

Novel Response-Adaptive Designs for Clinical Trials with Time-to-Event Outcomes

4th International Workshop in Sequential Methodologies
Alex Sverdlov¹

¹Acknowledgements/collaborators: Y. Ryznik (CS Ltd., Ukraine),
W. K. Wong (UCLA), W. F. Rosenberger (GMU), and Y. Tymofeyev (Janssen)
email: alex.sverdlov@novartis.com

July 18–21, 2013

Outline

Optimal response-adaptive randomization designs for multi-arm survival trials

- Exponential outcomes

- Weibull outcomes

- Statistical Software

Covariate-adjusted response-adaptive randomization for two-arm survival trials

Optimal dose-finding design problems for time-to-event outcomes

Response-Adaptive Randomization (RAR)

- ▶ **RAR** refers to sequential modification of treatment randomization probabilities based on accumulating data in the trial with the goal of assigning more patients to the better treatment while maintaining important statistical properties of the trial design¹.
- ▶ Modern research on RAR has been on development of **optimal RAR** designs for multi-arm and multi-objective clinical trials (phase II and III).
- ▶ Optimal RAR designs=balance between **individual ethics** (do the best for patients in the study) and **collective ethics** (do the best for future patients).

¹Hu F., Rosenberger, W.F. (2006). *The Theory of Response-Adaptive Randomization in Clinical Trials*. Wiley.

Motivation for RAR in Survival Trials

- ▶ **Statistical**: Censored heteroscedastic TTE outcomes \Rightarrow optimal allocation may be unbalanced across treatment arms.
- ▶ **Ethical**: Assign more patients to treatments that show benefit and minimize exposure of patients to inefficient treatments.
- ▶ **Logistical**: More patients may wish to participate in adaptive trials.

Three Steps to Develop Optimal RAR Designs²

1. Derive an optimal allocation to satisfy selected experimental objectives (e.g. minimize total expected hazard in the study subject to appropriate constraints on power of the test).
2. Construct a RAR procedure with minimal variability and high speed of convergence to the chosen optimal allocation.
3. Analyze clinical trial data following the chosen RAR procedure.

²Hu, F. and Rosenberger, W. F. (2003). Optimality, variability, power: evaluating response-adaptive randomization procedures for treatment comparisons. *Journal of the American Statistical Association* **98**, 671–678.

Optimal Allocation

- ▶ $K \geq 2$ treatment arms and TTE primary outcomes.
- ▶ n_k = sample size for treatment k , $n = \sum_{k=1}^K n_k$.
- ▶ Event times follow a parametric distribution with p.d.f. $f(t|\boldsymbol{\theta})$ and survivor function $S(t) = \int_t^\infty f(s|\boldsymbol{\theta})ds$.
- ▶ For the i th patient in group k , let
 - ▶ $T_{ik} > 0$ event time,
 - ▶ $C_{ik} > 0$ censoring time,
 - ▶ $t_{ik} = \min(T_{ik}, C_{ik})$ observed time,
 - ▶ $\delta_{ik} = \mathbf{1}\{T_{ik} \leq C_{ik}\}$ event indicator.
- ▶ The individual observations (t_{ik}, δ_{ik}) are independent for $i = 1, \dots, n_k$ and $k = 1, \dots, K$.

Likelihood and Fisher Information Matrix for θ

- ▶ Likelihood:

$$\mathcal{L}(\text{Data}|\theta) = \prod_{k=1}^K \prod_{i=1}^{n_k} \{f(t_{ik}|\theta)\}^{\delta_{ik}} \{S(t_{ik}|\theta)\}^{1-\delta_{ik}}.$$

- ▶ MLE of θ is found by solving the system of score equations

$$\frac{\partial}{\partial \theta} \log \mathcal{L}(\text{Data}|\theta) = \mathbf{0}.$$

- ▶ The Fisher information matrix for θ is

$$\mathbf{M}(\theta) = -\mathbb{E} \left\{ \frac{\partial^2}{\partial \theta \partial \theta^T} \log \mathcal{L}(\text{Data}|\theta) \right\},$$

whose inverse provides the lower bound on the variance of an unbiased estimator of θ .

Exponential Model

► Notations:

- $\boldsymbol{\theta} = (\theta_1, \dots, \theta_K)^T$: mean treatment survival times,
- $\varepsilon_k = E(\delta_{ik}) = \Pr(T_{ik} \leq C_{ik})$,
- $\rho_k =$ proportion for group k ($0 \leq \rho_k \leq 1$ and $\sum_{k=1}^K \rho_k = 1$),
- $\boldsymbol{\rho} = (\rho_1, \dots, \rho_K)^T$.

