

IMPACT OF COVARIATE ADAPTIVE ALLOCATION
PROCEDURES ON POWER AND VALIDITY IN
SMALL-SCALE CLINICAL STUDIES

by

Michael Flanagin

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Master of Science
(Biostatistics)

at the
UNIVERSITY OF WASHINGTON – SEATTLE
2018

Defended on July 4, 2018

Dissertation approved by the following members of the Final Oral Committee:
Amalia Magaret · Associate Professor of Biostatistics
Mike Leblanc · Associate Professor of Biostatistics

Abstract

This is where my abstract goes.

Acknowledgments

I wish to thank my thesis advisor Amalia for her support both in providing input on the direction as well as being a firm advocate on my part in navigating graduate school. It would be unimaginable for me to consider finishing without her support.

My thanks go to my parents Yen and Steve for supporting me throughout the graduate school process and "hanging in there" until the very end.

Contents

Abstract	i
Acknowledgments	ii
Contents	iii
List of Figures	vi
List of Tables	vii
1 Background	1
1.1 Randomization in RCTs	2
1.2 Restricted randomization	2
1.3 Dynamic (adaptive) randomization	3
1.4 Analysis methods	5
1.5 Aims	6
1.6 Measures to evaluate aims	7
2 Methods	8
2.1 Data generation	9
2.2 Allocation procedures	10
2.3 Varied conditions	10
2.4 Analysis approaches	11

3 Simulation	13
3.1 Binary Y, binary X	14
3.2 Binary Y, continuous X	26
3.3 Continuous Y, binary X	32
3.4 Continuous Y, continuous X	35
4 Results	38
4.1 Binary outcome setting	39
Binary predictors	39
Continuous predictors	39
Subsetted results	39
Bias	40
4.2 Continuous outcome setting	40
5 Discussion	41
5.1 Discussion	42
Points to address	42
Results we did not consider	42
6 Limitations	43
6.1 Simulation management	44
Issues	44
Lessons	46
6.2 Further development	47
7 Figures	49
7.1 Binary Y, binary X	50
7.2 Binary Y, continuous X	56

7.3 Continuous Y, binary X	62
7.4 Continuous Y, continuous X	65

List of Figures

7.1	Batch 1: Coverage probability	50
7.2	Batch 1 subset: Coverage probability	51
7.3	Batch 1: Median bias	52
7.4	Batch 1 subset: Median bias	53
7.5	Batch 1: Power	54
7.6	Batch 1 subset: Power	55
7.7	Batch 2: Coverage probability	56
7.8	Batch 2 subset: Coverage probability	57
7.9	Batch 2: Median bias	58
7.10	Batch 2 subset: Median bias	59
7.11	Batch 2: Power	60
7.12	Batch 2 subset: Power	61
7.13	Batch 3: Coverage probability	62
7.14	Batch 3: Median bias	63
7.15	Batch 3: Power	64
7.16	Batch 4: Coverage probability	65
7.17	Batch 4: Median bias	66
7.18	Batch 4: Power	67

List of Tables

3.1	Batch 1 (Binary Y, Binary X): Median bias	14
3.2	Batch 1 (Binary Y, Binary X): Power	16
3.3	Batch 1 (Binary Y, Binary X): Coverage probability	18
3.4	Batch 1 (Binary Y, Binary X): Median bias, subsetted	20
3.5	Batch 1 (Binary Y, Binary X): Power, subsetted	22
3.6	Batch 1 (Binary Y, Binary X): Coverage probability, subsetted	24
3.7	Batch 2 (Binary Y, Continuous X): Median bias	26
3.8	Batch 2 (Binary Y, Continuous X): Power	27
3.9	Batch 2 (Binary Y, Continuous X): Coverage probability	28
3.10	Batch 2 (Binary Y, Continuous X): Median bias, subsetted	29
3.11	Batch 2 (Binary Y, Continuous X): Power, subsetted	30
3.12	Batch 2 (Binary Y, Continuous X): Coverage probability, subsetted . . .	31
3.13	Batch 3 (Continuous Y, Binary X): Median bias	32
3.14	Batch 3 (Continuous Y, Binary X): Power	33
3.15	Batch 3 (Continuous Y, Binary X): Coverage probability	34
3.16	Batch 4 (Continuous Y, Continuous X): Median bias	35
3.17	Batch 4 (Continuous Y, Continuous X): Power	36
3.18	Batch 4 (Continuous Y, Continuous X): Coverage probability	37

1 *Background*

1.1 Randomization in clinical trials

Research to identify effective and/or efficacious interventions requires balance between scientific and logistic constraints. Randomization in clinical trials allows one to infer causation from associations in the presence of an appropriate experimental design, appreciating the limitations in identifying true causal relationships.

Oftentimes it is of interest to measure and adjust for known variables predictive of the outcome measurement (prognostic factors), to reduce confounding and to increase efficiency. On average, simple (or complete) randomization ensures the distribution of prognostic factors, measured and unmeasured, are balanced across groups. Chance imbalance in important prognostic factors may be seen as impacting the credibility of observed treatment effect estimates. The likelihood of imbalance increases as the sample size decreases or the number of prognostic variables increases. Especially in small studies, at any given point in the randomization process there could be a substantial imbalance in the number of patients assigned to each group. For this reason, restricted randomization techniques are often used to guarantee the sequential and overall imbalance in the number of participants in each group is controlled.

1.2 Restricted randomization

Blocked randomization is a restricted randomization approach that ensures current and overall imbalance in treatment assignments are controlled. Sequential balance is achieved by specifying a block size for which the sequence of (usually equal numbers of) treatment assignments is permuted and assigned to patients as they are enrolled. However, blocked randomization does not by itself guarantee overall balance in known prognostic variables of interest. For this reason, blocking is often combined with stratification to make study groups comparable with regard to specified stratifying

factors.

Stratified randomization is a procedure that separates the recruitment population into smaller subgroups (strata) where randomization, either simple or blocked, is performed. This property can be useful in multi-center trials, for instance, where it is of interest to account for between-center variability in patient outcomes due to unmeasured or unimportant factors *a priori*. While the randomization ratio is guaranteed to hold within pre-specified blocks of enrolled subjects, in small trials with many stratification factors one cannot assure accrued patients will fill the block for each subgroup, and randomization within strata alone will not ensure balance.

1.3 Dynamic (adaptive) randomization

Stratified block randomization is considered a static randomization method, as the probability of treatment assignment is not conditional on information on patients already enrolled. In contrast, adaptive (or dynamic) randomization approaches control imbalance by dynamically altering the randomization probability based on accrued patient information. In this thesis we consider covariate adaptive randomization procedures, which are a natural comparison to static randomization strategies intended to control imbalance of baseline prognostic factors across treatment groups. These procedures have been increasingly used as an alternative to stratified block randomization, particularly in small scale clinical trials with many prognostic factors.

Initial developments in covariate adaptive methods aimed to reduce the probability of undesirable, albeit unlikely, allocation sequences which result in both overall treatment group imbalances and imbalances within subgroups defined by important prognostic factors. In this subsection we follow the historical development of covariate adaptive approaches with a brief discussion of the characteristics and performance of a few

selected methods.

Biased coin randomization introduced by Efron (1971) was the first randomization method to change the probability of assignment dynamically based on observed covariate values of accrued patients. Simple randomization is performed until the disparity reaches a prespecified limit, at which time the group with the least subjects is biased to have a greater probability of assignment.

Taves (1974) extended Efron's biased coin design to the context of small scale clinical trials, where it is of interest to constrain imbalance in multiple prognostic factors across treatment groups. Briefly, the method sequentially allocates incoming patients deterministically to the treatment category that minimizes the overall unweighted sum of covariate imbalance given the new assignment. The assignment is performed deterministically: assignment is randomized only when assignment to either treatment category results in the same imbalance. Pocock and Simon (1975) further generalized Taves' method to incorporate relative importance of prognostic factors by introducing weighting of covariate imbalances into the overall imbalance metric.

Signorini et. al 1993 extended earlier methods in order to induce balance both overall and within strata while avoiding investigator bias through unblinding. He proposed a tree-based method of dynamic balancing randomization (DBR) that evaluates imbalance for each prognostic factor in a nested fashion by their prespecified order of importance. The method flexibly allows for different levels of imbalance in different strata and ensures conditional balance, meaning that within each subgroup the ratio of treatment assignment is constrained within prespecified bounds. However, the method does not guarantee balanced group assignments will be achieved within each prognostic factor considered separately.

Heritier et. al (2005) modify Signorini et. al's DBR method to control imbalance marginally within each prognostic factor. For each accrued patient, the potential

imbalance for each treatment assignment is considered sequentially within each prognostic factor in decreasing order of importance. If the potential observed imbalance exceeds a prespecified threshold, assignment is performed deterministically (or forced) to the group which minimizes the imbalance. Heritier et. al suggested including non-deterministic allocation to reduce the number of forced allocations and prevent investigator unblinding.

Model-based approaches are another alternative approach to dynamic randomization, where the probability of treatment assignment is chosen to minimize the variance of the estimated treatment effect. Model-based methods can flexibly incorporate continuous prognostic factors without the need to dichotomize into groups, and can include interaction terms and balance prognostic factors even when the number of variables is large. Aickin (1998, 2001, 2009) proposed a model-based approach to covariate adaptive randomization, where a subjects' treatment assignment is based on maximizing the log-likelihood of the model.

1.4 Analysis methods

Analysis of trials using a covariate adaptive allocation (CAA) scheme must account for the randomization scheme to recover the precision gains conferred by inducing more balanced treatment groups with respect to chosen balancing factors. To obtain the correct variance term and significance level for the test statistic, one must consider all possible sequences of assignments which could have been made in repeated trials assuming no group differences in mean response. In most cases, ignoring the randomization procedure and using standard regression methods that implicitly assume complete randomization lead to larger variance estimates and conservative inference. Since CAA modifies the randomization scheme to induce similarity across treatment arms

relative to within arms, the efficiency gain can be realized using a nonparametric re-randomization approach for estimating standard errors (Simon and Simon 2011). Briefly, observed values and entry order are fixed, treatment assignments are reshuffled and the test statistic computed for each permutation.

1.5 Aims

The goal of the thesis is to address in both the binary and continuous outcome setting if covariate adaptive randomization (CAR) followed by standard asymptotic tests yield valid inference, and if so, to quantify the gains in precision relative to simple randomization (SR) or stratified block randomization (SBR). We will compare Heritier's modified DBR scheme to stratified block randomization and complete randomization, while comparing re-randomization based permutation tests to standard asymptotic tests in a simulation study. We will consider the setting of equal allocation to treatment assignment, no temporal trend (drift), binary predictors, and two outcome types (binary and continuous). Our objective is to identify any scenarios, if any, where minimization improves power relative to SBR or simple randomization. Contour plots of effect size by sample size will compare power across methods for various outcome types and conditions.

We are also interested if and when the answers to the above questions change when the effect size of prognostic factors is varied relative to the treatment effect, the baseline prevalence varies from 5% to 50% in the binary outcome setting, inference on treatment effect is performed using none (or a subset) of the prognostic variables, and when the sample size is varied. It is well known that ignoring the minimization design tends to yield conservative inference, and that adjusting for covariates used in the randomization scheme (balancing factors) recovers type I error rate to nominal

significance levels (Xu, Proschan, Lee 2016). Through comparison of estimated marginal and conditional treatment effects we seek to confirm this finding. We consider different sample sizes ranging from $N=32$ to 108 to compare CAR to SBR as small scale trials are the setting in which alternative randomization methods are considered. We seek to identify the specific conditions by which CAR confers a precision advantage, if any, relative to other methods to offset the operational complexities involved in implementing an adaptive allocation procedure. Our intent is to provide guidance to clinical researchers for determining under what settings covariate adaptive allocation provides precision gains relative to competing approaches as well as which analysis method yields valid tests with the most power.

Chapter 2 will introduce the notation used throughout the thesis, and Chapter 3 will discuss the design of the simulation study in further detail. The tables of simulation results will be presented in Chapter 4 and the key observations will be discussed in Chapter 5.

1.6 Measures to evaluate aims

For each combination of randomization scheme and analysis approach, we assess validity by estimating the nominal significance level of the test under the null hypothesis. We evaluate accuracy by estimating any potential bias and the coverage probability of confidence intervals, comparing those generated with standard regression methods (Wald-type) to permutation test quantile-based confidence intervals. Average standard error estimates will also be reported for analyses using standard regression methods. We evaluate efficiency by computing mean squared error (MSE) and power as a function of the true treatment effect size.

2 *Methods*

2.1 Data generation

Simulations were conducted of a two-arm randomized clinical trial with equal allocation (1:1 treatment:control). We conducted simulations for both continuous and binary outcome and covariate types, and varied the overall trial size from 32 to 96.

The outcome measure (Y) was simulated with a marginal prevalence of 10% or 50% in the binary setting to evaluate the potential impact of low numbers of observed outcomes on inference. Continuous outcomes were simulated as normally distributed with constant variance, with mean as a linear combination of the treatment assignment (Z), pre-specified risk factors (X), and observed entry time (T).

Binary risk factors were simulated such that their marginal prevalence was either 25% or 50%. Continuous risk factors (X) were generated under a standard normal distribution. The risk factors are modeled as independent. The effect sizes for treatment assignment (Z) and prognostic variables (X) were separately varied from none, low, medium, and high. Balancing factors refer to risk factors used in adaptive allocation procedures, for which it is desired to have comparability either within or between treatment groups. The exact type of balance desired informs the choice of imbalance metric minimized at each sequential allocation step. For instance, it may be of interest to ensure within-strata subgroups have approximately proportional treatment and control assignments (conditional balance), or that treatment groups are otherwise comparable with respect to pre-specified balancing factors (marginal balance).

Observed patient entry times occurred following a uniform distribution. In subsequent re-randomization analysis, patient entry order is considered fixed (see Section 2aiv on re-randomization tests). Temporal trends in outcome prevalence over the course of a study (drift) were modeled by varying the entry time effect size from none, mild,

to severe. For the binary response setting under severe drift, for instance, the drift effect size was chosen such that marginal outcome prevalence varies three-fold over the study period.

2.2 Allocation procedures

For each simulated set of observed entry times and prognostic factors, treatment group assignments were determined using three allocation procedures: complete randomization, stratified permuted block randomization with fixed block sizes equal to overall trial size divided by number of strata (defined by all combinations of balancing factor levels), and an adapted form of covariate adaptive randomization proposed by Heritier et. al (2005).

For the covariate-adaptive randomization procedure, the maximum imbalance of treatment to control assignments (overall and within strata defined by each balancing factor level, considered separately) was set to 2. The allocation biasing probability, or the probability of assigning patient to treatment minimizing the imbalance measure when prospective imbalance meets or exceeds a prespecified threshold, was chosen as 0.7 to minimize the effect of non-deterministic and forced allocations on inference.

2.3 Varied conditions

The following table describes the conditions varied in the binary outcome setting. Note: in the continuous outcome setting, the effect sizes are modified to represent comparable differences in means to the given odds ratios.

2.4 Analysis approaches

For each allocation procedure we estimated the treatment effect, adjusted and unadjusted for balancing factors. To evaluate power, coverage, and type I error control, we report the associated linear and logistic regression model-based p-values, standard errors and Wald-type confidence intervals, based on a two-sided type I error threshold (alpha) of 0.05. For balancing purposes, continuous balancing factors were first dichotomized by their population median, and later parameterized as continuous in the adjusted analysis.

The simulation model is of the form

The adjusted regression model is

The unadjusted model

To compare the bias and power based on re-randomization analysis to that based on standard regression techniques, we compute power, coverage, and level and conducted re-randomization based inference for each simulated trial following covariate adaptive randomization. Re-randomization is a permutation-based method for estimating uncertainty and follows from the generally accepted sentiment to ‘analyze as you randomize’. The approach considers the outcomes (Y), balancing factors (X), and observed entry time (T) as fixed and repeats the allocation procedure multiple times, each generating a new sequence of treatment assignments (. Regression estimates are computed under each re-randomized treatment allocation sequence, and re-randomization based 95% confidence intervals are generated using the 2.5th and 97.5th quantile of the re-randomization-based treatment effect estimates. Re-randomization based p-values are estimated using the observed proportion of re-randomized allocation sequences yielding treatment effect estimates as or more extreme than the observed treatment effect.

Each simulation model configuration was simulated 10,000 trials, for which the re-randomization procedure was repeated 500 times.

