

Web Appendix for “An evaluation of constrained randomization for the design and analysis of group-randomized trials with binary outcomes”

Fan Li*, Elizabeth L. Turner, Patrick J. Heagerty, David M. Murray, William M. Vollmer and Elizabeth R. DeLong
*frank.li@duke.edu

A. Derivation of test statistic (7)

From GLMM (3), we have $\pi_{ij} = e^{\mu + \delta T_i + \mathbf{x}'_{ij}\beta + \gamma_i} / (1 + e^{\mu + \delta T_i + \mathbf{x}'_{ij}\beta + \gamma_i})$, and $\partial\pi_{ij}/\partial\delta = T_i\pi_{ij}(1 - \pi_{ij})$. Assuming the regularity conditions hold and so we can interchange integration and differentiation, we have

$$\begin{aligned} & \frac{\partial}{\partial\delta} \int \prod_{j=1}^{n_j} \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{1-y_{ij}} \phi(\gamma_i|0, \sigma_\gamma^2) d\gamma_i = \int \frac{\partial}{\partial\delta} \prod_{j=1}^{n_j} \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{1-y_{ij}} \phi(\gamma_i|0, \sigma_\gamma^2) d\gamma_i \\ &= \int \frac{\partial}{\partial\delta} \exp \left[\sum_{j=1}^{n_j} y_{ij} \log \pi_{ij} + \sum_{j=1}^{n_j} (1 - y_{ij}) \log(1 - \pi_{ij}) + \sum_{j=1}^{n_j} \log \phi(\gamma_i|0, \sigma_\gamma^2) \right] d\gamma_i \\ &= \int \left[T_i \sum_{j=1}^{n_j} \frac{y_{ij}}{\pi_{ij}} \pi_{ij} (1 - \pi_{ij}) - T_i \sum_{j=1}^{n_j} \frac{1 - y_{ij}}{1 - \pi_{ij}} \pi_{ij} (1 - \pi_{ij}) \right] \prod_{j=1}^{n_j} \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{1-y_{ij}} \phi(\gamma_i|0, \sigma_\gamma^2) d\gamma_i \\ &= T_i \int \sum_{j=1}^{n_j} (y_{ij} - \pi_{ij}) \prod_{j=1}^{n_j} \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{1-y_{ij}} \phi(\gamma_i|0, \sigma_\gamma^2) d\gamma_i \end{aligned}$$

Then the likelihood P-test statistic is

$$\begin{aligned} S_{\text{likelihood}} &= \frac{\partial}{\partial\delta} l(\theta, \sigma_\gamma^2) \Big|_{\delta=0} = \sum_{i=1}^{2g} \frac{\partial}{\partial\delta} \frac{\int \prod_{j=1}^{n_j} \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{1-y_{ij}} \phi(\gamma_i|0, \sigma_\gamma^2) d\gamma_i}{\int \prod_{j=1}^{n_j} \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{1-y_{ij}} \phi(\gamma_i|0, \sigma_\gamma^2) d\gamma_i} \Big|_{\delta=0} \\ &= \sum_{i=1}^{2g} T_i \frac{\int \sum_{j=1}^{n_j} (y_{ij} - \pi_{ij}) \prod_{j=1}^{n_j} \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{1-y_{ij}} \phi(\gamma_i|0, \sigma_\gamma^2) d\gamma_i}{\int \prod_{j=1}^{n_j} \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{1-y_{ij}} \phi(\gamma_i|0, \sigma_\gamma^2) d\gamma_i} \Big|_{\delta=0}. \end{aligned}$$

B. Matrix expressions for balance scores

We provide simple matrix expressions for simultaneously calculating the balance scores corresponding to the enumerated or simulated randomization schemes are presented as follows. Let R represent the number of all possible randomization schemes or the number of simulated unique schemes, and let n_T, n_C be the number of groups assigned to treatment and control arms, respectively. The total number of groups is denoted by $n = n_T + n_C$. For each data replicate, we specify the $R \times n$ matrix \mathbf{M} , so that each row $[\mathbf{M}]_{r\cdot} = (m_{r1}, m_{r2}, \dots, m_{rn})$ indicates a unique randomization scheme, where the (r, i) th element m_{ri} is 1 if group i is randomized to treatment according to the r th scheme and 0 otherwise. When K covariates are balanced by constrained randomization, we further write a $n \times K$ matrix \mathbf{X} of covariate means aggregated at the group level, so that the (i, k) th element x_{ik} is the i th group mean of the k th covariate. Particularly, if the k th covariate is a group-level covariate as in our simulation scenarios, x_{ik} simply takes the value of that group-level covariate from the i th group because all individuals from that group have identical values with respect to that covariate. Here we present a more general case allowing for balancing individual-level covariates, if available.