► Likelihood:

$$\mathcal{L}(\text{Data}|\boldsymbol{\theta}) = \prod_{k=1}^K \theta_k^{-\Delta_k} \exp(-\mathcal{T}_k/\theta_k),$$

where $\Delta_k = \sum_{i=1}^{n_k} \delta_{ik}$ and $\mathcal{T}_k = \sum_{i=1}^{n_k} t_{ik}$.

► The Fisher information matrix for $\boldsymbol{\theta}$ using design $\boldsymbol{\rho}$ is

$$\mathbf{M}(\boldsymbol{\rho}, \boldsymbol{\theta}) = n \cdot \text{diag} \left\{ \frac{\rho_1 \varepsilon_1}{\theta_1^2}, \dots, \frac{\rho_K \varepsilon_K}{\theta_K^2} \right\}.$$

Optimization Problems³

- ▶ Let $\boldsymbol{\theta}_c = \mathbf{A}^T \boldsymbol{\theta} = (\theta_2 - \theta_1, \dots, \theta_K - \theta_1)^T$ and $\boldsymbol{\Sigma}_n = \text{var}(\hat{\boldsymbol{\theta}}_c)$.
- ▶ D_A -optimal design: maximize $\log(\det\{\boldsymbol{\Sigma}_n\})$ s.t. $\sum_{k=1}^K \rho_k = 1$.
- ▶ Nonlinear programming optimal design:

$$\left\{ \begin{array}{ll} \text{minimize} & \sum_{k=1}^K w_k n_k \\ \text{subject to} & n_k / \sum_{j=1}^K n_j \geq B, \quad \sum_{k=1}^K n_k = n, \\ \text{and} & \boldsymbol{\theta}_c \boldsymbol{\Sigma}_n^{-1} \boldsymbol{\theta}_c \geq \eta, \end{array} \right.$$

where $\mathbf{w} = (w_1, \dots, w_K)^T$ and $B \in [0, 1/K]$ are user-defined, and $\eta > 0$ (optimal solution **will not** depend on η).

- ▶ $\mathbf{w} = (1, \dots, 1)^T \Rightarrow$ max.power of Wald test for a sample size n .
- ▶ $\mathbf{w} = (\theta_1^{-1}, \dots, \theta_K^{-1})^T \Rightarrow$ minimize mean total study hazard subject to power constraints.

³Sverdlov, O., Tymofyeyev, Y., Wong, W. K. (2011). Optimal response-adaptive randomized designs for multi-armed survival trials. *Statistics in Medicine* **30**, 2890–2910.

Weibull Model

- ▶ Assume $\log T_{ik} = \mu_k + bW_{ik}$, $W_{ik} \stackrel{\text{i.i.d.}}{\sim} f(w) = e^w \exp(-e^w)$.
- ▶ $\theta = (\mu_1, \dots, \mu_K, b)^\top$.
- ▶ Likelihood:

$$\mathcal{L}(\text{Data}|\theta) = \prod_{k=1}^K \prod_{i=1}^{n_k} \{b^{-1} e^{z_{ik}} \exp(-e^{z_{ik}})\}^{\delta_{ik}} \{\exp(-e^{z_{ik}})\}^{1-\delta_{ik}},$$

where $z_{ik} = (\log t_{ik} - \mu_k)/b$.

- ▶ The Fisher information for θ is

$$\mathbf{M}(\rho, \theta) = \frac{n}{b^2} \begin{pmatrix} \text{diag}\{\rho_1 \varepsilon_1, \dots, \rho_K \varepsilon_K\} & \mathbf{x} \\ \mathbf{x}^\top & \sum_{k=1}^K \rho_k (\varepsilon_k + c_k) \end{pmatrix},$$

- ▶ $\mathbf{x} = (\rho_1 a_1, \dots, \rho_K a_K)^\top$,
- ▶ $\varepsilon_k = \Pr(\delta_{ik} = 1)$, $a_k = E(z_{ik} e^{z_{ik}})$, $c_k = E(z_{ik}^2 e^{z_{ik}})$,
- ▶ ε_k , a_k and c_k are functions of θ and the censoring mechanism.

