Supporting information for "Time-to-Event Calibration-Free Odds Design: A Robust Efficient Design for Phase I Trials with Late-onset Outcomes" by Huaqing Jin and Guosheng Yin

A Odds ratio and threshold value of CFO

Suppose there are K dose levels under investigation in a trial, and let p_k be the corresponding toxicity probability and ϕ be the target DLT rate. After enrolling n cohorts of patients, we observe the cumulative data, $D_n = \{(x_k, m_k)\}_{k=1}^K$, where (x_k, m_k) represent the numbers of observed DLTs and patients at dose level k, respectively. Given the nth cohort treated at dose level d_n, the DLT rates at dose levels $(d_n - 1, d_n, d_n + 1)$ are denoted as (p_L, p_C, p_R) based on their left, central (current), and right positions, and (x_L, x_C, x_R) and (m_L, m_C, m_R) are the corresponding number of DLTs and number of patients, respectively.

For $k = L, C, R$, the odds of $p_k > \phi$ is calculated as

$$O_k = \frac{\Pr(p_k > \phi | x_k, m_k)}{\Pr(p_k \leq \phi | x_k, m_k)}$$

for $k = L, C, R$ corresponding to left, current and right doses. The reciprocal $\bar{O}_k = 1/O_k$ represents the odds of $p_k \leq \phi$. Under the Bayesian paradigm, a noninformative Beta$(\phi, 1-\phi)$ prior distribution is adopted for each DLT probability p_k.

To calculate the odds ratio $O_C/\bar{O}_L$, we further take the monotonic relationship $p_L < p_C$ into consideration. By accounting for such monotonicity, the marginal posterior density functions for p_L and p_C can be derived as

$$\begin{aligned} f_L(p_L | x_L, x_C) &\propto f_\beta(p_L; a_L, b_L) \int_{p_L}^1 f_\beta(p_C; a_C, b_C)\, \mathrm{d}p_C \\ f_C(p_C | x_L, x_C) &\propto f_\beta(p_C; a_C, b_C) \int_0^{p_C} f_\beta(p_L; a_L, b_L)\, \mathrm{d}p_L, \end{aligned}$$

where $f_\beta(\cdot; a_k, b_k)$ is the density function of Beta(a_k, b_k), with $a_k = \phi + x_k$ and $b_k = 1 - \phi + m_k - x_k$ for $k = L, C$, i.e., the posterior distribution of p_k given the data (x_k, m_k) without incorporating the monotonic relationship.

Let p_{0L} and p_{0C} denote the true values of p_L and p_C, respectively. To obtain the threshold value γ_L, we propose to minimize the probability of the incorrect vote,

$$\begin{aligned} & V_L(\gamma_L) \\ = \; & \Pr(O_C/\bar{O}_L > \gamma_L | p_{0C} = \phi, p_{0L} < \phi) + \Pr(O_C/\bar{O}_L \leq \gamma_L | p_{0L} = \phi, p_{0C} > \phi) \\ = \; & \sum_{i=0}^{m_C} \sum_{j=0}^{m_L} I(O_C/\bar{O}_L > \gamma_L) \Pr(x_C = i | p_{0C} = \phi) \Pr(x_L = j | p_{0L} < \phi) \\ & + \sum_{i=0}^{m_C} \sum_{j=0}^{m_L} I(O_C/\bar{O}_L \leq \gamma_L) \Pr(x_C = i | p_{0C} > \phi) \Pr(x_L = j | p_{0L} = \phi), \end{aligned}$$

where $I(\cdot)$ is the indicator function. Similar discussions apply to the odds ratio $\bar{O}_C/O_R$ and computation of γ_R (Jin and Yin, 2022).

B Simulation Details

B.1 Existing designs compared with TITE-CFO

The detailed settings of TITE-BOIN (Yuan et al., 2018) and TITE-CRM (Cheung and Chappell, 2000) used in the simulation studies are described as follows.

