	0.50	0.70	**0.80**
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	76.9	**82.5**		63.3	65.9	64.7	68.8		54.0	65.1	76.9	**82.5**
	$\phi_S^{true}(d, g)$		0.20	0.30	0.40	**0.60**		0.05	0.10	0.15	0.20		0.20	0.30	0.40	**0.60**
Weibull	Selection %	22.2	1.8	2.7	5.5	**67.8**	**91.3**	1.3	0.4	0.6	6.4	22.9	1.0	1.9	5.5	**68.7**
	Patients		8.9	11.1	12.7	**12.3**		8.6	9.4	11.3	11.4		12.1	14.4	17.3	**16.2**
Log-normal	Selection %	17.6	0.5	0.5	10.6	**70.8**	**83.7**	0.6	0.3	1.3	14.1	16.6	0.1	0.6	10.0	**72.7**
	Patients		8.3	10.2	13.3	**12.9**		8.5	9.4	11.4	11.4		11.2	13.4	17.8	**17.5**
Gamma	Selection %	22.4	0.9	1.4	9.5	**65.8**	**83.2**	0.8	0.1	1.0	14.9	20.9	0.6	1.1	8.5	**68.9**
	Patients		8.1	10.9	13.5	**12.3**		8.3	9.8	11.6	11.1		10.9	14.1	18.1	**16.6**
Scenario 4																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**1**					**1**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.10	0.12	**0.14**	0.16		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	0.30	**0.20**	0.15		0.30	0.30	**0.20**	0.15		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.35	0.50	**0.70**	0.80		0.35	0.50	**0.70**	0.80		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5
	$\phi_S^{true}(d, g)$		0.40	0.50	**0.70**	0.60		0.40	0.50	**0.70**	0.60		0.40	0.50	**0.70**	0.60
Weibull	Selection %	2.3	4.1	6.4	**78.8**	8.4	2.1	3.6	4.7	**81.1**	8.5	2.2	4.5	4.4	**81.4**	7.5
	Patients		8.1	11.1	**13.6**	11.5		8.5	11.1	**13.5**	11.5		11.0	14.5	**18.1**	15.2
Log-normal	Selection %	2.3	2.8	4.1	**76.9**	13.9	2.3	2.5	4.3	**77.7**	13.2	2.4	1.3	3.8	**80.2**	12.3
	Patients		8.3	10.9	**13.7**	11.5		8.3	10.9	**13.5**	11.6		11.0	14.4	**18.0**	15.4
Gamma	Selection %	2.7	2.3	4.6	**74.4**	16.0	2.4	3.2	5.6	**72.1**	16.7	2.2	1.8	4.3	**75.9**	15.8
	Patients		8.2	10.8	**13.3**	11.2		8.4	10.8	**13.7**	11.6		11.2	14.7	**18.2**	15.2

Table S.5: (Continued).

Scenario 5																
	Subgroup			**g=1**					**g=2**					**g=3**		
	z_g			**1**					**2**					**1**		
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.08	0.08	0.10	**0.10**		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	0.30	**0.20**	0.15		0.40	0.40	0.40	**0.35**		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.35	0.50	**0.70**	0.80		0.55	0.55	0.55	**0.60**		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	**76.9**	82.5		73.3	73.3	72.6	**75.0**		54.0	65.1	**76.9**	82.5
	$\phi_S^{true}(d, g)$		0.40	0.50	**0.70**	0.60		0.20	0.35	0.50	**0.75**		0.40	0.50	**0.70**	0.60
Weibull	Selection %	2.4	7.7	10.3	**67.3**	12.3	5.3	2.2	4.7	24.0	**63.8**	2.3	7.7	10.1	**67.1**	12.8
	Patients		9.5	11.6	**12.8**	10.6		9.5	11.4	13.1	**10.8**		12.7	15.2	**17.3**	14.1
Log-normal	Selection %	3.3	4.1	6.0	**67.6**	19.0	6.9	2.5	4.4	25.9	**60.3**	2.7	3.7	4.8	**69.6**	19.2
	Patients		9.2	11.4	**13.0**	10.7		9.4	11.4	12.9	**10.7**		12.4	14.9	**17.4**	14.1
Gamma	Selection %	3.0	5.0	7.8	**63.6**	20.6	6.6	2.3	6.1	26.8	**58.2**	2.2	4.2	6.8	**65.0**	21.8
	Patients		9.5	11.8	**12.8**	10.2		9.7	11.7	12.8	**10.4**		12.9	15.7	**17.0**	13.8
Scenario 6																
	Subgroup			**g=1**					**g=2**					**g=3**		
	z_g			**1**					**2**					**1**		
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	**0.12**	0.14	0.16		0.08	0.08	0.10	**0.10**		0.10	**0.12**	0.14	0.16
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	**0.20**	0.15	0.10		0.40	0.40	0.40	**0.35**		0.30	**0.20**	0.15	0.10
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.50	**0.70**	0.80	0.90		0.55	0.55	0.55	**0.60**		0.50	**0.70**	0.80	0.90
	$\phi_{ET}^{true}(d, g)$		65.7	**77.7**	83.2	88.8		73.3	73.3	72.6	**75.0**		65.7	**77.7**	83.2	88.8
	$\phi_S^{true}(d, g)$		0.40	**0.70**	0.60	0.50		0.20	0.35	0.50	**0.75**		0.40	**0.70**	0.60	0.50
Weibull	Selection %	0.5	3.9	**80.8**	13.2	1.6	8.2	4.5	12.4	29.0	**45.9**	0.4	3.0	**83.0**	12.6	1.0
	Patients		11.7	**14.1**	12.1	7.2		10.9	13.9	12.3	**7.4**		15.8	**18.7**	16.0	9.5
Log-normal	Selection %	1.9	2.5	**71.4**	20.5	3.7	10.1	3.7	17.1	31.6	**37.5**	1.7	1.6	**72.8**	20.3	3.6
	Patients		11.6	**14.3**	12.2	6.7		11.5	14.2	12.3	**6.7**		15.6	**18.8**	16.1	8.9
Gamma	Selection %	0.7	3.6	**71.9**	19.8	4.0	10.1	4.2	18.3	27.8	**39.6**	0.8	1.4	**74.9**	18.6	4.3
	Patients		11.7	**14.3**	12.0	6.9		11.4	14.1	12.3	**6.9**		15.8	**19.0**	16.2	9.1

Table S.5: (Continued).