Optimization Problems for ($K = 2$)-arm Trials^{4,5}

- ▶ **Optimal**⁴ allocation minimizing mean total hazard in the study subject to a constraint on power.
- ▶ **D-optimal**⁵ allocation for most accurate estimation of $\theta = (\mu_1, \mu_2, b)^\top$.
- ▶ **D_A-optimal**⁵ allocation for most accurate estimation of $\Delta = \mu_2 - \mu_1$.
- ▶ **Hazard ratio-optimal**⁵ allocation for most accurate estimation of $\log(\text{HR}) = (\mu_2 - \mu_1)/b$.

⁴Zhang, L., Rosenberger, W. F. (2007). Response-adaptive randomization for survival trials: the parametric approach. *Applied Statistics* **56**, 153–165.

⁵Sverdlov, O., Ryznik, Y. and Wong, W. K. (2012). Doubly-adaptive biased coin designs for balancing competing objectives in time-to-event trials. *Statistics and Its Interface* **5(4)**, 401–413.

Optimization Problems for ($K > 2$)-arm Trials⁶

- ▶ **D-optimal** allocation: $\rho^* = \arg \min_{\rho} \{-\log(\det\{\mathbf{M}(\rho, \theta)\})\}$.
- ▶ **Compound optimal** allocation:
 - ▶ Let $\Phi_1(\rho)$ and $\Phi_2(\rho)$ be convex functionals for Objective 1 (D-optimality) and Objective 2 (hazard pattern estimation).
 - ▶ For a user-defined α ($0 \leq \alpha \leq 1$), solve:

$$\begin{cases} \text{minimize} & \alpha\Phi_1(\rho) + (1 - \alpha)\Phi_2(\rho) \\ \text{subject to} & \sum_{k=1}^K \rho_k = 1. \end{cases}$$

- ▶ **Weighted distance optimal** allocation:
 - ▶ Two objectives: **inferential**, implemented by ρ_I , and **ethical**, implemented by ρ_E .
 - ▶ $\lambda(\rho, \tilde{\rho})$ = distance metric between two probability vectors.
 - ▶ For a user-defined α ($0 \leq \alpha \leq 1$), determine

$$\rho^* = \arg \min_{\rho} \{\alpha\lambda(\rho, \rho_I) + (1 - \alpha)\lambda(\rho, \rho_E)\}.$$

⁶Sverdlov, O., Ryznik, Y. and Wong, W. K. (2013). Efficient and ethical response-adaptive randomization designs for multi-arm clinical trials with Weibull time-to-event outcomes. *Journal of Biopharmaceutical Statistics*, accepted.

Response-Adaptive Randomization

- ▶ Optimal allocation designs are functions of unknown θ and cannot be implemented directly.
- ▶ One can estimate θ sequentially using accumulating data from patients in the trial. This leads to the introduction of RAR into the trial design.
- ▶ In survival trials outcomes are naturally delayed.
 - ▶ RAR is applicable only when responses occur “not too far out” in the accrual pattern.
 - ▶ $\geq 60\%$ of study patients should contribute data throughout the recruitment phase for RAR to be meaningful.

Doubly Adaptive Biased Coin Design (DBCD)⁷

- ▶ DBCD is a RAR procedure that can be used to target a selected allocation $\boldsymbol{\rho} = (\rho_1(\boldsymbol{\theta}), \dots, \rho_K(\boldsymbol{\theta}))^T$.
- ▶ Under widely satisfied conditions, DBCD has established statistical properties:
 - ▶ The MLE $\hat{\boldsymbol{\theta}}_n$ is strongly consistent for $\boldsymbol{\theta}$ and is asymptotically normal.
 - ▶ The vector of allocation proportions $\mathbf{N}_n/n = (N_1(n)/n, \dots, N_K(n)/n)^T$ is strongly consistent for $\boldsymbol{\rho}$ and is asymptotically normal.
- ▶ Therefore, DBCD has similar asymptotic properties to fixed randomization designs, and standard asymptotic inference procedures should apply.

⁷Hu, F., Zhang, L. X. (2004). Asymptotic properties of doubly adaptive biased coin designs for multitreatment clinical trials. *The Annals of Statistics* **32**, 268–301.

Practical Considerations

1. How frequently should randomization probabilities be updated?
 - ▶ **Fully sequential procedure** utilizes all available data; however the study is completely unblinded.
 - ▶ **Two-stage design**: At Stage 1 randomize Km_0 patients equally among K treatments; as Stage 2 randomize remaining $(n - Km_0)$ patients using RAR based on data from Stage 1.
 - ▶ **Multi-stage design**: Recalculate randomization probabilities after cohorts of patients.
2. Which data should be looked at when recalculating randomization probabilities at an interim analysis (IA)?
 - ▶ Include data **only** from those patients whose outcomes have been observed before the time of IA. **Or**
 - ▶ Include data from all patients; some patients will be censored by the time of IA.

