Determining a Maximal Period of Administration for an Experimental Agent

Thomas Braun
U. of Michigan School of Public Health
Department of Biostatistics
Peter Thall
U. of Texas M.D. Anderson Cancer Center
Department of Biomathematics
Zheng Yuan
U. of Michigan School of Public Health
Department of Biostatistics

Motivating Example

- Agent under investigation:

  $\Rightarrow$ Recombinant human keratinocyte growth factor (KGF)

- Patient group:

  $\Rightarrow$ Allogeneic bone marrow transplant (BMT) recipients

  - Experience acute graft-versus-host disease (aGVHD)
  - Occurs within 100 days of BMT
  - Includes severe GI distress, $i.e.$ vomiting, diarrhea, bleeding

Motivating Example

- Potential treatment:

  $\Rightarrow$ Reduction of chemotherapy- and radiation-induced injury to the mucosal lining of the lower gastro-intestinal (GI) tract

- Potential toxicities:

  $\Rightarrow$ mild to moderate skin-related events such as rash, reddening, and edema

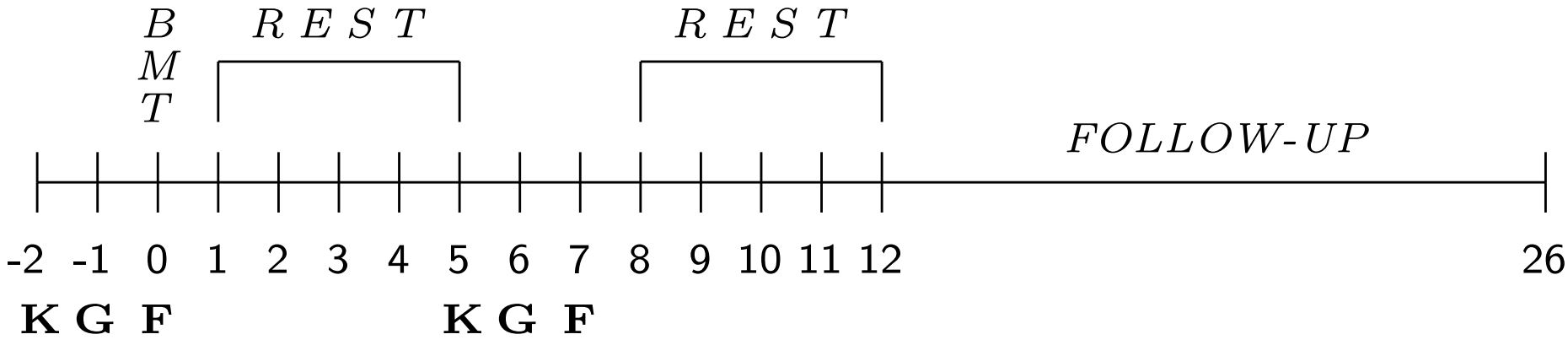
  $\Rightarrow$ reversible increases in amylase and lipase indicative of pancreas dysfunction

Motivating Example

- Prior Phase I trials
  - $\Rightarrow$ 60 mg/kg/day of IV KGF associated with mimimal toxicity in BMT recipients
  - $\Rightarrow$ administered on days -2, -1, 0, 5, 6, 7 post-BMT
    - $*$ 3-days-on/4-days-off schedule for two weeks

- Subjects followed additional two weeks after last administration of KGF.
  - $\Rightarrow$ Implicitly assumes toxicity risk for single administration ends after 18 days.

Motivating Example



Question:

$\Rightarrow$ How often can we repeat this two-week course and keep cumulative probability of toxicity close to a pre-determined threshold?

Study Design Considerations

- Sequentially enroll subjects as they are eligible
- Each two weeks of administrations = one course
- Number of courses (schedule) assigned to subject i determined from information collected on subjects $1, 2, \ldots, i-1$
- Comparing schedules:
  - Use data from enrolled subjects to compute posterior probabilities of toxicity
  - Schedule best satisfying pre-defined criteria is MTS

Study Design Considerations

- Naive approach - use standard Phase I trial design:
  - Treat each two-week course as a "dose"

    Braun, Ferrara, Levine (2003) CCT
  - Use TiTE-CRM (Cheung & Chappell, 2000, *Biometrics*) to weight information of subjects who have not completed their "dose"

- Better approach
  - Directly model hazard of toxicity for every single administration

Preliminary Notation

- Global study values:

  t^* = any given time from start of the trial when evaluation is made

  n^* = number of patients enrolled up to t^*

  τ = duration of observation for each subject

  p_τ = acceptable level of cumulative toxicity by τ

- Subject-specific values ($i = 1, \cdots, n^*$):

  e_i = *study* time when subject i enters the trial

  $\boldsymbol{s}_i = \{s_{i,1}, \ldots\ldots, s_{i,m_i}\}$ denotes successive *patient* times at which the i^{th} subject receives the agent, where $s_{i,1}$ coincides with study entry.

Preliminary Notation

Reconciling study time and subject time:

- Subject i receives the agent at $study$ times

$$0 < e_i + s_{i,1} < e_i + s_{i,2} < \cdots < e_i + s_{i,m_i}$$

Notation is very general:

- Agent may be administered whenever and as frequently as desired to each patient
- Allows for an arbitrary number of different treatment sequences to be studied in the trial
- A subject's actual administration times can deviate from scheduled times

Preliminary Notation

We restrict focus to k treatment sequences $\boldsymbol{s}^{(1)}, \cdots, \boldsymbol{s}^{(k)}$:

- $\boldsymbol{s}^{(j)} = (s_1, s_2, \ldots, s_{m(j)})$
- j^{th} sequence has a total of $m(j)$ administrations at i^{th} *patient* times $s_{i,1}, \ldots, s_{i,m(j)}$
- $\boldsymbol{s}^{(j)} \subset \boldsymbol{s}^{(j+1)}$ for each $j = 1, \ldots, k-1$
- $m(1) < m(2) \cdots < m(k)$

