Bayesian Design of Clinical Trials Using Joint Models for Recurrent and Terminating Events

Joseph G. Ibrahim

Department of Biostatistics, Gillings School of Global Public Health

The University of North Carolina at Chapel Hill

Joint work with Jiawei Xu and Matthew A. Psioda

ENAR 2022

Background

- Recurrent event data are increasingly common in clinical trials. Designs based on more time-to-event measurements (i.e., recurrent events) provide greater efficiency and power compared to designs using only one time-to-event endpoint (Chen et al., 2014).
- In situations with recurrent events and a terminating event (e.g., death), a patient's clinical experience will be characterized by both event processes and the two processes are generally dependent on one another.
- The censorship of recurrent events is no longer non-informative, but instead depends on the terminating event (Rondeau et al., 2007).

Summary

- We develop a Bayesian clinical trial design focused on evaluating an investigational product's (IP's) effect on both recurrent and terminating event processes considered as multiple primary endpoints, using a multi-frailty joint model.
- Inferences for the multiple primary outcomes are based on posterior model probabilities corresponding to mutually exclusive hypotheses regarding the benefit of IP with respect to the two event processes.
- We demonstrate the methodology by designing a colorectal cancer clinical trial with a goal of demonstrating the IP causes a favorable effect on *at least one* of the two outcomes but no harm on either.

Notation

- Using notation similar to Rondeau et al. (2007), assume the *ith* patient has a terminating event at time D_i or is censored at time C_i and X_{ij} is the *jth* recurrent event for $j = 1, ..., n_i$, where n_i is the total number of observed recurrent events or censoring for patient i.
- Define $T_{ij} = \min(X_{ij}, D_i, C_i)$ as the observed recurrent event times with $\delta_{ij} = I(t_{ij} = X_{ij})$ indicating whether the *jth* recurrent event occurred.
- Define $T_i^* = \min(D_i, C_i)$ as the last observed time for the *ith* patient, which is either a time of terminating event or censoring. Let $\delta_i^* = I(T_I^* = D_i)$ denote the indicator for whether the patient is censored or not.
- The gap time (i.e., time-between-events) is then given by $S_{ij} = T_{ij} - T_{i,j-1}$ with $T_{i0} = 0$ for the recurrent event hazard function.

Multi-frailty Joint Model

- For the *ith* patient, the multi-frailty joint model of hazard functions for the recurrent event and terminating event using gap times can be written as equations (1) and (2), respectively,

$$r_i(s|\mu_i,\nu_i) = \mu_i\nu_i r_0(s)\exp(z_i\gamma_r + x'_{ir}\beta_r) = \mu_i\nu_i r_i(s) \quad (1)$$

$$\lambda_i(t|\mu_i) = \mu_i\lambda_0(t)\exp(z_i\gamma_\lambda + x'_{i\lambda}\beta_\lambda) = \mu_i\lambda_i(t) \quad (2)$$

- z_i is the IP indicator, x_{ir} and $x_{i\lambda}$ are the vectors of covariates for the recurrent and terminating event hazard functions, respectively.
- γ_r, β_r and $r_0(t)$ are treatment effect, vector of coefficient parameters and piecewise constant baseline hazard function for the recurrent event model.
- γ_λ, β_λ and $\lambda_0(t)$ are the analogous quantities for the terminating event model.

Multi-frailty Joint Model

- The frailty $\mu_i \sim \text{Gamma}(1/\theta, 1/\theta)$ accounts for the dependence between the two event processes and the frailty $\nu_i \sim \text{Gamma}(1/\eta, 1/\eta)$ accounts for dependence between recurrent event times. The two frailties are assumed to be independent from each other.
- Conditional on the frailty ν_i, the gap times for the same patient are mutually independent.
- The association parameter θ controls the strength of dependence between the two event processes and, conditional on the variance parameter η, a larger θ reflects a stronger dependence between the recurrent and terminating events.

Likelihood

- Let **D** be the observed data for n patients. Denote $\psi = (\gamma, \beta, \lambda_0, r_0, \theta, \eta)$ as the full set of parameters, where $\boldsymbol{\gamma} = (\gamma_r, \gamma_\lambda)$ and $\boldsymbol{\beta} = (\beta_r, \beta_\lambda)$.
- For the *ith* patient, the likelihood contribution conditional on μ_i, ν_i takes the form:

$$\begin{aligned} L_i(\psi|\mu_i, \nu_i, \mathbf{D}) &= \prod_{j=1}^{n_i} \left\{ \mu_i \nu_i r_0(T_{ij}) e^{z_i\gamma_r + x'_{ir}\beta_r} \right\}^{\delta_{ij}} \exp\{-\mu_i \nu_i R_0(T_i^*) e^{z_i\gamma_r + x'_{ir}\beta_r}\} \\ &\times \{\mu_i \lambda_0(T_i^*) e^{z_i\gamma_\lambda + x'_{i\lambda}\beta_\lambda}\}^{\delta_i^*} \exp\{-\mu_i \Lambda_0(T_i^*) e^{z_i\gamma_\lambda + x'_{i\lambda}\beta_\lambda}\} \end{aligned}$$

- where $R_0(t)$ and $\Lambda_0(t)$ are the cumulative piecewise constant baseline hazard functions corresponding to $r_0(t)$ and $\lambda_0(t)$, respectively.
- The complete data likelihood contribution for the *ith* patient is obtained by multiplying the likelihood above by the distribution for the frailties.