We obtain a “column-standardized” matrix $\tilde{\mathbf{X}}$ by the following two steps: (i) subtracting the column-specific mean from each column of \mathbf{X} , and (ii) scaling each column \mathbf{X} by the covariate- or column-specific weight $\omega_k^{1/2}$. Mathematically, this

amounts to the following matrix operation:

$$\tilde{\mathbf{X}} = (\mathbf{I} - n^{-1}\mathbf{J})\mathbf{X}\mathbf{W}^{1/2}, \tag{B.1}$$

where \mathbf{I} is the n -dimensional unit matrix, \mathbf{J} is the n -dimensional matrix of ones and $\mathbf{W} = \text{diag}(\omega_1, \dots, \omega_k)$. Then the $R \times 1$ vector of l_2 balance scores is calculated efficiently in one step by:

$$\mathbf{B}_{(l_2)} = \left(\frac{n_T n_C}{n}\right)^{-2} (\mathbf{M}\tilde{\mathbf{X}})^{\circ 2} \mathbf{1}_{K \times 1}, \tag{B.2}$$

where the operation ‘ $\circ 2$ ’ is the Hadamard power (element-wise operation on matrices) and $\mathbf{1}_{K \times 1}$ is a $K \times 1$ column vector of ones. The justification for the expression (B.2) is the following. First, write the average of the n group-level means by $\bar{x}_{.k} = n^{-1} \sum_{i=1}^n x_{ik}$. Then the (i, k) th element of $\tilde{\mathbf{X}}$ is

$$\tilde{x}_{ik} = \sqrt{\omega_k}(x_{ik} - \bar{x}_{.k}) = \sqrt{\omega_k}\left(x_{ik} - \frac{\sum_{i=1}^n x_{ik}}{n}\right) = \sqrt{\omega_k}\left(x_{ik} - \frac{n_T \bar{x}_{T_k} + n_C \bar{x}_{C_k}}{n}\right),$$

where $\bar{x}_{T_k}, \bar{x}_{C_k}$ denote the average of the k th group-level variable means from two arms. The r th element of $\mathbf{B}_{(l_2)}$ is

$$\begin{aligned} [\mathbf{B}_{(l_2)}]_r &= \left(\frac{n_T n_C}{n}\right)^{-2} \sum_{k=1}^K \left(\sum_{i=1}^n m_{ri} \tilde{x}_{ik}\right)^2 \\ &= \left(\frac{n_T n_C}{n}\right)^{-2} \sum_{k=1}^K \left[\sum_{i:m_{ri}=1} \sqrt{\omega_k} \left(x_{ik} - \frac{n_T \bar{x}_{T_k} + n_C \bar{x}_{C_k}}{n}\right)\right]^2 \\ &= \left(\frac{n_T n_C}{n}\right)^{-2} \sum_{k=1}^K \omega_k \left(n_T \bar{x}_{T_k} - \frac{n_T^2}{n} \bar{x}_{T_k} - \frac{n_T n_C}{n} \bar{x}_{C_k}\right)^2 \\ &= \left(\frac{n_T n_C}{n}\right)^{-2} \sum_{k=1}^K \omega_k \left(\frac{n_T n_C}{n} (\bar{x}_{T_k} - \bar{x}_{C_k})\right)^2 = \sum_{k=1}^K \omega_k (\bar{x}_{T_k} - \bar{x}_{C_k})^2. \end{aligned}$$

For the default choice of weights, we let $\omega_k = 1/s_k^2$ in equation (B.1), where the variance of the k th covariate is given by $s_k^2 = (n-1)^{-1} \sum_{i=1}^n (x_{ik} - \bar{x}_{.k})^2$.