In the KGF trial:

- $\boldsymbol{s}^{(1)} = (0,\ 1,\ 2,\ 7,\ 8,\ 9)$
- $\boldsymbol{s}^{(2)} = (0,\ 1,\ 2,\ 7,\ 8,\ 9,\ 14,\ 15,\ 16,\ 21,\ 22,\ 23)$

Preliminary Notation

- $Y_i =$ actual, possibly yet unobserved, amount of time after study entry at which patient i experiences toxicity

- At interim study time t^* :

$$(Y_i^o, \delta_i) = \begin{cases} (Y_i, 1) & \text{if} \quad e_i + Y_i \leq t^* \\ (t^* - e_i, 0) & \text{if} \quad e_i + Y_i > t^*, \end{cases}$$

- $m_i =$ index of the last administration received by patient i at t^*

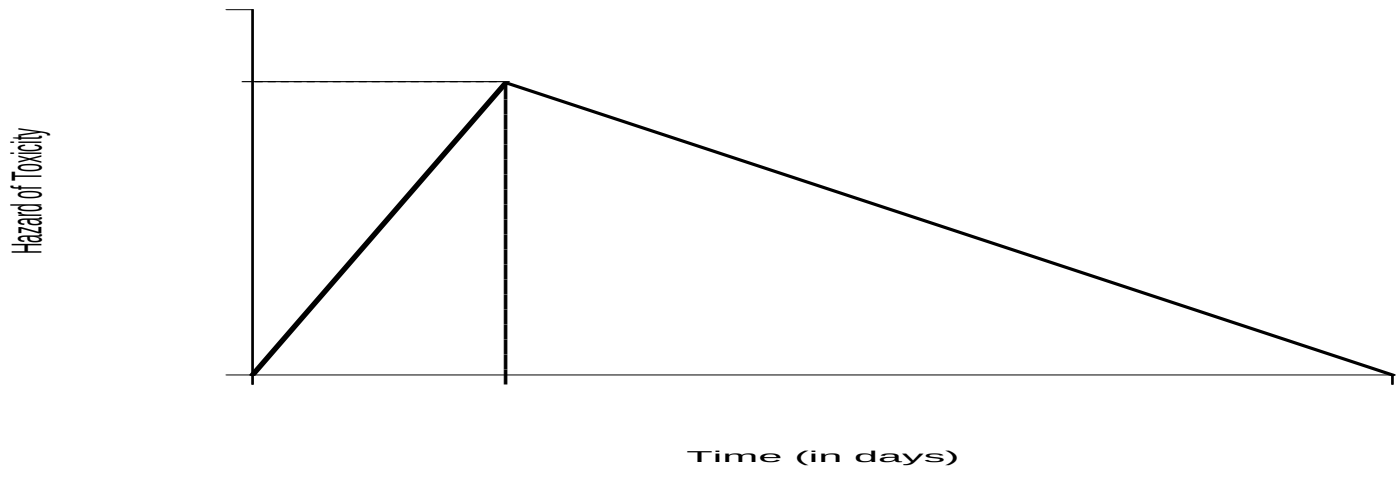
  $m_i \leq m(j)$ either due to

  (a) administrative censoring

  (b) patient i had toxicity at study time $e_i + s_{i,m_i}$ (further administration is stopped)

Modelling Single Administration Hazard

- Assume risk of toxicity increases after administration of KGF, reaches a peak, then decays to zero
- Model as a piecewise linear function:



- Hazard is identical for each administration

Modelling Single Administration Hazard

- Mathematically:

$$h(u;\boldsymbol{\theta}) = \begin{cases} \frac{\theta_2}{\theta_1} u & ; \quad 0 \leq u \leq \theta_1 \\ \\ \frac{\theta_2}{\theta_3 - \theta_1}(\theta_3 - u) & ; \quad \theta_1 < u \leq \theta_3 \end{cases}$$

θ_1 represents the point in time when $h(u;\boldsymbol{\theta})$ reaches maximum

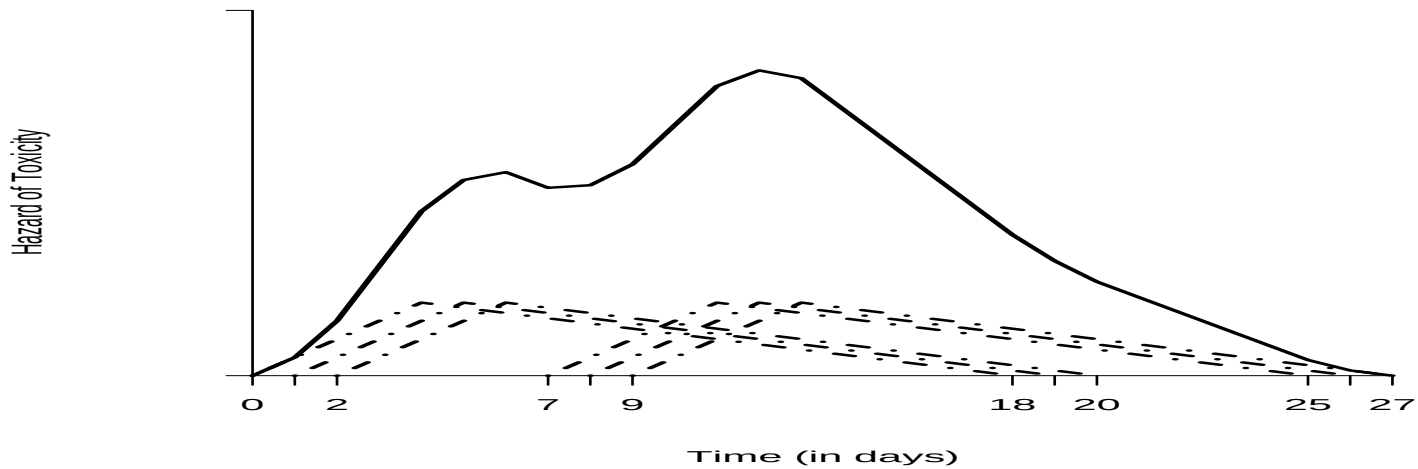
θ_2 is the hazard at $s = \theta_1$

θ_3 represents the point at which hazard vanishes

Total Hazard Function

- Using the single administration hazard function $h(u; \boldsymbol{\theta})$, we define the total hazard function for each subject:

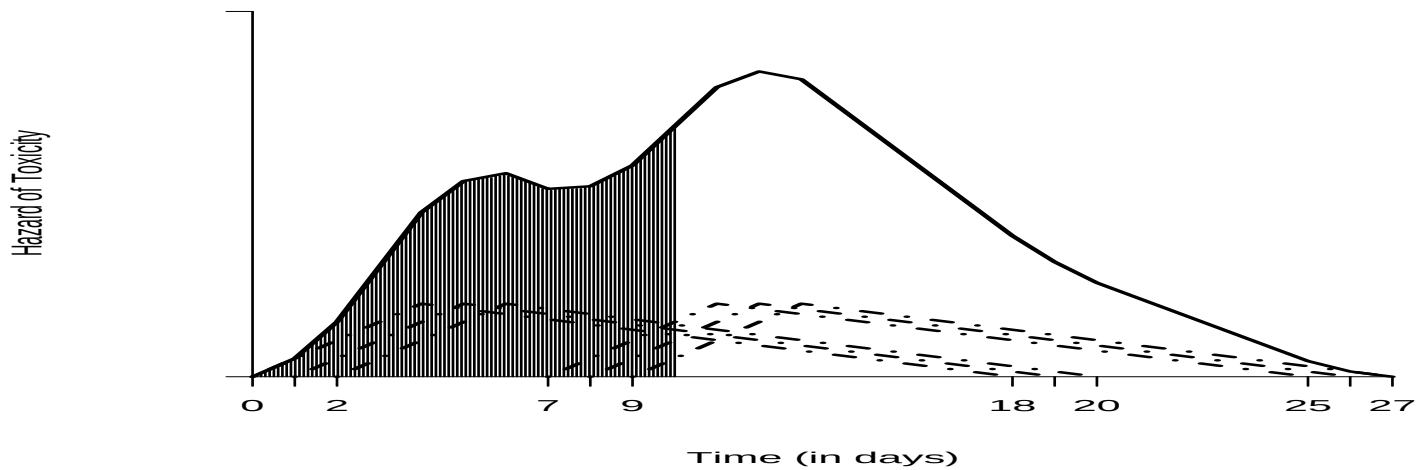
$$\lambda(t^* \mid \boldsymbol{\theta}, \boldsymbol{s}, Y^o) = \sum_{\ell=1}^{m} h(Y^o - s_\ell \mid \boldsymbol{\theta})$$



Cumulative Hazard Function

- Subject's cumulative hazard function is

$$\Lambda(t^* \mid \boldsymbol{\theta}, \boldsymbol{s}, Y^o) \quad = \quad \int_0^{Y^o} \sum_{\ell=1}^{m} h(u - s_\ell \mid \boldsymbol{\theta}) \, du$$



Functions of CHF

- Cumulative distribution function (CDF) of Y is

$$\begin{aligned} \Pr(Y \leq t^* \mid \boldsymbol{\theta}, \boldsymbol{s}, Y^o) &= F(t^* \mid \boldsymbol{\theta}, \boldsymbol{s}, Y^o) \\ &= 1 - \exp\{-\Lambda(t^* \mid \boldsymbol{\theta}, \boldsymbol{s}, Y^o)\} \end{aligned}$$

- Probability density function (PDF) is

$$f(t^* \mid \boldsymbol{\theta}, \boldsymbol{s}, Y^o) = \lambda(t^* \mid \boldsymbol{\theta}, \boldsymbol{s}, Y^o) \exp\{-\Lambda(t^* \mid \boldsymbol{\theta}, \boldsymbol{s}, Y^o)\}$$

Likelihood for θ

- At interim time t^*, we have enrolled n^* subjects
- For subject $i = 1, 2, \ldots, n^*$, we have information

$$\mathcal{D}_i = (\boldsymbol{s}_i, Y_i^0, \delta_i)$$

- Total information is $\mathcal{D}^* = (t^*, \mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_{n^*})$
- Likelihood for $\boldsymbol{\theta}$ at t^* is

$$\begin{aligned} \mathcal{L}(\mathcal{D}^* \mid \boldsymbol{\theta}) &= \prod_{i=1}^{n^*} \{f(t^* \mid \boldsymbol{\theta}, \boldsymbol{s}_i, Y_i^o)\}^{\delta_i} \\ &x \quad \prod_{i=1}^{n^*} \{1 - F(t^* \mid \boldsymbol{\theta}, \boldsymbol{s}_i, Y_i^o)\}^{1-\delta_i} \end{aligned}$$

Posterior Distribution for θ

- Denoting the prior by $p(\boldsymbol{\theta})$, the posterior of $\boldsymbol{\theta}$ is

$$g(\boldsymbol{\theta} \mid \mathcal{D}^*) \quad = \quad \frac{\mathcal{L}(\mathcal{D}^* \mid \boldsymbol{\theta}) p(\boldsymbol{\theta})}{\int \mathcal{L}(\mathcal{D}^* \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}}.$$

- Above integral cannot be obtained analytically under our assumed model
  - Compute posterior quantities via Markov chain Monte Carlo (MCMC) methods