Extensive Simulation Studies: Summary

- ▶ **Convergence to the target:** Optimal RAR designs work as intended when $\geq 60\%$ of data are observed throughout the recruitment stage.
 - ▶ Allocation is generally skewed towards more variable arms with longer survival times.
 - ▶ RAR designs have higher variability of allocation proportions than completely randomized design (CRD).
- ▶ **Statistical characteristics:** Both parametric (Wald) and log-rank tests were studied.
 - ▶ Simulated type I error rate is very close to the nominal; a 1% inflation is observed in small and moderate samples ($n = 150$), both for RAR and CRD.
 - ▶ Optimal RAR designs have similar or higher (1% – 2% and up to 4%) power compared to CRD.
- ▶ **Ethical characteristics:** RAR designs result in higher average total survival time and have **modest but consistent** reductions in the number of deaths compared to CRD.

Example: Redesigning a Phase III Survival Trial⁸

- ▶ A randomized phase III clinical trial in patients with locally advanced head and neck cancer treated with standard fractionated RT alone (Treatment 1) or RT+cisplatin (Treatment 2) or RT+carboplatin (Treatment 3).
- ▶ The reported intent-to-treat median overall survival times for treatment groups 1, 2, and 3 are 12.2, 48.6, and 24.5 months.
- ▶ Assume Weibull event times and consider 3 choices of $\theta = (\mu_1, \mu_2, \mu_3, b)^T$ (Scenarios A, B, and C) to match the reported treatment effects:

Scenario	μ_1	μ_2	μ_3	b	Hazard Ratio	
					2 vs. 1	3 vs. 1
A	2.81	4.20	3.51	0.85	0.20	0.44
B	2.87	4.25	3.57	1.00	0.25	0.50
C	2.96	4.34	3.66	1.25	0.33	0.57

⁸Fountzilas, G., Ciuleanu, E., Dafni, U., et al. (2004). Concomitant radiochemotherapy vs radiotherapy alone in patients with head and neck cancer. *Medical Oncology* 21(2), 95–107.

Simulation Study Setup

- ▶ To match the reported trial, assume:
 - ▶ Recruitment $R = 55$ months; study duration $D = 96$ months.
 - ▶ Patient enrollment follows a Poisson process over $(0, R)$.
 - ▶ Patient drop-out time is $\text{Uniform}(0, D)$; patients who are alive and have not dropped out by the end of the study are administratively censored.
- ▶ Target allocation $\boldsymbol{\rho}^* = (\rho_1^*, \rho_2^*, \rho_3^*)^\top$

$$\rho_k^* = 0.5\rho_{Ik} + 0.5\rho_{Ek}, \quad k = 1, 2, 3,$$

where $(\rho_{I1}, \rho_{I2}, \rho_{I3})^\top$ is D-optimal and $(\rho_{E1}, \rho_{E2}, \rho_{E3})^\top$ is “ethical” allocation:

$$\rho_{Ek} = \frac{\{\exp(\mu_k/b)\}^2}{\sum_{j=1}^3 \{\exp(\mu_j/b)\}^2}, \quad k = 1, 2, 3. \quad (1)$$

- ▶ DBCD with two interim analyses (after $n/3$ and after $2n/3$ patients) is used to target $\boldsymbol{\rho}^*$. CRD is simulated as the reference procedure.

Simulation Study Results (10,000 simulation runs)

	Scenario A ($n = 66$)		Scenario B ($n = 72$)		Scenario C ($n = 123$)	
	CRD	RAR*	CRD	RAR†	CRD	RAR§
ρ_1	0.335 (0.041)	0.281 (0.083)	0.334 (0.032)	0.264 (0.078)	0.334 (0.025)	0.265 (0.072)
ρ_2	0.331 (0.041)	0.390 (0.117)	0.332 (0.032)	0.409 (0.099)	0.333 (0.026)	0.404 (0.080)
ρ_3	0.334 (0.041)	0.330 (0.103)	0.334 (0.032)	0.328 (0.093)	0.333 (0.025)	0.331 (0.080)
$M(D)$	0.995	0.947	0.997	0.944	0.998	0.953
$M(D_A)$	0.999	0.915	0.999	0.930	0.999	0.945
Power	0.916	0.917	0.903	0.901	0.883	0.881
TD (SD)	39 (4)	37 (4)	50 (4)	48 (5)	75 (5)	71 (6)
TH (SD)	305 (107)	281 (104)	283 (70)	258 (68)	295 (49)	274 (48)
TT (SD)	1331 (178)	1375 (153)	1797 (167)	1874 (185)	2757 (220)	2841 (241)