Scenario 7																
	Subgroup			**g=1**					**g=2**					**g=3**		
	z_g			**1**					**2**					**1**		
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.08	0.08	**0.10**	0.10		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	0.30	**0.20**	0.15		0.40	0.40	**0.40**	0.35		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.35	0.50	**0.70**	0.80		0.55	0.55	**0.55**	0.60		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5
	$\phi_S^{true}(d, g)$		0.40	0.50	**0.70**	0.60		0.20	0.35	**0.60**	0.40		0.40	0.50	**0.70**	0.60
Weibull	Selection %	2.3	7.3	9.8	**70.5**	10.1	6.9	4.4	8.2	**69.3**	11.2	2.1	6.7	11.5	**71.2**	8.5
	Patients		9.5	11.7	**12.8**	10.6		9.5	11.5	**13.0**	10.7		12.6	15.2	**17.4**	14.1
Log-normal	Selection %	3.4	3.2	6.0	**73.5**	13.9	8.8	3.5	6.0	**73.4**	8.3	2.7	3.8	5.1	**74.9**	13.5
	Patients		9.2	11.5	**13.0**	10.7		9.4	11.4	**12.9**	10.5		12.3	15.1	**17.4**	14.0
Gamma	Selection %	2.8	4.2	8.7	**69.8**	14.5	8.1	4.8	7.2	**71.1**	8.8	2.3	3.9	7.2	**72.1**	14.5
	Patients		9.5	11.8	**12.7**	10.3		9.8	11.7	**12.9**	10.3		12.9	15.7	**17.1**	13.9
Scenario 8																
	Subgroup			**g=1**					**g=2**					**g=3**		
	z_g			**1**					**2**					**3**		
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.16	**0.20**	0.26	0.30		0.08	0.08	0.10	**0.10**
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.20	0.20	**0.20**	0.15		0.35	**0.30**	0.30	0.30		0.30	0.30	0.25	**0.20**
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.40	0.45	**0.55**	0.65		0.55	**0.65**	0.65	0.65		0.30	0.45	0.60	**0.70**
	$\phi_{ET}^{true}(d, g)$		54.9	58.2	**65.3**	70.9		69.2	**74.0**	72.0	70.6		50.7	62.4	72.0	**78.4**
	$\phi_S^{true}(d, g)$		0.30	0.40	**0.70**	0.40		0.30	**0.60**	0.40	0.35		0.20	0.30	0.35	**0.75**
Weibull	Selection %	12.0	2.5	9.6	**70.0**	5.9	14.4	5.0	**64.2**	11.1	5.3	18.3	0.8	2.7	5.7	**72.5**
	Patients		9.4	11.3	**12.3**	10.7		9.6	**11.8**	12.0	10.8		11.5	14.2	16.3	**16.6**
Log-normal	Selection %	9.6	1.3	9.5	**63.7**	15.9	16.8	4.3	**53.5**	16.7	8.7	16.4	0.6	6.7	6.8	**69.5**
	Patients		9.1	10.9	**12.2**	11.4		9.3	**11.9**	12.0	11.0		11.5	14.0	15.9	**16.1**
Gamma	Selection %	8.9	2.3	10.2	**64.9**	13.7	12.7	4.5	**52.7**	16.6	13.5	15.3	1.2	8.5	8.5	**66.5**
	Patients		9.6	10.8	**12.5**	11.1		9.9	**11.6**	12.0	10.8		11.8	14.2	16.2	**16.1**

Table S.6: Simulation results for sensitivity of the PGen I-II design to the maximum overall sample size n_{123} = 90, 120, or 150.

Scenario 1																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**1**					**1**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.12	0.22	0.36	0.50		0.12	0.22	0.36	0.50		0.12	0.22	0.36	0.50
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.35	0.30	0.20	0.15		0.35	0.30	0.20	0.15		0.35	0.30	0.20	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.45	0.60	0.75	0.80		0.45	0.60	0.75	0.80		0.45	0.60	0.75	0.80
	$\phi_{ET}^{true}(d, g)$		62.7	69.5	72.9	70.1		62.7	69.5	72.9	70.1		62.7	69.5	72.9	70.1
n_{123}	$\phi_S^{true}(d, g)$		0.10	0.20	0.30	0.45		0.10	0.20	0.30	0.45		0.10	0.20	0.30	0.45
90	Selection %	**87.1**	0.8	0.3	2.9	8.9	**86.6**	0.7	0.6	3.0	9.1	**87.0**	0.8	0.1	2.5	9.6
	Patients		7.0	7.8	7.0	2.0		7.0	8.1	7.0	2.1		9.3	10.5	9.0	2.7
120	Selection %	**90.8**	0.4	0.0	1.5	7.3	**89.7**	0.5	0.1	1.8	7.9	**92.6**	0.2	0.1	0.9	6.2
	Patients		9.4	9.7	8.6	3.0		9.5	9.8	8.8	3.0		12.9	12.9	11.7	3.5
150	Selection %	**93.2**	0.1	0.0	1.6	5.1	**92.5**	0.1	0.0	1.5	5.9	**94.3**	0.0	0.0	1.4	4.3
	Patients		12.5	11.8	10.5	3.7		12.7	11.9	10.6	3.8		17.2	15.9	13.6	4.3
Scenario 2																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**2**					**1**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.14	0.18	0.24		0.10	0.12	**0.14**	0.16		0.10	0.14	0.18	0.24
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.35	0.35	0.35	0.30		0.30	0.30	**0.20**	0.15		0.35	0.35	0.35	0.30
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.45	0.50	0.50	0.60		0.35	0.50	**0.70**	0.80		0.45	0.50	0.50	0.60
	$\phi_{ET}^{true}(d, g)$		63.3	65.9	64.7	68.8		54.0	65.1	**76.9**	82.5		63.3	65.9	64.7	68.8
n_{123}	$\phi_S^{true}(d, g)$		0.10	0.20	0.25	0.30		0.20	0.30	**0.70**	0.40		0.10	0.20	0.25	0.30
90	Selection %	**74.2**	2.4	3.2	10.0	10.2	26.7	4.4	5.8	**59.3**	3.8	**75.4**	2.4	2.5	9.9	9.8
	Patients		5.1	6.7	7.3	6.3		6.2	7.3	**8.8**	6.1		7.0	8.9	9.7	8.3
120	Selection %	**80.1**	2.1	0.7	7.1	10.0	25.3	3.5	2.8	**64.9**	3.5	**83.6**	1.3	1.0	5.8	8.3
	Patients		6.9	8.3	9.6	8.5		8.5	10.0	**11.5**	8.2		9.4	11.1	12.8	11.0
150	Selection %	**83.9**	1.8	0.6	6.1	7.6	22.3	2.1	2.3	**70.5**	2.8	**86.1**	1.1	0.8	4.6	7.4
	Patients		8.9	10.4	11.7	10.1		10.6	12.8	**14.7**	10.4		12.2	13.7	15.6	13.7

