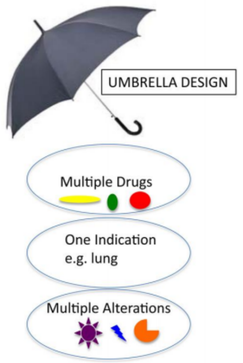
Robust Blocked Response-Adaptive Randomization (RAR) Designs

Thevaa Chandereng[1*], Rick Chappell[1,2]

[1]Dept. of Statistics, University of Wisconsin, [2]Dept. of Biostatistics & Medical Informatics, University of Wisconsin

[*]1300 University Ave, Madison, WI 53706 Email: **chandereng@wisc.edu**

A Motivating Trial

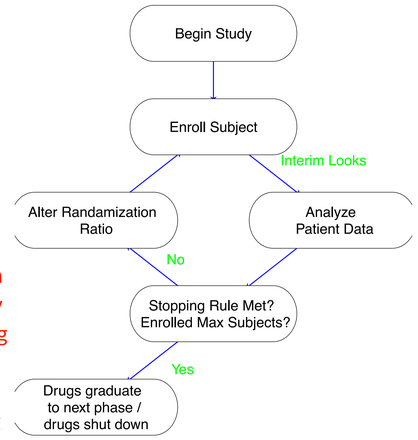


Biomarker-integrated Approaches of Targeted Therapy for Lung Cancer Elimination (BATTLE - 1)

- an umbrella design;
- 4 treatments, 5 biomarkers;
- primary end point, 8-week disease control rate;
- treat more patients in promising groups based on biomarker profile;
- suspend ineffective groups early

Design of BATTLE-1 Trial

- 1st interim look: at least one patient in each treatment and biomarker groups
- ? All subsequent interim looks: patient by patient



Conclusion

- "More smokers enrolled in the latter part of the study compared to the beginning of the study".
- Time trends are nearly universally ignored among RAR proponents.

Time-trends and How it affect RAR?

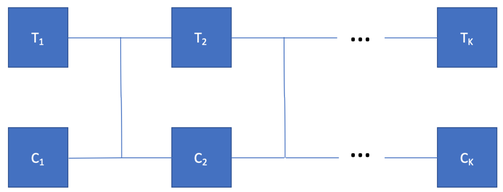
- Thall et al. (2015) investigated type-I error under a linear time-trend induced in the traditional RAR design and showed that the type-I error is significantly above the nominal level.

Traditional RAR

- fixed allocation for burn-in period (first *k* patients)
- alter randomization ratio per patient

Blocked RAR Designs

- patients are assigned in blocks and the randomization ratio is recomputed for blocks
- the final analysis is stratified by blocks



Research Objectives

To provide a flexible and practical approach for RAR designs:

- Provide a design that control the type I error under time-trends
- Compare the operating characteristics for different number of blocks
- Bayesian and frequentist approaches
- Reduce the non-trivial risk of assigning more subjects to the inferior treatment

Methods

Approach	Frequentist	Bayesian
Algorithm	Modified RPW rule	BAR(n/2N)
Non-stratified analysis	Chi-square test	Beta-binomial conjugate prior
Stratified analysis	CMH test	Binary regression
Early stopping rules	Alpha-spending	success: $P_{B>A}(data) > .99$, failure: $P_{B>A}(data) < .01$
Yates correction	Not applied	-

Modified RPW rule

The probability of randomizing subjects to treatment A in block j, $\pi_{j,A}$ is defined as

$$\pi_{j,A} = \frac{\frac{y_{A,j-1}+1}{N_{A,j-1}+2}}{\frac{y_{A,j-1}+1}{N_{A,j-1}+2} + \frac{y_{B,j-1}+1}{N_{B,j-1}+2}},$$

where $y_{A,j-1}$ and $N_{A,j-1}$ are the numbers of events and subjects in treatment A and $y_{B,j-1}$ and $N_{B,j-1}$ are the numbers of events and subjects in treatment B after block j − 1.

BAR(n/2N)

$$\pi_{j,A} = \frac{(p_{A>B,j-1}(data))^{n/2N}}{(p_{A>B,j-1}(data))^{n/2N} + (p_{B>A,j-1}(data))^{n/2N}},$$

where $p_{A>B,j-1}$(data) is the posterior probability that treatment A has a higher success rate than treatment B, $p_{A>B,j-1}$(data)) = 1 − $p_{B>A,j-1}$(data)) after the j − 1st block, n is the number of accrued patients and N is the maximum sample size.