Replacing (B.1) by

$$\tilde{\mathbf{X}} = (\mathbf{I} - n^{-1}\mathbf{J})\mathbf{X}\tilde{\mathbf{W}}, \tag{B.3}$$

where $\tilde{\mathbf{W}} = \text{diag}(\tilde{\omega}_1, \dots, \tilde{\omega}_k)$, we can similarly observe that the $R \times 1$ vector of l_1 balance scores is given by:

$$\mathbf{B}_{(l_1)} = \left(\frac{n_T n_C}{n}\right)^{-1} \{(\mathbf{M}\tilde{\mathbf{X}})^{\circ 2}\}^{\circ \frac{1}{2}} \mathbf{1}_{K \times 1}. \tag{B.4}$$

To see this, we check that the r th element of $\mathbf{B}_{(l_1)}$ is

$$[\mathbf{B}_{(l_1)}]_r = \left(\frac{n_T n_C}{n}\right)^{-1} \sum_{k=1}^K \left|\sum_{i=1}^n m_{ri} \tilde{x}_{ik}\right| = \left(\frac{n_T n_C}{n}\right)^{-1} \sum_{k=1}^K \tilde{\omega}_k \left|\frac{n_T n_C}{n} (\bar{x}_{T_k} - \bar{x}_{C_k})\right| = \sum_{k=1}^K \tilde{\omega}_k |\bar{x}_{T_k} - \bar{x}_{C_k}|.$$

The default choice of $\tilde{\omega}_k = 1/s_k$.

C. Example SAS code for model-based tests

Linearization F-test

```
/* longdata: a dataset containing the following variables
   Y: binary response variable (0/1)
   t: binary treatment indicator (0/1)
   x1, x2: group- or individual-level covariates (binary or continuous)
   cluster: cluster identifier */

proc glimmix data = longdata;
class cluster;
model Y(event='1') = t x1 x2 / solution ddfm=betwithin dist=bin link=logit;
random intercept / subject=cluster;
run;
quit;
```

Likelihood F-test

```
/* we could initialize values for PROC BLMIXED using GEE from PROC GENMOD,
   as suggested by Zhang et al. (2011) On fitting generalized linear mixed-effects
   models for binary responses using different statistical packages; Stat Med */

* obtain initial values;
proc genmod data=longdata descending;
class cluster;
model Y = t x1 x2 / dist=bin link=logit;
repeated subject=cluster / type=exch covb;
run;
* initial values: beta0=-0.40 delta=-0.04 beta1=0.43 beta2=1.18 s2u=0.16;

proc nlmixed data = longdata qpoints = 10 df = 12;
*manually specify the between-within df;
parms beta0=-0.40 delta=-0.04 beta1=0.43 beta2=1.18 s2u=0.16;
eta = beta0 + delta*t + beta1*x1 + beta2*x2 + u;
p = exp(eta)/(1+exp(eta));
model y ~ binomial(1, p);
random u ~ normal(0,s2u) subject=cluster;
run;
quit;
```

Bias-corrected GEE t-tests

```
/* the "empirical = " option controls the type of bias-correction
   to the sandwich variance estimator
   the MD correction is provided by "empirical = FIRORES"
   the KC correction is provided by "empirical = ROOT"
   the FG correction is provided by "empirical = FIROEEQ" */
```

```

proc glimmix data=longdata empirical=FIRORES;
class cluster;
model Y(event='1') = t x1 x2 / solution ddfm=betwithin dist=bin link=logit ;
random _residual_ / subject=cluster type=cs;
run;

proc glimmix data=longdata empirical=ROOT;
class cluster;
model Y(event='1') = t x1 x2 / solution ddfm=betwithin dist=bin link=logit ;
random _residual_ / subject=cluster type=cs;
run;

proc glimmix data=longdata empirical=FIROEEQ;
class cluster;
model Y(event='1') = t x1 x2 / solution ddfm=betwithin dist=bin link=logit ;
random _residual_ / subject=cluster type=cs;
run;