3 *Simulation*

3.1 Batch 1: Binary Outcome, Binary Predictors

Table 3.1: Batch 1 (Binary Y, Binary X): Median bias

n	Pr(Y)	Pr(X)	bZ	bX	Model-based				Rerandomization		
					CR		SBR		CAA		CAA
					adj	unadj	adj	unadj	adj	unadj	adj
32	0.1	0.25	1.0	1.1	0.000	0.000	0.000	0.000	0.000	0.000	0.000
32	0.1	0.25	1.0	3.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
32	0.1	0.25	1.1	1.1	-0.019	-0.044	-0.095	-0.095	-0.065	-0.095	-0.065
32	0.1	0.25	1.1	3.0	-0.007	0.038	-0.045	-0.095	-0.028	-0.095	-0.028
32	0.1	0.25	3.0	1.1	0.297	0.167	0.196	0.143	0.237	0.143	0.237
32	0.1	0.25	3.0	3.0	0.316	0.065	0.227	0.000	0.282	0.000	0.282
32	0.1	0.50	1.0	1.1	0.000	0.000	0.000	0.000	0.000	0.000	0.000
32	0.1	0.50	1.0	3.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
32	0.1	0.50	1.1	1.1	-0.013	-0.065	-0.095	-0.095	-0.079	-0.095	-0.079
32	0.1	0.50	1.1	3.0	0.028	0.022	-0.011	-0.095	-0.031	-0.095	-0.031
32	0.1	0.50	3.0	1.1	0.289	0.143	0.173	0.065	0.284	0.143	0.284
32	0.1	0.50	3.0	3.0	0.283	0.065	0.222	0.059	0.229	0.059	0.229
32	0.5	0.25	1.0	1.1	0.012	0.000	0.000	0.000	0.000	0.000	0.000
32	0.5	0.25	1.0	3.0	0.001	0.000	0.000	0.000	-0.003	0.000	-0.003
32	0.5	0.25	1.1	1.1	0.021	-0.008	0.008	-0.047	0.011	-0.031	0.011
32	0.5	0.25	1.1	3.0	0.030	-0.026	0.009	-0.047	-0.006	-0.047	-0.006
32	0.5	0.25	3.0	1.1	0.131	0.031	0.145	0.031	0.128	-0.011	0.128
32	0.5	0.25	3.0	3.0	0.176	-0.049	0.117	-0.059	0.148	-0.059	0.148
32	0.5	0.50	1.0	1.1	-0.007	0.000	-0.008	0.000	0.012	0.000	0.012
32	0.5	0.50	1.0	3.0	-0.003	0.000	-0.014	0.000	0.000	0.000	0.000
32	0.5	0.50	1.1	1.1	-0.002	-0.031	-0.014	-0.047	0.012	-0.047	0.012
32	0.5	0.50	1.1	3.0	0.006	-0.047	-0.004	-0.047	-0.010	-0.080	-0.010
32	0.5	0.50	3.0	1.1	0.136	0.031	0.168	0.031	0.133	0.000	0.133
32	0.5	0.50	3.0	3.0	0.144	-0.087	0.153	-0.077	0.149	-0.077	0.149
96	0.1	0.25	1.0	1.1	0.004	0.000	0.001	0.000	0.002	0.000	0.002
96	0.1	0.25	1.0	3.0	0.001	0.000	-0.002	0.000	-0.005	0.000	-0.005
96	0.1	0.25	1.1	1.1	0.009	0.002	-0.004	-0.004	-0.007	-0.013	-0.007
96	0.1	0.25	1.1	3.0	-0.001	-0.004	-0.001	-0.002	0.006	-0.013	0.006
96	0.1	0.25	3.0	1.1	0.081	0.047	0.084	0.047	0.077	0.047	0.077
96	0.1	0.25	3.0	3.0	0.065	-0.045	0.071	-0.047	0.062	-0.047	0.062
96	0.1	0.50	1.0	1.1	0.007	0.000	0.005	0.000	0.000	0.000	0.000

96	0.1	0.50	1.0	3.0	-0.001	0.000	0.018	0.000	0.001	0.000	0.001
96	0.1	0.50	1.1	1.1	0.009	0.000	0.002	-0.002	-0.009	-0.035	-0.009
96	0.1	0.50	1.1	3.0	0.006	-0.013	0.017	0.002	0.011	0.000	0.011
96	0.1	0.50	3.0	1.1	0.085	0.054	0.070	0.047	0.093	0.054	0.093
96	0.1	0.50	3.0	3.0	0.062	-0.037	0.065	-0.036	0.077	-0.036	0.077
96	0.5	0.25	1.0	1.1	0.013	0.007	0.002	0.000	0.001	0.000	0.001
96	0.5	0.25	1.0	3.0	0.016	0.009	0.000	0.000	0.007	0.000	0.007
96	0.5	0.25	1.1	1.1	0.006	-0.001	-0.004	-0.011	-0.001	-0.011	-0.001
96	0.5	0.25	1.1	3.0	0.021	0.007	-0.002	-0.012	0.015	-0.010	0.015
96	0.5	0.25	3.0	1.1	0.039	0.013	0.037	0.013	0.038	0.013	0.038
96	0.5	0.25	3.0	3.0	0.046	-0.084	0.039	-0.078	0.036	-0.078	0.036
96	0.5	0.50	1.0	1.1	-0.008	-0.005	-0.005	0.000	0.005	0.000	0.005
96	0.5	0.50	1.0	3.0	-0.007	0.000	-0.008	0.000	0.005	0.000	0.005
96	0.5	0.50	1.1	1.1	-0.002	-0.008	-0.001	-0.009	0.001	-0.008	0.001
96	0.5	0.50	1.1	3.0	0.013	-0.007	0.011	-0.012	-0.002	-0.012	-0.002
96	0.5	0.50	3.0	1.1	0.042	0.013	0.041	0.017	0.037	0.013	0.037
96	0.5	0.50	3.0	3.0	0.037	-0.112	0.032	-0.143	0.041	-0.136	0.041

Table 3.2: Batch 1 (Binary Y, Binary X): Power

n	Pr(Y)	Pr(X)	bZ	bX	Model-based				Rerandomization		
					CR		SBR		CAA		CAA
					adj	unadj	adj	unadj	adj	unadj	adj
32	0.1	0.25	1.0	1.1	0.009	0.001	0.006	0.000	0.007	0.000	0.075
32	0.1	0.25	1.0	3.0	0.015	0.001	0.017	0.000	0.018	0.001	0.071
32	0.1	0.25	1.1	1.1	0.009	0.001	0.007	0.000	0.008	0.000	0.077
32	0.1	0.25	1.1	3.0	0.016	0.001	0.017	0.000	0.020	0.001	0.070
32	0.1	0.25	3.0	1.1	0.015	0.004	0.012	0.003	0.013	0.003	0.147
32	0.1	0.25	3.0	3.0	0.024	0.007	0.025	0.004	0.026	0.007	0.157
32	0.1	0.50	1.0	1.1	0.009	0.000	0.008	0.000	0.005	0.000	0.078
32	0.1	0.50	1.0	3.0	0.018	0.001	0.022	0.001	0.020	0.001	0.070
32	0.1	0.50	1.1	1.1	0.009	0.001	0.007	0.000	0.004	0.000	0.079
32	0.1	0.50	1.1	3.0	0.019	0.002	0.021	0.000	0.022	0.001	0.071
32	0.1	0.50	3.0	1.1	0.015	0.004	0.009	0.002	0.013	0.004	0.149
32	0.1	0.50	3.0	3.0	0.042	0.007	0.040	0.003	0.045	0.006	0.165
32	0.5	0.25	1.0	1.1	0.028	0.029	0.031	0.026	0.028	0.025	0.057
32	0.5	0.25	1.0	3.0	0.019	0.023	0.022	0.017	0.025	0.022	0.056
32	0.5	0.25	1.1	1.1	0.025	0.027	0.027	0.022	0.033	0.029	0.063
32	0.5	0.25	1.1	3.0	0.023	0.025	0.024	0.017	0.026	0.022	0.059
32	0.5	0.25	3.0	1.1	0.210	0.230	0.232	0.230	0.222	0.219	0.326
32	0.5	0.25	3.0	3.0	0.172	0.192	0.197	0.181	0.198	0.182	0.302
32	0.5	0.50	1.0	1.1	0.026	0.029	0.026	0.025	0.026	0.025	0.062
32	0.5	0.50	1.0	3.0	0.022	0.025	0.028	0.016	0.024	0.019	0.059
32	0.5	0.50	1.1	1.1	0.029	0.030	0.024	0.023	0.024	0.023	0.064
32	0.5	0.50	1.1	3.0	0.027	0.031	0.028	0.018	0.028	0.022	0.061
32	0.5	0.50	3.0	1.1	0.202	0.226	0.223	0.223	0.223	0.228	0.327
32	0.5	0.50	3.0	3.0	0.161	0.177	0.190	0.164	0.180	0.170	0.300
96	0.1	0.25	1.0	1.1	0.026	0.018	0.024	0.019	0.026	0.018	0.058
96	0.1	0.25	1.0	3.0	0.038	0.026	0.041	0.026	0.036	0.019	0.051
96	0.1	0.25	1.1	1.1	0.025	0.019	0.027	0.020	0.028	0.018	0.063
96	0.1	0.25	1.1	3.0	0.038	0.027	0.040	0.025	0.039	0.022	0.055
96	0.1	0.25	3.0	1.1	0.226	0.210	0.213	0.205	0.224	0.207	0.358
96	0.1	0.25	3.0	3.0	0.270	0.245	0.274	0.241	0.278	0.242	0.376
96	0.1	0.50	1.0	1.1	0.024	0.018	0.025	0.017	0.026	0.018	0.060
96	0.1	0.50	1.0	3.0	0.045	0.024	0.044	0.021	0.048	0.026	0.061
96	0.1	0.50	1.1	1.1	0.025	0.019	0.023	0.018	0.027	0.019	0.061
96	0.1	0.50	1.1	3.0	0.047	0.026	0.045	0.021	0.049	0.025	0.060

96	0.1	0.50	3.0	1.1	0.217	0.212	0.219	0.207	0.232	0.218	0.363
96	0.1	0.50	3.0	3.0	0.308	0.267	0.328	0.277	0.318	0.267	0.394
96	0.5	0.25	1.0	1.1	0.043	0.045	0.044	0.048	0.038	0.045	0.050
96	0.5	0.25	1.0	3.0	0.046	0.047	0.045	0.041	0.043	0.041	0.055
96	0.5	0.25	1.1	1.1	0.046	0.049	0.049	0.052	0.046	0.052	0.057
96	0.5	0.25	1.1	3.0	0.050	0.051	0.051	0.048	0.048	0.047	0.060
96	0.5	0.25	3.0	1.1	0.725	0.744	0.730	0.745	0.741	0.764	0.752
96	0.5	0.25	3.0	3.0	0.680	0.649	0.691	0.675	0.683	0.672	0.704
96	0.5	0.50	1.0	1.1	0.044	0.047	0.045	0.050	0.048	0.053	0.059
96	0.5	0.50	1.0	3.0	0.042	0.048	0.042	0.036	0.046	0.044	0.058
96	0.5	0.50	1.1	1.1	0.045	0.047	0.049	0.054	0.050	0.056	0.059
96	0.5	0.50	1.1	3.0	0.046	0.048	0.047	0.042	0.046	0.040	0.057
96	0.5	0.50	3.0	1.1	0.725	0.741	0.740	0.757	0.739	0.758	0.752
96	0.5	0.50	3.0	3.0	0.661	0.629	0.676	0.647	0.669	0.648	0.694

Table 3.3: Batch 1 (Binary Y, Binary X): Coverage probability

n	Pr(Y)	Pr(X)	bZ	bX	Model-based				Rerandomization		
					CR		SBR		CAA		CAA
					adj	unadj	adj	unadj	adj	unadj	adj
32	0.1	0.25	1.0	1.1	0.989	0.999	0.993	1.000	0.991	0.999	1.000
32	0.1	0.25	1.0	3.0	0.983	0.998	0.979	0.999	0.979	0.999	1.000
32	0.1	0.25	1.1	1.1	0.989	0.999	0.992	1.000	0.990	0.999	0.968
32	0.1	0.25	1.1	3.0	0.982	0.998	0.979	0.999	0.978	0.999	0.982
32	0.1	0.25	3.0	1.1	0.958	0.968	0.962	0.971	0.957	0.968	0.977
32	0.1	0.25	3.0	3.0	0.969	0.979	0.959	0.975	0.957	0.975	0.987
32	0.1	0.50	1.0	1.1	0.988	0.999	0.991	0.999	0.993	0.998	1.000
32	0.1	0.50	1.0	3.0	0.979	0.998	0.974	0.998	0.975	0.998	1.000
32	0.1	0.50	1.1	1.1	0.988	0.999	0.992	0.999	0.992	0.999	0.967
32	0.1	0.50	1.1	3.0	0.977	0.997	0.975	0.999	0.974	0.998	0.983
32	0.1	0.50	3.0	1.1	0.954	0.966	0.959	0.967	0.954	0.967	0.976
32	0.1	0.50	3.0	3.0	0.956	0.977	0.960	0.977	0.958	0.976	0.988
32	0.5	0.25	1.0	1.1	0.963	0.954	0.958	0.948	0.964	0.956	1.000
32	0.5	0.25	1.0	3.0	0.973	0.960	0.967	0.959	0.965	0.957	1.000
32	0.5	0.25	1.1	1.1	0.965	0.964	0.966	0.968	0.959	0.961	1.000
32	0.5	0.25	1.1	3.0	0.969	0.967	0.966	0.971	0.965	0.967	1.000
32	0.5	0.25	3.0	1.1	0.967	0.962	0.966	0.961	0.967	0.962	0.997
32	0.5	0.25	3.0	3.0	0.975	0.967	0.978	0.972	0.969	0.969	0.999
32	0.5	0.50	1.0	1.1	0.964	0.955	0.964	0.954	0.963	0.953	1.000
32	0.5	0.50	1.0	3.0	0.969	0.956	0.964	0.968	0.969	0.965	1.000
32	0.5	0.50	1.1	1.1	0.963	0.960	0.969	0.968	0.969	0.968	1.000
32	0.5	0.50	1.1	3.0	0.967	0.960	0.966	0.975	0.962	0.968	1.000
32	0.5	0.50	3.0	1.1	0.969	0.963	0.968	0.965	0.967	0.961	0.997
32	0.5	0.50	3.0	3.0	0.973	0.959	0.972	0.966	0.973	0.968	0.999
96	0.1	0.25	1.0	1.1	0.973	0.979	0.974	0.980	0.973	0.980	1.000
96	0.1	0.25	1.0	3.0	0.959	0.971	0.956	0.971	0.960	0.979	1.000
96	0.1	0.25	1.1	1.1	0.974	0.978	0.974	0.979	0.971	0.980	1.000
96	0.1	0.25	1.1	3.0	0.962	0.971	0.959	0.974	0.959	0.978	1.000
96	0.1	0.25	3.0	1.1	0.972	0.974	0.975	0.977	0.972	0.974	0.935
96	0.1	0.25	3.0	3.0	0.963	0.966	0.968	0.972	0.962	0.969	0.933
96	0.1	0.50	1.0	1.1	0.974	0.980	0.974	0.981	0.971	0.980	1.000
96	0.1	0.50	1.0	3.0	0.951	0.974	0.953	0.977	0.950	0.972	1.000
96	0.1	0.50	1.1	1.1	0.972	0.978	0.974	0.980	0.972	0.977	1.000
96	0.1	0.50	1.1	3.0	0.950	0.972	0.955	0.977	0.949	0.973	1.000

96	0.1	0.50	3.0	1.1	0.974	0.974	0.979	0.978	0.976	0.974	0.931
96	0.1	0.50	3.0	3.0	0.961	0.966	0.958	0.970	0.961	0.974	0.913
96	0.5	0.25	1.0	1.1	0.954	0.953	0.952	0.950	0.958	0.954	1.000
96	0.5	0.25	1.0	3.0	0.951	0.951	0.952	0.957	0.955	0.957	1.000
96	0.5	0.25	1.1	1.1	0.957	0.956	0.954	0.952	0.956	0.952	1.000
96	0.5	0.25	1.1	3.0	0.950	0.952	0.952	0.957	0.953	0.957	1.000
96	0.5	0.25	3.0	1.1	0.955	0.956	0.954	0.954	0.957	0.954	0.003
96	0.5	0.25	3.0	3.0	0.952	0.944	0.952	0.952	0.957	0.951	0.010
96	0.5	0.50	1.0	1.1	0.953	0.951	0.952	0.948	0.949	0.946	1.000
96	0.5	0.50	1.0	3.0	0.955	0.950	0.956	0.963	0.951	0.956	1.000
96	0.5	0.50	1.1	1.1	0.956	0.955	0.951	0.950	0.954	0.951	1.000
96	0.5	0.50	1.1	3.0	0.960	0.952	0.956	0.964	0.956	0.961	1.000
96	0.5	0.50	3.0	1.1	0.952	0.953	0.953	0.953	0.952	0.954	0.001
96	0.5	0.50	3.0	3.0	0.953	0.937	0.954	0.953	0.950	0.946	0.015