Simulation Setup

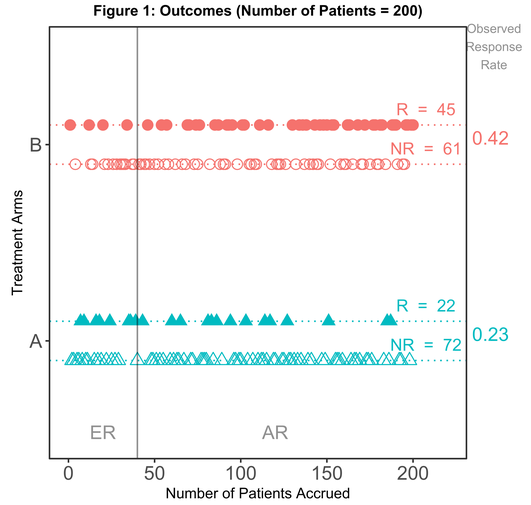
- N = 200 subjects
- K = 1, 200 (non-stratified analysis)
- K = 2, 4, 5, 10, 20, 100 (stratified analysis)
- ?10,000 independent simulations
- p_A is set to 0.25, p_B is set to 0.25 (null case), 0.45 (alternative case)
- Burn-in period of 40 patients (fixed randomization)
- Allocation probability bounded between 0.2 and 0.8
- To examine the effects of time-trends, we increased both p_A and p_B linearly from their initial values to a final value of 0.25 larger

An Illustration: Operating Characteristics

p_B	No. of Block(s)	Power	Bias	$E(N)$	π_{20}	$N_B - N_A$
0.25	1	0.06	0.00	176.73	0.06	0 (-28, 26)
	2	0.06	0.00	179.09	0.15	0.29 (-44, 44)
	4	0.06	0.00	184.24	0.23	0.06 (-60, 58)
	5	0.06	0.00	186.01	0.25	-0.3 (-64, 62)
	10	0.07	0.00	195.50	0.29	-0.83 (-72, 70)
	20	0.07	0.01	200.00	0.31	-0.54 (-76, 74)
	100	0.01	0.00	200.00	0.33	-0.87 (-80, 78)
	200	0.16	0.00	174.68	0.25	0.12 (-60, 60)
0.45	1	0.91	0.07	97.05	0.02	0.02 (-20, 21)
	2	0.90	0.07	112.34	0.01	11.07 (-14, 56)
	4	0.88	0.06	142.83	0.00	32.7 (-9, 82)
	5	0.88	0.06	151.83	0.00	39.78 (-8, 88)
	10	0.87	0.03	183.78	0.00	58.94 (-4, 100)
	20	0.86	0.01	200.00	0.00	69.6 (18, 106)
	100	0.33	-0.14	200.00	0.00	71.83 (16, 108)
	200	0.94	0.07	94.87	0.00	16.04 (-8, 64)

RAR using Bayesian approach with early stopping criteria and with 0.25 drift. p_A is set to 0.25 for all cases. E(N) represents the mean sample size. Bias = δ−δ. π_{20} = P(NA−NB > 20), N_A and N_B denotes the number of patients assigned to treatment A and B. The mean (2.5%, 97.5%) of $N_B - N_A$ is reported in the last column. 10,000 simulations were done for each case.

Figure 1: Outcomes (Number of Patients = 200)



A single simulation for RAR with p_A = 0.25, p_B = 0.45 with N of 200 and 5 number of blocks (frequentist design).

Discussion and Summary

- Blocks with fewer patients can reduce the power of the trial because if patients are not randomized to both treatments in a block, the block becomes uninformative.
- Large number of blocks should be clearly avoided for both ethical reason and poor design.
- Small number of blocks (K = 2, 4 and 5) has a good tradeoff between efficiency and ethically treating patients to the best known superior treatment.
- RAR proponents cannot ignore time-trends.
- Time-trend can significantly impact the type-I error rate.
- ?R package blockRAR
- Choice of algorithm: low probability of assigning more subjects to the inferior arm and moderate scheme to alter randomization ratio
- RAR carries a non-trivial risk of creating a large sample size imbalance in favor of the inferior treatment.