Study Design

- We consider a design to demonstrate the superiority of an IP compared to a control with respect to recurrent and terminating events as multiple primary outcomes.
- We follow patients for both recurrent and terminating events starting at baseline.
- For each patient, recurrent events are documented until a fixed time or the occurrence of the terminating event. If the recurrent and terminating events occur at the same time, only the terminal event is recorded.

Superiority Test

- In order to test whether the IP has benefit to at least one of the two event processes and no harm to either, we consider the following group of hypotheses:

$$
\begin{array}{lccc}
H_1: & \exp(\gamma_r) > \delta_r & \text{or} & \exp(\gamma_\lambda) > \delta_\lambda \\
H_2: & \exp(\gamma_r) > \delta_r & \text{and} & \exp(\gamma_\lambda) = \delta_\lambda \\
H_3: & \exp(\gamma_r) = \delta_r & \text{and} & \exp(\gamma_\lambda) > \delta_\lambda \\
H_4: & \exp(\gamma_r) = \delta_r & \text{and} & \exp(\gamma_\lambda) = \delta_\lambda \\
H_5: & \exp(\gamma_r) < \delta_r & \text{and} & \exp(\gamma_\lambda) = \delta_\lambda \\
H_6: & \exp(\gamma_r) = \delta_r & \text{and} & \exp(\gamma_\lambda) < \delta_\lambda \\
H_7: & \exp(\gamma_r) < \delta_r & \text{and} & \exp(\gamma_\lambda) < \delta_\lambda
\end{array}
$$

- We consider the union of H_1, H_2, H_3 and H_4 as the null hypothesis (i.e. $H_0 = H_1 \cup H_2 \cup H_3 \cup H_4$) with the alternative as the union of H_5, H_6 and H_7 (i.e. $H_a = H_5 \cup H_6 \cup H_7$).

Superiority Test

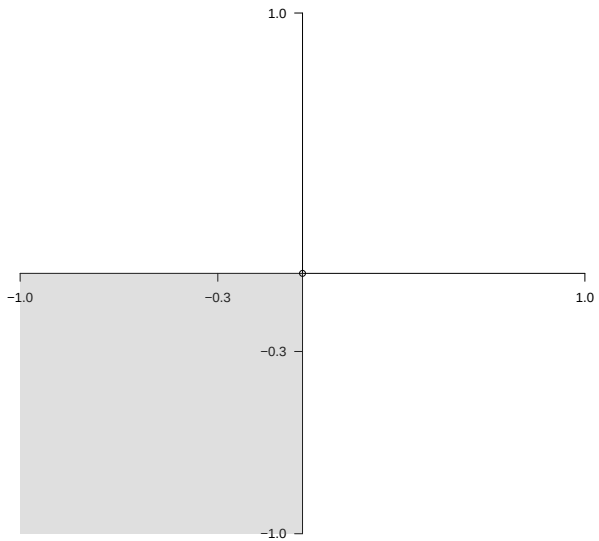

Figure 1: Depiction of the null and alternative hypothesis space.

The gray shaded region (including the coordinate axis boundaries but not the origin) corresponds to the alternative and the complement of the aforementioned space (including the origin) corresponds to the null hypothesis. The hypothesized treatment effects are -0.3 on both recurrent and terminating event processes.

Basis and *Hypothesis* Models

- Assume $\tilde{M}_j$ is the model corresponding to hypotheses H_j, for $j = 1, ..., 7$.
- Consider the model space of the treatment parameter vector $(\gamma_r, \gamma_\lambda)$, let B_1 denote the full model (i.e., neither treatment parameter fixed), B_2 and B_3 define the models with γ_λ and γ_r fixed, respectively, and B_4 denotes the model having both parameters fixed.
- We refer to models $B_i, i = 1, ..., 4$ as the *Basis* models and $\tilde{M}_j, j = 1, ..., 7$ as the *Hypothesis* models.

Basis and *Hypothesis* Models

Table 1: Relationship between *Basis* and *Hypothesis* models.