Table S.6: (Continued).

Scenario 3																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**2**					**1**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	0.14	**0.16**		0.10	0.14	0.18	0.24		0.10	0.12	0.14	**0.16**
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	0.30	0.20	**0.15**		0.35	0.35	0.35	0.30		0.30	0.30	0.20	**0.15**
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.35	0.50	0.70	**0.80**		0.45	0.50	0.50	0.60		0.35	0.50	0.70	**0.80**
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	76.9	**82.5**		63.3	65.9	64.7	68.8		54.0	65.1	76.9	**82.5**
n_{123}	$\phi_S^{true}(d, g)$		0.20	0.30	0.40	**0.60**		0.05	0.10	0.15	0.20		0.20	0.30	0.40	**0.60**
90	Selection %	31.8	3.8	5.9	8.2	**50.3**	**78.0**	3.0	1.8	2.2	15.0	30.7	3.7	6.0	6.6	**53.0**
	Patients		5.4	6.7	7.5	**7.4**		5.0	5.8	7.1	7.3		7.2	8.5	10.2	**9.8**
120	Selection %	25.3	2.8	4.0	6.3	**61.6**	**86.4**	1.7	0.8	1.5	9.6	24.4	1.7	2.9	6.6	**64.4**
	Patients		6.9	9.0	10.1	**9.9**		6.6	7.6	9.3	9.7		9.5	11.2	13.6	**13.1**
150	Selection %	22.2	1.8	2.7	5.5	**67.8**	**91.3**	1.3	0.4	0.6	6.4	22.9	1.0	1.9	5.5	**68.7**
	Patients		8.9	11.1	12.7	**12.3**		8.6	9.4	11.3	11.4		12.1	14.4	17.3	**16.2**
Scenario 4																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**1**					**1**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.10	0.12	**0.14**	0.16		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	0.30	**0.20**	0.15		0.30	0.30	**0.20**	0.15		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.35	0.50	**0.70**	0.80		0.35	0.50	**0.70**	0.80		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5
n_{123}	$\phi_S^{true}(d, g)$		0.40	0.50	**0.70**	0.60		0.40	0.50	**0.70**	0.60		0.40	0.50	**0.70**	0.60
90	Selection %	2.5	13.6	15.8	**57.1**	11.0	2.5	14.4	15.8	**56.4**	10.9	2.0	14.7	13.2	**61.9**	8.2
	Patients		4.8	6.6	**8.2**	6.9		5.0	6.6	**8.2**	7.1		6.6	8.6	**11.0**	9.2
120	Selection %	2.0	8.0	10.4	**70.3**	9.3	2.1	7.3	11.6	**69.5**	9.5	2.1	7.5	9.3	**73.9**	7.2
	Patients		6.4	9.0	**10.7**	9.2		6.7	8.8	**10.9**	9.3		8.8	11.5	**14.5**	12.4
150	Selection %	2.3	4.1	6.4	**78.8**	8.4	2.1	3.6	4.7	**81.1**	8.5	2.2	4.5	4.4	**81.4**	7.5
	Patients		8.1	11.1	**13.6**	11.5		8.5	11.1	**13.5**	11.5		11.0	14.5	**18.1**	15.2

Table S.6: (Continued).

Scenario 5

	Subgroup	**g=1**					**g=2**					**g=3**				
	z_g	**1**					**2**					**1**				
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.08	0.08	0.10	**0.10**		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	0.30	**0.20**	0.15		0.40	0.40	0.40	**0.35**		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.35	0.50	**0.70**	0.80		0.55	0.55	0.55	**0.60**		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	**76.9**	82.5		73.3	73.3	72.6	**75.0**		54.0	65.1	**76.9**	82.5
n_{123}	$\phi_S^{true}(d, g)$		0.40	0.50	**0.70**	0.60		0.20	0.35	0.50	**0.75**		0.40	0.50	**0.70**	0.60
90	Selection %	2.6	13.2	16.8	**53.8**	13.6	6.3	5.3	7.8	28.1	**52.5**	2.0	13.7	16.6	**55.7**	12.0
	Patients		5.5	6.9	**7.7**	6.5		5.4	6.7	7.9	**6.8**		7.5	9.1	**10.5**	8.6
120	Selection %	1.7	11.1	14.4	**60.7**	12.1	5.9	4.4	6.0	25.4	**58.3**	1.8	9.8	12.7	**63.8**	11.9
	Patients		7.4	9.4	**10.2**	8.7		7.5	9.1	10.5	**8.7**		10.0	12.0	**14.1**	11.4
150	Selection %	2.4	7.7	10.3	**67.3**	12.3	5.3	2.2	4.7	24.0	**63.8**	2.3	7.7	10.1	**67.1**	12.8
	Patients		9.5	11.6	**12.8**	10.6		9.5	11.4	13.1	**10.8**		12.7	15.2	**17.3**	14.1