```

D. Web Tables

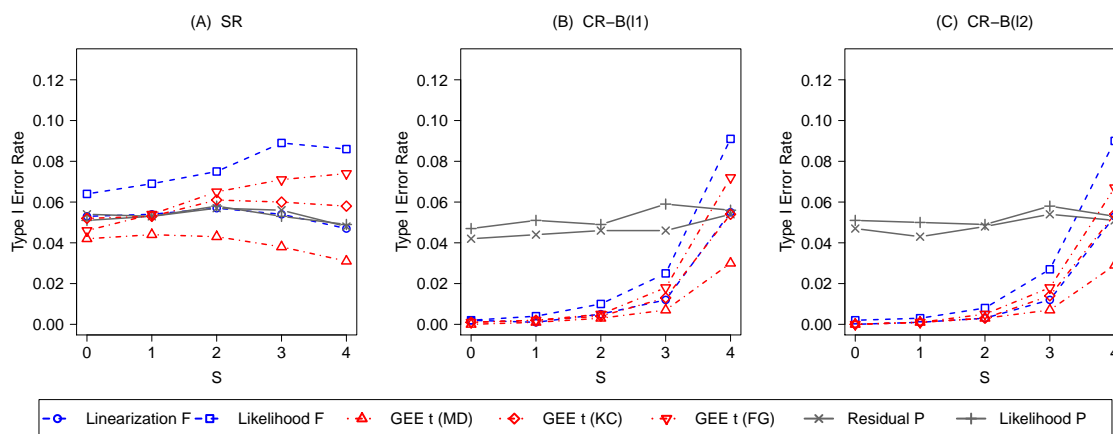
Web Table 1. Type I error rate for the unadjusted ($S = 0$) and fully-adjusted ($S = 4$) tests under simple versus constrained randomization with $g = 8$. The l_2 balance metric is used by constrained randomization; the group-level covariates are assumed prognostic.

ICC	Randomization	q	Linearization F		GEE t (KC)		Residual P		Likelihood P	
			$S = 0$	$S = 4$	$S = 0$	$S = 4$	$S = 0$	$S = 4$	$S = 0$	$S = 4$
$\rho = 0.001$	Constrained	0.01	0.000	0.028	0.000	0.057	0.045	0.050	0.048	0.050
		0.05	0.000	0.025	0.000	0.053	0.053	0.046	0.049	0.046
		0.10	0.000	0.030	0.000	0.057	0.054	0.054	0.043	0.053
		0.30	0.000	0.024	0.001	0.059	0.049	0.051	0.048	0.051
		0.50	0.003	0.023	0.003	0.059	0.049	0.048	0.044	0.049
	Simple	1.00	0.051	0.028	0.047	0.068	0.051	0.059	0.049	0.058
$\rho = 0.01$	Constrained	0.01	0.000	0.053	0.001	0.055	0.048	0.049	0.047	0.048
		0.05	0.000	0.046	0.000	0.054	0.054	0.048	0.057	0.047
		0.10	0.000	0.046	0.000	0.053	0.050	0.045	0.052	0.045
		0.30	0.001	0.046	0.001	0.050	0.044	0.040	0.052	0.043
		0.50	0.004	0.043	0.003	0.047	0.049	0.042	0.045	0.041
	Simple	1.00	0.050	0.050	0.048	0.068	0.050	0.053	0.056	0.052
$\rho = 0.05$	Constrained	0.01	0.001	0.058	0.002	0.061	0.047	0.054	0.044	0.054
		0.05	0.001	0.043	0.001	0.051	0.050	0.044	0.051	0.043
		0.10	0.003	0.047	0.002	0.051	0.047	0.045	0.044	0.043
		0.30	0.005	0.051	0.004	0.044	0.046	0.054	0.053	0.051
		0.50	0.007	0.046	0.008	0.052	0.044	0.044	0.051	0.047
	Simple	1.00	0.048	0.051	0.046	0.072	0.047	0.051	0.054	0.053