Eliciting Prior Distribution for θ

- We denote the prior distribution of $\boldsymbol{\theta}$ as

$$p(\boldsymbol{\theta}) = p_3(\theta_3)\ p_1(\theta_1|\theta_3)\ p_2(\theta_2)$$

- For $p_3(\theta_3)$, the investigator must identify:
  - (a) $[T_\ell, T_u]$, plausible range for θ_3
  - (b) μ_{θ_3}, the expected value of θ_3

- For $p_1(\theta_1|\theta_3)$, the investigator must identify a 95% credible interval for θ_1, denoted $m \pm d$

- For $p_2(\theta_2)$, the investigator must identify the $a\ priori$ optimal schedule, denoted $\mathbf{s}^*$

Eliciting Prior Distribution for θ_3

- We assume θ_3 has a generalized beta distribution

$$p_3(u) = \frac{(u - T_\ell)^{a_3 - 1}(T_u - u)^{b_3 - 1}}{B(a_3, b_3)(T_u - T_\ell)^{a_3 + b_3 - 1}}, \qquad T_\ell \leq u \leq T_u,$$

  where

$$\begin{aligned} B(a_3, b_3) &= \int_0^1 x^{a_3 - 1}(1 - x)^{b_3 - 1} dx \\ a_3 &= k_3(\mu_{\theta_3} - T_\ell) \\ b_3 &= k_3(T_u - \mu_{\theta_3}) \end{aligned}$$

- Tuning constant k_3 modulates variance of $p_3(u)$, which decreases as k_3 increases

Eliciting Prior Distribution for θ_1

- We also assume θ_1, conditional upon θ_3, follows a generalized beta distribution

$$p_1(u|\theta_3) = \frac{u^{a_1-1}(\theta_3 - u)^{b_1-1}}{\theta_3^{a_1+b_1-1}B(a_1, b_1)}, \qquad 0 \leq u \leq \theta_3$$

- From the 95% credible interval, $m \pm d$, we have:

$$\begin{aligned} a_1 &= \frac{m}{\theta_3}\left[\frac{4m(\theta_3 - m)}{d^2} - 1\right] \\ b_1 &= \frac{\theta_3 - m}{\theta_3}\left[\frac{4m(\theta_3 - m)}{d^2} - 1\right], \end{aligned}$$

assuming approximate symmetry of $p_1(\theta_1|\theta_3)$ about m

Eliciting Prior Distribution for θ_2

- Investigator believes schedule $\mathbf{s}^*$ is MTS

  $\Rightarrow$ Schedule $\mathbf{s}^*$ has cumulative probability of toxicity by τ closest to p_τ

- Determine θ_2^* so that

$$E\{F(\tau \mid \theta_2^*, \theta_1, \theta_3^*, \boldsymbol{s}^*, 1)\} = p_\tau,$$

  where θ_3^* is an assumed value for θ_3.

- Example:

  $\Rightarrow$ In the KGF trial, investigators believed that schedule 2 (4 weeks) was optimal assuming $\theta_3^* = 18$

  $\Rightarrow$ Each administration has a cumulative hazard of $9\theta_2^*$

  $\Rightarrow$ Entire schedule has a cumulative hazard of $108\theta_2^*$

  $\Rightarrow$ $\theta_2^* = -log(1 - p_\tau)/108 \quad -- > \ \mu_{\theta_2}$

Eliciting Prior Distribution for θ_2

- θ_2 describes the height of the single-administration hazard

  $\Rightarrow$ θ_2 is quantitatively different than θ_1 and θ_3

  $\Rightarrow$ Generalized beta distribution inappropriate choice for $p_2(u)$

- We assume θ_2 has a prior Gamma distribution

$$p_2(u) = b_2^{a_2} u^{a_2 - 1} exp\{-b_2 u\}/\Gamma(a_2), \qquad u \geq 0,$$

  where $a_2 = k_2$, $b_2 = k_2/\mu_{\theta_2}$, and $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$

- Tuning constant k_2 modulates the variance of $p_2(u)$, which decreases as k_2 increases

Fine-Tuning Prior Distribution for θ

- Variances of $p_3(\theta_3)$, $p_1(\theta_1|\theta_3)$, and $p_2(\theta_2)$ heavily influence the ability of the data to influence $p(\boldsymbol{\theta})$

- Exhaustive sensitivity analysis of (d, k_2, k_3) is necessary before implementing the design
  - Simulate toxicity times for a 3-5 subjects
  - Apply design using various values of d, k_2, k_3
  - Compare the prior means for $\boldsymbol{\theta}$ to their respective posterior values

Visualizing Prior for $F(\tau \mid \boldsymbol{\theta}, \boldsymbol{s}^{(j)}, 1)$

- From $p(\boldsymbol{\theta})$, we have implied a specific prior distribution for the cumulative probability of toxicity for each schedule

- To visualize the prior distribution for $F(\tau \mid \boldsymbol{\theta}, \boldsymbol{s}^{(j)}, 1)$ for each schedule j:
  - Draw B samples from the prior distribution of $\boldsymbol{\theta}$
  - Compute B prior estimates of $F(\tau \mid \boldsymbol{\theta}, \boldsymbol{s}^{(j)}, 1)$ for each schedule j
  - Plot histograms to determine if there is any undue prior influence on those estimates, i.e, placing too much mass at 0 or 1