Table 3.4: Batch 1 (Binary Y, Binary X): Median bias, subsetted

Avg. nsims	n	Pr(Y)	Pr(X)	bZ	bX	Model-based				Rerandomization		
						CR		SBR		CAA		CAA
						adj	unadj	adj	unadj	adj	unadj	adj
3452 (68.9%)	32	0.1	0.25	1.0	1.1	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3818 (76.2%)	32	0.1	0.25	1.0	3.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3451 (68.9%)	32	0.1	0.25	1.1	1.1	-0.095	-0.095	-0.095	-0.095	-0.095	-0.095	-0.095
3804 (75.9%)	32	0.1	0.25	1.1	3.0	-0.078	-0.065	-0.095	-0.095	-0.080	-0.095	-0.080
3062 (61.1%)	32	0.1	0.25	3.0	1.1	-0.389	-0.435	-0.403	-0.470	-0.402	-0.405	-0.402
3408 (68%)	32	0.1	0.25	3.0	3.0	-0.235	-0.405	-0.278	-0.405	-0.256	-0.336	-0.256
3445 (68.8%)	32	0.1	0.50	1.0	1.1	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3868 (77.2%)	32	0.1	0.50	1.0	3.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3443 (68.7%)	32	0.1	0.50	1.1	1.1	-0.087	-0.095	-0.095	-0.095	-0.095	-0.095	-0.095
3879 (77.4%)	32	0.1	0.50	1.1	3.0	-0.028	-0.095	-0.074	-0.095	-0.068	-0.095	-0.068
3053 (60.9%)	32	0.1	0.50	3.0	1.1	-0.391	-0.435	-0.391	-0.434	-0.400	-0.470	-0.400
3482 (69.5%)	32	0.1	0.50	3.0	3.0	-0.223	-0.351	-0.218	-0.336	-0.234	-0.336	-0.234
5008 (100%)	32	0.5	0.25	1.0	1.1	0.012	0.000	0.000	0.000	0.000	0.000	0.000
5009 (100%)	32	0.5	0.25	1.0	3.0	0.001	0.000	0.000	0.000	-0.002	0.000	-0.002
5009 (100%)	32	0.5	0.25	1.1	1.1	0.021	-0.008	0.008	-0.047	0.011	-0.031	0.011
5009 (100%)	32	0.5	0.25	1.1	3.0	0.030	-0.026	0.009	-0.047	-0.006	-0.047	-0.006
5001 (99.8%)	32	0.5	0.25	3.0	1.1	0.128	0.015	0.144	0.031	0.126	-0.022	0.126
5001 (99.8%)	32	0.5	0.25	3.0	3.0	0.172	-0.053	0.116	-0.059	0.147	-0.059	0.147
5008 (100%)	32	0.5	0.50	1.0	1.1	-0.007	0.000	-0.008	0.000	0.012	0.000	0.012
5009 (100%)	32	0.5	0.50	1.0	3.0	-0.003	0.000	-0.014	0.000	0.000	0.000	0.000
5009 (100%)	32	0.5	0.50	1.1	1.1	-0.003	-0.031	-0.014	-0.047	0.012	-0.047	0.012
5009 (100%)	32	0.5	0.50	1.1	3.0	0.006	-0.047	-0.004	-0.047	-0.010	-0.080	-0.010
5000 (99.8%)	32	0.5	0.50	3.0	1.1	0.133	0.031	0.167	0.031	0.132	0.000	0.132
5004 (99.9%)	32	0.5	0.50	3.0	3.0	0.141	-0.087	0.153	-0.077	0.146	-0.077	0.146
4944 (98.7%)	96	0.1	0.25	1.0	1.1	0.003	0.000	0.000	0.000	0.000	0.000	0.000
4985 (99.5%)	96	0.1	0.25	1.0	3.0	0.001	0.000	-0.001	0.000	-0.003	0.000	-0.003
4939 (98.6%)	96	0.1	0.25	1.1	1.1	0.004	-0.002	-0.009	-0.006	-0.015	-0.047	-0.015
4983 (99.5%)	96	0.1	0.25	1.1	3.0	-0.001	-0.004	-0.002	-0.003	0.006	-0.013	0.006
4726 (94.3%)	96	0.1	0.25	3.0	1.1	0.032	0.000	0.019	0.000	0.018	0.000	0.018
4877 (97.3%)	96	0.1	0.25	3.0	3.0	0.038	-0.062	0.041	-0.070	0.036	-0.070	0.036
4950 (98.8%)	96	0.1	0.50	1.0	1.1	0.006	0.000	0.004	0.000	0.000	0.000	0.000
4992 (99.6%)	96	0.1	0.50	1.0	3.0	-0.002	-0.002	0.017	0.000	0.001	0.000	0.001
4947 (98.7%)	96	0.1	0.50	1.1	1.1	0.007	-0.003	-0.001	-0.002	-0.013	-0.047	-0.013
4991 (99.6%)	96	0.1	0.50	1.1	3.0	0.005	-0.013	0.014	0.000	0.011	0.000	0.011

4752 (94.9%)	96	0.1	0.50	3.0	1.1	0.034	0.000	0.021	0.000	0.044	0.013	0.044
4899 (97.8%)	96	0.1	0.50	3.0	3.0	0.044	-0.050	0.049	-0.047	0.060	-0.045	0.060
5010 (100%)	96	0.5	0.25	1.0	1.1	0.013	0.007	0.002	0.000	0.001	0.000	0.001
5010 (100%)	96	0.5	0.25	1.0	3.0	0.016	0.009	0.000	0.000	0.007	0.000	0.007
5010 (100%)	96	0.5	0.25	1.1	1.1	0.006	-0.001	-0.004	-0.011	-0.001	-0.011	-0.001
5010 (100%)	96	0.5	0.25	1.1	3.0	0.021	0.007	-0.002	-0.012	0.015	-0.010	0.015
5010 (100%)	96	0.5	0.25	3.0	1.1	0.039	0.013	0.037	0.013	0.038	0.013	0.038
5010 (100%)	96	0.5	0.25	3.0	3.0	0.046	-0.084	0.039	-0.078	0.036	-0.078	0.036
5010 (100%)	96	0.5	0.50	1.0	1.1	-0.008	-0.005	-0.005	0.000	0.005	0.000	0.005
5010 (100%)	96	0.5	0.50	1.0	3.0	-0.007	0.000	-0.008	0.000	0.005	0.000	0.005
5010 (100%)	96	0.5	0.50	1.1	1.1	-0.002	-0.008	-0.001	-0.009	0.001	-0.008	0.001
5010 (100%)	96	0.5	0.50	1.1	3.0	0.013	-0.007	0.011	-0.012	-0.002	-0.012	-0.002
5010 (100%)	96	0.5	0.50	3.0	1.1	0.042	0.013	0.041	0.017	0.037	0.013	0.037
5010 (100%)	96	0.5	0.50	3.0	3.0	0.037	-0.112	0.032	-0.143	0.041	-0.136	0.041

Table 3.5: Batch 1 (Binary Y, Binary X): Power, subsetted

Avg. nsims	n	Pr(Y)	Pr(X)	bZ	bX	Model-based				Rerandomization		
						CR		SBR		CAA		CAA
						adj	unadj	adj	unadj	adj	unadj	adj
3452 (68.9%)	32	0.1	0.25	1.0	1.1	0.006	0.001	0.003	0.000	0.003	0.001	0.017
3818 (76.2%)	32	0.1	0.25	1.0	3.0	0.008	0.001	0.007	0.000	0.009	0.001	0.025
3451 (68.9%)	32	0.1	0.25	1.1	1.1	0.007	0.001	0.003	0.000	0.003	0.000	0.017
3804 (75.9%)	32	0.1	0.25	1.1	3.0	0.008	0.001	0.008	0.001	0.010	0.001	0.024
3062 (61.1%)	32	0.1	0.25	3.0	1.1	0.018	0.007	0.012	0.005	0.012	0.005	0.044
3408 (68%)	32	0.1	0.25	3.0	3.0	0.024	0.010	0.022	0.006	0.025	0.010	0.065
3445 (68.8%)	32	0.1	0.50	1.0	1.1	0.009	0.001	0.009	0.000	0.005	0.000	0.021
3868 (77.2%)	32	0.1	0.50	1.0	3.0	0.021	0.002	0.025	0.001	0.022	0.001	0.022
3443 (68.7%)	32	0.1	0.50	1.1	1.1	0.009	0.001	0.008	0.001	0.004	0.000	0.021
3879 (77.4%)	32	0.1	0.50	1.1	3.0	0.021	0.002	0.024	0.001	0.024	0.001	0.022
3053 (60.9%)	32	0.1	0.50	3.0	1.1	0.019	0.007	0.013	0.003	0.018	0.006	0.046
3482 (69.5%)	32	0.1	0.50	3.0	3.0	0.055	0.010	0.052	0.004	0.057	0.009	0.068
5008 (100%)	32	0.5	0.25	1.0	1.1	0.028	0.029	0.031	0.026	0.028	0.025	0.057
5009 (100%)	32	0.5	0.25	1.0	3.0	0.019	0.023	0.022	0.017	0.025	0.022	0.056
5009 (100%)	32	0.5	0.25	1.1	1.1	0.025	0.027	0.027	0.022	0.033	0.029	0.063
5009 (100%)	32	0.5	0.25	1.1	3.0	0.023	0.025	0.024	0.017	0.026	0.022	0.059
5001 (99.8%)	32	0.5	0.25	3.0	1.1	0.210	0.231	0.232	0.230	0.222	0.219	0.326
5001 (99.8%)	32	0.5	0.25	3.0	3.0	0.173	0.192	0.197	0.182	0.199	0.182	0.301
5008 (100%)	32	0.5	0.50	1.0	1.1	0.026	0.029	0.026	0.025	0.026	0.025	0.062
5009 (100%)	32	0.5	0.50	1.0	3.0	0.022	0.025	0.028	0.016	0.024	0.019	0.059
5009 (100%)	32	0.5	0.50	1.1	1.1	0.029	0.030	0.024	0.023	0.024	0.023	0.064
5009 (100%)	32	0.5	0.50	1.1	3.0	0.027	0.031	0.028	0.018	0.028	0.022	0.061
5000 (99.8%)	32	0.5	0.50	3.0	1.1	0.203	0.227	0.223	0.224	0.223	0.228	0.326
5004 (99.9%)	32	0.5	0.50	3.0	3.0	0.161	0.177	0.190	0.164	0.180	0.170	0.299
4944 (98.7%)	96	0.1	0.25	1.0	1.1	0.026	0.018	0.024	0.019	0.026	0.018	0.050
4985 (99.5%)	96	0.1	0.25	1.0	3.0	0.039	0.026	0.041	0.026	0.036	0.019	0.049
4939 (98.6%)	96	0.1	0.25	1.1	1.1	0.026	0.019	0.027	0.020	0.028	0.018	0.054
4983 (99.5%)	96	0.1	0.25	1.1	3.0	0.038	0.027	0.040	0.025	0.039	0.022	0.052
4726 (94.3%)	96	0.1	0.25	3.0	1.1	0.238	0.222	0.227	0.218	0.238	0.220	0.324
4877 (97.3%)	96	0.1	0.25	3.0	3.0	0.277	0.252	0.282	0.248	0.285	0.248	0.360
4950 (98.8%)	96	0.1	0.50	1.0	1.1	0.024	0.018	0.025	0.017	0.027	0.018	0.053
4992 (99.6%)	96	0.1	0.50	1.0	3.0	0.045	0.024	0.044	0.021	0.048	0.026	0.058
4947 (98.7%)	96	0.1	0.50	1.1	1.1	0.025	0.019	0.023	0.018	0.027	0.019	0.054
4991 (99.6%)	96	0.1	0.50	1.1	3.0	0.047	0.026	0.045	0.021	0.049	0.025	0.057

4752 (94.9%)	96	0.1	0.50	3.0	1.1	0.230	0.224	0.231	0.218	0.244	0.230	0.334
4899 (97.8%)	96	0.1	0.50	3.0	3.0	0.316	0.274	0.334	0.282	0.325	0.273	0.381
5010 (100%)	96	0.5	0.25	1.0	1.1	0.043	0.045	0.044	0.048	0.038	0.045	0.050
5010 (100%)	96	0.5	0.25	1.0	3.0	0.046	0.047	0.045	0.041	0.043	0.041	0.055
5010 (100%)	96	0.5	0.25	1.1	1.1	0.046	0.049	0.049	0.052	0.046	0.052	0.057
5010 (100%)	96	0.5	0.25	1.1	3.0	0.050	0.051	0.051	0.048	0.048	0.047	0.060
5010 (100%)	96	0.5	0.25	3.0	1.1	0.725	0.744	0.730	0.745	0.741	0.764	0.752
5010 (100%)	96	0.5	0.25	3.0	3.0	0.680	0.649	0.691	0.675	0.683	0.672	0.704
5010 (100%)	96	0.5	0.50	1.0	1.1	0.044	0.047	0.045	0.050	0.048	0.053	0.059
5010 (100%)	96	0.5	0.50	1.0	3.0	0.042	0.048	0.042	0.036	0.046	0.044	0.058
5010 (100%)	96	0.5	0.50	1.1	1.1	0.045	0.047	0.049	0.054	0.050	0.056	0.059
5010 (100%)	96	0.5	0.50	1.1	3.0	0.046	0.048	0.047	0.042	0.046	0.040	0.057
5010 (100%)	96	0.5	0.50	3.0	1.1	0.725	0.741	0.740	0.757	0.739	0.758	0.752
5010 (100%)	96	0.5	0.50	3.0	3.0	0.661	0.629	0.676	0.647	0.669	0.648	0.694