* Target allocation is (0.203, 0.551, 0.246).
 † Target allocation is (0.211, 0.526, 0.262).
 § Target allocation is (0.224, 0.495, 0.281).
 $M(D)$, median D -efficiency; $M(D_A)$, median D_A -efficiency;
 TD, total number of deaths; TH, total hazard; TT, total time; SD, standard deviation.

- ▶ RAR generates skewed allocations favoring the arms with longer survival times but it has more variable allocation proportions than CRD.
- ▶ Due to delayed responses, RAR does not completely attain its targets.
- ▶ RAR has $\geq 94\%$ median D -efficiency, $\geq 92\%$ median D_A -efficiency, and the same average power as CRD.
- ▶ RAR has **7% to 9% lower** average total hazard, **2 to 4 fewer** average deaths and longer average total survival time than CRD.

Statistical Software⁹

- ▶ A user-friendly software interface **RARtool** was developed in MATLAB, to facilitate the design of randomized comparative clinical trials with time-to-event outcomes.
 - ▶ It implements state-of-the-art RAR methodology from several recently published papers.
 - ▶ It can compute optimal allocation designs and values of different statistical efficiency criteria for user-selected sets of experimental parameters.
 - ▶ It can perform Monte-Carlo simulations of RAR procedures targeting selected optimal allocations under a variety of scenarios.
- ▶ **RARtool** is intended to fill the gap between methodology and implementation of optimal RAR designs in time-to-event trials.

⁹Ryeznic, Y., Sverdlov, O., and Wong, W. K. (2013). RARtool—a MATLAB software package for designing response-adaptive randomized clinical trials with time-to-event outcomes, submitted.

Covariate-Adjusted Response-Adaptive (CARA) Randomization

- ▶ **CARA** randomization is an extension of RAR: treatment randomization probabilities are modified based on history of previous patients' treatment assignments, responses and covariates, and the covariate vector of the current patient¹⁰.
- ▶ Two main reasons to consider CARA:
 - ▶ **Statistical**: For nonlinear and heteroscedastic models, optimal allocation may not be balanced across treatment arms.
 - ▶ **Ethical**: The degree and direction of treatment effect may differ for patient subgroups within a treatment \Rightarrow increase probability of assigning the treatment that is most efficacious given the patient's covariate profile (**personalized treatment**).

¹⁰Rosenberger, W. F. and Sverdlov, O. (2008). Handling covariates in the design of clinical trials. *Statistical Science* **23**(3), 404–419

CARA Randomization for Survival Trials¹¹

- ▶ n patients enroll sequentially and are randomized to A or B.
- ▶ For the i th patient, survival time T_{ik} , conditional on covariates \mathbf{z}_i is exponential with mean

$$\lambda_k(\mathbf{z}_i) = \exp(\boldsymbol{\theta}_k^T \mathbf{z}_i), \quad k = A, B,$$

where $\boldsymbol{\theta}_k = (\theta_{k0}, \theta_{k1}, \dots, \theta_{kp})^T$ and $\mathbf{z}_i = (1, z_{1i}, \dots, z_{pi})^T$.

- ▶ T_{ik} is subject to independent right-censoring with $C_i > 0$.
 - ▶ $t_{ik} = \min(T_{ik}, C_i)$,
 - ▶ $\delta_{ik} = \mathbf{1}\{t_{ik} = T_{ik}\}$,
 - ▶ (t_{ik}, δ_{ik}) are independent, $i = 1, \dots, n_k$, $k = A, B$,
 - ▶ $\varepsilon_k(\mathbf{z}_i) = E(\delta_{ik}) = \Pr(T_{ik} \leq C_i | \boldsymbol{\theta}_k, \mathbf{z}_i)$.

¹¹Sverdlov O., Rosenberger W. F., and Ryznik Y. (2013). Utility of covariate-adjusted response-adaptive randomization in survival trials. *Statistics in Biopharmaceutical Research* **5**(1), 38–53.