Scenario 6

	Subgroup	**g=1**					**g=2**					**g=3**				
	z_g	**1**					**2**					**1**				
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	**0.12**	0.14	0.16		0.08	0.08	0.10	**0.10**		0.10	**0.12**	0.14	0.16
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	**0.20**	0.15	0.10		0.40	0.40	0.40	**0.35**		0.30	**0.20**	0.15	0.10
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.50	**0.70**	0.80	0.90		0.55	0.55	0.55	**0.60**		0.50	**0.70**	0.80	0.90
	$\phi_{ET}^{true}(d, g)$		65.7	**77.7**	83.2	88.8		73.3	73.3	72.6	**75.0**		65.7	**77.7**	83.2	88.8
n_{123}	$\phi_S^{true}(d, g)$		0.40	**0.70**	0.60	0.50		0.20	0.35	0.50	**0.75**		0.40	**0.70**	0.60	0.50
90	Selection %	0.7	10.4	**64.0**	20.5	4.4	5.8	7.6	20.2	26.7	**39.7**	0.6	9.1	**67.6**	17.2	5.5
	Patients		6.8	**8.8**	7.1	4.3		6.3	8.5	7.3	**4.4**		9.3	**11.6**	9.5	5.7
120	Selection %	0.7	6.7	**75.1**	14.8	2.7	8.1	5.4	17.7	25.4	**43.4**	0.4	5.8	**77.5**	13.5	2.8
	Patients		9.1	**11.4**	9.7	5.8		8.8	11.1	9.8	**6.0**		12.7	**15.1**	12.7	7.6
150	Selection %	0.5	3.9	**80.8**	13.2	1.6	8.2	4.5	12.4	29.0	**45.9**	0.4	3.0	**83.0**	12.6	1.0
	Patients		11.7	**14.1**	12.1	7.2		10.9	13.9	12.3	**7.4**		15.8	**18.7**	16.0	9.5

Table S.6: (Continued).

Scenario 7

	Subgroup	**g=1**					**g=2**					**g=3**				
	z_g	**1**					**2**					**1**				
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.08	0.08	**0.10**	0.10		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.30	0.30	**0.20**	0.15		0.40	0.40	**0.40**	0.35		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.35	0.50	**0.70**	0.80		0.55	0.55	**0.55**	0.60		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5
n_{123}	$\phi_S^{true}(d, g)$		0.40	0.50	**0.70**	0.60		0.20	0.35	**0.60**	0.40		0.40	0.50	**0.70**	0.60
90	Selection %	2.5	12.6	18.0	**55.5**	11.4	6.7	8.9	13.7	**57.8**	12.9	2.5	14.8	17.3	**56.9**	8.5
	Patients		5.5	7.0	**7.7**	6.6		5.4	6.6	**8.1**	6.6		7.5	9.0	**10.6**	8.5
120	Selection %	2.2	9.8	14.4	**63.8**	9.8	6.9	6.3	8.2	**65.8**	12.8	2.2	8.5	12.2	**68.8**	8.3
	Patients		7.5	9.4	**10.2**	8.6		7.5	9.1	**10.5**	8.6		10.0	12.1	**14.1**	11.4
150	Selection %	2.3	7.3	9.8	**70.5**	10.1	6.9	4.4	8.2	**69.3**	11.2	2.1	6.7	11.5	**71.2**	8.5
	Patients		9.5	11.7	**12.8**	10.6		9.5	11.5	**13.0**	10.7		12.6	15.2	**17.4**	14.1

Scenario 8

	Subgroup	**g=1**					**g=2**					**g=3**				
	z_g	**1**					**2**					**3**				
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.16	**0.20**	0.26	0.30		0.08	0.08	0.10	**0.10**
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.20	0.20	**0.20**	0.15		0.35	**0.30**	0.30	0.30		0.30	0.30	0.25	**0.20**
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.40	0.45	**0.55**	0.65		0.55	**0.65**	0.65	0.65		0.30	0.45	0.60	**0.70**
	$\phi_{ET}^{true}(d, g)$		54.9	58.2	**65.3**	70.9		69.2	**74.0**	72.0	70.6		50.7	62.4	72.0	**78.4**
n_{123}	$\phi_S^{true}(d, g)$		0.30	0.40	**0.70**	0.40		0.30	**0.60**	0.40	0.35		0.20	0.30	0.35	**0.75**
90	Selection %	11.1	4.8	13.1	**64.0**	7.0	16.9	6.7	**53.6**	14.4	8.4	17.6	2.4	6.4	8.4	**65.2**
	Patients		5.5	6.7	**7.3**	6.6		5.6	**7.1**	7.2	6.6		6.8	8.5	9.7	**10.2**
120	Selection %	9.9	3.9	10.8	**68.2**	7.2	14.8	6.3	**58.0**	14.2	6.7	17.8	1.0	3.9	7.0	**70.3**
	Patients		7.5	8.8	**10.1**	8.6		7.5	**9.6**	9.5	8.8		9.0	11.4	13.1	**13.4**
150	Selection %	12.0	2.5	9.6	**70.0**	5.9	14.4	5.0	**64.2**	11.1	5.3	18.3	0.8	2.7	5.7	**72.5**
	Patients		9.4	11.3	**12.3**	10.7		9.6	**11.8**	12.0	10.8		11.5	14.2	16.3	**16.6**

Table S.7: Simulation results for sensitivity to the assumed distribution for Y_S' in PGen I-II design design (the data-generating distribution is Weibull).