Hypothesis	*Hypothesis* Model	*Basis* Model
H_0	$\tilde{M}_1$	B_1
	$\tilde{M}_2$	B_2
	$\tilde{M}_3$	B_3
	$\tilde{M}_4$	B_4
H_a	$\tilde{M}_5$	B_2
	$\tilde{M}_6$	B_3
	$\tilde{M}_7$	B_1

Posterior Model Probabilities (PMP's)

- Let E_j be an indicator that γ takes a value in the parameter space associated with hypothesis H_j, for $j = 1, ..., 7$.
- One can show that the posterior probabilities for the null and alternative hypotheses, respectively, take the form

$$\begin{aligned}
P(H_0|\mathbf{D}) &= P(\tilde{M}_1|\mathbf{D}) + P(\tilde{M}_2|\mathbf{D}) + P(\tilde{M}_3|\mathbf{D}) + P(\tilde{M}_4|\mathbf{D}) \\
&= P(B_1|\mathbf{D}) \cdot P(E_1|B_1, \mathbf{D}) + P(B_2|\mathbf{D}) \cdot P(E_2|B_2, \mathbf{D}) \\
&+ P(B_3|\mathbf{D}) \cdot P(E_3|B_3, \mathbf{D}) + P(B_4|\mathbf{D}) \\
P(H_a|\mathbf{D}) &= P(\tilde{M}_5|\mathbf{D}) + P(\tilde{M}_6|\mathbf{D}) + P(\tilde{M}_7|\mathbf{D}) \\
&= P(B_2|\mathbf{D}) \cdot P(E_5|B_2, \mathbf{D}) + P(B_3|\mathbf{D}) \cdot P(E_6|B_3, \mathbf{D}) \\
&+ P(B_1|\mathbf{D}) \cdot P(E_7|B_1, \mathbf{D})
\end{aligned}$$

- where $P(E_j|B_k, \mathbf{D})$ can be easily obtained by computing the proportion of MCMC samples that satisfy H_j given B_k.

Posterior Model Probabilities (PMP's)

- We adapt the methods of Chen (1994) and Chen and Shao (1997) to estimate $P(B_k|\mathbf{D}), k = 1, 2, 3, 4$.
- For $\phi = (\psi, \mu, \nu)$, let $\phi^{(B_k)}$ be the vector of parameters in *Basis* model B_k that are free to vary. Write $\phi = (\phi^{(-B_k)}, \phi^{(B_k)})$ where $\phi^{(-B_k)}$ is the complementary set of parameters that are fixed under model B_k.
- Based on the MCMC sample $\{\phi_{(i)}, i = 1, ..., N\}$ from the full model B_1, the posterior probability of model B_k can be estimated as

$$\hat{p}(B_k|\mathbf{D}) = \frac{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{L(\phi_{(i)}^{(B_k)})\pi(\phi_{(i)}^{(B_k)})w(\phi_{(i)}^{(-B_k)}|\phi_{(i)}^{(B_k)})}{L(\phi_{(i)})\pi(\phi_{(i)})}\right)p(B_k)}{\sum_{p=1}^{2^P}\frac{1}{N}\sum_{i=1}^{N}\left(\frac{L(\phi_{(i)}^{(B_p)})\pi(\phi_{(i)}^{(B_p)})w(\phi_{(i)}^{(-B_p)}|\phi_{(i)}^{(B_p)})}{L(\phi_{(i)})\pi(\phi_{(i)})}\right)p(B_p)}$$

Prior Model Probabilities

- Under the assumption that $\gamma_r \perp\!\!\!\perp \gamma_\lambda$, we have the prior distribution $\pi(\boldsymbol{\gamma}) = \pi(\gamma_r)\pi(\gamma_\lambda)$.
- Both $\pi(\gamma_r)$ and $\pi(\gamma_\lambda)$ are specified as mixture distributions with $\pi(\gamma_r) = \pi_r \cdot 1(\gamma_r = \Delta_{0r}) + (1 - \pi_r) \cdot f_r(\gamma_r)$ and $\pi(\gamma_\lambda) = \pi_\lambda \cdot 1(\gamma_\lambda = \Delta_{0\lambda}) + (1 - \pi_\lambda) \cdot f_\lambda(\gamma_\lambda)$, where $\Delta_{0r} = \log(\delta_r)$ and $\Delta_{0\lambda} = \log(\delta_\lambda)$, $f_r(\cdot)$ and $f_\lambda(\cdot)$ are Normal(ω_r, σ_r^2) and Normal$(\omega_\lambda, \sigma_\lambda^2)$, respectively.
- The prior model probabilities for the *Basis* models are:

$$\begin{aligned} P(B_1) &= (1-\pi_r)(1-\pi_\lambda) \\ P(B_2) &= (1-\pi_r)\pi_\lambda \\ P(B_3) &= \pi_r(1-\pi_\lambda) \\ P(B_4) &= \pi_r\pi_\lambda \end{aligned}$$