CARA Randomization Designs

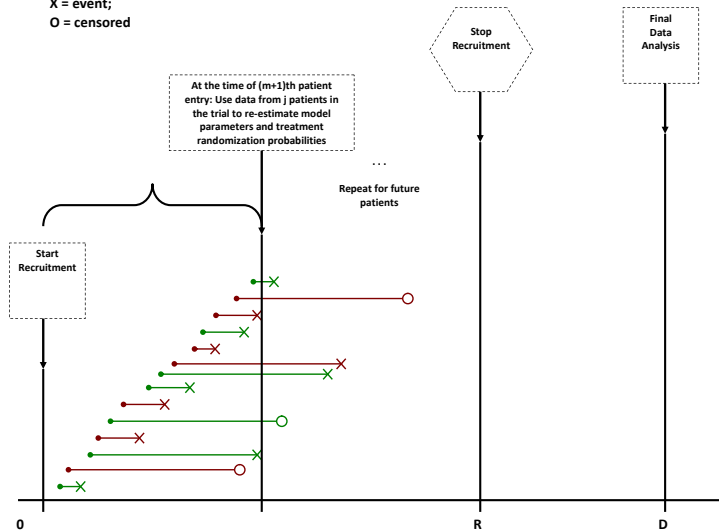
- ▶ Randomize initial $2m_0$ patients (m_0 is small positive integer) equally between A and B.
- ▶ Suppose $m \geq 2m_0$ patients have been randomized and outcome data (from some of them) are available. Compute $(\hat{\theta}_{A,m}, \hat{\theta}_{B,m})$, estimates of (θ_A, θ_B) .
- ▶ The $(m + 1)$ th patient with covariate vector \mathbf{z}_{m+1} is randomized to A with probability

$$\phi_{m+1} = g(\hat{\theta}_{A,m}, \hat{\theta}_{B,m}, \mathbf{z}_{m+1}),$$

where $0 \leq g(\cdot) \leq 1$ is an appropriately chosen allocation function, skewed in favor of the “better” treatment arm.

CARA Randomization Schematic

X = event;
O = censored



Proposed Designs

1 CARA designs with targets:

- ▶ Derive an optimal allocation (OA) for a model without covariates, and use a covariate-adjusted version of the OA as the target.
- ▶ Use CARA DBCD procedure¹² to sequentially allocate patients.

2 Weighted Optimality CARA designs:

- ▶ Treatment randomization probabilities for a patient are obtained by maximizing a utility function that combines “inferential” (D-optimality) and “ethical” criteria¹³.
- ▶ Pre-specified tradeoff parameter $\gamma \in [0, \infty)$ allows to achieve balance between the objectives: $\gamma = 0$ is “maximum information” design; $\gamma \rightarrow \infty$ is “most ethical” design.

¹²Zhang, L-X. and Hu, F. (2009). A new family of covariate-adjusted response-adaptive designs and their properties. *Applied Mathematics—A Journal of Chinese Universities*, **24** 1–13.

¹³Atkinson, A. C. and Biswas, A. (2005). Bayesian adaptive biased-coin designs for clinical trials with normal responses. *Biometrics*, **61** 118–125.

Extensive Simulation Studies

- ▶ Ten competing designs were compared:
 - ▶ 2 balanced randomization procedures (CRD and Pocock-Simon's method).
 - ▶ 6 CARA randomization designs.
 - ▶ 2 RAR designs (only main treatment effects are estimated in the design).
- ▶ Operating characteristics:
 - ▶ Allocation proportion $N_A(n)/n$ and its variability.
 - ▶ Power and type I error.
 - ▶ Relative efficiency in estimation.
 - ▶ Total number of events in the trial.
- ▶ Both correctly specified exponential model and uniform recruitment and various violations of these assumptions.

Key Findings from Simulations

- Allocation proportion**
- ▷ Under $H_0 : \theta_A = \theta_B$, all designs result in equal allocation.
 - ▷ Under Alternatives, CARA and RAR designs result in skewed allocation to the better treatment.
 - ▷ CARA and RAR are more variable than balanced designs.
 - ▷ $N_A(n)/n$ is normally distributed (consistent with theory).
- Type I error and power**
- ▷ For a range of sample sizes from 100 to 400, the type I error for all designs ranges from 0.048 to 0.064.
 - ▷ CARA, RAR, and balanced designs have very similar power.
- Estimation efficiency**
- ▷ CARA and RAR designs are at least 96% as efficient as balanced designs.
 - ▷ All designs have consistent and normally distributed M.L.E's.
- Number of events**
- ▷ CARA and RAR designs result in up to 3–6 fewer events (on average) compared to balanced designs.