- **TITE-BOIN:** Following Yuan et al. (2018), we set $\phi_1 = 0.6\phi$ and $\phi_2 = 1.4\phi$. We eliminate dose level k and all dose levels above k from the trial if $\Pr(p_k > \phi | x_k, m_k \geq 3) > 0.95$. If the posterior probability of the first dose level satisfies $\Pr(p_1 > \phi | x_1, m_1 \geq 3) > 0.95$, then we terminate the entire trial for safety. However, unlike the original paper (Yuan et al., 2018), we do not include the suspending rule for TITE-BOIN in our simulation studies. Due to the fact that the suspending rule elongates the trial duration, if we include such a rule, the comparison between TITE-BOIN and other methods would be unfair.

- **TITE-CRM:** We adopt the power model $p_j = \pi_j^{\exp(\alpha)}$ with the model skeleton selected by the method of Lee and Cheung (2009). We choose an initial guess of the MTD at dose level $\lceil K/2 \rceil$ while the halfwidth of the indifference interval is set as 0.05 following Lin and Yin (2017, 2018). If $\Pr(p_1 > \phi|\text{data}) > 0.95$, the early stopping rule would be triggered to terminate the trial.

B.2 Fixed scenarios

We have incorporated the 16 fixed scenarios presented in Yuan et al. (2018) into our study. Out of these scenarios, eight are associated with a target DLT rate of 0.2 while the remaining eight have a target DLT rate of 0.3. There are 7 dose levels in total. We use 12 cohorts with cohort size 3. Specifically, 16 scenarios are listed in Table A.1.

Table A.1: The 16 fixed scenarios in our numerical studies.

Scenario	Dose level						
	1	2	3	4	5	6	7
	Target DLT 0.2						
1	0.05	**0.20**	0.46	0.50	0.60	0.70	0.80
2	0.02	0.05	**0.20**	0.28	0.34	0.40	0.44
3	0.01	0.05	0.10	**0.20**	0.32	0.50	0.70
4	0.01	0.04	0.07	0.10	0.50	0.70	0.90
5	0.01	0.05	0.10	0.14	**0.20**	0.26	0.34
6	0.01	0.02	0.03	0.05	**0.20**	0.40	0.50
7	0.01	0.04	0.07	0.10	0.15	**0.20**	0.25
8	0.01	0.02	0.03	0.04	0.05	**0.20**	0.45
	Target DLT 0.3						
1	**0.30**	0.40	0.50	0.60	0.70	0.80	0.90
2	0.14	**0.30**	0.39	0.48	0.56	0.64	0.70
3	0.07	**0.23**	0.41	0.49	0.62	0.68	0.73
4	0.05	0.15	**0.30**	0.40	0.50	0.60	0.70
5	0.05	0.12	0.20	**0.30**	0.38	0.49	0.56
6	0.01	0.04	0.08	0.15	**0.30**	0.36	0.43
7	0.02	0.04	0.08	0.10	0.20	**0.30**	0.40
8	0.01	0.03	0.05	0.07	0.09	**0.30**	0.50

B.3 Generation of random scenarios

We generate random scenarios to assess the performance of the phase I designs in Sections 3.1 and 3.2 following Paoletti et al. (2004). Specifically, the procedure is detailed as follows.

1. We randomly select, with equal probabilities, one of the K dose levels as the MTD and denote that dose level as k_{MTD}.

2. Let Φ be the cumulative distribution function (CDF) of the standard normal distribution. The probability of the MTD is $p_{k_{\text{MTD}}} = \Phi(\epsilon_{\text{MTD}})$ with $\epsilon_{\text{MTD}} \sim N(\Phi^{-1}(\phi), \sigma_0^2)$, where ϕ is the target toxicity probability.