Prior Model Probabilities

- The induced prior model probabilities for the *Hypothesis* models are then defined as

$$
\begin{aligned}
P(\tilde{M}_1) &= P(B_1)\{1 - F_\lambda(\Delta_{0\lambda})F_r(\Delta_{0r})\} \\
P(\tilde{M}_2) &= P(B_2)\{1 - F_r(\Delta_{0r})\} \\
P(\tilde{M}_3) &= P(B_3)\{1 - F_\lambda(\Delta_{0\lambda})\} \\
P(\tilde{M}_4) &= P(B_4) \\
P(\tilde{M}_5) &= P(B_2)F_r(\Delta_{0r}) \\
P(\tilde{M}_6) &= P(B_3)F_\lambda(\Delta_{0\lambda}) \\
P(\tilde{M}_7) &= P(B_1)F_r(\Delta_{0r})F_\lambda(\Delta_{0\lambda})
\end{aligned}
$$

- where $F_r(\Delta_{0r}) = \int 1(\gamma_r < \Delta_{0r}) f_r(\gamma_r) d\gamma_r = \Phi(\frac{\Delta_{0r} - \omega_r}{\sigma_r})$ and $F_r(\Delta_{0\lambda}) = \int 1(\gamma_\lambda < \Delta_{0\lambda}) f_\lambda(\gamma_\lambda) d\gamma_\lambda = \Phi(\frac{\Delta_{0\lambda} - \omega_\lambda}{\sigma_\lambda})$ where $\Phi(\cdot)$ is the CDF of the standard normal distribution.

Prior Model Probabilities

- Without loss of generality, we assume $\pi_r = \pi_\lambda = \pi$ with both $f_r(\cdot)$ and $f_\lambda(\cdot)$ centered at 0 (i.e., $\omega_r = \omega_\lambda = 0$). Thus, we have $F_r(\Delta_{0r} = 0) = F_r(\Delta_{0\lambda} = 0) = \frac{1}{2}$ and for the alternative hypothesis,

$$\begin{aligned} P(\tilde{M}_5) + P(\tilde{M}_6) + P(\tilde{M}_7) &= \pi(1-\pi)/2 + \pi(1-\pi)/2 + (1-\pi)^2/4 \\ &= -\frac{3}{4}\pi^2 + \frac{1}{3}\pi + \frac{1}{4} \end{aligned}$$

  which is maximized at $(\frac{1}{3}, \frac{1}{3})$.

- Therefore, the maximum weight that the alternative hypothesis can obtain is $\frac{1}{3}$, which occurs when $\pi_r = \pi_\lambda = \frac{1}{3}$.

Estimation and Inference

- We consider a weighted average Bayesian type I error rate and power, with weights determined by user-specified null and alternative sampling prior distributions.
- The null sampling prior gives non-zero weight to values of $\boldsymbol{\psi}$ such that H_0 is satisfied and the alternative sampling prior such that H_a is satisfied.
- We only consider point-mass sampling priors such that $\pi_0^{(s)}(\boldsymbol{\psi}) = 1(\boldsymbol{\psi} = \boldsymbol{\psi}_0)$ and $\pi_1^{(s)}(\boldsymbol{\psi}) = 1(\boldsymbol{\psi} = \boldsymbol{\psi}_1)$.
- A general description for how to elicit non-degenerate null and alternative sampling priors from historical data can be found in Psioda and Ibrahim (2018).

Estimation and Inference

- For a fixed value of ψ, the null hypothesis rejection rate is defined as $r(\psi) = E[1\{P(H_a|\mathbf{D}) \geq p_0\}|\psi]$.
- Then the Bayesian type I error rate and power are defined as

$$\alpha^{(s)} = E[r(\psi)|\pi_0^{(s)}] \qquad \text{and} \qquad 1 - \beta^{(s)} = E[r(\psi)|\pi_1^{(s)}]$$

which are weighted averages of $r(\xi)$ with weights determined by $\pi_0^{(s)}(\psi)$ and $\pi_1^{(s)}(\psi)$.

Sample Size Determination (SSD)

- We want to determine the smallest v such that the Bayesian type I error rate and power are controlled at level $\alpha^{(s)}$ and $1 - \beta^{(s)}$, respectively.
- The number of patients enrolled in the trial may be chosen to obtain a specified number of events in a specified interval of time on average.
- Let the sample size and number of events be given by n and v, respectively. We consider an approach that fixes the ratio $r = \frac{n}{v}$ but varies the number of events.

Example: Bayesian Clinical Design for Colorectal Cancer

- We consider a colorectal cancer study conducted at Hospital Universitary in L'Hospitalet, Spain (González et al., 2005).
- The study investigated sex based inequalities in hospital readmission among patients diagnosed with colorectal cancer.
- There were 403 patients diagnosed between January 1996 and December 1998 and they were actively followed up until 2002.
- Hospital readmission times related to colorectal cancer after surgery were collected, with mortality also recorded during follow up.