Robustness to Model Misspecification

- ▶ CARA designs are robust to misspecification of the exponential model (e.g. when event times follow Weibull, log-logistic, or log-normal distribution), **provided that final data are analyzed using the correctly specified model.**
- ▶ If the final model is misspecified, estimates are biased, type I error may be inflated and power may be lost; yet this is common to both CARA and balanced randomization designs.
- ▶ Misspecification of the recruitment pattern (e.g. recruitment times are not uniform) has little impact on statistical properties of CARA designs.

Example: Redesigning a Survival Trial

- ▶ Karapetis et al.¹⁴ reported the results of cetuximab trial in advanced colorectal cancer.
- ▶ In a 21-month period, $n = 572$ eligible patients were randomized at a 1 : 1 ratio among TRT A (cetuximab plus best supportive care) and TRT B (best supportive care alone).
- ▶ The primary endpoint was overall survival (OS).
- ▶ Effectiveness of cetuximab was significantly associated with *K-ras* mutation status:
 - ▶ Patients with **wild-type *K-ras* tumors** benefited from cetuximab (median OS, 9.5 vs. 4.8 months; HR for death, 0.55).
 - ▶ Patients with a colorectal tumor bearing **mutated *K-ras*** did not benefit from cetuximab (median OS, 4.6 vs. 4.5 months; HR for death, 0.98).

¹⁴Karapetis, C. S., Khambata-Ford, S., Jonker, D. J., et al. (2008). *K-ras* mutations and benefit from cetuximab in advanced colorectal cancer. *The New England Journal of Medicine*, **359** 1757–1765.

Simulation Study Results (10,000 simulation runs)

$n = 572$	Pocock-Simon	CARA	RAR
N_A/n (S.D.)	0.500 (0.002)	0.588 (0.037)	0.583 (0.039)
$N_{A0} N_{B0}$ (S.D.)	169 169 (1)	211 127 (16)	197 141 (14)
$N_{A1} N_{B1}$ (S.D.)	117 117 (1)	125 109 (13)	137 97 (11)
Deaths (S.D.)	372 (11)	362 (12)	366 (12)
Total Time (S.D.)	3076 (106)	3155 (113)	3132 (112)
$\hat{\theta}_{A0}$ (S.D.)	2.62 (0.11)	2.62 (0.10)	2.62 (0.10)
$\hat{\theta}_{A1}$ (S.D.)	-0.68 (0.16)	-0.68 (0.15)	-0.68 (0.15)
$\hat{\theta}_{B0}$ (S.D.)	1.87 (0.09)	1.86 (0.11)	1.87 (0.10)
$\hat{\theta}_{B1}$ (S.D.)	0.02 (0.14)	0.03 (0.16)	0.02 (0.16)

- ▶ Because of treatment-covariate interaction, CARA resulted in greater skewing to A in the wild-type *K-ras* subgroup than in the mutated *K-ras* subgroup, whereas RAR had similar degree of skewing in the subgroups.
- ▶ CARA and RAR had, on average, 10 and 6 fewer deaths and greater total survival time than the Pocock-Simon design.
- ▶ All three procedures had the same power and very similar M.L.E.'s.

CARA Randomization: Conclusions

- ▶ CARA designs generate skewed allocations according to covariate-specific treatment differences and can result in fewer events in the trial, while having similar statistical properties (type I error/power/estimation efficiency) to balanced randomization designs.
- ▶ **Exponential regression** is a reasonable **working model** to facilitate design adaptations. However, it is crucial that final data are analyzed using correctly specified model.
- ▶ Delayed responses slow down convergence to target allocation ⇒ substantial amount of patient outcome data must be observed during the recruitment phase.
- ▶ The number of covariates in the model at the design stage must be limited (e.g. most predictive genetic biomarkers).

Optimal Dose Finding in TTE trials

- ▶ Potential applications to medical studies in
 - ▶ **Virology** (duration of viral shedding);
 - ▶ **Dentistry** (time to onset and/or duration of anesthesia);
 - ▶ **Oncology** (progression-free survival, overall survival).
- ▶ Assume a quadratic dose-response relationship:

$$\log T = \beta_0 + \beta_1 x + \beta_2 x^2 + bW,$$

where $x = \log(\text{Dose})$, $W \sim f(w)$ with support on $(-\infty, \infty)$ and $S(w) = \int_w^\infty f(u)du$.

