SUPPLEMENTARY MATERIAL for Three-Outcome Dual-Criterion Randomized Phase II Clinical Trial Design

Yujia Wang[1,*], Xiaohan Chi[2,*], Ruitao Lin[2]

[1]Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, U.S.A.
[2]Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, U.S.A.
*Corresponding authors: ywang74@mdanderson.org and xchi@mdanderson.org

1 3-by-2 design

The TDR design can be extended to the 3-by-2 decision boundaries to allow finer control of inconclusive regions. The difference between a 3-by-2 design and a 2-by-2 design is that two difference boundaries, a lower boundary r and an upper boundary s, are employed to define an inconclusive region where the difference in treatment and control responses is borderline. Therefore, the definitions of α, β, η, and γ change accordingly as below:

$$\begin{aligned}
\alpha &= \sum_{D_1 \cap D_3} B(y_E, n_E, p_E \mid H_0) B(y_C, n_C, p_C \mid H_0), \\
\beta &= \sum_{(D_1' \cap D_3') \cup D_2} B(y_E, n_E, p_E \mid H_a) B(y_C, n_C, p_C \mid H_a), \\
\eta &= \sum_{((D_1' \cap D_2') \cap D_3) \cup (D_1 \cap D_3')} B(y_E, n_E, p_E \mid H_0) B(y_C, n_C, p_C \mid H_0), \\
\gamma &= \sum_{((D_1' \cap D_2') \cap D_3) \cup (D_1 \cap D_3')} B(y_E, n_E, p_E \mid H_a) B(y_C, n_C, p_C \mid H_a),
\end{aligned} \tag{1}$$

where $D_1 = \{(y_E, y_C) : y_E - y_C \geq s\}, D_2 = \{(y_E, y_C) : y_E - y_C \leq r\}, D_3 = \{y_E : y_E \geq m\}$. The definition for clinical significance boundary m remains unchanged. Note that as the difference boundaries change, m may deviate from what is observed in a 2-by-2 design.

2 Tables and figures

Table 1: Determination of γ_{max} and λ_{max} using loss function for TDR one-stage 2-by-2 design at $p_C = 0.35, p_E = 0.55, \alpha \leq 0.20, \beta \leq 0.20$

π	N	γ_{max}	λ_{max}	w_{40}	w_{50}	w_{60}
0.8	56	0.01	0.05	0.44025	0.44025	0.45621
0.79	56	0.02	0.1	0.44246	0.44246	0.45807
0.8	46	0.03	0.1	0.41192	0.41192	0.42038
0.78	42	0.05	0.15	0.40367	0.40367	0.40872
0.81	40	0.07	0.15	0.39281	0.39281	0.39727
0.77	38	0.1	0.2	0.39512	0.39512	0.39685

γ_{max}: design constraint for γ; λ_{max}: design constraint for λ; N: total sample size; π: power; w: weight parameter for loss function.

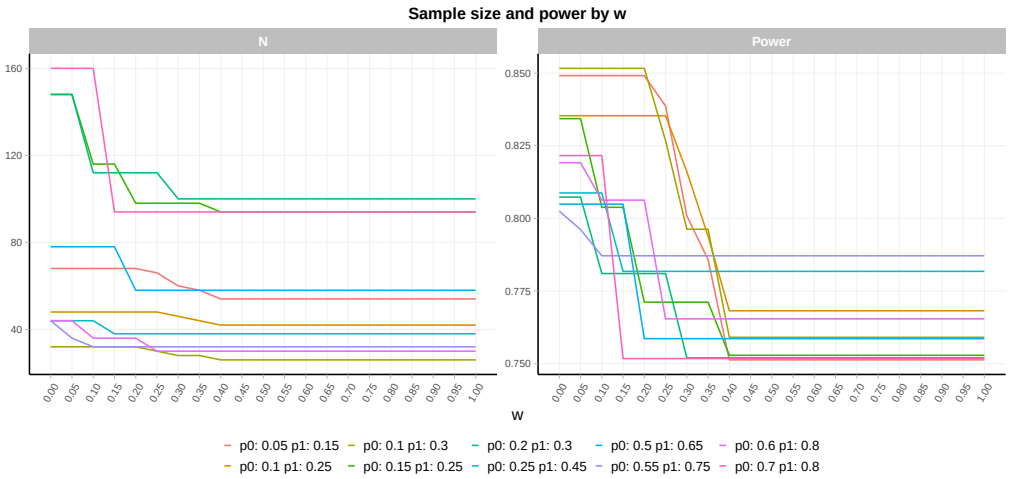

Figure 1: Sensitivity analysis of loss function parameter w. Ten cases from the one-stage 2-by-2 TDR design were randomly selected to compute loss scores across w. The optimal sample sizes N and powers π were plotted.