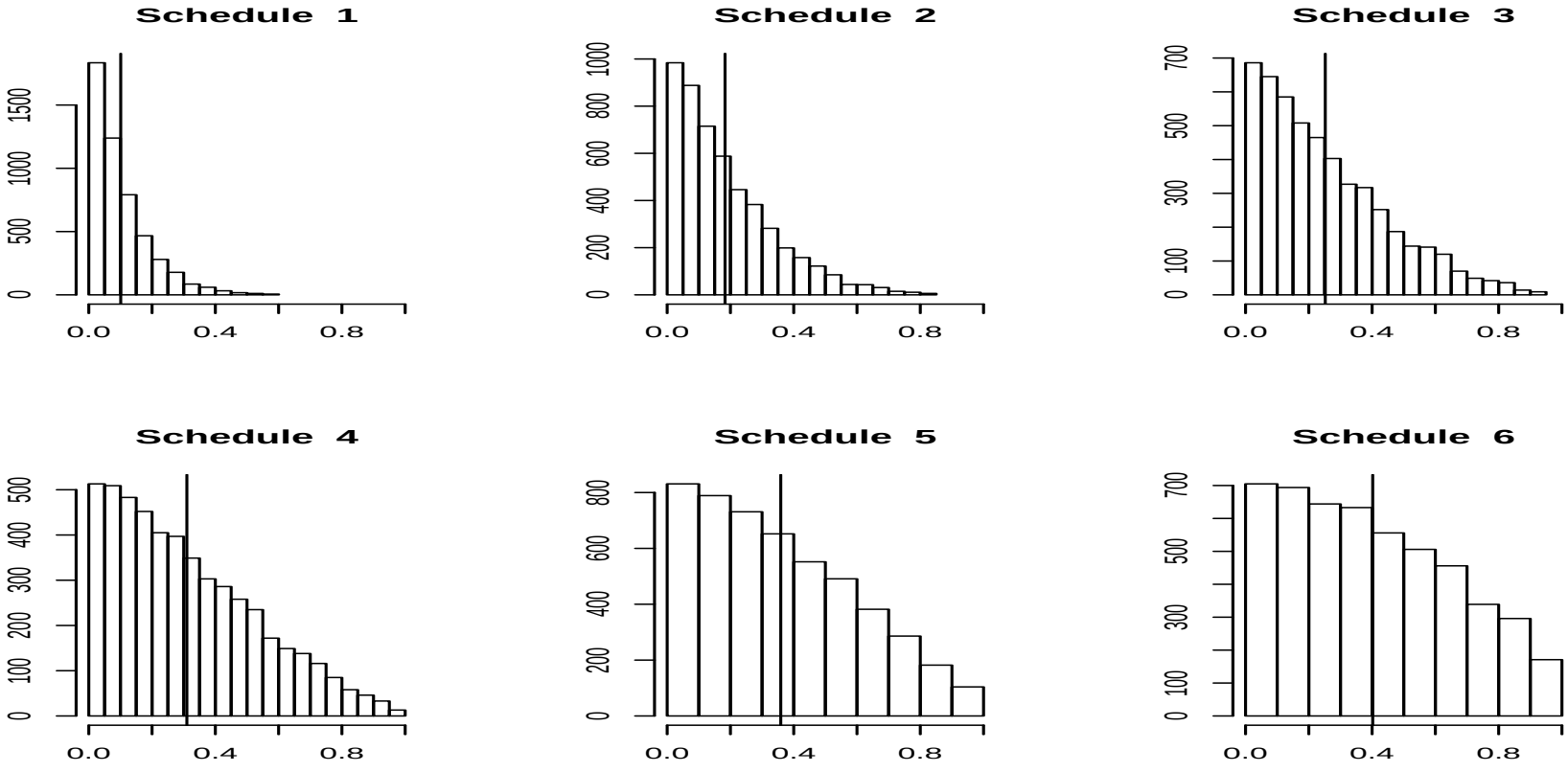
Visualizing Prior for $F(\tau \mid \boldsymbol{\theta}, \boldsymbol{s}^{(j)}, 1)$

- Example:
  - θ_3 ranges over $[4, 50]$ days, with a mean of 18 days
  - θ_1 occurs 2 ± 2 days after administration
  - θ_2 has mean 0.0021 if $p_\tau = 0.20$, Schedule 2 is the MTS and $\theta_3 = 18$
  - $k_2 = k_3 = 1$

- Mean cumulative probability of toxicity for Schedule 2 is closest to 0.20

- Variation of $F(\tau \mid \boldsymbol{\theta}, \boldsymbol{s}^{(j)}, 1)$ increases with j due the cumulative nature of the schedules

Visualizing Prior for $F(\tau \mid \boldsymbol{\theta}, \boldsymbol{s}^{(j)}, 1)$



Trial Conduct Issues

- $N =$ maximum number of patients to be enrolled in the trial
- First patient is assigned the shortest sequence, $\boldsymbol{s}^{(1)}$
- Only incremental schedule escalation is permitted
- At least M subjects must have received schedules $\mathbf{s}^{(1)}, ..., \mathbf{s}^{(j-1)}$ before schedule $\mathbf{s}^{(j)}$ is assigned
- Schedule de-escalation is permitted without any constraint

Determining Schedule Assignments

- Given p_τ, we will consider two alternative criteria for choosing each subject's sequence

- Criterion 1: At time t^*, for each $j = 1, ..., k$, compute

$$F_j^*(\tau) = E\{F(\tau \mid \boldsymbol{\theta}, \boldsymbol{s}^{(j)}, 1) \mid \mathcal{D}^*\}$$

- The best sequence is defined as that having $F_j^*(\tau)$ closest to p_τ, that is, minimizing $|F_j^*(\tau) - p_\tau|$

Determining Schedule Assignments

- Criterion 2: At time t^*, for each $j = 1, ..., k$, compute

$$\phi_j(\tau) = \Pr\{F(\tau \mid \boldsymbol{\theta}, \boldsymbol{s}^{(j)}, 1) > p_\tau \mid \mathcal{D}^*\}$$

- Given a fixed upper limit, $\bar{p}$, the best sequence is defined as the longest sequence for which $\phi_j(\tau) < \bar{p}$, that is, for which the risk of toxicity is acceptable

- Under either Criterion 1 or 2:

  Best sequence is assigned to next patient

  MTS is defined as the best sequence using information from all N subjects at the end of the trial