Example: Bayesian Clinical Design for Colorectal Cancer

- We consider a design evaluating an IP (e.g., chemotherapy) with respect to hospital readmission times and mortality as multiple primary outcomes.
- For each patient, readmission to a hospital was recorded until some fixed time (e.g., 6 years) or the occurrence of the terminating event.
- Patients were randomized to two treatment arms using a 1:1 allocation scheme and the accrual rate was simulated to be uniform over a 1 year period.
- Censoring (i.e., dropout) was assumed to follow a mixture distribution with a 0.05 probability of dropout and, conditional on being a dropout, the time to dropout was uniform over a 6 year period.
- We included a binary covariate (male vs female) which was simulated such that approximately 50% of the subjects were females.

Example: Bayesian Clinical Design for Colorectal Cancer

- We assumed piecewise constant baseline hazards for both the recurrent and terminating hazard functions.
- For the treatment effects, we proposed the hypothesized effects on both hospital readmission and mortality as $\Delta_{1r} = \Delta_{1\lambda} = -0.3$ but also allowed various sampling priors.
- The prior distributions for the treatment effects on both event processes are the same as $\pi(\gamma) = \frac{1}{3} \cdot 1(\gamma = 0) + \frac{2}{3} \cdot \mathrm{N}(0, 0.6)$, where γ is used here to represent the treatment effect on the recurrent or terminating event process.

Example: Bayesian Clinical Design for Colorectal Cancer

- We evaluate the performance of the proposed approach against a frequentist approach through type I error rate and power estimates.
- The frequentist approach is implemented based on a one-sided superiority test using the Cox model, with both the recurrent and terminating event hypotheses taken to be co-primary. The dependence between recurrent event times is accounted for using the marginal approach of Wei et al. (1989).
- The frequentist approach does not perfectly align with the proposed Bayesian approach, as it doesn't have the capability to evaluate whether the IP is beneficial to at least one of the event processes and not harmful on either, whereas the Bayesian approach does.

Results: Type I Error Rate

Table 2: Type 1 error rate estimates

Model	Bayes	Freq	Rec	Ter
Type I error rate	0.002	0.006	0.053	0.047

Bayes: Bayesian testing approach based on multiple primary outcomes.
Freq: Frequentist testing approach for co-primary outcomes.
Rec: Test of treatment effect on the recurrent event process.
Ter: Test of treatment effect on the terminating event process.

- Table 2 shows the estimates under the scenario where there is no treatment effect on either the recurrent or terminating event hazard (i.e., $\gamma_r = \gamma_\lambda = 0$).

Results: Bayesian Type I Error Rate

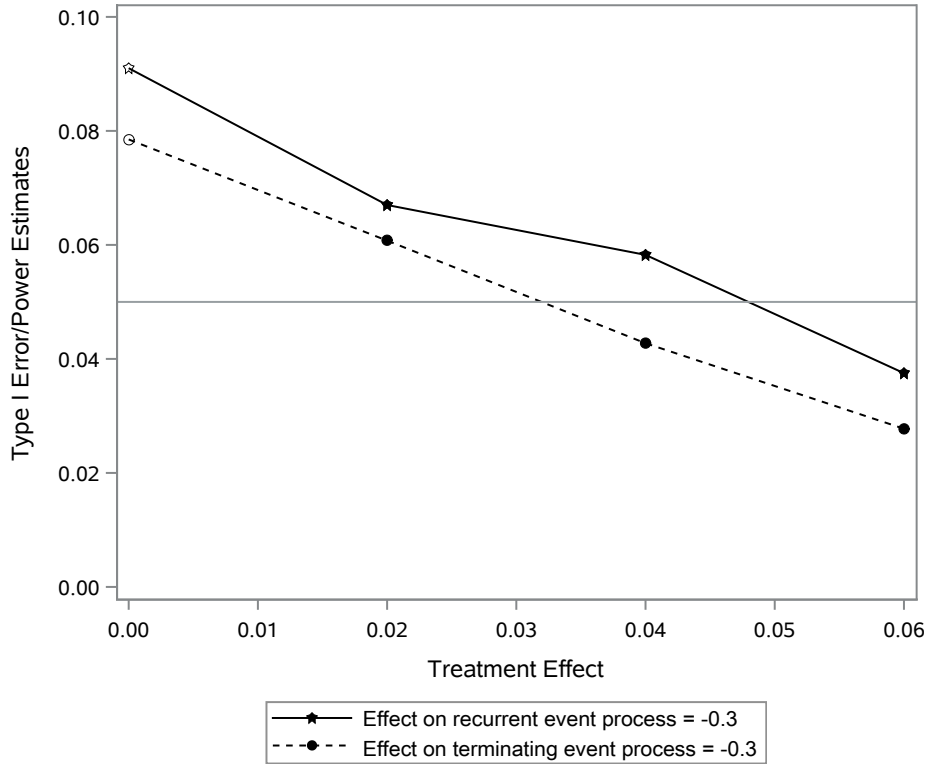

Figure 2: Bayesian type I error rate curves.

Figure 2 presents the Bayesian estimated type I error rate curves when the IP has a favorable effect on one event process but has no effect or causes a modest degree of harm on the other.