3. For $\{p_k\}_{k=1}^{k_{\text{MTD}}-1}$, we generate

$$p_{k-1} = \Phi\left[\Phi^{-1}(p_k) - \left\{\Phi^{-1}(p_k) - \Phi^{-1}(2\phi - p_k)\right\} I\left\{\Phi^{-1}(p_k) > \Phi^{-1}(\phi)\right\} - \epsilon_{k-1}^2\right],$$

   where $I(\cdot)$ is the indicator function and $\epsilon_{k-1} \sim N(\mu_1, \sigma_1^2)$.

4. For $\{p_k\}_{k=k_{\text{MTD}}+1}^{K}$, we generate

$$p_{k+1} = \Phi\left[\Phi^{-1}(p_k) + \left\{\Phi^{-1}(2\phi - p_k) - \Phi^{-1}(p_k)\right\} I\left\{\Phi^{-1}(p_k) < \Phi^{-1}(\phi)\right\} + \epsilon_{k+1}^2\right],$$

   where $\epsilon_{k+1} \sim N(\mu_2, \sigma_2^2)$.

Following Liu and Yuan (2015), we choose $\sigma_0 = 0.05$ and $\sigma_1 = \sigma_2 = 0.35$, and tune the parameters $\mu_1 = \mu_2$ to achieve the desirable Δ, i.e., the average probability difference around the target.

C More results

C.1 Results when target DLT rate is 0.2

The simulation results with a target DLT rate of 0.2 are shown in Figure A.1 (fixed) and Figure A.2 (random). All configurations remain consistent with the simulation studies discussed in Section 3 of the manuscript.