Assessing Performance via Simulation

- $k = 6$ schedules of KGF are investigated: 2, 4, 6, 8, 10 or 12 weeks

- Six scenarios studied: Schedule $\mathbf{s}^{(j)}$ optimal under the jth scenario
  - $\theta_1 = 2$ days; $\theta_3 = 18$ days
  - $1000\theta_2 \in [4.13, 2.07, 1.38, 1.03, 0.83, 0.69]$

- $N = 30$

- $\tau = 100$ days

- $p_\tau = 0.20$

Assessing Performance via Simulation

- Prior distributions:

  $\Rightarrow$ θ_3 has mean of 18 days; range of [4, 100] days

  $\Rightarrow$ θ_1 between 0-4 days (2 ± 2 days)

  $\Rightarrow$ Schedule 6 is MTS

  θ_2 has a prior mean 0.00069

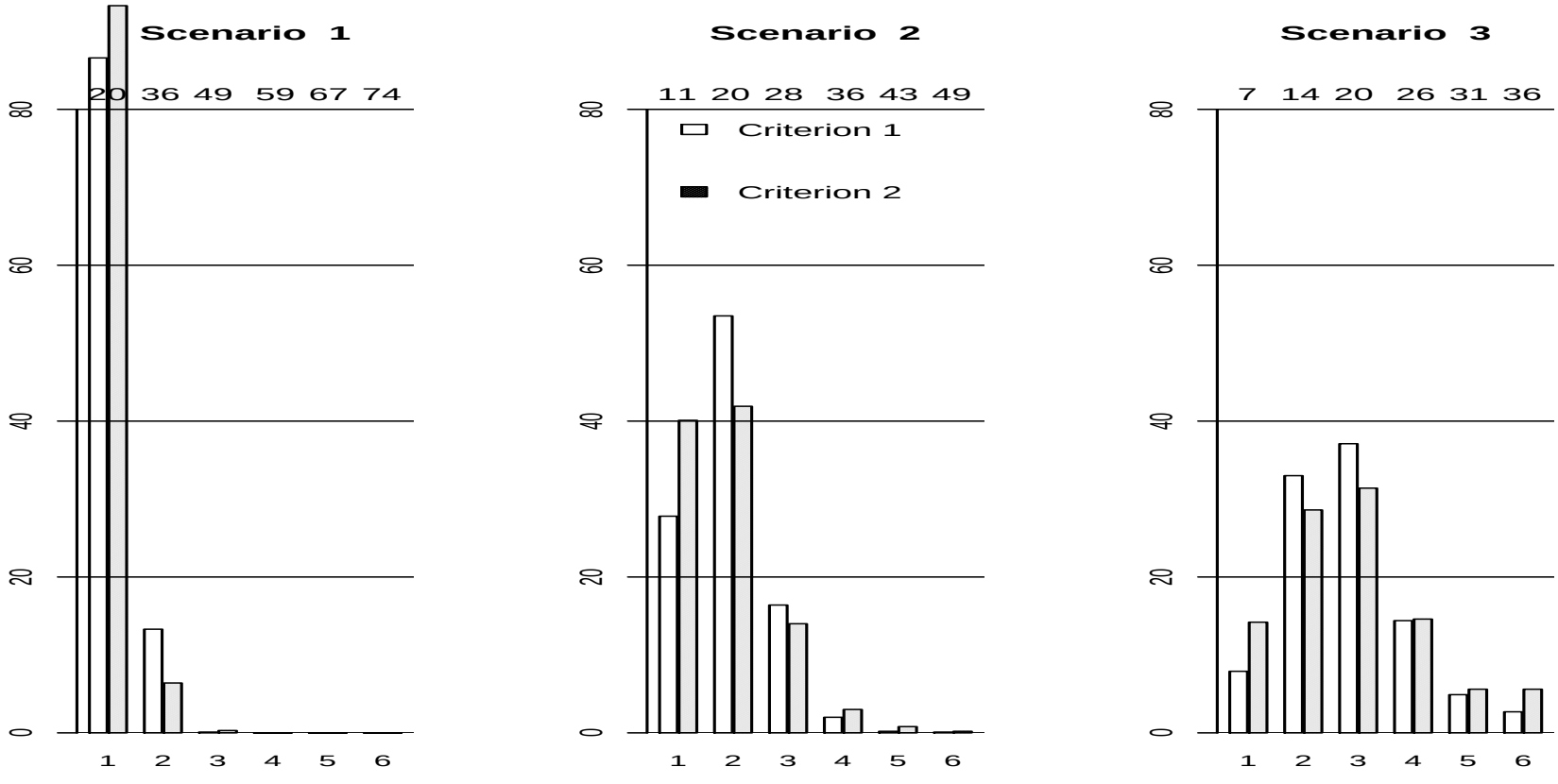
  $\Rightarrow$ $k_3 = 0.1$; $k_2 = 0.2$