Scenario 1															
	Subgroup		**g=1**					**g=2**					**g=3**		
	z_g		**1**					**1**					**1**		
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.12	0.22	0.36	0.50		0.12	0.22	0.36	0.50		0.12	0.22	0.36	0.50
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.35	0.30	0.20	0.15		0.35	0.30	0.20	0.15		0.35	0.30	0.20	0.15
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.45	0.60	0.75	0.80		0.45	0.60	0.75	0.80		0.45	0.60	0.75	0.80
	$\phi_{ET}^{true}(d, g)$		62.7	69.5	72.9	70.1		62.7	69.5	72.9	70.1		62.7	69.5	72.9	70.1
	$\phi_S^{true}(d, g)$		0.10	0.20	0.30	0.45		0.10	0.20	0.30	0.45		0.10	0.20	0.30	0.45
PE	Selection %	**93.2**	0.1	0.0	1.6	5.1	**92.5**	0.1	0.0	1.5	5.9	**94.3**	0.0	0.0	1.4	4.3
	Patients		12.5	11.8	10.5	3.7		12.7	11.9	10.6	3.8		17.2	15.9	13.6	4.3
Weibull	Selection %	**92.9**	0.0	0.0	2.0	5.1	**96.0**	0.0	0.0	1.0	3.0	**94.9**	0.0	0.0	2.0	3.0
	Patients		12.5	11.0	11.3	4.1		12.4	11.6	11.7	3.8		17.2	15.3	14.0	3.9
Scenario 2															
	Subgroup		**g=1**					**g=2**					**g=3**		
	z_g		**1**					**2**					**1**		
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}\ (1 \mid d, g)$		0.10	0.14	0.18	0.24		0.10	0.12	**0.14**	0.16		0.10	0.14	0.18	0.24
	$\pi_E^{\text{true}}\ (1 \mid d, g)$		0.35	0.35	0.35	0.30		0.30	0.30	**0.20**	0.15		0.35	0.35	0.35	0.30
	$\pi_E^{\text{true}}\ (2 \mid d, g)$		0.45	0.50	0.50	0.60		0.35	0.50	**0.70**	0.80		0.45	0.50	0.50	0.60
	$\phi_{ET}^{true}(d, g)$		63.3	65.9	64.7	68.8		54.0	65.1	**76.9**	82.5		63.3	65.9	64.7	68.8
	$\phi_S^{true}(d, g)$		0.10	0.20	0.25	0.30		0.20	0.30	**0.70**	0.40		0.10	0.20	0.25	0.30
PE	Selection %	**83.9**	1.8	0.6	6.1	7.6	22.3	2.1	2.3	**70.5**	2.8	**86.1**	1.1	0.8	4.6	7.4
	Patients		8.9	10.4	11.7	10.1		10.6	12.8	**14.7**	10.4		12.2	13.7	15.6	13.7
Weibull	Selection %	**81.8**	0.0	1.0	8.1	9.1	18.2	1.0	0.0	**75.8**	5.1	**80.8**	0.0	3.0	5.1	11.1
	Patients		7.8	9.3	11.7	11.2		10.3	13.5	**16.1**	10.9		10.7	13.2	15.8	14.7

Table S.7: (Continued).

Scenario 3															
	Subgroup		**g=1**					**g=2**					**g=3**		
	z_g		**1**					**2**					**1**		
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}(1 \mid d, g)$		0.10	0.12	0.14	**0.16**		0.10	0.14	0.18	0.24		0.10	0.12	0.14	**0.16**
	$\pi_E^{\text{true}}(1 \mid d, g)$		0.30	0.30	0.20	**0.15**		0.35	0.35	0.35	0.30		0.30	0.30	0.20	**0.15**
	$\pi_E^{\text{true}}(2 \mid d, g)$		0.35	0.50	0.70	**0.80**		0.45	0.50	0.50	0.60		0.35	0.50	0.70	**0.80**
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	76.9	**82.5**		63.3	65.9	64.7	68.8		54.0	65.1	76.9	**82.5**
	$\phi_S^{true}(d, g)$		0.20	0.30	0.40	**0.60**		0.05	0.10	0.15	0.20		0.20	0.30	0.40	**0.60**
PE	Selection %	22.2	1.8	2.7	5.5	**67.8**	**91.3**	1.3	0.4	0.6	6.4	22.9	1.0	1.9	5.5	**68.7**
	Patients		8.9	11.1	12.7	**12.3**		8.6	9.4	11.3	11.4		12.1	14.4	17.3	**16.2**
Weibull	Selection %	23.2	0.0	5.1	9.1	**62.6**	**94.9**	0.0	2.0	1.0	2.0	24.2	1.0	4.0	7.1	**63.6**
	Patients		9.6	11.5	12.4	**11.0**		8.1	9.3	11.0	10.1		11.8	14.4	17.5	**15.8**
Scenario 4															
	Subgroup		**g=1**					**g=2**					**g=3**		
	z_g		**1**					**1**					**1**		
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}(1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.10	0.12	**0.14**	0.16		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}(1 \mid d, g)$		0.30	0.30	**0.20**	0.15		0.30	0.30	**0.20**	0.15		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}(2 \mid d, g)$		0.35	0.50	**0.70**	0.80		0.35	0.50	**0.70**	0.80		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5
	$\phi_S^{true}(d, g)$		0.40	0.50	**0.70**	0.60		0.40	0.50	**0.70**	0.60		0.40	0.50	**0.70**	0.60
PE	Selection %	2.3	4.1	6.4	**78.8**	8.4	2.1	3.6	4.7	**81.1**	8.5	2.2	4.5	4.4	**81.4**	7.5
	Patients		8.1	11.1	**13.6**	11.5		8.5	11.1	**13.5**	11.5		11.0	14.5	**18.1**	15.2
Weibull	Selection %	6.1	1.0	5.1	**78.8**	9.1	6.1	1.0	4.0	**79.8**	9.1	5.1	2.0	6.1	**76.8**	10.1
	Patients		8.5	10.8	**12.8**	11.5		8.4	10.5	**13.4**	10.8		11.3	14.4	**17.3**	15.3

Table S.7: (Continued).