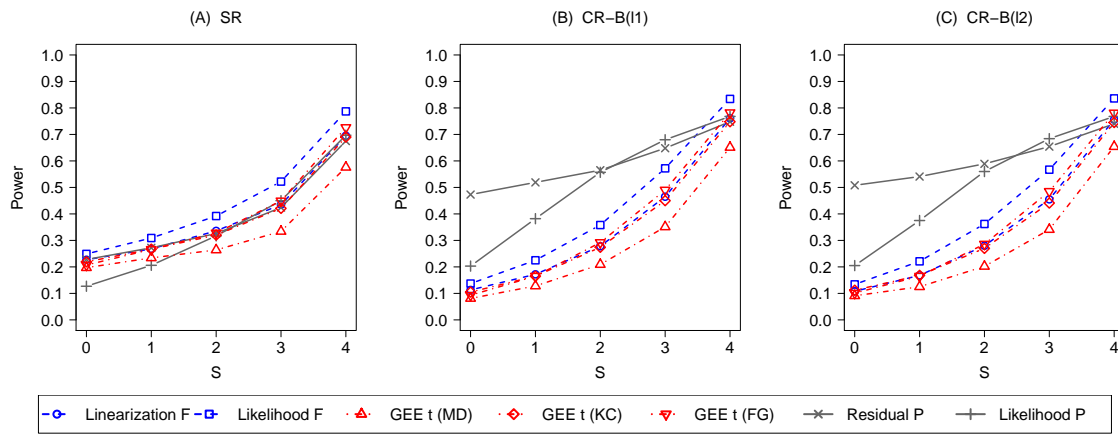
Web Table 2. Power for the unadjusted ($S = 0$) and fully-adjusted ($S = 4$) tests under simple versus constrained randomization with $g = 8$. The l_2 balance metric is used by constrained randomization; the group-level covariates are assumed prognostic.

ICC	Randomization	q	Linearization F		GEE t (KC)		Residual P		Likelihood P	
			$S = 0$	$S = 4$	$S = 0$	$S = 4$	$S = 0$	$S = 4$	$S = 0$	$S = 4$
$\rho = 0.001$	Constrained	0.01	0.014	1.000	0.015	1.000	0.505	0.987	0.300	0.986
		0.05	0.021	1.000	0.021	0.999	0.521	1.000	0.316	1.000
		0.10	0.032	1.000	0.034	0.999	0.439	1.000	0.266	1.000
	Simple	0.30	0.073	1.000	0.071	0.999	0.349	0.998	0.212	0.999
		0.50	0.098	1.000	0.093	0.999	0.280	0.997	0.183	0.999
		1.00	0.159	0.996	0.150	0.994	0.157	0.966	0.125	0.967
$\rho = 0.01$	Constrained	0.01	0.017	0.961	0.016	0.959	0.455	0.929	0.267	0.936
		0.05	0.024	0.968	0.027	0.962	0.465	0.959	0.270	0.964
		0.10	0.031	0.957	0.034	0.954	0.394	0.951	0.222	0.957
	Simple	0.30	0.073	0.948	0.073	0.942	0.330	0.936	0.193	0.943
		0.50	0.098	0.942	0.094	0.928	0.268	0.922	0.165	0.935
		1.00	0.156	0.901	0.148	0.895	0.153	0.830	0.115	0.843
$\rho = 0.05$	Constrained	0.01	0.031	0.521	0.030	0.519	0.287	0.484	0.151	0.498
		0.05	0.042	0.505	0.039	0.503	0.305	0.488	0.145	0.506
		0.10	0.044	0.505	0.042	0.495	0.273	0.488	0.134	0.510
	Simple	0.30	0.074	0.470	0.074	0.474	0.236	0.464	0.133	0.484
		0.50	0.091	0.462	0.085	0.461	0.206	0.456	0.104	0.477
		1.00	0.145	0.420	0.135	0.443	0.143	0.405	0.086	0.422