- Subject inter-arrival times $\sim U[12, 16]$ days

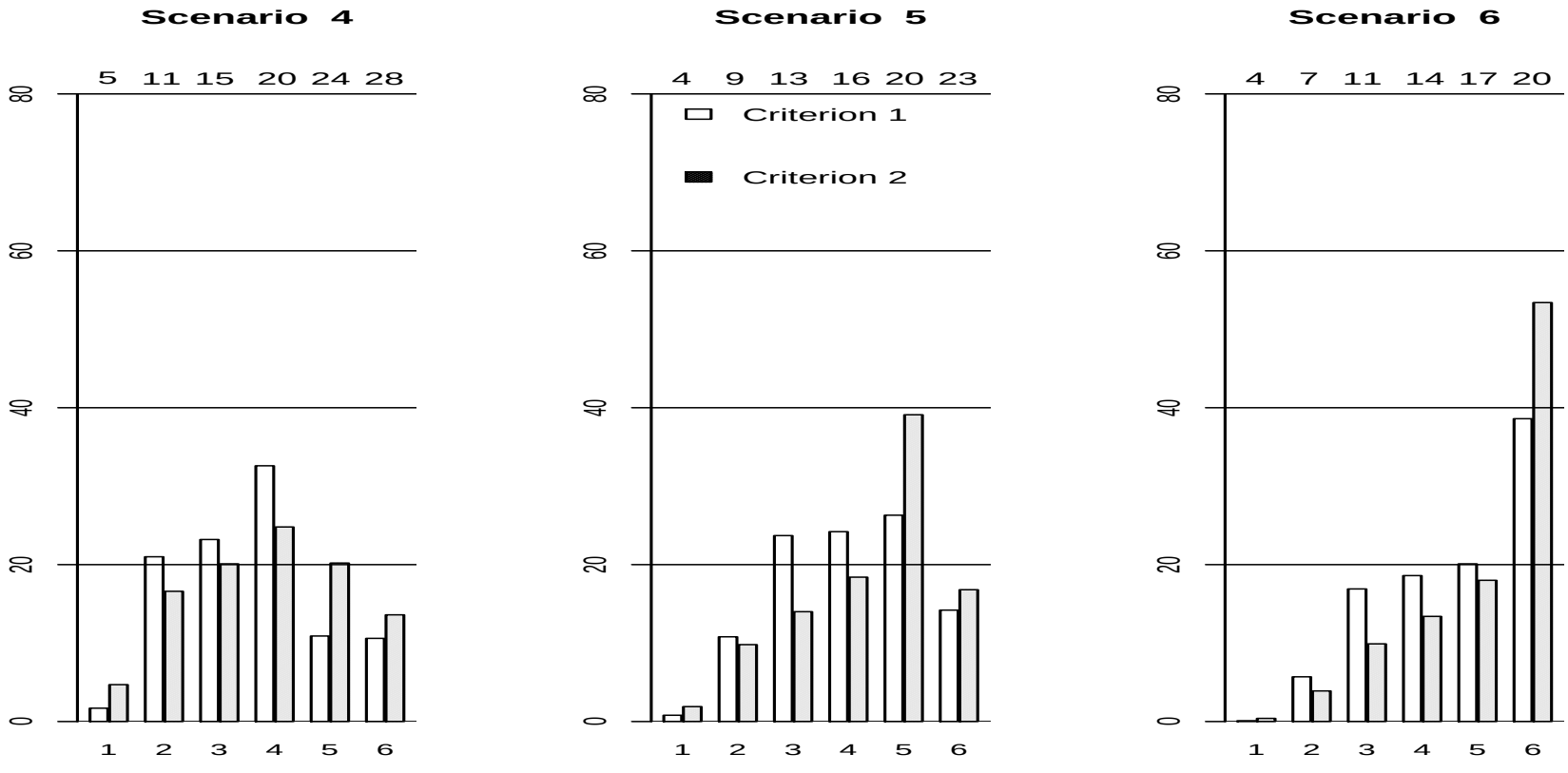
- Posteriors based upon 2000 samples; burn-in of 500 samples