Results: Bayesian Power

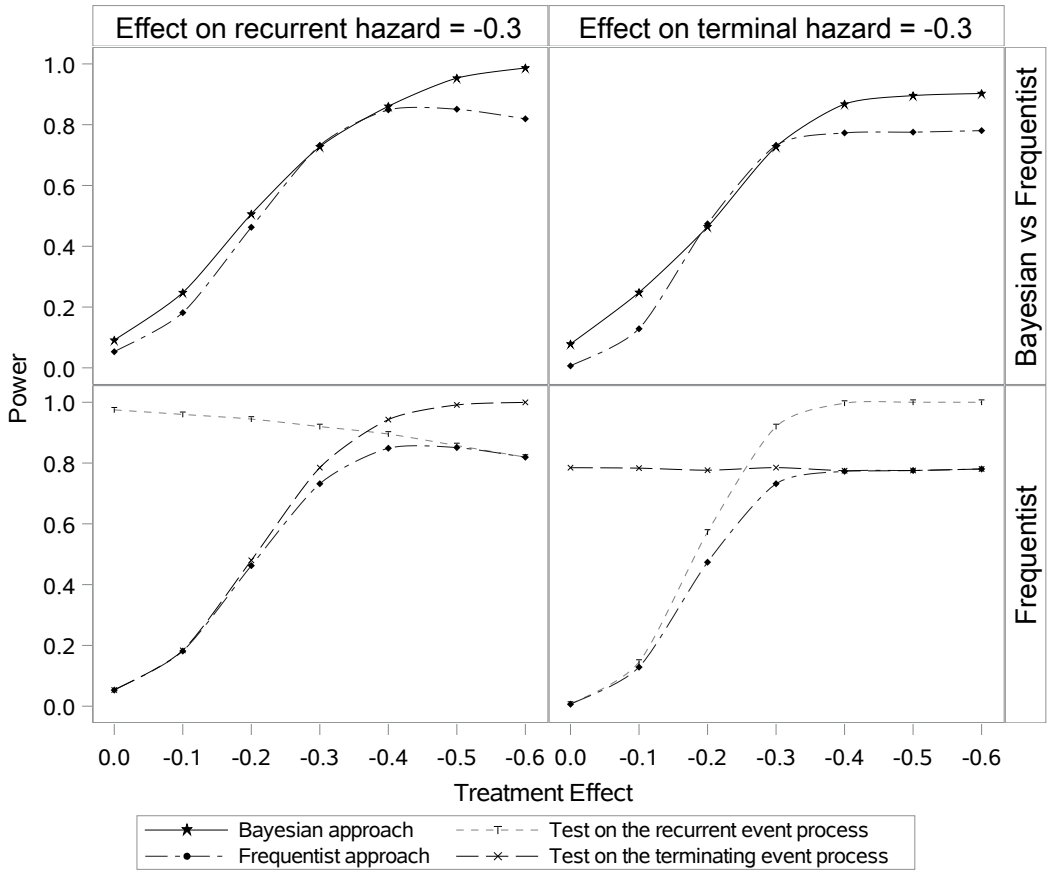

Figure 3: Bayesian estimated power curves.

We took the treatment effect on one event process equal to the hypothesized level and varied the effect on the other event process.