Table 3.6: Batch 1 (Binary Y, Binary X): Coverage probability, subsetted

Avg. nsims	n	Pr(Y)	Pr(X)	bZ	bX	Model-based				Rerandomization		
						CR		SBR		CAA		CAA
						adj	unadj	adj	unadj	adj	unadj	adj
3452 (68.9%)	32	0.1	0.25	1.0	1.1	0.991	0.999	0.996	1.000	0.994	0.999	1.000
3818 (76.2%)	32	0.1	0.25	1.0	3.0	0.989	0.998	0.989	0.999	0.988	0.998	1.000
3451 (68.9%)	32	0.1	0.25	1.1	1.1	0.991	0.999	0.995	1.000	0.995	0.999	0.954
3804 (75.9%)	32	0.1	0.25	1.1	3.0	0.989	0.998	0.988	0.999	0.988	0.999	0.977
3062 (61.1%)	32	0.1	0.25	3.0	1.1	0.936	0.947	0.946	0.954	0.937	0.948	0.962
3408 (68%)	32	0.1	0.25	3.0	3.0	0.966	0.969	0.954	0.964	0.950	0.964	0.981
3445 (68.8%)	32	0.1	0.50	1.0	1.1	0.986	0.998	0.989	0.999	0.991	0.997	1.000
3868 (77.2%)	32	0.1	0.50	1.0	3.0	0.975	0.997	0.970	0.998	0.971	0.997	1.000
3443 (68.7%)	32	0.1	0.50	1.1	1.1	0.987	0.999	0.989	0.999	0.991	0.999	0.953
3879 (77.4%)	32	0.1	0.50	1.1	3.0	0.973	0.996	0.971	0.999	0.970	0.998	0.979
3053 (60.9%)	32	0.1	0.50	3.0	1.1	0.928	0.943	0.935	0.947	0.928	0.946	0.961
3482 (69.5%)	32	0.1	0.50	3.0	3.0	0.943	0.966	0.947	0.968	0.946	0.966	0.984
5008 (100%)	32	0.5	0.25	1.0	1.1	0.963	0.954	0.958	0.948	0.964	0.956	1.000
5009 (100%)	32	0.5	0.25	1.0	3.0	0.973	0.960	0.967	0.959	0.965	0.957	1.000
5009 (100%)	32	0.5	0.25	1.1	1.1	0.965	0.964	0.966	0.968	0.959	0.961	1.000
5009 (100%)	32	0.5	0.25	1.1	3.0	0.969	0.967	0.966	0.971	0.965	0.967	1.000
5001 (99.8%)	32	0.5	0.25	3.0	1.1	0.967	0.962	0.966	0.961	0.967	0.962	0.997
5001 (99.8%)	32	0.5	0.25	3.0	3.0	0.975	0.967	0.978	0.972	0.969	0.969	0.999
5008 (100%)	32	0.5	0.50	1.0	1.1	0.964	0.955	0.964	0.954	0.963	0.953	1.000
5009 (100%)	32	0.5	0.50	1.0	3.0	0.969	0.956	0.964	0.968	0.969	0.965	1.000
5009 (100%)	32	0.5	0.50	1.1	1.1	0.963	0.960	0.969	0.968	0.969	0.968	1.000
5009 (100%)	32	0.5	0.50	1.1	3.0	0.967	0.960	0.966	0.975	0.962	0.968	1.000
5000 (99.8%)	32	0.5	0.50	3.0	1.1	0.969	0.963	0.968	0.965	0.967	0.961	0.997
5004 (99.9%)	32	0.5	0.50	3.0	3.0	0.973	0.959	0.972	0.966	0.973	0.968	0.999
4944 (98.7%)	96	0.1	0.25	1.0	1.1	0.973	0.979	0.973	0.979	0.972	0.980	1.000
4985 (99.5%)	96	0.1	0.25	1.0	3.0	0.959	0.971	0.956	0.971	0.960	0.979	1.000
4939 (98.6%)	96	0.1	0.25	1.1	1.1	0.974	0.978	0.973	0.978	0.971	0.980	1.000
4983 (99.5%)	96	0.1	0.25	1.1	3.0	0.962	0.971	0.959	0.974	0.959	0.978	1.000
4726 (94.3%)	96	0.1	0.25	3.0	1.1	0.971	0.973	0.974	0.976	0.971	0.972	0.932
4877 (97.3%)	96	0.1	0.25	3.0	3.0	0.962	0.965	0.967	0.971	0.961	0.968	0.932
4950 (98.8%)	96	0.1	0.50	1.0	1.1	0.973	0.980	0.973	0.980	0.971	0.980	1.000
4992 (99.6%)	96	0.1	0.50	1.0	3.0	0.951	0.974	0.953	0.977	0.950	0.972	1.000
4947 (98.7%)	96	0.1	0.50	1.1	1.1	0.972	0.977	0.974	0.979	0.972	0.977	1.000
4991 (99.6%)	96	0.1	0.50	1.1	3.0	0.950	0.972	0.955	0.977	0.949	0.973	1.000

4752 (94.9%)	96	0.1	0.50	3.0	1.1	0.972	0.973	0.978	0.977	0.975	0.973	0.928
4899 (97.8%)	96	0.1	0.50	3.0	3.0	0.960	0.965	0.958	0.969	0.960	0.973	0.911
5010 (100%)	96	0.5	0.25	1.0	1.1	0.954	0.953	0.952	0.950	0.958	0.954	1.000
5010 (100%)	96	0.5	0.25	1.0	3.0	0.951	0.951	0.952	0.957	0.955	0.957	1.000
5010 (100%)	96	0.5	0.25	1.1	1.1	0.957	0.956	0.954	0.952	0.956	0.952	1.000
5010 (100%)	96	0.5	0.25	1.1	3.0	0.950	0.952	0.952	0.957	0.953	0.957	1.000
5010 (100%)	96	0.5	0.25	3.0	1.1	0.955	0.956	0.954	0.954	0.957	0.954	0.003
5010 (100%)	96	0.5	0.25	3.0	3.0	0.952	0.944	0.952	0.952	0.957	0.951	0.010
5010 (100%)	96	0.5	0.50	1.0	1.1	0.953	0.951	0.952	0.948	0.949	0.946	1.000
5010 (100%)	96	0.5	0.50	1.0	3.0	0.955	0.950	0.956	0.963	0.951	0.956	1.000
5010 (100%)	96	0.5	0.50	1.1	1.1	0.956	0.955	0.951	0.950	0.954	0.951	1.000
5010 (100%)	96	0.5	0.50	1.1	3.0	0.960	0.952	0.956	0.964	0.956	0.961	1.000
5010 (100%)	96	0.5	0.50	3.0	1.1	0.952	0.953	0.953	0.953	0.952	0.954	0.001
5010 (100%)	96	0.5	0.50	3.0	3.0	0.953	0.937	0.954	0.953	0.950	0.946	0.015

3.2 Batch 2: Binary Outcome, Continuous Predictors

Table 3.7: Batch 2 (Binary Y, Continuous X): Median bias

n	Pr(Y)	bZ	bX	Model-based				Rerandomization		
				CR		SBR		CAA		CAA
				adj	unadj	adj	unadj	adj	unadj	adj
32	0.1	1.0	1.1	0.000	0.000	0.000	0.000	0.000	—	0.000
32	0.1	1.0	3.0	0.014	0.000	-0.019	0.000	0.047	—	0.047
32	0.1	1.1	1.1	-0.048	-0.065	-0.065	-0.095	0.005	—	0.005
32	0.1	1.1	3.0	0.031	-0.049	0.007	-0.095	0.044	—	0.044
32	0.1	3.0	1.1	0.354	0.143	0.311	0.143	0.366	—	0.366
32	0.1	3.0	3.0	0.332	-0.198	0.332	-0.236	0.354	—	0.354
32	0.5	1.0	1.1	0.012	0.000	-0.014	0.000	-0.012	—	-0.012
32	0.5	1.0	3.0	0.017	0.000	-0.002	0.000	0.016	—	0.016
32	0.5	1.1	1.1	0.011	-0.026	0.010	-0.031	0.001	—	0.001
32	0.5	1.1	3.0	0.047	-0.031	0.040	-0.080	0.021	—	0.021
32	0.5	3.0	1.1	0.144	0.036	0.175	0.041	0.152	—	0.152
32	0.5	3.0	3.0	0.158	-0.310	0.194	-0.325	0.165	—	0.165
96	0.1	1.0	1.1	0.014	0.000	0.003	0.000	0.002	—	0.002
96	0.1	1.0	3.0	0.013	0.000	-0.004	0.000	0.013	—	0.013
96	0.1	1.1	1.1	0.010	0.004	0.002	-0.002	0.015	—	0.015
96	0.1	1.1	3.0	0.015	-0.026	-0.001	-0.044	0.027	—	0.027
96	0.1	3.0	1.1	0.076	0.039	0.090	0.047	0.075	—	0.075
96	0.1	3.0	3.0	0.075	-0.251	0.078	-0.260	0.078	—	0.078
96	0.5	1.0	1.1	0.009	0.000	-0.014	-0.002	0.001	—	0.001
96	0.5	1.0	3.0	-0.001	0.000	-0.006	0.000	0.000	—	0.000
96	0.5	1.1	1.1	0.013	-0.001	-0.014	-0.012	0.013	—	0.013
96	0.5	1.1	3.0	0.008	-0.015	0.000	-0.018	0.005	—	0.005
96	0.5	3.0	1.1	0.037	0.011	0.026	0.013	0.033	—	0.033
96	0.5	3.0	3.0	0.057	-0.323	0.034	-0.336	0.061	—	0.061

Table 3.8: Batch 2 (Binary Y, Continuous X): Power

n	Pr(Y)	bZ	bX	Model-based				Rerandomization		
				CR		SBR		CAA		CAA
				adj	unadj	adj	unadj	adj	unadj	adj
32	0.1	1.0	1.1	0.068	0.001	0.073	0.000	0.068	—	0.061
32	0.1	1.0	3.0	0.126	0.004	0.126	0.002	0.125	—	0.061
32	0.1	1.1	1.1	0.067	0.001	0.074	0.000	0.069	—	0.062
32	0.1	1.1	3.0	0.124	0.004	0.123	0.003	0.128	—	0.060
32	0.1	3.0	1.1	0.067	0.006	0.068	0.003	0.067	—	0.137
32	0.1	3.0	3.0	0.177	0.021	0.178	0.012	0.179	—	0.146
32	0.5	1.0	1.1	0.030	0.028	0.029	0.026	0.027	—	0.054
32	0.5	1.0	3.0	0.050	0.029	0.046	0.013	0.046	—	0.054
32	0.5	1.1	1.1	0.028	0.028	0.026	0.023	0.027	—	0.057
32	0.5	1.1	3.0	0.051	0.029	0.050	0.014	0.050	—	0.055
32	0.5	3.0	1.1	0.219	0.235	0.242	0.235	0.216	—	0.289
32	0.5	3.0	3.0	0.181	0.116	0.192	0.097	0.174	—	0.210
96	0.1	1.0	1.1	0.025	0.018	0.022	0.016	0.026	—	0.063
96	0.1	1.0	3.0	0.065	0.032	0.057	0.024	0.062	—	0.061
96	0.1	1.1	1.1	0.028	0.019	0.023	0.017	0.027	—	0.068
96	0.1	1.1	3.0	0.065	0.035	0.063	0.026	0.064	—	0.064
96	0.1	3.0	1.1	0.215	0.204	0.218	0.214	0.209	—	0.346
96	0.1	3.0	3.0	0.390	0.272	0.400	0.268	0.392	—	0.398
96	0.5	1.0	1.1	0.045	0.049	0.046	0.053	0.043	—	0.063
96	0.5	1.0	3.0	0.049	0.047	0.050	0.035	0.046	—	0.065
96	0.5	1.1	1.1	0.047	0.050	0.050	0.055	0.053	—	0.073
96	0.5	1.1	3.0	0.056	0.053	0.055	0.037	0.051	—	0.069
96	0.5	3.0	1.1	0.718	0.727	0.721	0.742	0.714	—	0.718
96	0.5	3.0	3.0	0.568	0.447	0.560	0.442	0.562	—	0.563

Table 3.9: Batch 2 (Binary Y, Continuous X): Coverage probability

n	Pr(Y)	bZ	bX	Model-based				Rerandomization		
				CR		SBR		CAA		CAA
				adj	unadj	adj	unadj	adj	unadj	adj
32	0.1	1.0	1.1	0.930	0.999	0.924	1.000	0.929	—	1.000
32	0.1	1.0	3.0	0.862	0.993	0.860	0.997	0.863	—	1.000
32	0.1	1.1	1.1	0.932	0.999	0.925	1.000	0.930	—	0.966
32	0.1	1.1	3.0	0.867	0.994	0.865	0.996	0.861	—	0.997
32	0.1	3.0	1.1	0.913	0.970	0.909	0.970	0.910	—	0.975
32	0.1	3.0	3.0	0.868	0.967	0.872	0.978	0.870	—	0.999
32	0.5	1.0	1.1	0.960	0.956	0.958	0.948	0.965	—	1.000
32	0.5	1.0	3.0	0.940	0.955	0.941	0.971	0.943	—	1.000
32	0.5	1.1	1.1	0.963	0.962	0.965	0.968	0.963	—	1.000
32	0.5	1.1	3.0	0.938	0.960	0.939	0.978	0.941	—	1.000
32	0.5	3.0	1.1	0.966	0.964	0.966	0.963	0.966	—	0.994
32	0.5	3.0	3.0	0.935	0.940	0.941	0.961	0.934	—	0.999
96	0.1	1.0	1.1	0.973	0.980	0.976	0.981	0.971	—	1.000
96	0.1	1.0	3.0	0.931	0.968	0.939	0.975	0.933	—	1.000
96	0.1	1.1	1.1	0.972	0.980	0.976	0.982	0.970	—	1.000
96	0.1	1.1	3.0	0.934	0.967	0.937	0.973	0.933	—	1.000
96	0.1	3.0	1.1	0.978	0.978	0.980	0.978	0.976	—	0.897
96	0.1	3.0	3.0	0.935	0.927	0.934	0.933	0.932	—	0.861
96	0.5	1.0	1.1	0.951	0.947	0.949	0.945	0.954	—	1.000
96	0.5	1.0	3.0	0.947	0.951	0.945	0.963	0.951	—	1.000
96	0.5	1.1	1.1	0.954	0.951	0.954	0.951	0.949	—	1.000
96	0.5	1.1	3.0	0.943	0.951	0.943	0.963	0.953	—	1.000
96	0.5	3.0	1.1	0.951	0.952	0.954	0.954	0.951	—	0.081
96	0.5	3.0	3.0	0.947	0.879	0.946	0.891	0.944	—	0.385

Table 3.10: Batch 2 (Binary Y, Continuous X): Median bias, subsetted

Avg. nsims	n	Pr(Y)	bZ	bX	Model-based				Rerandomization		
					CR		SBR		CAA		
					adj	unadj	adj	unadj	adj	unadj	
3472 (69.3%)	32	0.1	1.0	1.1	0.000	0.000	0.000	0.000	0.000	–	0.000
4336 (86.5%)	32	0.1	1.0	3.0	0.016	0.000	-0.021	0.000	0.044	–	0.044
3464 (69.1%)	32	0.1	1.1	1.1	-0.095	-0.095	-0.095	-0.095	-0.070	–	-0.070
4334 (86.5%)	32	0.1	1.1	3.0	0.012	-0.049	-0.013	-0.095	0.019	–	0.019
3049 (60.9%)	32	0.1	3.0	1.1	-0.349	-0.435	-0.336	-0.405	-0.346	–	-0.346
4138 (82.6%)	32	0.1	3.0	3.0	0.044	-0.336	0.091	-0.336	0.041	–	0.041
5008 (100%)	32	0.5	1.0	1.1	0.012	0.000	-0.014	0.000	-0.013	–	-0.013
4999 (99.8%)	32	0.5	1.0	3.0	0.017	0.000	-0.002	0.000	0.016	–	0.016
5010 (100%)	32	0.5	1.1	1.1	0.011	-0.026	0.010	-0.031	0.001	–	0.001
4997 (99.7%)	32	0.5	1.1	3.0	0.047	-0.031	0.041	-0.080	0.021	–	0.021
4999 (99.8%)	32	0.5	3.0	1.1	0.141	0.031	0.173	0.036	0.150	–	0.150
4984 (99.5%)	32	0.5	3.0	3.0	0.151	-0.310	0.189	-0.325	0.157	–	0.157
4935 (98.5%)	96	0.1	1.0	1.1	0.014	0.000	0.004	0.000	-0.001	–	-0.001
5009 (100%)	96	0.1	1.0	3.0	0.013	0.000	-0.004	0.000	0.013	–	0.013
4930 (98.4%)	96	0.1	1.1	1.1	0.007	0.001	-0.002	-0.004	0.010	–	0.010
5009 (100%)	96	0.1	1.1	3.0	0.015	-0.026	-0.001	-0.044	0.027	–	0.027
4747 (94.8%)	96	0.1	3.0	1.1	0.026	0.000	0.040	0.013	0.019	–	0.019
5001 (99.8%)	96	0.1	3.0	3.0	0.073	-0.260	0.077	-0.260	0.077	–	0.077
5010 (100%)	96	0.5	1.0	1.1	0.009	0.000	-0.014	-0.002	0.001	–	0.001
5010 (100%)	96	0.5	1.0	3.0	-0.001	0.000	-0.006	0.000	0.000	–	0.000
5010 (100%)	96	0.5	1.1	1.1	0.013	-0.001	-0.014	-0.012	0.013	–	0.013
5010 (100%)	96	0.5	1.1	3.0	0.008	-0.015	0.000	-0.018	0.005	–	0.005
5010 (100%)	96	0.5	3.0	1.1	0.037	0.011	0.026	0.013	0.033	–	0.033
5010 (100%)	96	0.5	3.0	3.0	0.057	-0.323	0.034	-0.336	0.061	–	0.061