- $M = 1000$ simulations

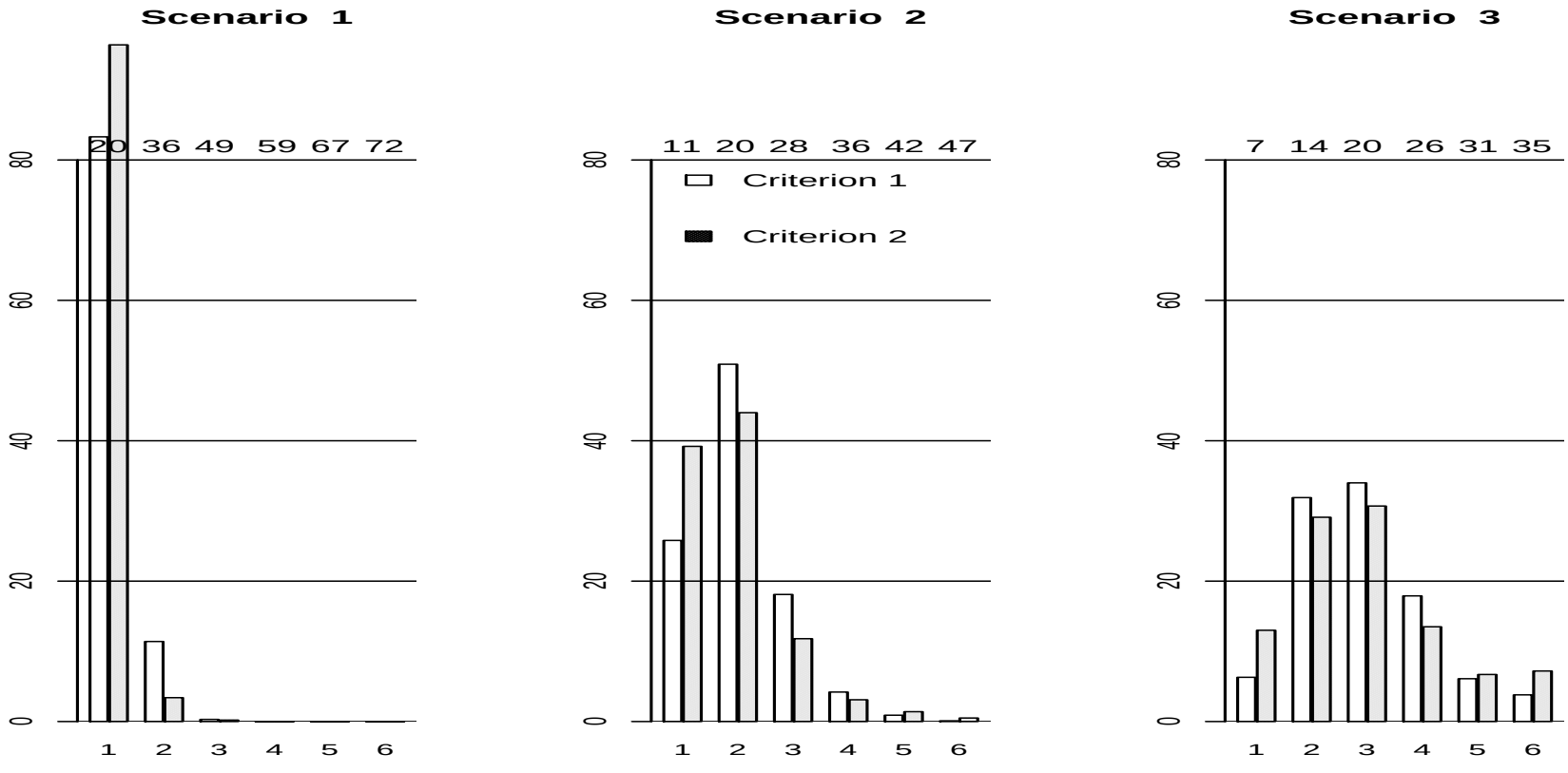
Identifying the MTS: $\theta_3 = 18$



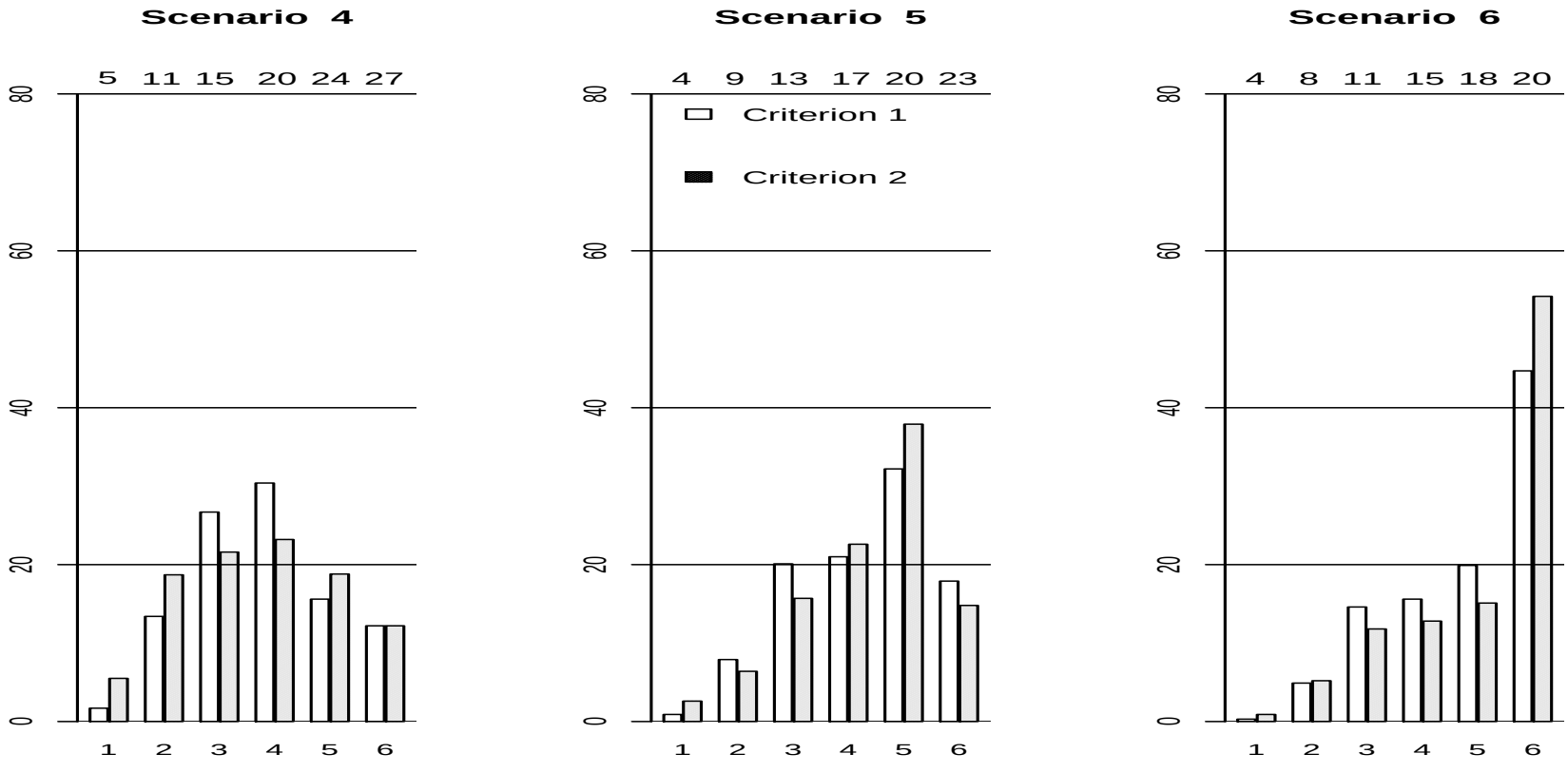
Identifying the MTS: $\theta_3 = 18$



Identifying the MTS: $\theta_3 = 50$



Identifying the MTS: $\theta_3 = 50$



Concluding Remarks

- Overall determination of MTS relatively insensitive to $p_1(\theta_1)$:
  - θ_1 influences the time at which each patient experiences toxicity
  - Has more influence on schedule assignments during the study
- Strong homogeneity assumption:
  - Patients who receive the shortest sequence $\boldsymbol{s}^{(1)}$ provide information about the toxicity of schedules $\boldsymbol{s}^{(j)}, j \geq 1$
- Extensions to Phase II & III studies of cumulative dosing
  - Adjusting for between-patient variability