Results: SSD Example

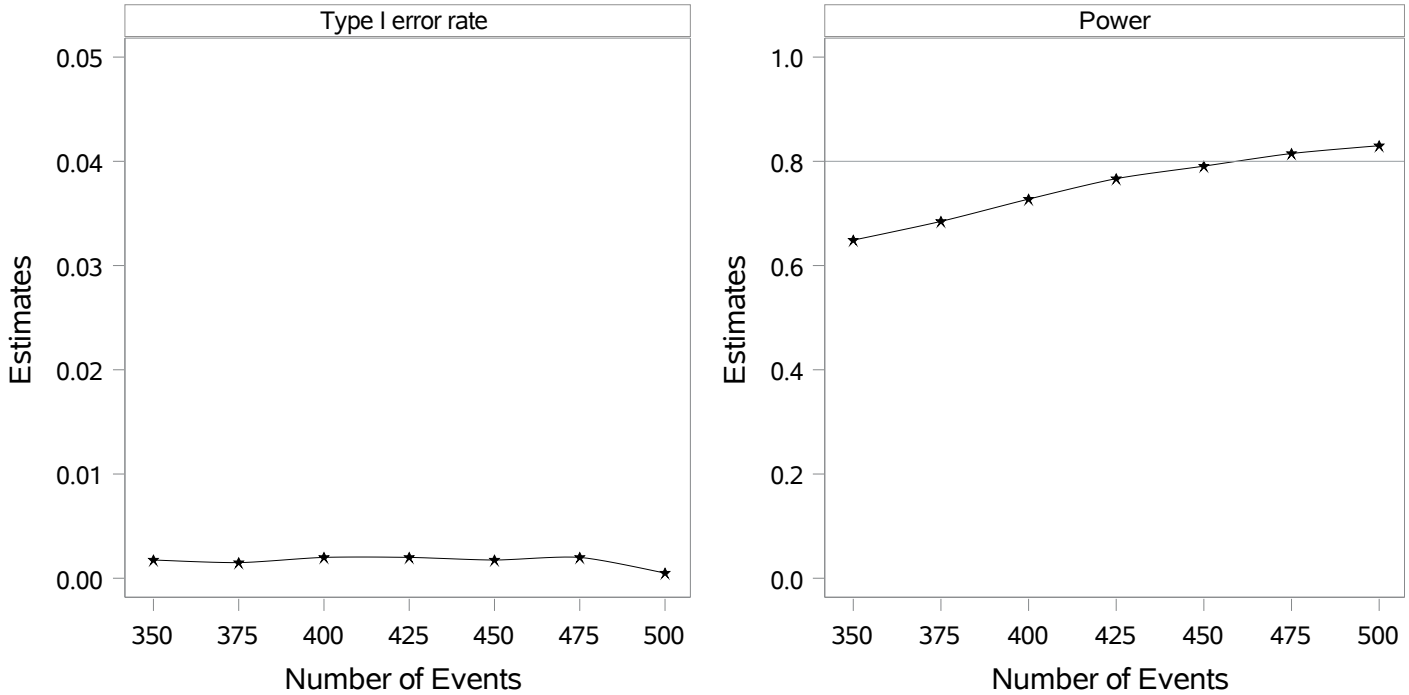

Figure 4: Bayesian Type I error rate and power curves based on varying numbers of required terminal events.

- We considered the number of terminating events $v = 350$ to 500 in increments of 25.
- The treatment effects on both event processes were assumed to equal the hypothesized level.

Discussion: Comparisons with Frequentist Approach

- For the colorectal cancer study, we compared the performance of the proposed approach to a frequentist approach that treated both events as co-primary.
- An alternative frequentist approach is to consider the events as multiple primary endpoints with an appropriate multiplicity correction. The hypothesis test becomes whether the IP shows a beneficial effect on either one of the event processes (regardless of whether there is a harmful effect on the other).
- The proposed Bayesian approach could be reformulated for this setting but an IP that provides benefit to one event process but harm on the other is not easily interpretable as beneficial.

Discussion: Use of Sampling Priors

- We used point-mass sampling priors based on parameter estimates from an analysis of the colorectal cancer data.
- The Bayesian framework for power and type I error evaluation is applicable for non-degenerate sampling priors on the parameters as well.
- For more extensive discussion on the use of non-degenerate sampling priors, we refer the reader to the work of Psioda and Ibrahim (2018, 2019) and the references cited therein.
- General advice for how to choose the point-mass sampling priors in the the joint model setting is also given in Xu et al. (2020).

Discussion: Choice of π

- We proposed $\pi_r = \pi_\lambda = \pi = \frac{1}{3}$ to specify the prior distributions for treatment effects which help to indirectly elicit the prior probabilities for the *Basis* and *Hypothesis* models.
- $\pi_r = \pi_\lambda = \pi = \frac{1}{3}$ was proposed as the default when there is little information suggesting a more appropriate choice for π_r and π_λ.
- The alternative space for the proposed approach includes parameter values consistent with IP benefit on (1) *both* event processes and (2) one and no effect on the other.
- Our choice of $\pi = \frac{1}{3}$ reflects a compromise between these two scenarios and reflects a priori uncertainty regarding on which event process the IP will have an effect in case (2).

Discussion: Choice of π

Table 3: Bayesian Type I error rate and power estimates with different choices of π

γ_r	γ_λ	$\pi = 0.5$	$\pi = 1/3$	$\pi = 0.05$
0	0	0.001	0.002	0.009
0	-0.3	0.133	0.079	0.057
0	-0.6	0.490	0.214	0.077
-0.3	0	0.175	0.091	0.054
-0.6	0	0.271	0.123	0.060
-0.3	-0.3	0.666	0.725	0.794

γ_r: Treatment effect on recurrent event process.
γ_λ: Treatment effect on terminating event process.

References I

Ming-Hui Chen. Importance-Weighted Marginal Bayesian Posterior Density Estimation. *Journal of the American Statistical Association*, 89(427):818–824, may 1994. ISSN 01621459. doi: 10.2307/2290907. URL http://www.jstor.org/stable/2290907.

Ming-Hui Chen and Qi-Man Shao. On Monte Carlo Methods for Estimating Ratios of Normalizing Constants. *The Annals of Statistics*, 25(4):1563–1594, may 1997. ISSN 00905364. URL http://www.jstor.org/stable/2959063.