Scenario 5																
	Subgroup			**g=1**					**g=2**					**g=3**		
	z_g			**1**					**2**					**1**		
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}(1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.08	0.08	0.10	**0.10**		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}(1 \mid d, g)$		0.30	0.30	**0.20**	0.15		0.40	0.40	0.40	**0.35**		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}(2 \mid d, g)$		0.35	0.50	**0.70**	0.80		0.55	0.55	0.55	**0.60**		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	**76.9**	82.5		73.3	73.3	72.6	**75.0**		54.0	65.1	**76.9**	82.5
	$\phi_S^{true}(d, g)$		0.40	0.50	**0.70**	0.60		0.20	0.35	0.50	**0.75**		0.40	0.50	**0.70**	0.60
PE	Selection %	2.4	7.7	10.3	**67.3**	12.3	5.3	2.2	4.7	24.0	**63.8**	2.3	7.7	10.1	**67.1**	12.8
	Patients		9.5	11.6	**12.8**	10.6		9.5	11.4	13.1	**10.8**		12.7	15.2	**17.3**	14.1
Weibull	Selection %	3.0	2.0	13.1	**70.7**	11.1	9.1	0.0	3.0	29.3	**58.6**	3.0	1.0	13.1	**68.7**	14.1
	Patients		9.5	12.2	**12.8**	10.6		8.8	11.3	13.1	**11.8**		12.1	15.3	**16.6**	14.6
Scenario 6																
	Subgroup			**g=1**					**g=2**					**g=3**		
	z_g			**1**					**2**					**1**		
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}(1 \mid d, g)$		0.10	**0.12**	0.14	0.16		0.08	0.08	0.10	**0.10**		0.10	**0.12**	0.14	0.16
	$\pi_E^{\text{true}}(1 \mid d, g)$		0.30	**0.20**	0.15	0.10		0.40	0.40	0.40	**0.35**		0.30	**0.20**	0.15	0.10
	$\pi_E^{\text{true}}(2 \mid d, g)$		0.50	**0.70**	0.80	0.90		0.55	0.55	0.55	**0.60**		0.50	**0.70**	0.80	0.90
	$\phi_{ET}^{true}(d, g)$		65.7	**77.7**	83.2	88.8		73.3	73.3	72.6	**75.0**		65.7	**77.7**	83.2	88.8
	$\phi_S^{true}(d, g)$		0.40	**0.70**	0.60	0.50		0.20	0.35	0.50	**0.75**		0.40	**0.70**	0.60	0.50
PE	Selection %	0.5	3.9	**80.8**	13.2	1.6	8.2	4.5	12.4	29.0	**45.9**	0.4	3.0	**83.0**	12.6	1.0
	Patients		11.7	**14.1**	12.1	7.2		10.9	13.9	12.3	**7.4**		15.8	**18.7**	16.0	9.5
Weibull	Selection %	0.0	2.0	**80.6**	16.3	1.0	8.2	0.0	13.3	32.7	**45.9**	0.0	1.0	**82.7**	15.3	1.0
	Patients		10.7	**14.7**	13.0	7.0		9.3	14.4	13.3	**8.1**		13.8	**18.3**	17.2	10.2

Table S.7: (Continued).

Scenario 7																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**2**					**1**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}(1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.08	0.08	**0.10**	0.10		0.10	0.12	**0.14**	0.16
	$\pi_E^{\text{true}}(1 \mid d, g)$		0.30	0.30	**0.20**	0.15		0.40	0.40	**0.40**	0.35		0.30	0.30	**0.20**	0.15
	$\pi_E^{\text{true}}(2 \mid d, g)$		0.35	0.50	**0.70**	0.80		0.55	0.55	**0.55**	0.60		0.35	0.50	**0.70**	0.80
	$\phi_{ET}^{true}(d, g)$		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5		54.0	65.1	**76.9**	82.5
	$\phi_S^{true}(d, g)$		0.40	0.50	**0.70**	0.60		0.20	0.35	**0.60**	0.40		0.40	0.50	**0.70**	0.60
PE	Selection %	2.3	7.3	9.8	**70.5**	10.1	6.9	4.4	8.2	**69.3**	11.2	2.1	6.7	11.5	**71.2**	8.5
	Patients		9.5	11.7	**12.8**	10.6		9.5	11.5	**13.0**	10.7		12.6	15.2	**17.4**	14.1
Weibull	Selection %	4.0	2.0	13.1	**74.7**	6.1	17.2	0.0	5.1	**70.7**	7.1	3.0	1.0	12.1	**76.8**	7.1
	Patients		9.4	12.3	**12.2**	10.4		8.6	11.2	**13.1**	11.4		12.2	16.0	**17.1**	14.6
Scenario 8																
	Subgroup		**g=1**					**g=2**					**g=3**			
	z_g		**1**					**2**					**3**			
	Dose(d)	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$	$\mathbf{d_0}$	$\mathbf{d_1}$	$\mathbf{d_2}$	$\mathbf{d_3}$	$\mathbf{d_4}$
	$\pi_T^{\text{true}}(1 \mid d, g)$		0.10	0.12	**0.14**	0.16		0.16	**0.20**	0.26	0.30		0.08	0.08	0.10	**0.10**
	$\pi_E^{\text{true}}(1 \mid d, g)$		0.20	0.20	**0.20**	0.15		0.35	**0.30**	0.30	0.30		0.30	0.30	0.25	**0.20**
	$\pi_E^{\text{true}}(2 \mid d, g)$		0.40	0.45	**0.55**	0.65		0.55	**0.65**	0.65	0.65		0.30	0.45	0.60	**0.70**
	$\phi_{ET}^{true}(d, g)$		54.9	58.2	**65.3**	70.9		69.2	**74.0**	72.0	70.6		50.7	62.4	72.0	**78.4**
	$\phi_S^{true}(d, g)$		0.30	0.40	**0.70**	0.40		0.30	**0.60**	0.40	0.35		0.20	0.30	0.35	**0.75**
PE	Selection %	12.0	2.5	9.6	**70.0**	5.9	14.4	5.0	**64.2**	11.1	5.3	18.3	0.8	2.7	5.7	**72.5**
	Patients		9.4	11.3	**12.3**	10.7		9.6	**11.8**	12.0	10.8		11.5	14.2	16.3	**16.6**
Weibull	Selection %	12.6	2.1	16.8	**62.1**	6.3	21.1	0.0	**58.9**	16.8	3.2	23.2	0.0	5.3	6.3	**65.3**
	Patients		8.8	11.7	**12.2**	10.0		7.6	**11.9**	12.8	10.0		9.8	15.6	17.1	**16.2**