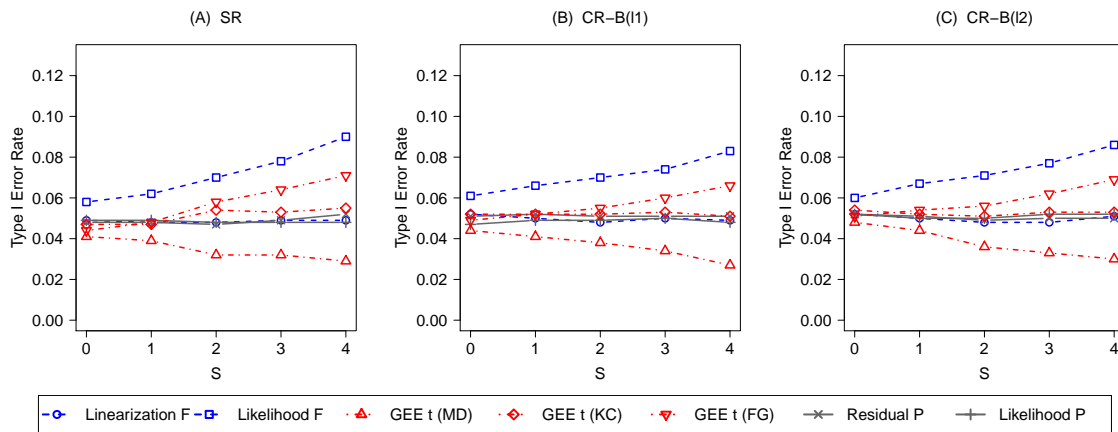
E. Web Figures



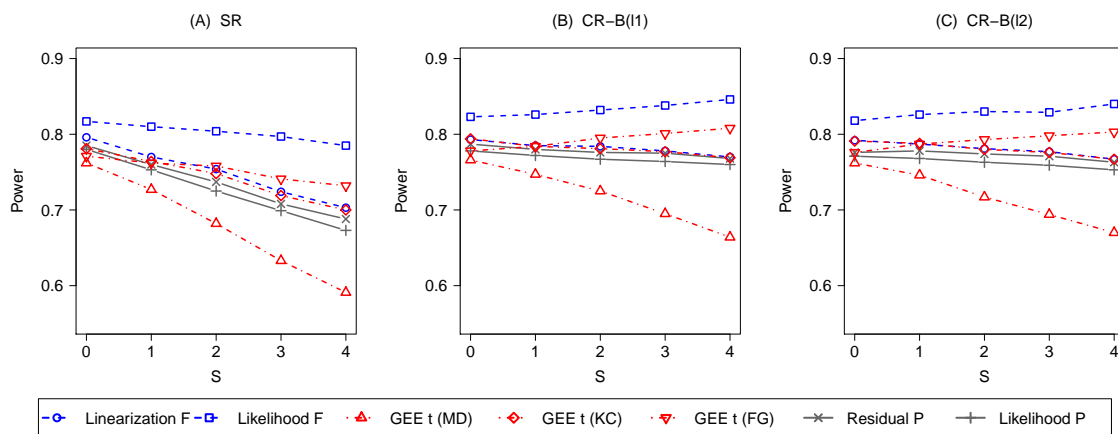
Web Figure 1. Type I error rates with varying degrees (S) of analysis-based adjustment for prognostic group-level covariates under constrained randomization (CR) with 2 balance metrics ($B_{(l1)}$ and $B_{(l2)}$) versus simple randomization (SR); $g = 13$, ICC = 0.05; $q = 0.1$ under CR.



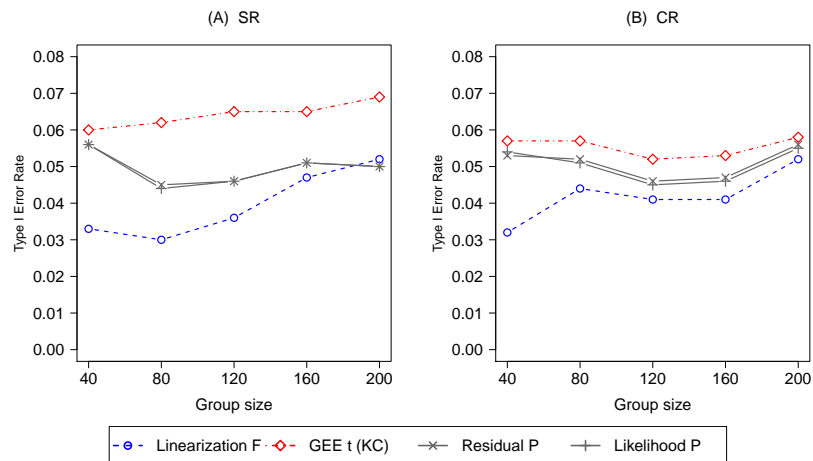
Web Figure 2. Power with varying degrees (S) of analysis-based adjustment for prognostic group-level covariates under constrained randomization (CR) with 2 balance metrics ($B_{(11)}$ and $B_{(12)}$) versus simple randomization (SR); $g = 13$, $ICC = 0.05$; $q = 0.1$ under CR.