References II

Ming-Hui Chen, Joseph G Ibrahim, Donglin Zeng, Kuolung Hu, and Catherine Jia. Bayesian design of superiority clinical trials for recurrent events data with applications to bleeding and transfusion events in myelodyplastic syndrome. *Biometrics*, 70 (4):1003–1013, dec 2014. ISSN 1541-0420. doi: 10.1111/biom.12215. URL https://www.ncbi.nlm.nih.gov/pubmed/25041037https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4276515/.

Juan Ramon González, Esteve Fernandez, Víctor Moreno, Josepa Ribes, Mercè Peris, Matilde Navarro, Maria Cambray, and Josep Maria Borràs. Sex differences in hospital readmission among colorectal cancer patients. *Journal of Epidemiology and Community Health*, 59(6):506 LP – 511, jun 2005. doi: 10.1136/jech.2004.028902. URL http://jech.bmj.com/content/59/6/506.abstract.

References III

Matthew A. Psioda and Joseph G. Ibrahim. Bayesian design of a survival trial with a cured fraction using historical data. *Statistics in Medicine*, 37(26):3814–3831, 2018.

Matthew A Psioda and Joseph G Ibrahim. Bayesian clinical trial design using historical data that inform the treatment effect. *Biostatistics*, 20(3):400–415, 2019. doi: 10.1093/biostatistics/kxy009.

Virginie Rondeau, Simone Mathoulin-Pelissier, Hélène Jacqmin-Gadda, Véronique Brouste, and Pierre Soubeyran. Joint Frailty Models for Recurring Events and Death Using Maximum Penalized Likelihood Estimation: Application on Cancer Events. *Biostatistics (Oxford, England)*, 8:708–721, nov 2007. doi: 10.1093/biostatistics/kxl043.

References IV

L J Wei, D Y Lin, and L Weissfeld. Regression Analysis of Multivariate Incomplete Failure Time Data by Modeling Marginal Distributions. *Journal of the American Statistical Association*, 84(408):1065–1073, may 1989. ISSN 01621459. doi: 10.2307/2290084. URL `http://www.jstor.org/stable/2290084`.

Jiawei Xu, Matthew A Psioda, and Joseph G Ibrahim. Bayesian design of clinical trials using joint models for longitudinal and time-to-event data. *Biostatistics*, nov 2020. ISSN 1465-4644. doi: 10.1093/biostatistics/kxaa044. URL `https://doi.org/10.1093/biostatistics/kxaa044`.

Posterior Model Probabilities (PMP's)

- Following Chen (1994), the weight function $w(\phi^{(-B_k)}|\phi^{(B_k)})$ is a completely known conditional density of $\phi^{(-B_k)}|\phi^{(B_k)}$, with the optimal choice of $w(\phi^{(-B_k)}|\phi^{(B_k)}) = p(\phi^{(-B_k)}|\phi^{(B_k)}, \mathbf{D})$.
- Since a closed form for $p(\phi^{(-B_k)}|\phi^{(B_k)}, \mathbf{D})$ is typically not available, Chen provided an empirical procedure to select $w(\phi^{(-B_k)}|\phi^{(B_k)})$.
- Specifically, we first compute the sample mean and covariance matrix $(\tilde{\phi}, \tilde{\Sigma})$ based on the MCMC samples $\{\phi_{(i)}, i = 1, ..., N\}$. Then $w(\phi^{(-B_k)}|\phi^{(B_k)})$ can be approximated using the conditional density of $\phi^{(-B_k)}|\phi^{(B_k)}$ based on a normal approximation.

Prior Model Probabilities

- For the standard deviation in $f_r(\cdot)$ and $f_\lambda(\cdot)$, we propose using $\sigma_r = \rho \cdot |\Delta_{1r}|$ and $\sigma_\lambda = \rho \cdot |\Delta_{1\lambda}|$ where Δ_{1r} and $\Delta_{1\lambda}$ are the hypothesized treatment effects on the recurrent and terminating event processes, respectively.

Table 4: Power estimates with different choices of ρ

ρ	1	2	3	4	5
Power	0.701	0.725	0.706	0.687	0.688

- We took $\rho = 2$ for modeling fitting because it allows the highest Bayesian power estimates under the default priors.
- The Bayesian power estimate is not sensitive to the choice of ρ.

Colorectal Cancer Data Analysis

Table 5: Nuisance parameter estimates for the colorectal cancer data.

Parameter Estimates			
Model	Description	Parameter	Posterior Mode
Frailty Models	Recurrent Event Frailty	η	0.046
	Association Parameter	θ	1.269
Recurrent Event Model	Gender	β_r	0.481
	Baseline Hazard	$\log r_1$	-5.314
		$\log r_2$	-5.951
		$\log r_3$	-6.715
		$\log r_4$	-6.872
		$\log r_5$	-7.259
Terminating Event Model	Gender	β_λ	0.260
	Baseline Hazard	$\log \lambda_1$	-8.064
		$\log \lambda_2$	-7.900
		$\log \lambda_3$	-8.299
		$\log \lambda_4$	-8.224
		$\log \lambda_5$	-8.493