Table 2: TDR one-stage 2-by-2 design at 30% confidence interval of p_C, $\alpha \leq 0.20, \beta \leq 0.20$

p_C	p_E	γ_{max}	λ_{max}	s	m	N	π	β	α	α_{max}	γ	η	λ
0.1	0.25	0.12	0.25	2	4	50	0.79	0.2	0.15	0.19	0.02	0.08	0.05
0.1	0.3	0.08	0.2	1	3	28	0.8	0.13	0.14	0.2	0.08	0.23	0.15
0.1	0.35	0.12	0.2	1	3	22	0.77	0.11	0.08	0.13	0.11	0.27	0.19
0.2	0.35	0.13	0.25	1	9	62	0.77	0.12	0.14	0.19	0.11	0.3	0.21
0.2	0.4	0.1	0.15	2	6	40	0.77	0.19	0.14	0.19	0.04	0.13	0.09
0.2	0.45	0.1	0.2	1	4	22	0.76	0.14	0.14	0.2	0.1	0.25	0.17
0.3	0.45	0.17	0.3	0	14	70	0.76	0.08	0.13	0.19	0.16	0.42	0.29
0.3	0.5	0.12	0.25	1	9	42	0.77	0.12	0.13	0.19	0.11	0.3	0.21
0.35	0.5	0.1	0.2	2	17	78	0.76	0.16	0.14	0.19	0.08	0.22	0.15
0.35	0.55	0.12	0.25	1	11	46	0.78	0.11	0.13	0.18	0.11	0.31	0.21
0.35	0.6	0.1	0.15	2	8	32	0.77	0.18	0.12	0.16	0.05	0.17	0.11
0.4	0.6	0.1	0.15	2	14	54	0.79	0.14	0.12	0.17	0.07	0.22	0.14
0.4	0.65	0.11	0.2	1	8	28	0.77	0.12	0.13	0.19	0.11	0.29	0.2
0.5	0.65	0.1	0.2	2	23	78	0.76	0.16	0.14	0.19	0.08	0.23	0.15
0.5	0.7	0.1	0.15	2	14	46	0.78	0.17	0.15	0.2	0.05	0.18	0.11
0.55	0.7	0.1	0.2	2	25	78	0.77	0.15	0.14	0.18	0.08	0.23	0.16
0.55	0.75	0.1	0.15	2	14	42	0.76	0.19	0.15	0.19	0.05	0.17	0.11
0.6	0.85	0.1	0.1	2	11	30	0.81	0.17	0.15	0.19	0.02	0.14	0.08
0.65	0.85	0.05	0.15	2	16	42	0.81	0.16	0.15	0.19	0.03	0.17	0.1
0.7	0.85	0.11	0.25	1	21	52	0.77	0.12	0.14	0.2	0.11	0.3	0.2

γ_{max}: design constraint for γ; λ_{max}: design constraint for λ; s: statistical difference boundary; m: clinical relevance boundary; N: total sample size; π: power; α: type I error; α_{max}: maximum type I error at 30% confidence interval of p_C; β: type II error; γ: inconclusive probability under H_a; η: inconclusive probability under H_0; λ: average inconclusive probability under H_0 and H_a.

Table 3: TDR one-stage 3-by-2 design under $\alpha \leq 0.20, \beta \leq 0.20$

p_C	p_E	γ_{max}	λ_{max}	r	s	m	N	π	β	α	γ	η	λ
0.1	0.25	0.13	0.2	-5	1	4	44	0.79	0.08	0.15	0.12	0.27	0.2
0.1	0.3	0.12	0.2	-5	1	3	28	0.8	0.09	0.14	0.12	0.25	0.18
0.1	0.35	0.14	0.25	-4	1	3	22	0.77	0.09	0.08	0.14	0.28	0.21
0.2	0.35	0.15	0.25	-7	1	8	56	0.77	0.08	0.16	0.14	0.3	0.22
0.2	0.4	0.14	0.25	-6	1	5	32	0.78	0.09	0.17	0.13	0.27	0.2
0.2	0.45	0.17	0.2	-5	1	4	24	0.81	0.07	0.17	0.12	0.26	0.19
0.3	0.45	0.15	0.25	-8	1	11	58	0.77	0.09	0.2	0.14	0.28	0.21
0.3	0.5	0.15	0.25	-6	1	7	34	0.77	0.09	0.19	0.14	0.27	0.21
0.35	0.5	0.15	0.25	-8	1	13	60	0.76	0.09	0.19	0.15	0.29	0.22
0.35	0.55	0.13	0.2	-7	1	9	40	0.81	0.07	0.2	0.12	0.27	0.2
0.35	0.6	0.14	0.25	-5	1	6	24	0.78	0.08	0.18	0.14	0.27	0.2
0.4	0.6	0.16	0.25	-7	1	10	38	0.76	0.09	0.16	0.15	0.3	0.22
0.4	0.65	0.13	0.2	-5	1	7	26	0.81	0.07	0.19	0.13	0.27	0.2
0.5	0.65	0.15	0.25	-9	1	20	68	0.78	0.08	0.17	0.15	0.31	0.23
0.5	0.7	0.16	0.25	-6	1	12	38	0.77	0.08	0.16	0.15	0.3	0.23
0.55	0.7	0.15	0.25	-8	1	19	60	0.78	0.08	0.2	0.14	0.29	0.21
0.55	0.75	0.15	0.25	-6	1	12	36	0.8	0.07	0.19	0.13	0.28	0.21
0.6	0.85	0.16	0.25	-4	1	7	18	0.77	0.08	0.19	0.15	0.26	0.2
0.65	0.85	0.15	0.25	-5	1	11	28	0.78	0.07	0.18	0.15	0.27	0.21
0.7	0.85	0.17	0.25	-6	1	19	48	0.79	0.07	0.19	0.14	0.28	0.21