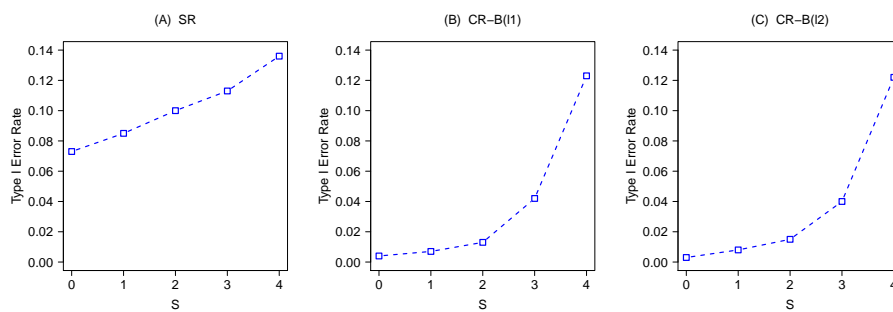
Web Figure 3. Type I error rates with varying degrees (S) of analysis-based adjustment for non-prognostic group-level covariates under constrained randomization (CR) with 2 balance metrics ($B_{(11)}$ and $B_{(12)}$) versus simple randomization (SR); $g = 13$, $ICC = 0.05$; $q = 0.1$ under CR.



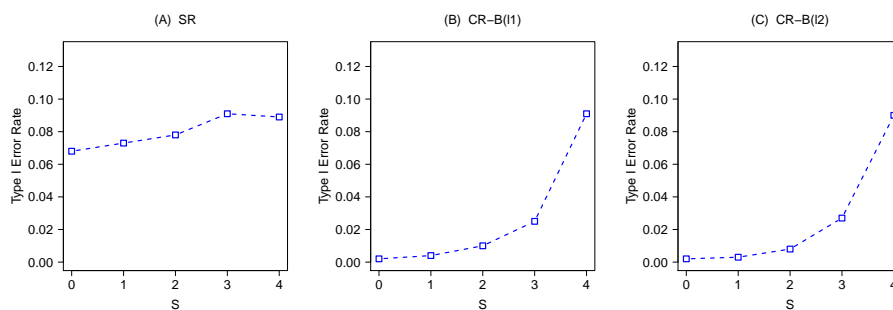
Web Figure 4. Power with varying degrees (S) of analysis-based adjustment for non-prognostic group-level covariates under constrained randomization (CR) with 2 balance metrics ($B_{(11)}$ and $B_{(12)}$) versus simple randomization (SR); $g = 13$, $ICC = 0.05$; $q = 0.1$ under CR.



Web Figure 5. Type I error rates for the fully-adjusted tests ($S = 4$) under both simple randomization (SR) and constrained randomization (CR) at $g = 8$, $ICC = 0.01$ with increasing group sizes. Under CR, the $l1$ metric was used and $q = 0.1$. All four group-level covariates are associated with the outcomes.



Web Figure 6. Type I error rates of the likelihood ratio test with varying degrees (S) of analysis-based adjustment for prognostic group-level covariates under constrained randomization (CR) with 2 balance metrics ($B_{(l1)}$ and $B_{(l2)}$) versus simple randomization (SR); $g = 8$, $ICC = 0.05$; $q = 0.1$ under CR.



Web Figure 7. Type I error rates of the likelihood ratio test with varying degrees (S) of analysis-based adjustment for prognostic group-level covariates under constrained randomization (CR) with 2 balance metrics ($B_{(l1)}$ and $B_{(l2)}$) versus simple randomization (SR); $g = 13$, $ICC = 0.05$; $q = 0.1$ under CR.