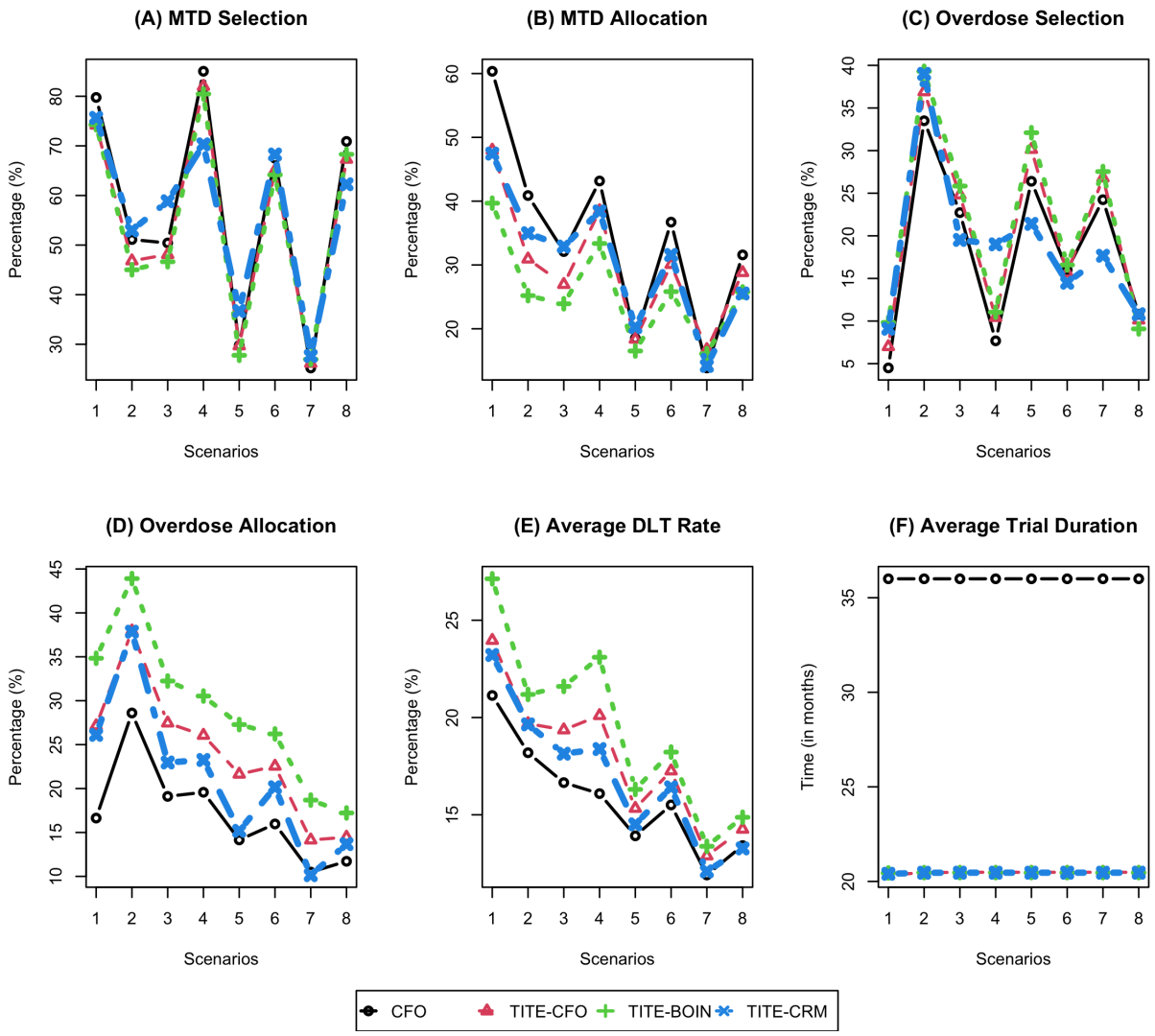

Figure A.1: Simulation results of the CFO, TITE-CFO, TITE-BOIN and TITE-CRM designs with the target DLT rate 0.2 and sample size 36 under eight fixed scenarios. For MTD selection and allocation, a higher value is preferred. For overdose selection and allocation, the average DLT rate and average trial duration, a lower value is preferred.

C.2 Threhold values

The threshold values of (γ_L, γ_R) under the target DLT rates 0.2, 0.25 and 0.3. are shown in Figure A.3.