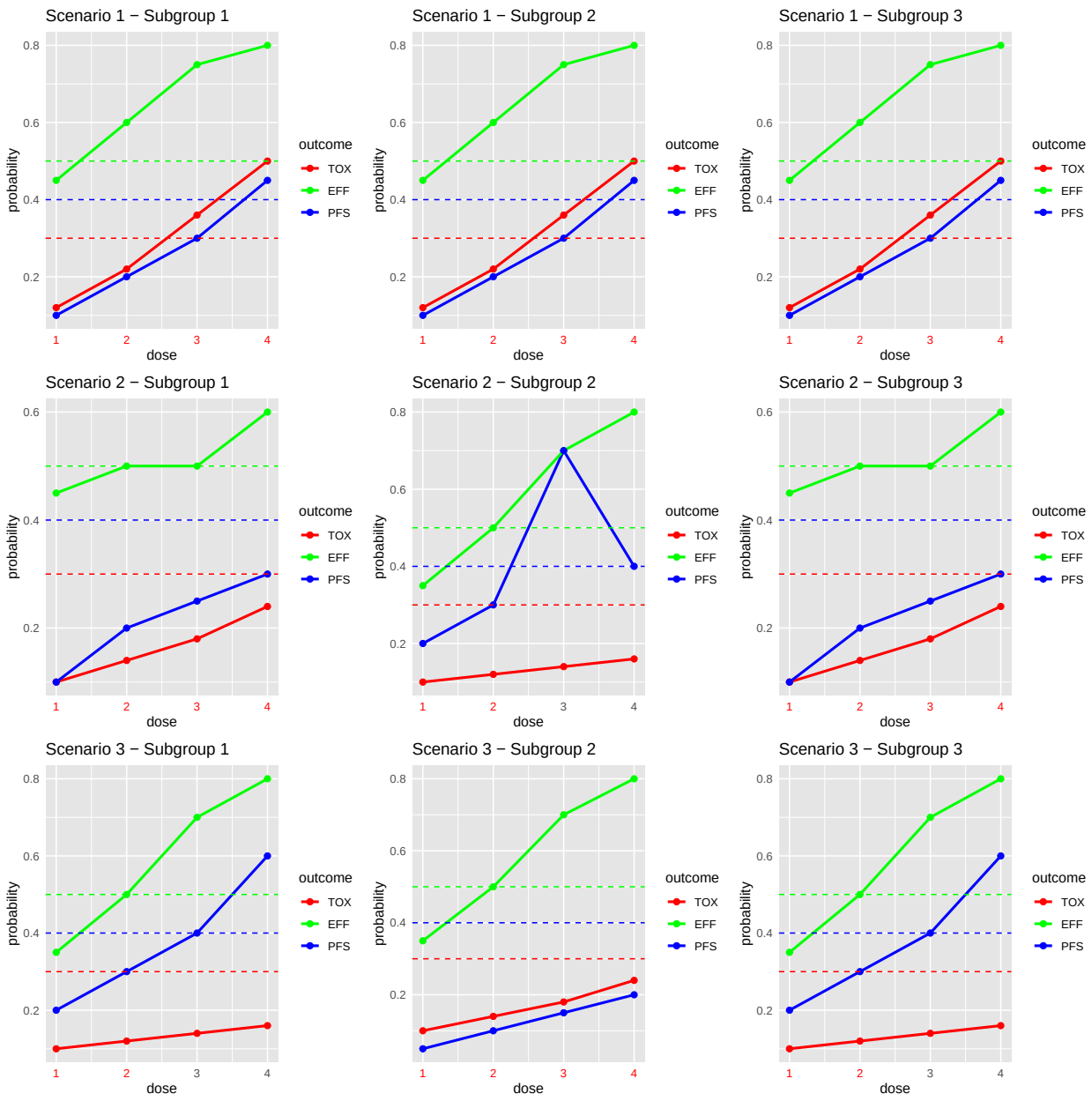

Figure S.1: Dose-outcome curves for the scenarios studied in the simulations. Plots of $\pi_T^{\text{true}}(1 \mid d, g)$ are given in red, $\pi_E^{\text{true}}(2 \mid d, g)$ in green, and $\phi_S^{\text{true}}(d_j, g)$ in blue. Horizontal dashed lines represent the fixed acceptability limits $\underline{\pi}_E = .50$ (green), $\overline{\pi}_T = .30$ (red) and $\underline{\pi}_S = .40$ (blue). Indices of unacceptable doses are given in red on the horizontal axis.