Table 3.11: Batch 2 (Binary Y, Continuous X): Power, subsetted

Avg. nsims	n	Pr(Y)	bZ	bX	Model-based				Rerandomization		
					CR		SBR		CAA		
					adj	unadj	adj	unadj	adj	unadj	
3472 (69.3%)	32	0.1	1.0	1.1	0.011	0.001	0.011	0.000	0.012	—	0.014
4336 (86.5%)	32	0.1	1.0	3.0	0.058	0.004	0.060	0.002	0.059	—	0.027
3464 (69.1%)	32	0.1	1.1	1.1	0.011	0.001	0.012	0.000	0.013	—	0.014
4334 (86.5%)	32	0.1	1.1	3.0	0.057	0.005	0.060	0.003	0.063	—	0.027
3049 (60.9%)	32	0.1	3.0	1.1	0.025	0.009	0.019	0.005	0.028	—	0.037
4138 (82.6%)	32	0.1	3.0	3.0	0.127	0.025	0.131	0.014	0.134	—	0.088
5008 (100%)	32	0.5	1.0	1.1	0.030	0.028	0.029	0.026	0.027	—	0.054
4999 (99.8%)	32	0.5	1.0	3.0	0.047	0.029	0.043	0.013	0.044	—	0.054
5010 (100%)	32	0.5	1.1	1.1	0.028	0.028	0.026	0.023	0.027	—	0.057
4997 (99.7%)	32	0.5	1.1	3.0	0.048	0.029	0.048	0.014	0.047	—	0.054
4999 (99.8%)	32	0.5	3.0	1.1	0.219	0.235	0.242	0.235	0.216	—	0.288
4984 (99.5%)	32	0.5	3.0	3.0	0.175	0.116	0.190	0.097	0.169	—	0.206
4935 (98.5%)	96	0.1	1.0	1.1	0.025	0.018	0.022	0.017	0.027	—	0.054
5009 (100%)	96	0.1	1.0	3.0	0.065	0.032	0.057	0.024	0.062	—	0.061
4930 (98.4%)	96	0.1	1.1	1.1	0.028	0.019	0.023	0.018	0.027	—	0.058
5009 (100%)	96	0.1	1.1	3.0	0.066	0.035	0.063	0.026	0.064	—	0.064
4747 (94.8%)	96	0.1	3.0	1.1	0.227	0.215	0.229	0.225	0.221	—	0.313
5001 (99.8%)	96	0.1	3.0	3.0	0.391	0.272	0.400	0.269	0.392	—	0.397
5010 (100%)	96	0.5	1.0	1.1	0.045	0.049	0.046	0.053	0.043	—	0.063
5010 (100%)	96	0.5	1.0	3.0	0.049	0.047	0.050	0.035	0.046	—	0.065
5010 (100%)	96	0.5	1.1	1.1	0.047	0.050	0.050	0.055	0.053	—	0.073
5010 (100%)	96	0.5	1.1	3.0	0.056	0.053	0.055	0.037	0.051	—	0.069
5010 (100%)	96	0.5	3.0	1.1	0.718	0.727	0.721	0.742	0.714	—	0.718
5010 (100%)	96	0.5	3.0	3.0	0.568	0.447	0.560	0.442	0.562	—	0.563

Table 3.12: Batch 2 (Binary Y, Continuous X): Coverage probability, subsetted

Avg. nsims	n	Pr(Y)	bZ	bX	Model-based				Rerandomization		
					CR		SBR		CAA		
					adj	unadj	adj	unadj	adj	unadj	
3472 (69.3%)	32	0.1	1.0	1.1	0.986	0.998	0.984	0.999	0.985	—	1.000
4336 (86.5%)	32	0.1	1.0	3.0	0.927	0.992	0.923	0.997	0.927	—	1.000
3464 (69.1%)	32	0.1	1.1	1.1	0.987	0.998	0.986	1.000	0.985	—	0.950
4334 (86.5%)	32	0.1	1.1	3.0	0.932	0.993	0.927	0.996	0.925	—	0.997
3049 (60.9%)	32	0.1	3.0	1.1	0.941	0.950	0.943	0.952	0.934	—	0.959
4138 (82.6%)	32	0.1	3.0	3.0	0.929	0.961	0.929	0.974	0.928	—	0.998
5008 (100%)	32	0.5	1.0	1.1	0.960	0.956	0.958	0.948	0.965	—	1.000
4999 (99.8%)	32	0.5	1.0	3.0	0.943	0.955	0.944	0.971	0.945	—	1.000
5010 (100%)	32	0.5	1.1	1.1	0.963	0.962	0.965	0.968	0.963	—	1.000
4997 (99.7%)	32	0.5	1.1	3.0	0.941	0.960	0.942	0.978	0.944	—	1.000
4999 (99.8%)	32	0.5	3.0	1.1	0.966	0.964	0.966	0.963	0.966	—	0.994
4984 (99.5%)	32	0.5	3.0	3.0	0.941	0.940	0.944	0.961	0.939	—	0.999
4935 (98.5%)	96	0.1	1.0	1.1	0.972	0.980	0.976	0.981	0.970	—	1.000
5009 (100%)	96	0.1	1.0	3.0	0.931	0.968	0.939	0.975	0.933	—	1.000
4930 (98.4%)	96	0.1	1.1	1.1	0.972	0.980	0.976	0.982	0.970	—	1.000
5009 (100%)	96	0.1	1.1	3.0	0.934	0.967	0.937	0.973	0.933	—	1.000
4747 (94.8%)	96	0.1	3.0	1.1	0.977	0.977	0.979	0.977	0.975	—	0.893
5001 (99.8%)	96	0.1	3.0	3.0	0.935	0.927	0.933	0.933	0.932	—	0.861
5010 (100%)	96	0.5	1.0	1.1	0.951	0.947	0.949	0.945	0.954	—	1.000
5010 (100%)	96	0.5	1.0	3.0	0.947	0.951	0.945	0.963	0.951	—	1.000
5010 (100%)	96	0.5	1.1	1.1	0.954	0.951	0.954	0.951	0.949	—	1.000
5010 (100%)	96	0.5	1.1	3.0	0.943	0.951	0.943	0.963	0.953	—	1.000
5010 (100%)	96	0.5	3.0	1.1	0.951	0.952	0.954	0.954	0.951	—	0.081
5010 (100%)	96	0.5	3.0	3.0	0.947	0.879	0.946	0.891	0.944	—	0.385

3.3 Batch 3: Continuous Outcome, Binary Predictors

Table 3.13: Batch 3 (Continuous Y, Binary X): Median bias

n	Pr(X)	bZ	bX	Model-based				Rerandomization		
				CR		SBR		CAA		CAA
				adj	unadj	adj	unadj	adj	unadj	adj
32	0.25	1	1.1	0.005	0.005	0.009	0.008	0.003	-0.003	0.003
32	0.25	1	3.0	0.005	-0.003	0.009	-0.003	0.003	-0.089	0.003
32	0.25	3	1.1	0.005	0.005	0.009	0.008	0.003	-0.003	0.003
32	0.25	3	3.0	0.005	-0.003	0.009	-0.003	0.003	-0.089	0.003
32	0.50	1	1.1	0.005	0.006	0.003	-0.003	0.004	0.005	0.004
32	0.50	1	3.0	0.005	-0.006	0.003	0.001	0.004	0.045	0.004
32	0.50	3	1.1	0.005	0.006	0.003	-0.003	0.004	0.005	0.004
32	0.50	3	3.0	0.005	-0.006	0.003	0.001	0.004	0.045	0.004
96	0.25	1	1.1	-0.002	-0.006	0.001	0.000	-0.002	-0.010	-0.002
96	0.25	1	3.0	-0.002	-0.007	0.001	0.001	-0.002	-0.142	-0.002
96	0.25	3	1.1	-0.002	-0.006	0.001	0.000	-0.002	-0.010	-0.002
96	0.25	3	3.0	-0.002	-0.007	0.001	0.001	-0.002	-0.142	-0.002
96	0.50	1	1.1	-0.004	-0.004	0.001	0.002	-0.002	-0.004	-0.002
96	0.50	1	3.0	-0.004	0.001	0.001	0.003	-0.002	0.018	-0.002
96	0.50	3	1.1	-0.004	-0.004	0.001	0.002	-0.002	-0.004	-0.002
96	0.50	3	3.0	-0.004	0.001	0.001	0.003	-0.002	0.018	-0.002

Table 3.14: Batch 3 (Continuous Y, Binary X): Power

Table 3.15: Batch 3 (Continuous Y, Binary X): Coverage probability

n	Pr(X)	bZ	bX	Model-based				Rerandomization		
				CR		SBR		CAA		
				adj	unadj	adj	unadj	adj	unadj	
32	0.25	1	1.1	0.937	0.937	0.933	0.934	0.938	0.940	0.999
32	0.25	1	3.0	0.937	0.945	0.933	0.982	0.938	0.931	0.999
32	0.25	3	1.1	0.937	0.937	0.933	0.934	0.938	0.940	0.090
32	0.25	3	3.0	0.937	0.945	0.933	0.982	0.938	0.931	0.090
32	0.50	1	1.1	0.935	0.935	0.941	0.941	0.934	0.935	1.000
32	0.50	1	3.0	0.935	0.941	0.941	0.996	0.934	0.928	1.000
32	0.50	3	1.1	0.935	0.935	0.941	0.941	0.934	0.935	0.001
32	0.50	3	3.0	0.935	0.941	0.941	0.996	0.934	0.928	0.001
96	0.25	1	1.1	0.948	0.950	0.950	0.948	0.951	0.947	0.999
96	0.25	1	3.0	0.948	0.943	0.950	0.987	0.951	0.924	0.999
96	0.25	3	1.1	0.948	0.950	0.950	0.948	0.951	0.947	0.005
96	0.25	3	3.0	0.948	0.943	0.950	0.987	0.951	0.924	0.005
96	0.50	1	1.1	0.948	0.947	0.955	0.957	0.951	0.951	1.000
96	0.50	1	3.0	0.948	0.942	0.955	0.997	0.951	0.901	1.000
96	0.50	3	1.1	0.948	0.947	0.955	0.957	0.951	0.951	0.000
96	0.50	3	3.0	0.948	0.942	0.955	0.997	0.951	0.901	0.000

3.4 Batch 4: Continuous Outcome, Continuous Predictors

Table 3.16: Batch 4 (Continuous Y, Continuous X): Median bias

n	bZ	bX	Model-based				Rerandomization		
			CR		SBR		CAA		CAA
			adj	unadj	adj	unadj	adj	unadj	adj
32	1.0	1.1	-0.013	-0.009	0.002	-0.006	-0.013	-0.009	-0.013
32	1.0	3.0	-0.013	-0.019	0.002	-0.002	-0.013	-0.019	-0.013
32	1.1	1.1	-0.013	-0.009	0.002	-0.006	-0.013	-0.009	-0.013
32	1.1	3.0	-0.013	-0.019	0.002	-0.002	-0.013	-0.019	-0.013
32	3.0	1.1	-0.013	-0.009	0.002	-0.006	-0.013	-0.009	-0.013
32	3.0	3.0	-0.013	-0.019	0.002	-0.002	-0.013	-0.019	-0.013
96	1.0	1.1	-0.003	-0.004	0.002	0.003	-0.003	-0.004	-0.003
96	1.0	3.0	-0.003	-0.013	0.002	-0.002	-0.003	-0.013	-0.003
96	1.1	1.1	-0.003	-0.004	0.002	0.003	-0.003	-0.004	-0.003
96	1.1	3.0	-0.003	-0.013	0.002	-0.002	-0.003	-0.013	-0.003
96	3.0	1.1	-0.003	-0.004	0.002	0.003	-0.003	-0.004	-0.003
96	3.0	3.0	-0.003	-0.013	0.002	-0.002	-0.003	-0.013	-0.003

Table 3.17: Batch 4 (Continuous Y, Continuous X): Power

n	bZ	bX	Model-based				Rerandomization		
			CR		SBR		CAA		CAA
			adj	unadj	adj	unadj	adj	unadj	adj
32	1.0	1.1	0.051	0.048	0.052	0.054	0.051	0.048	0.054
32	1.0	3.0	0.051	0.049	0.052	0.007	0.051	0.049	0.054
32	1.1	1.1	0.060	0.059	0.061	0.062	0.060	0.059	0.058
32	1.1	3.0	0.060	0.051	0.061	0.008	0.060	0.051	0.058
32	3.0	1.1	0.999	0.999	0.999	0.999	0.999	0.999	0.999
32	3.0	3.0	0.999	0.435	0.999	0.419	0.999	0.435	0.999
96	1.0	1.1	0.052	0.051	0.049	0.049	0.052	0.051	0.054
96	1.0	3.0	0.052	0.049	0.049	0.005	0.052	0.049	0.054
96	1.1	1.1	0.079	0.079	0.074	0.074	0.079	0.079	0.084
96	1.1	3.0	0.079	0.052	0.074	0.006	0.079	0.052	0.084
96	3.0	1.1	1.000	1.000	1.000	1.000	1.000	1.000	1.000
96	3.0	3.0	1.000	0.895	1.000	0.961	1.000	0.895	1.000

Table 3.18: Batch 4 (Continuous Y, Continuous X): Coverage probability

n	bZ	bX	Model-based				Rerandomization			
			CR		SBR		CAA		CAA	
			adj	unadj	adj	unadj	adj	unadj	adj	adj
32	1.0	1.1	0.938	0.941	0.939	0.937	0.938	0.941	1	
32	1.0	3.0	0.938	0.940	0.939	0.991	0.938	0.940	1	
32	1.1	1.1	0.938	0.941	0.939	0.937	0.938	0.941	1	
32	1.1	3.0	0.938	0.940	0.939	0.991	0.938	0.940	1	
32	3.0	1.1	0.938	0.941	0.939	0.937	0.938	0.941	0	
32	3.0	3.0	0.938	0.940	0.939	0.991	0.938	0.940	0	
96	1.0	1.1	0.946	0.947	0.949	0.948	0.946	0.947	1	
96	1.0	3.0	0.946	0.949	0.949	0.994	0.946	0.949	1	
96	1.1	1.1	0.946	0.947	0.949	0.948	0.946	0.947	1	
96	1.1	3.0	0.946	0.949	0.949	0.994	0.946	0.949	1	
96	3.0	1.1	0.946	0.947	0.949	0.948	0.946	0.947	0	
96	3.0	3.0	0.946	0.949	0.949	0.994	0.946	0.949	0	

4 *Results*

4.1 Binary outcome setting

Binary predictors

Median bias (Table 3.1) is small if the estimated treatment effect is null or when trial size is large ($n=96$).

Comparing these results to those subsetting on `glm()` convergence and at least one outcome occurring in each treatment group, median bias is larger in small sample size settings (see Table 3.4).

Subsetting on these two criterion led to excluding greater than 30 percent of simulations in low sample size scenarios.

Adjusting for prognostic factors used in the balancing procedure consistently resulted in increases

Table 3.2 Table 3.5 Table 3.3 Table 3.6

Continuous predictors

Table 3.1 Table 3.4

Table 3.2 Table 3.5 Table 3.3 Table 3.6

Subsetted results

Simulations with low observed event rates could result in large estimated treatment effects, often corresponding to cases where the `glm()` algorithm did not converge. For this reason, subsetted results are presented for simulation outcomes where both the `glm()` algorithm converged and at least one outcome was observed in each treatment arm.

This issue occurs more often when sample size is small ($n=32$), outcome prevalence ($Pr(Y)$) is small, and for small treatment and prognostic factor effect sizes (bZ and bX , respectively).

Bias

Median bias (Table 3.1) is small if the estimated treatment effect is null, or when sample size is large ($n=96$). Table 3.4

4.2 Continuous outcome setting

5 *Discussion*

5.1 Discussion

Points to address

Validity of treatment effect estimates on adjustment for balancing factors

As the literature notes,

Testing for PFs and validity the validity of the test (aka type I error control, or having the observed size be under the type I error threshold) is only achieved through adjusting for the variables included in the analysis

Results we did not consider

We did not examine the following:

- subset of PFs and validity The effect of adjusting for only a subset of prognostic factors, but more importantly, having prognostic factors that are not used in the balancing process.
- dichotomizing into groups We did not assess the impact of using the continuous values for adjustment in the analysis process, even though we had to dichotomize
- drift T The effect of drift β_T on bias and validity of treatment effect estimates. We did not examine the interaction between changes in block size or maximum allowed imbalance on this effect.
- max imbalance MI The effect of allocation method parameters on (unmeasured) metrics of 'balance' and variance of treatment effect estimates.
- of reference distribution We did not consider the number of re-randomizations to effectively model

6 *Limitations*

Due to many unforeseen factors, we ran into these complications:

6.1 Simulation management

For each simulation, we had to keep track of the following:

- Prognostic factors matrix \mathbf{X}
- Entry time vector \mathbf{T}
- Allocation sequence vectors \mathbf{Z} , each with their associated
- Outcome measure vector \mathbf{Y}
- Model-based regression estimates:
 - Estimate $\hat{\beta}$
 - Standard error $\hat{SE}(\hat{\beta})$
 - P-value p_T
 - T-statistic t_{obs}
 - Confidence interval $\hat{CI}(\hat{\beta})$
- Re-randomization based regression estimates, including:
 - Re-randomized allocation sequence vectors \mathbf{Z} , each with their associated estimate $\hat{\beta}$

We structured the simulation around using the ”**simulator**” R package, which handles file management (saving/loading data), RNG seeds for parallel processing, and sequence of simulation steps.