γ_{max}: design constraint for γ; λ_{max}: design constraint for λ; r: lower statistical difference boundary; s: upper statistical difference boundary; m: clinical relevance boundary; N: total sample size; π: power; α: type I error; β: type II error; γ: inconclusive probability under H_a; η: inconclusive probability under H_0; λ: average inconclusive probability under H_0 and H_a.

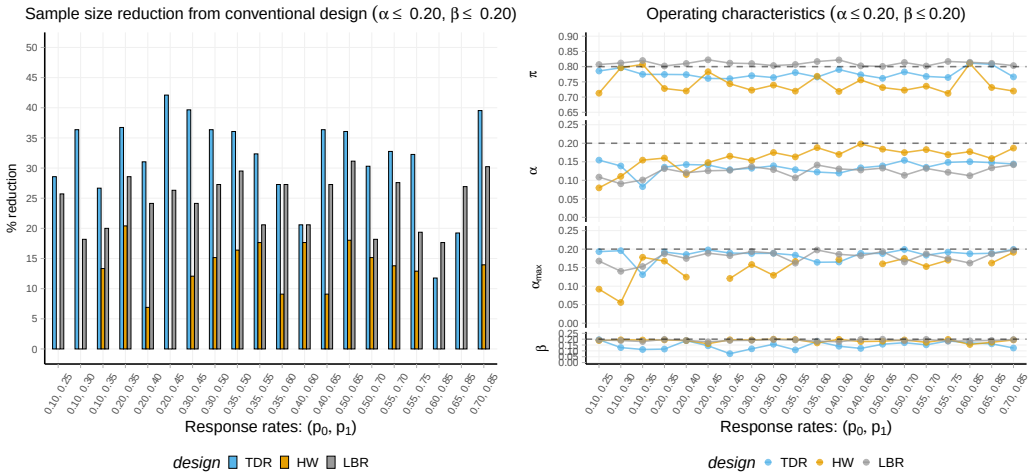

Figure 2: Robustness of the one-stage 2-by-2 TDR design at 30% confidence interval of p_C.(a) Sample size reduction with respect to the two-sample t-test at 30% confidence interval of p_C; (b) Operating characteristics power π, type I error α, maximum type I error α_{max} at 30% confidence interval of p_C, and type II error β.

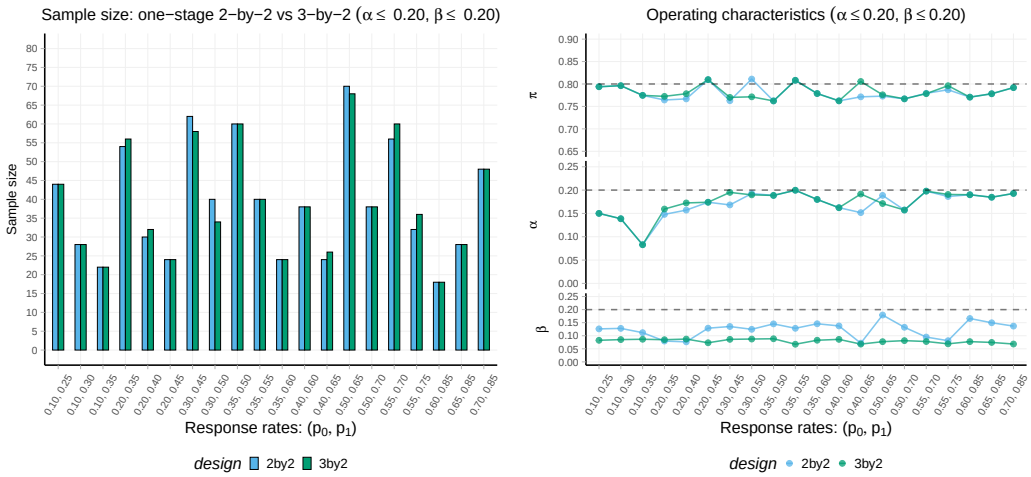

Figure 3: Comparison of one-stage 2-by-2 TDR design with 3-by-2 TDR design under $\alpha \leq 0.20, \beta \leq 0.20$. (a) Sample size reduction with respect to the two-sample t-test; (b) Operating characteristics power π, type I error α, and type II error β.