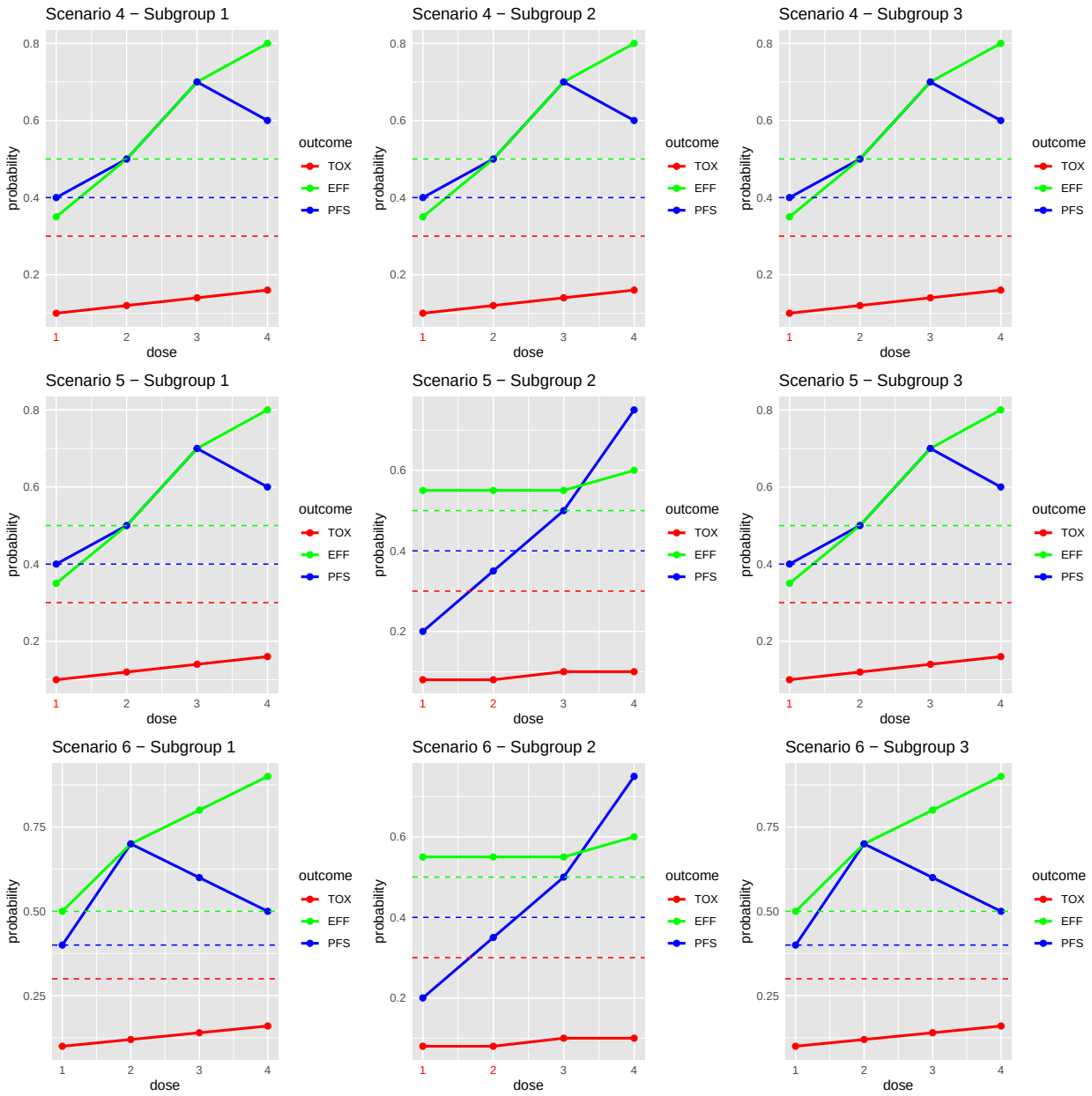

Figure S.1: (Continued).

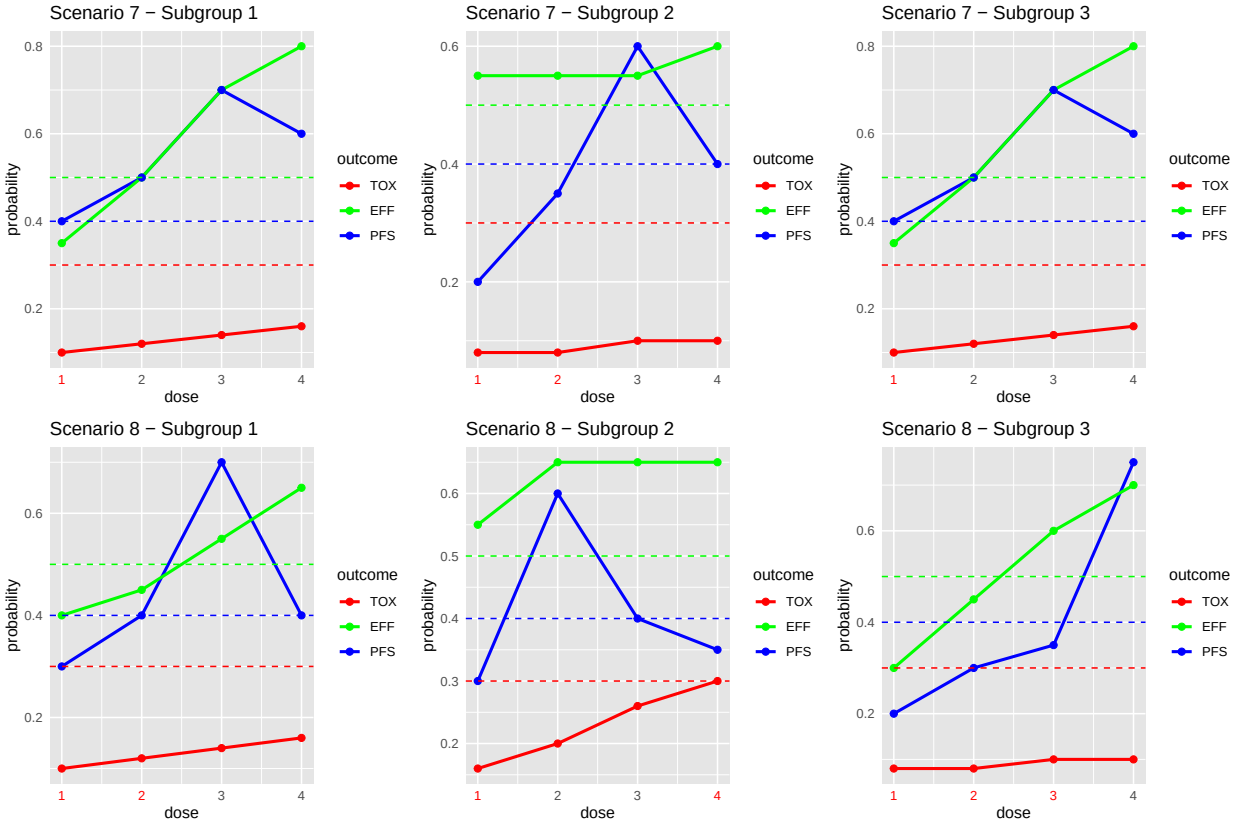

Figure S.1: (Continued).