Issues

Memory management was poor. The simulator stored data as so:

- each **Model** object contains the simulation parameters

- each **Draw** object contains the prognostic factor and entry times (**X**, **T**) for all subjects
- the first **Output** object contains the allocation sequence vector and outcome vector (**Z**, **Y**) for all subjects
- the second **Output** object contains the model-based regression estimates ($\hat{\beta}$, $\hat{SE}(\hat{\beta})$, p_T , t_{obs} , $\hat{CI}(\hat{\beta})$)
- the third **Output** object contains the re-randomization based regression estimates ($\hat{\beta}$, $\hat{SE}(\hat{\beta})$, p_T , t_{obs} , $\hat{CI}(\hat{\beta})$)

Each 'look' at the data took lots of time, and (to this day) we have not figured out how to efficiently run the code. Granted, the task was computationally expensive: on a grid of models, for each simulation, perform 500 re-allocations and estimate the parameters, and then compute summary statistics on the re-randomized procedures?

At this point, we had already invested significantly into the "simulator" package to consider alternative approaches.

Computing metrics was not as straightforward as the package authors made it seem.

We computed metrics on the output, but loading the required **Output** objects into memory was too computationally costly.

We ended up running a separate R script that loaded each "Output" object individually, computed the metrics we wanted, and then saving both the output as .csv and metrics as .csv files.

Memory management The problem posed challenges for both memory management, and code performance as the number of simulations grew.

Lessons

Know when to use packages, and when to make your own code. At first, using the "simulator" package was great because it handled the RNG generation, file management, and forced us to delineate the steps of the simulation process.

Where using the "simulator" package fell short was when issues arose and we wanted to diagnose problems in the code, as well as work on memory and performance speedups.

The file sizes were extremely large (even though they were stored as .Rdata files). The output of all functions had to be in list() format.

Create a roadmap for programming, including which tools (packages) you're planning to use. We originally set forth to plan each simulation separately, but due to concerns with memory management (and to a lesser extent, with handling RNGs under parallelization) we decided on using the 'simulator' package.

This proved to be costly to port the code as well as to learn the structure of a new package. Downstream, it cost more time to diagnose issues, as well as increased the difficulty for optimizing the code.

Properly define the scope of your project. We had to pare down the questions we were going to ask, because the scope of the questions we sought to ask was rapidly growing in complexity. We wanted to consider the impact of

- a time-dependent shift in outcome measure or prevalence (drift) by modeling entry time \mathbf{T} .
- allocation method parameters, including
 - block size \mathbf{B} for SBR and its' impact on controlling for drift,

- maximum imbalance bounds **MI** for CAA procedures and its' impact on inference metrics
-

6.2 Further development

Package for implementing rerandomization analysis It would be useful to address the gap between analysis methods proposed in the literature prior to advances in computing that make rerandomization an attractive approach for estimating uncertainty.

Visualizing the restrictions to allocation sequence hyperplane Part of the difficulty in understanding the effect of covariate adaptive allocation procedures on the validity (in the philosophical sense) of results is the disconnect between the trade-off between forcing/inducing balance in key prognostic factors across treatment groups and the implications for how 'random' we can consider the treatment assignment variable.

When stepping away from complete randomization to restricted randomization approaches, one loses the ability to treat all subjects' treatment assignments as independent. The validity of the test (hypothesis test rejection proportion under the null hypothesis, and whether it is below the prespecified type I error threshold) is recovered only when the test is adjusted for the prognostic factors used in the balancing process.

A visualization of all possible randomization schemes (perhaps colored based on the likelihood of observing each sequence) would be useful to:

1. Understand the impact of restricted randomization methods on restricting the size of the reference distribution,

2. Show that certain sequences are 'impossible' under restricted randomization methods,
3. Show the properties of SBR (i.e. allocation proportion equals allocation ratio after each filled block)
4. Shows control of balance sequentially, one advantage to such procedures.

Model-based approaches We did not consider a class of covariate-adaptive allocation methods that posit a model for the relationship between prognostic factors, treatment assignment, and outcome measure in order to bias treatment assignment towards sequences that minimize the variance of the estimated treatment effect. As noted in the optimal ratio of prognostic factors within subgroups is unlikely to be 50:50, meaning the ratio for minimizing the variance of the adjusted treatment effect depends on the strength of the prognostic factor on the outcome (the odds ratio).

7 *Figures*

7.1 Batch 1: Binary Outcome, Binary Predictors

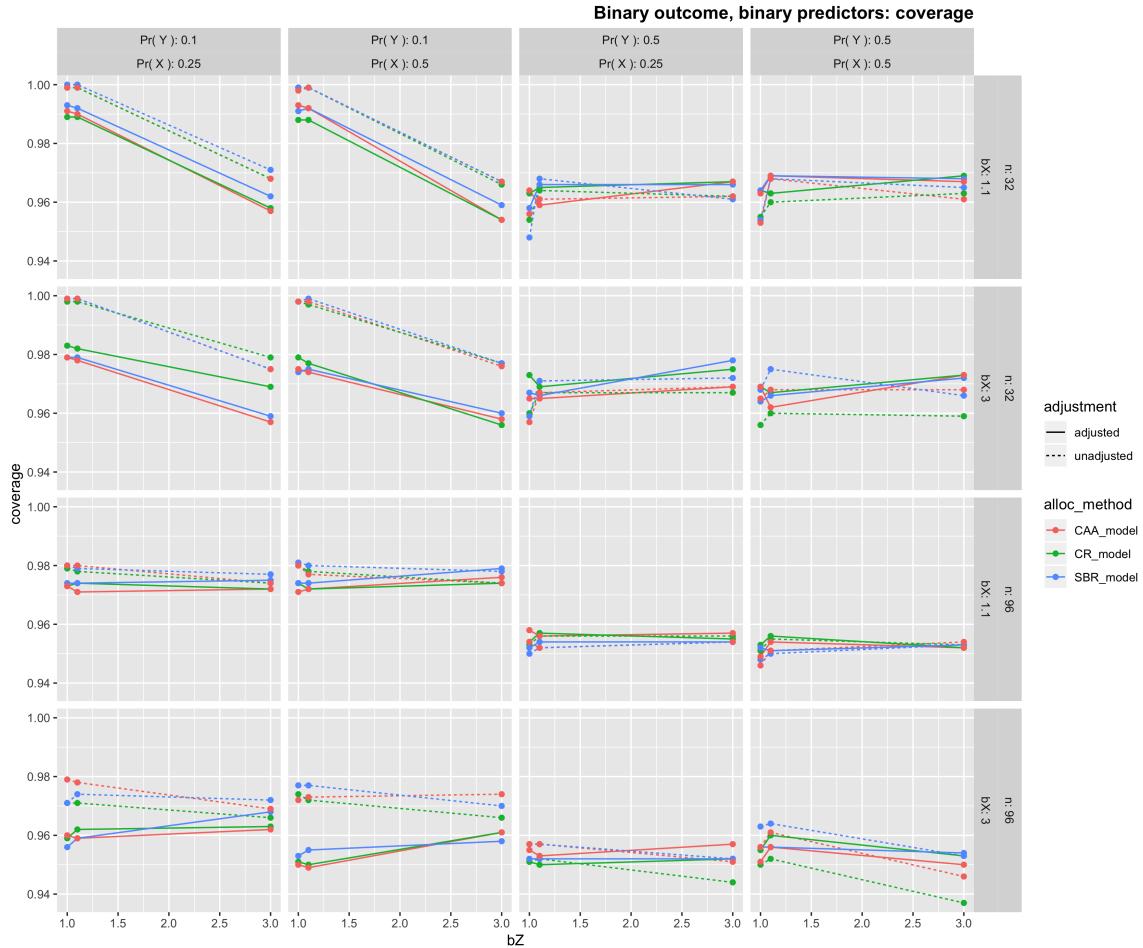


Figure 7.1: Batch 1: Coverage probability

Coverage probability nears 100 percent at low outcome prevalence and low trial size.

Coverage probability nears 95 percent when outcome prevalence is high and when sample size is large (n=96).

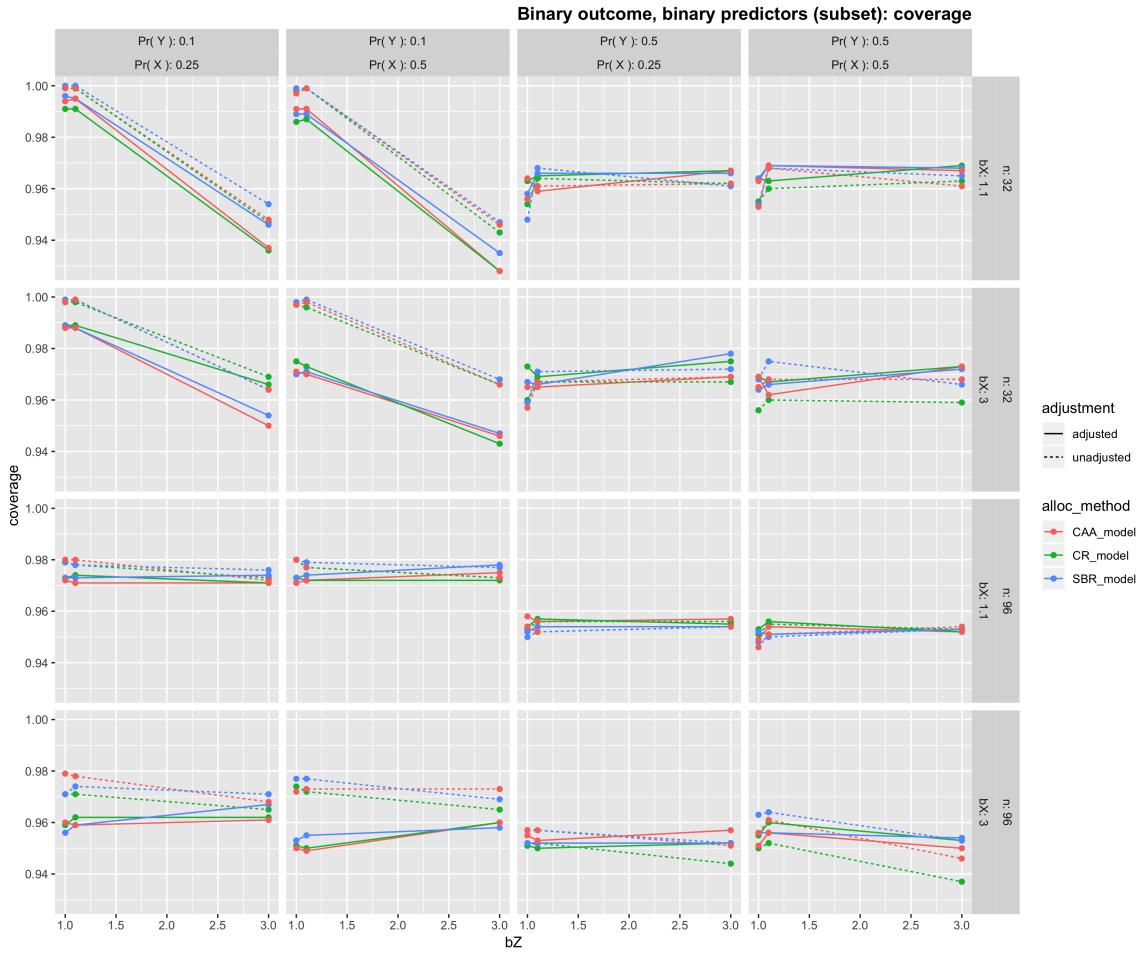


Figure 7.2: Batch 1 subset: Coverage probability

Subsetting did not substantially change the observed patterns in coverage probability over outcome prevalence, prognostic factor prevalence, sample size, or prognostic factor effect size.

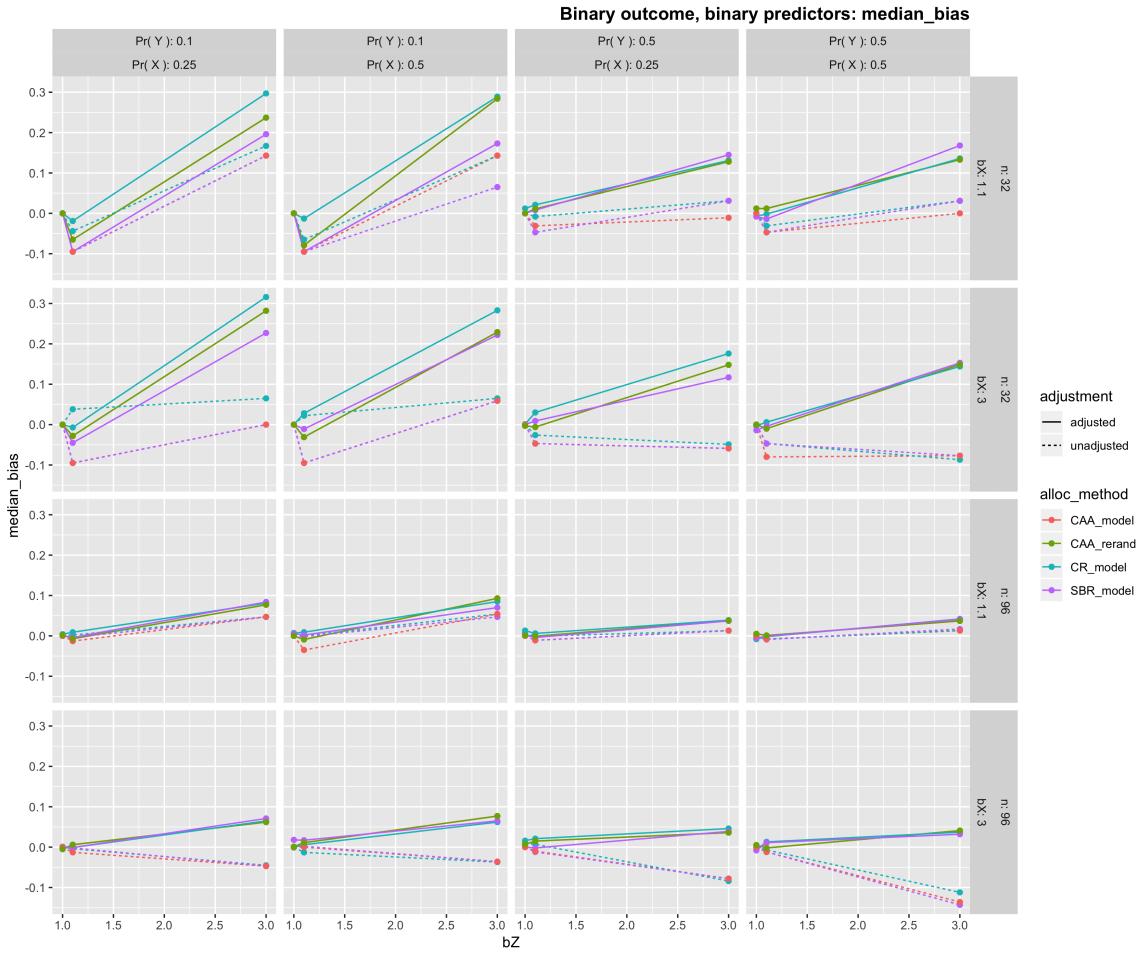


Figure 7.3: Batch 1: Median bias

Median bias of estimates remains near 0 when sample size is large ($n=96$) for both adjusted and unadjusted estimates. The unadjusted estimates become more biased as prognostic factor effect size increases.