- ▶ $f(w) = \exp(w - e^w) \Rightarrow T \sim \text{Weibull}$;
 - ▶ $f(w) = e^{-w}(1 + e^{-w})^{-2} \Rightarrow T \sim \text{loglogistic}$;
 - ▶ $f(w) = \frac{1}{\sqrt{2\pi}}e^{-w^2/2} \Rightarrow T \sim \text{lognormal}$.
- ▶ Data structure: $\{(t_i, \delta_i, x_i), i = 1, \dots, n\}$, where $t_i = \min(T_i, C_i)$, $\delta_i = \mathbf{1}\{T_i \leq C_i\}$ and $x_i = \log(\text{Dose}_i)$.

Likelihood and Fisher Information

- ▶ Likelihood:

$$\mathcal{L}(\text{Data}|\boldsymbol{\theta}) = \prod_{i=1}^n \{b^{-1}f(w_i)\}^{\delta_i} \{S(w_i)\}^{1-\delta_i},$$

where $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, b)^\top$ and $w_i = (\log t_i - \boldsymbol{\beta}^\top \mathbf{x}_i)/b$.

- ▶ Fisher information for $\boldsymbol{\theta}$ at \mathbf{x}_i is

$$\mathbf{I}(\mathbf{x}_i, \boldsymbol{\theta}) = \frac{1}{b^2} \begin{pmatrix} -E(A_i)\mathbf{x}_i\mathbf{x}_i^\top & -E(A_i w_i)\mathbf{x}_i \\ -E(A_i w_i)\mathbf{x}_i^\top & E(\delta_i) - E(A_i w_i^2)\mathbf{x}_i \end{pmatrix},$$

where $A_i = \delta_i \frac{\partial^2 \log f(w_i)}{\partial w_i^2} + (1 - \delta_i) \frac{\partial^2 \log S(w_i)}{\partial w_i^2}$.

- ▶ Design: $\xi = \{(x_i, \rho_i), 0 \leq \rho_i \leq 1, \sum_{i=1}^K \rho_i = 1\}$.
- ▶ Full design information matrix:

$$\mathbf{M}(\xi, \boldsymbol{\theta}) = n \sum_{i=1}^K \rho_i \mathbf{I}(\mathbf{x}_i, \boldsymbol{\theta}).$$

Optimal Design Problems (Ongoing Work)

- ▶ **Locally D-optimal** design: $\xi_D = \arg \min_{\xi} \{-\log |\mathbf{M}(\xi, \boldsymbol{\theta})|\}$.
- ▶ **Locally c-optimal** design: minimize variance of the MLE of $d = -\beta_1/2\beta_2$.
- ▶ **Bayesian D-optimal** design:
 $\xi_D^* = \arg \min_{\xi} \{-\int \log |\mathbf{M}(\xi, \boldsymbol{\theta})| g(\boldsymbol{\theta}) d\boldsymbol{\theta}\}$, where $g(\boldsymbol{\theta})$ is a prior density of $\boldsymbol{\theta}$.
- ▶ **Bayesian c-optimal** design.
- ▶ **Penalized D-optimal** design¹⁵.
- ▶ If there is no censoring, then $\mathbf{M}(\xi, \boldsymbol{\theta})$ does not depend on $(\beta_0, \beta_1, \beta_2)$. With censoring, optimal designs have complex structure and must be found using numerical optimization.

¹⁵ Dragalin, V. and Fedorov, V. (2006). Adaptive designs for dose-finding based on efficacy-toxicity response. *Journal of Statistical Planning and Inference* **136** 1800–1823.

Adaptive Designs (Ongoing Work)

- ▶ **Pilot design:** Initial m patients are allocated to dose levels according to a uniform allocation design.
- ▶ Given data $\mathcal{F}_m = \{(x_1, t_1, \delta_1), \dots, (x_m, t_m, \delta_m)\}$, obtain the estimate $\hat{\theta}_m$ and $\xi_m = \xi(\hat{\theta}_m)$.
- ▶ Stepwise allocation of the remaining $(n - m)$ patients is made to maximize an appropriate sensitivity function.
- ▶ **Points to be addressed:**
 - ▶ A comparison of optimal and adaptive designs with uniform design in terms of statistical efficiency and ethical criteria.
 - ▶ Robustness of the designs to misspecification of event time distribution.
 - ▶ Effect of delayed response and recruitment patterns on statistical properties of adaptive designs.

Thank You!

Questions?