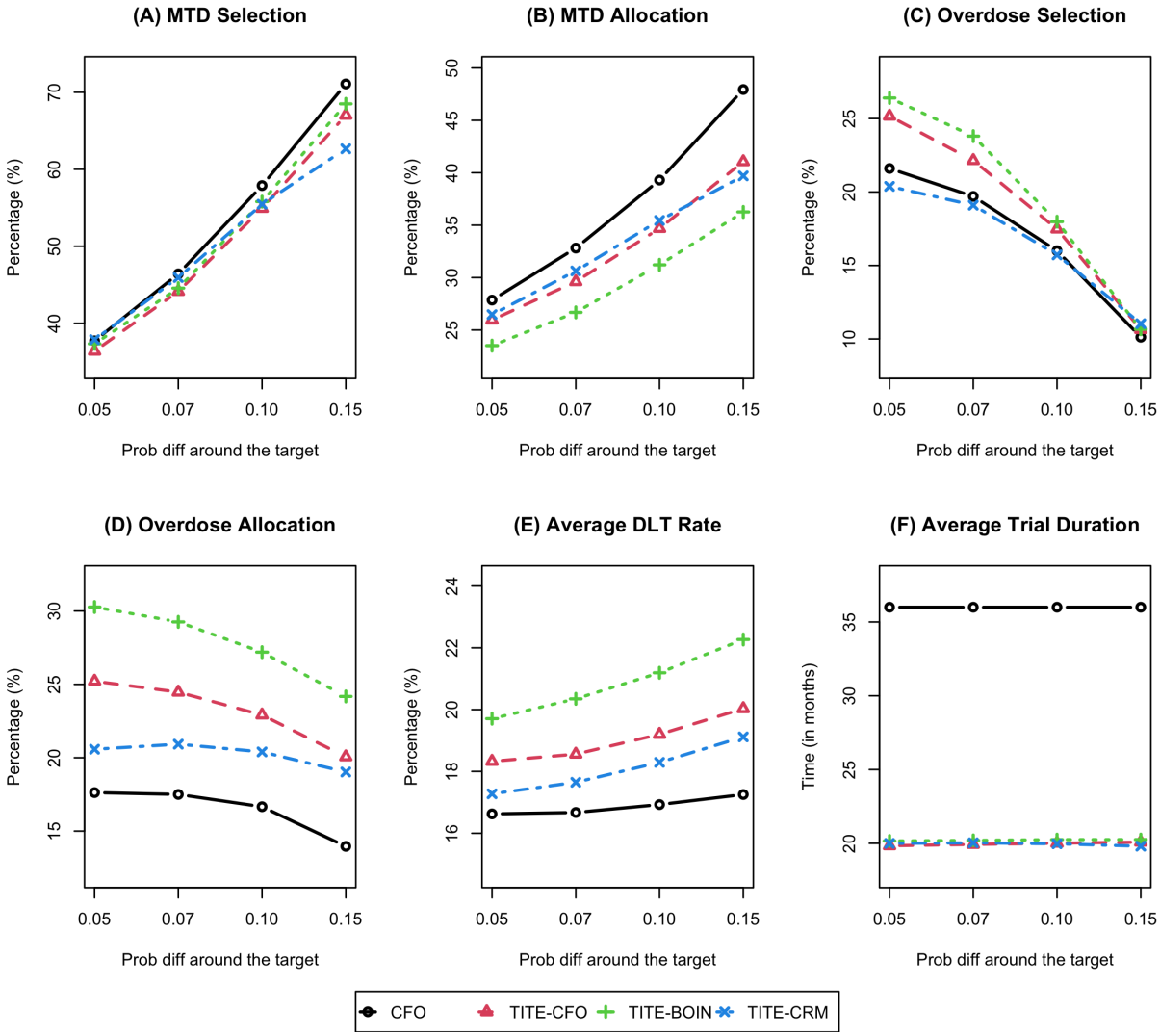

Figure A.2: Simulation results of the CFO, TITE-CFO, TITE-BOIN and TITE-CRM designs with the target DLT rate 0.2 and sample size 36 under random scenarios when the average probability difference around the target DLT rate varies from 0.05 to 0.15. For MTD selection and allocation, a higher value is preferred. For overdose selection and allocation, the average DLT rate and average trial duration, a lower value is preferred.