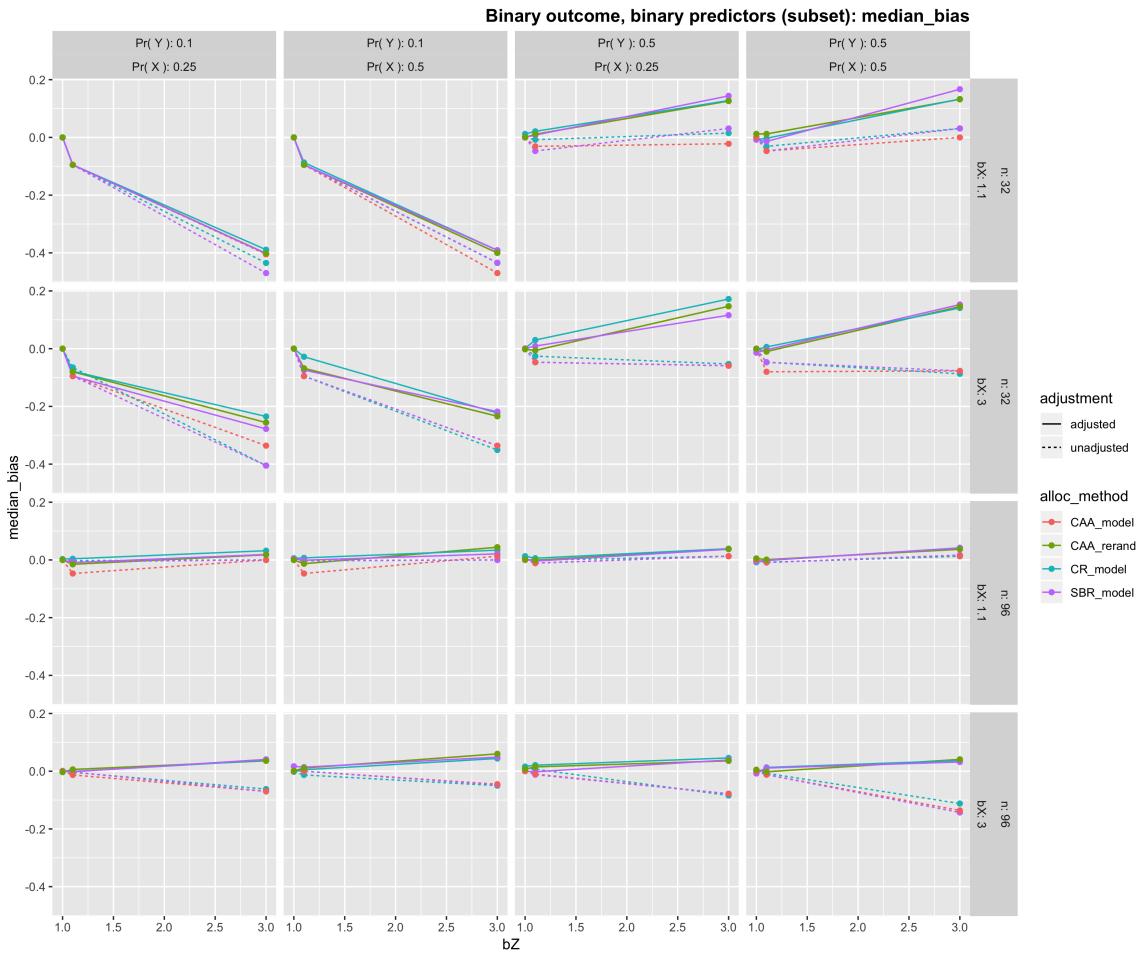


Figure 7.4: Batch 1 subset: Median bias

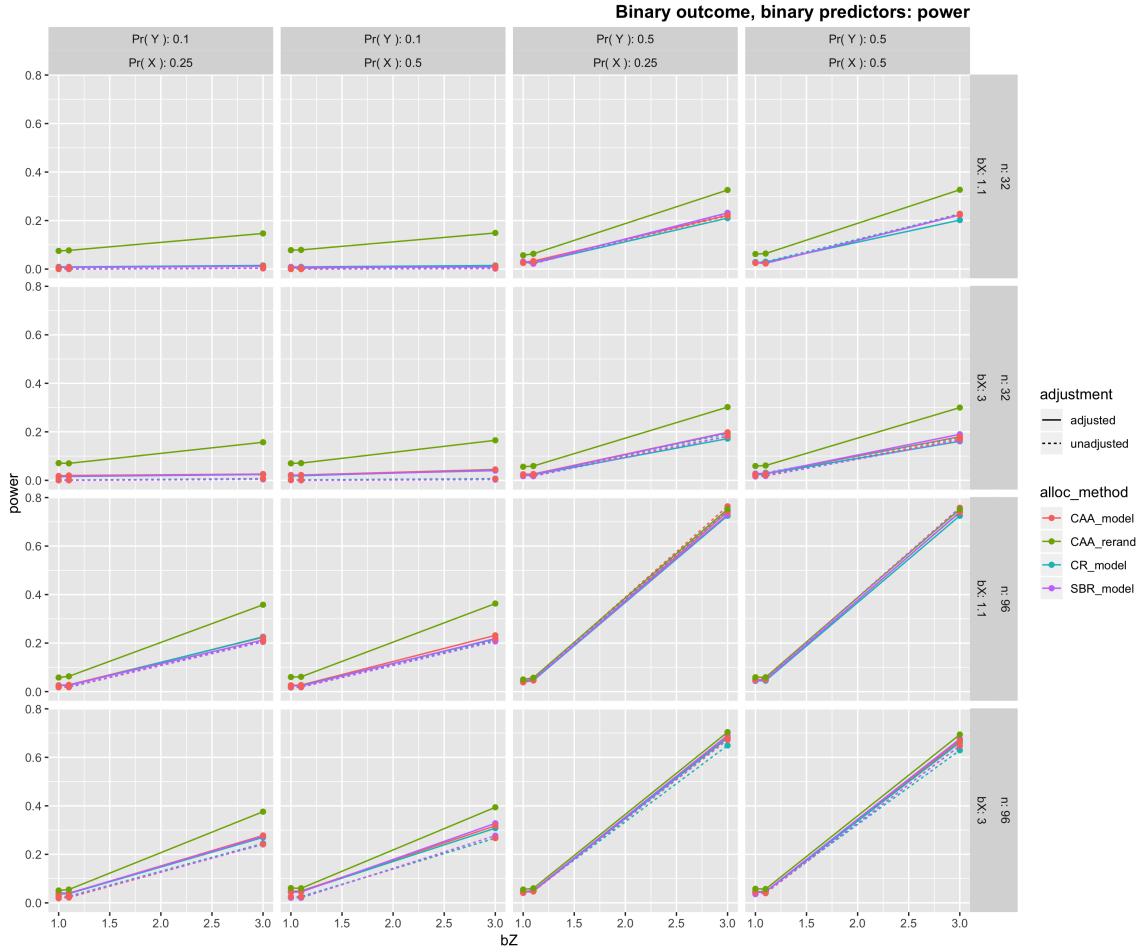


Figure 7.5: Batch 1: Power

Power is substantially higher using CAA followed by rerandomization when compared to all other methods, including CAA followed by model-based inference. The modest increase in power is seen in low sample size settings and low marginal outcome prevalence. The trends in power for detecting the unadjusted estimate do not appear to be affected by either prognostic factor prevalence ($\Pr(X)$) or effect size (bX).

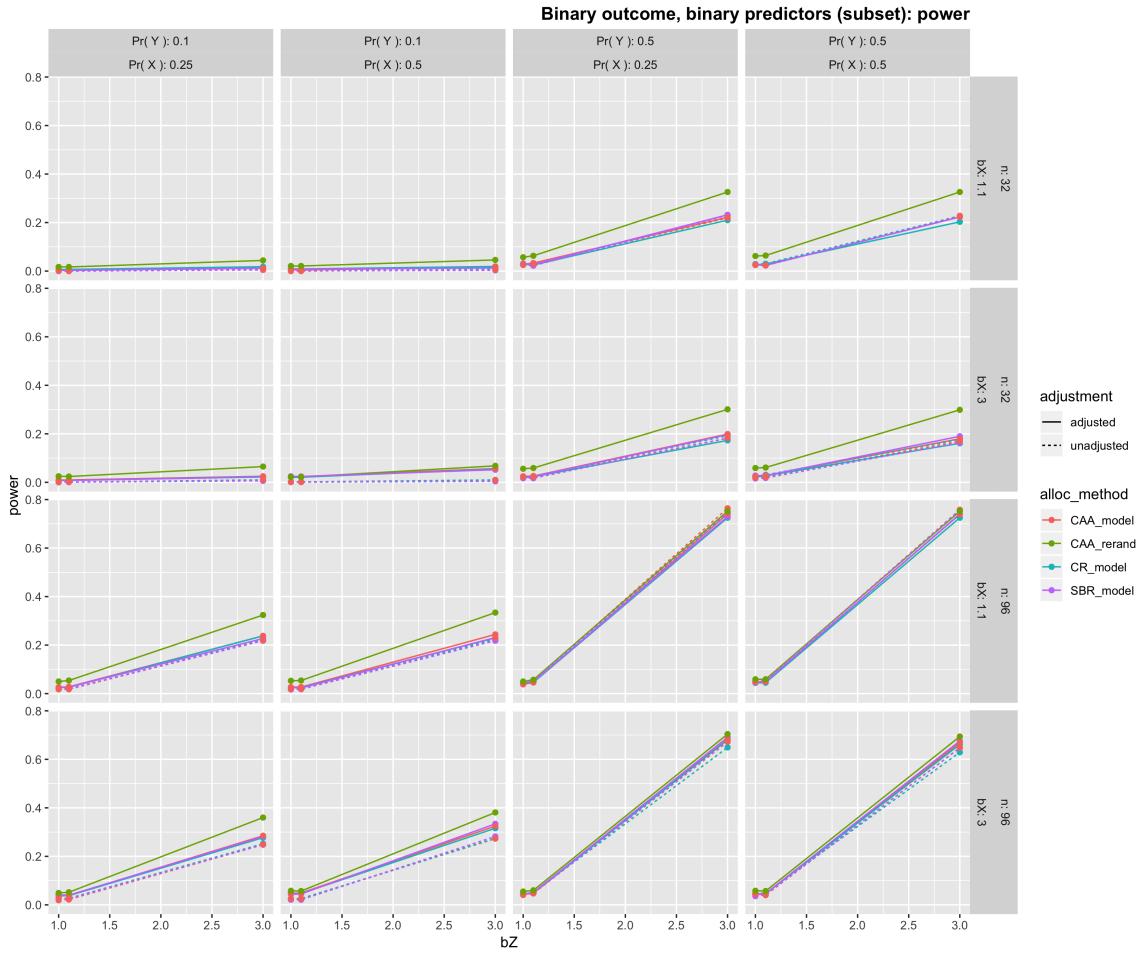


Figure 7.6: Batch 1 subset: Power

Subsetted power results show better type I error control for CAA followed by re-randomization overall. The modest power increase remains in large sample size ($n=96$) and low outcome prevalence settings, as well as small sample size ($n=32$) and high outcome prevalence settings ($\Pr(Y) = 0.5$).

7.2 Batch 2: Binary Outcome, Continuous Predictors

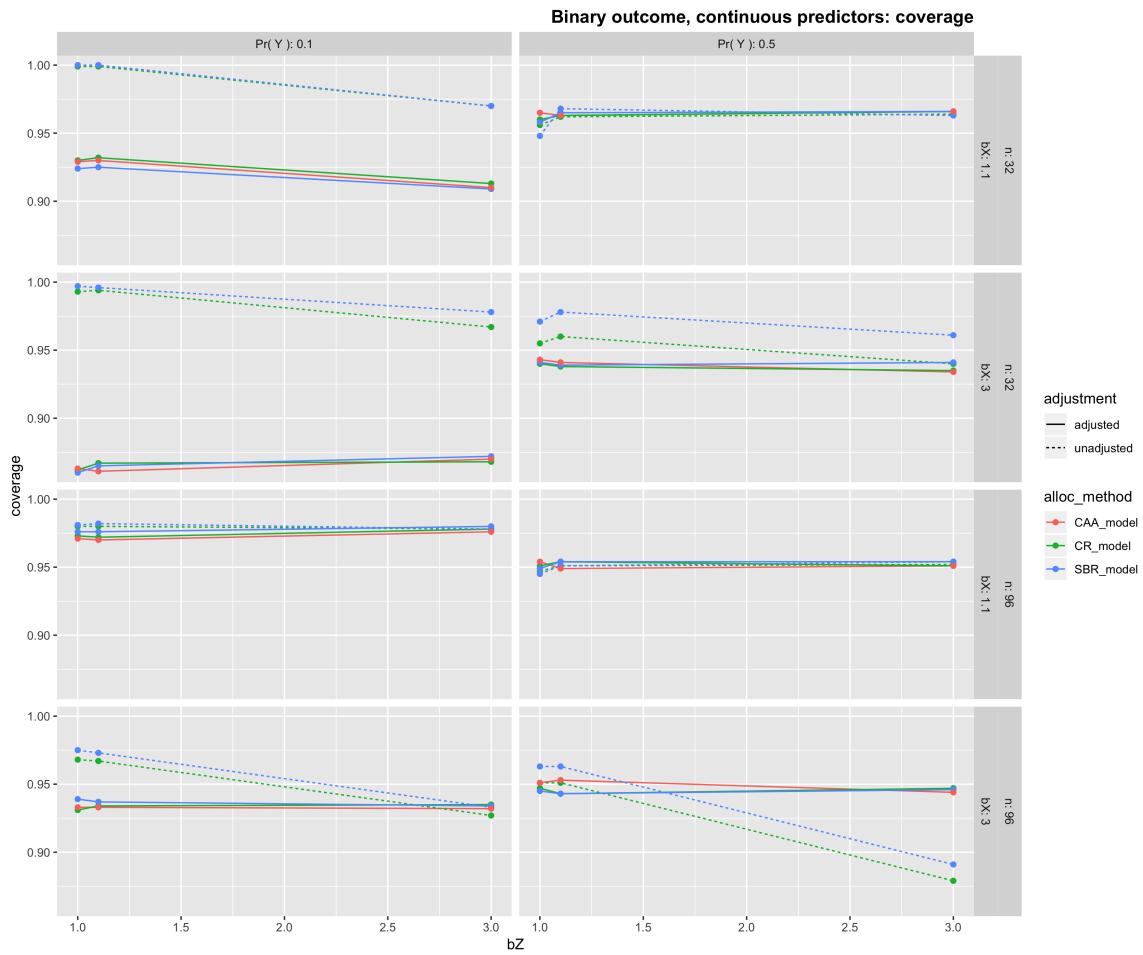


Figure 7.7: Batch 2: Coverage probability

In low outcome prevalence settings ($\text{Pr}(Y) = 0.1$) the coverage probability is less than 95 percent for adjusted estimates and greater than 95 percent for unadjusted estimates.

Coverage probability is not affected by treatment assignment effect size after adjusting for prognostic factors. Considering unadjusted estimates, treatment effect

coverage probability changes with prognostic factor effect size and is markedly greater when the effect size is large ($bX = 3$).

(batch 2): Adjusting for prognostic factors, especially when their effect size is large, is necessary to maintain proper coverage probability using model-based approaches.