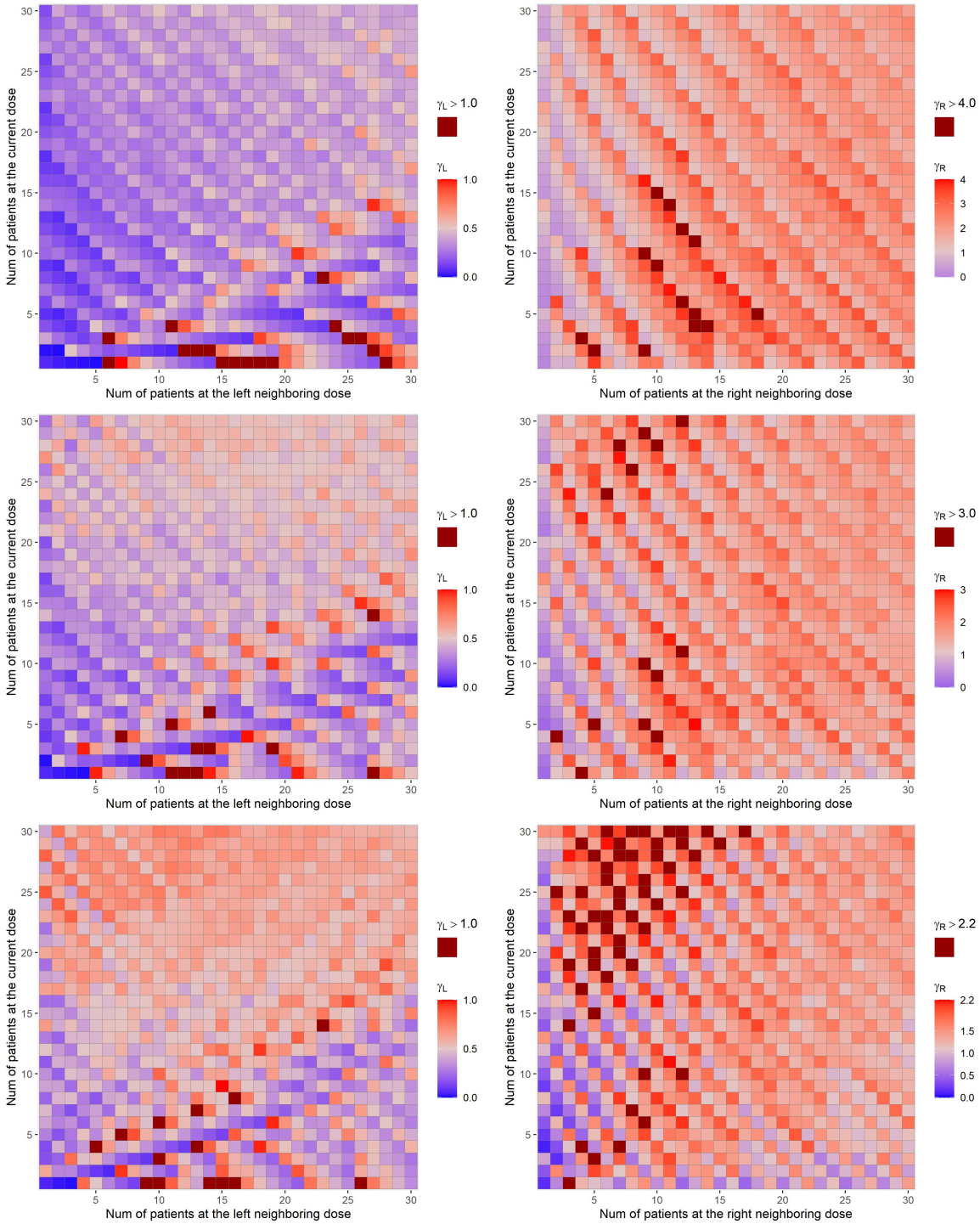

Figure A.3: The threshold values of (γ_L, γ_R) when the numbers of patients treated at the left, current, and right doses vary from 1 to 30 given the target DLT rates being 0.20 (top), 0.25 (middle) and 0.3 (bottom).

References

Y. K. Cheung and R. Chappell. Sequential designs for phase I clinical trials with late-onset toxicities. *Biometrics*, 56(4):1177–1182, 2000.

H. Jin and G. Yin. CFO: Calibration-free odds design for phase I/II clinical trials. *Statistical Methods in Medical Research*, page 09622802221079353, 2022.

S. M. Lee and Y. K. Cheung. Model calibration in the continual reassessment method. *Clinical Trials*, 6(3):227–238, 2009.

R. Lin and G. Yin. Nonparametric overdose control with late-onset toxicity in phase I clinical trials. *Biostatistics*, 18(1):180–194, 2017.

R. Lin and G. Yin. Uniformly most powerful bayesian interval design for phase I dose-finding trials. *Pharmaceutical Statistics*, 17(6):710–724, 2018.

S. Liu and Y. Yuan. Bayesian optimal interval designs for phase I clinical trials. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 64(3):507–523, 2015.

X. Paoletti, J. O'Quigley, and J. Maccario. Design efficiency in dose finding studies. *Computational Statistics & Data Analysis*, 45(2):197–214, 2004.

Y. Yuan, R. Lin, D. Li, L. Nie, and K. E. Warren. Time-to-event bayesian optimal interval design to accelerate phase I trials. *Clinical Cancer Research*, 24(20):4921–4930, 2018.