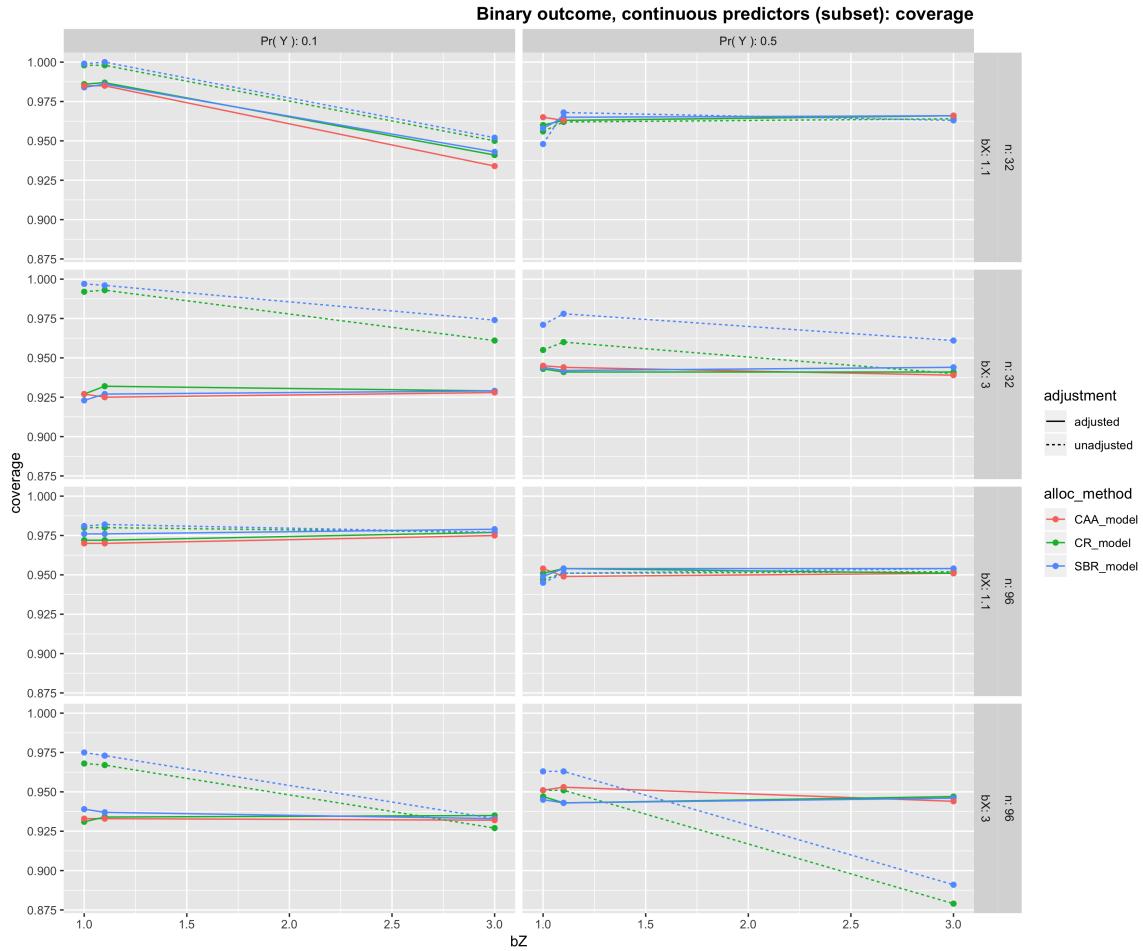


Figure 7.8: Batch 2 subset: Coverage probability

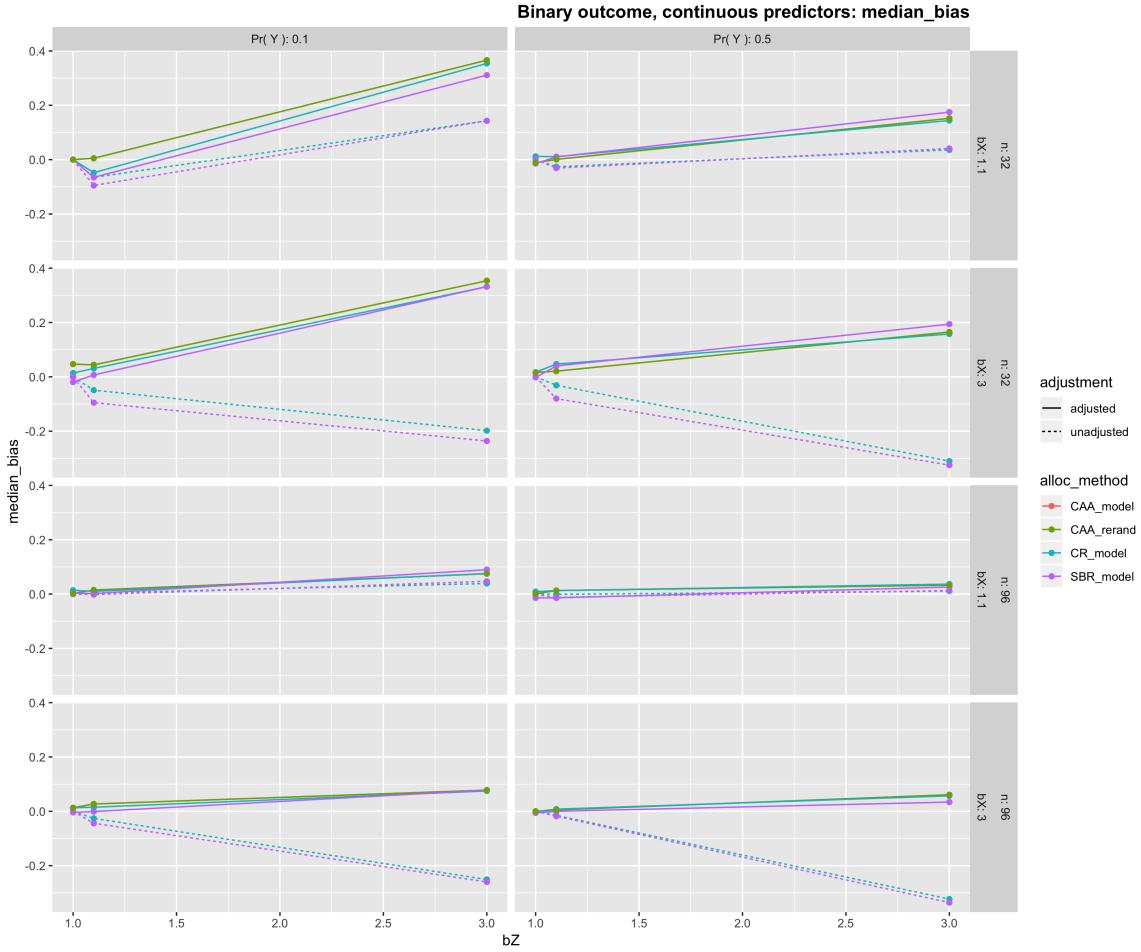


Figure 7.9: Batch 2: Median bias

Adjusting for prognostic factors used in the balancing method controls median bias, whereas the differences between unadjusted and adjusted estimates increases with treatment assignment effect size, and becomes greater as prognostic factor effect size increases.

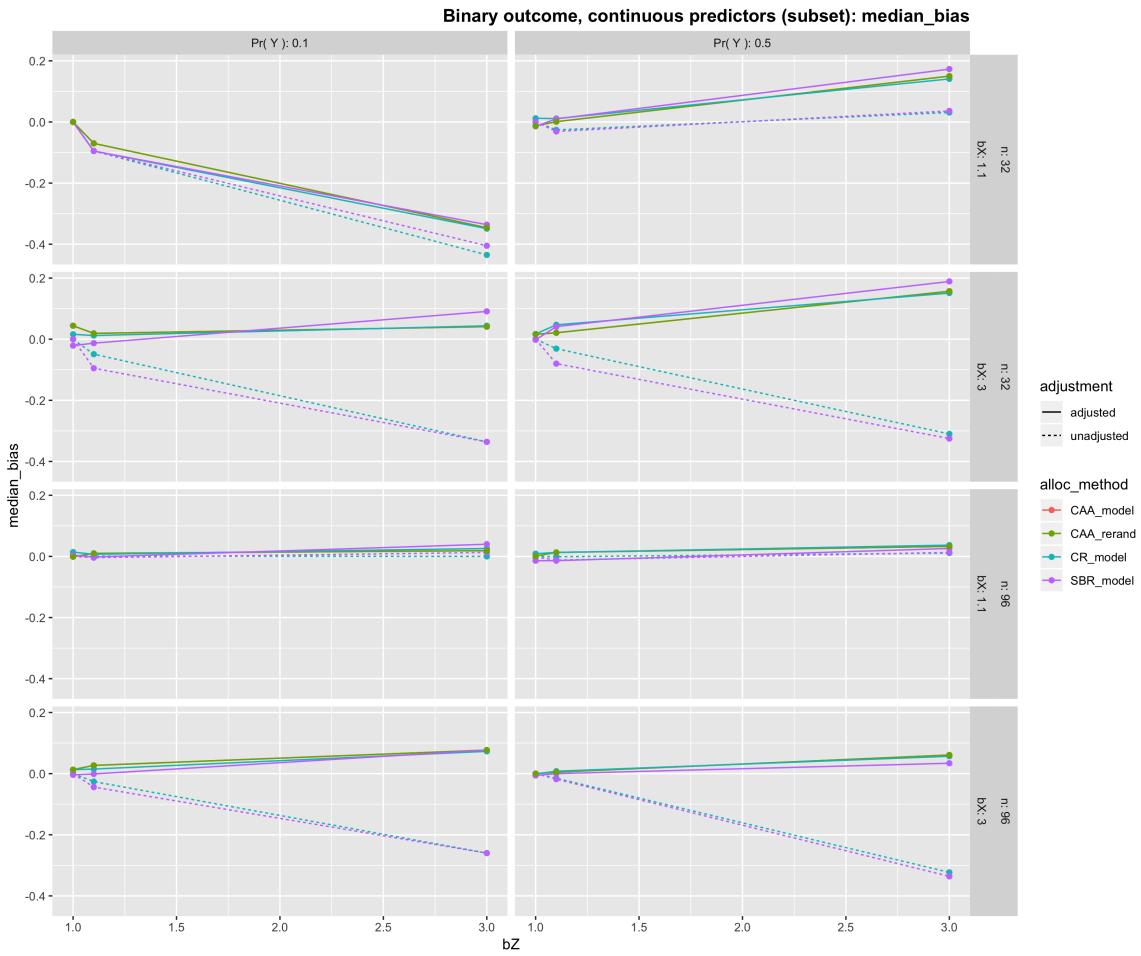


Figure 7.10: Batch 2 subset: Median bias

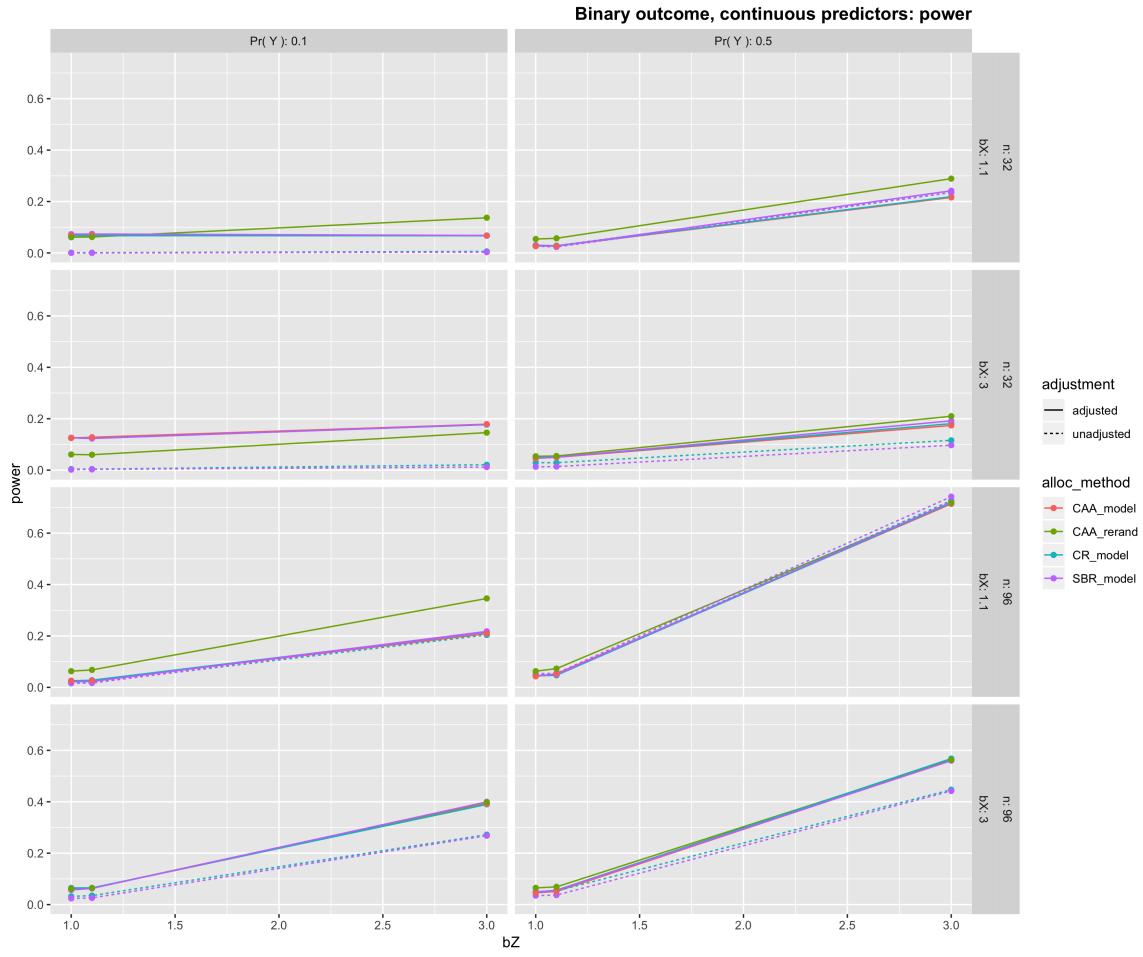


Figure 7.11: Batch 2: Power

In low outcome prevalence ($\text{Pr}(Y)=0.1$) under strong prognostic factor effect size ($bX = 3$), both CAA with model-based inference and SBR have higher power but fail to control size: the power under the null setting (type I error) is greater than 0.10.

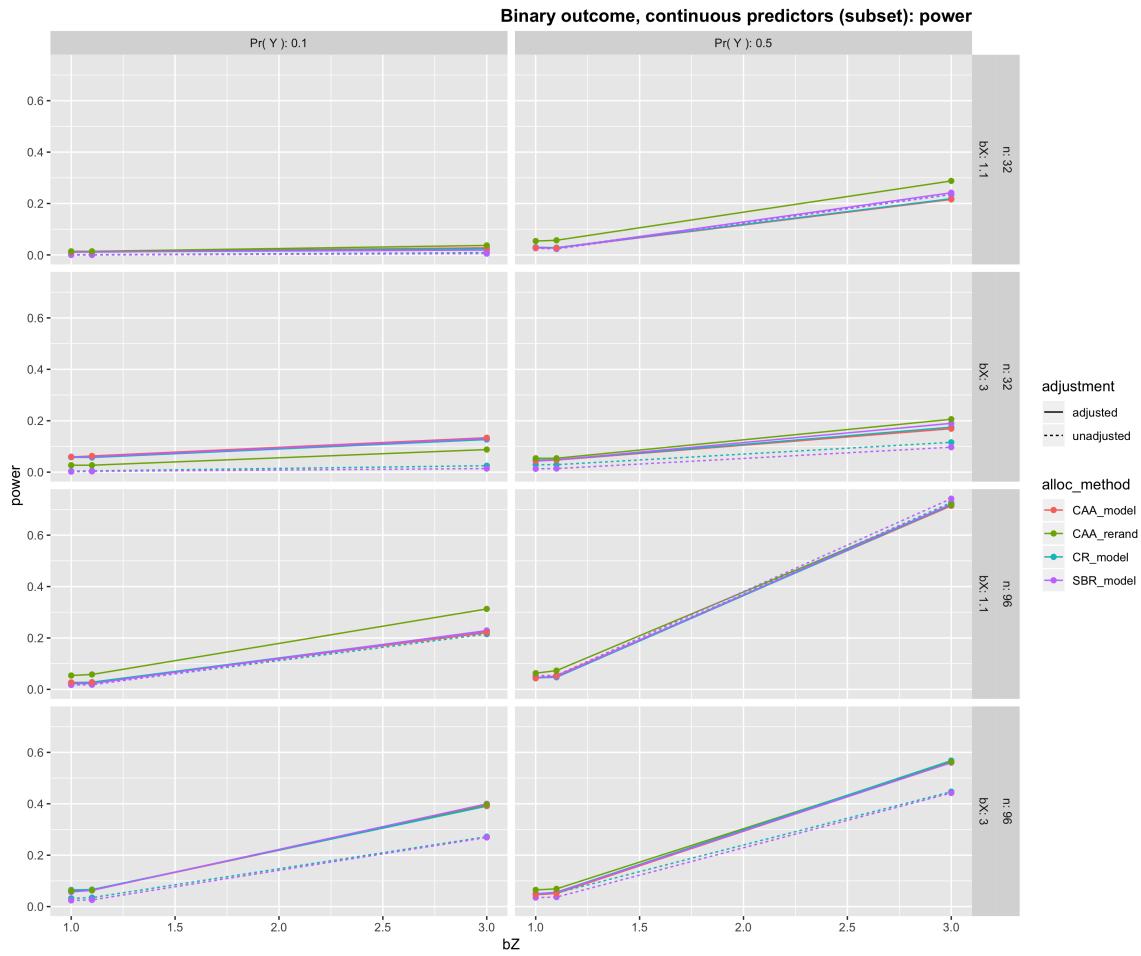


Figure 7.12: Batch 2 subset: Power

7.3 Batch 3: Continuous Outcome, Binary Predictors

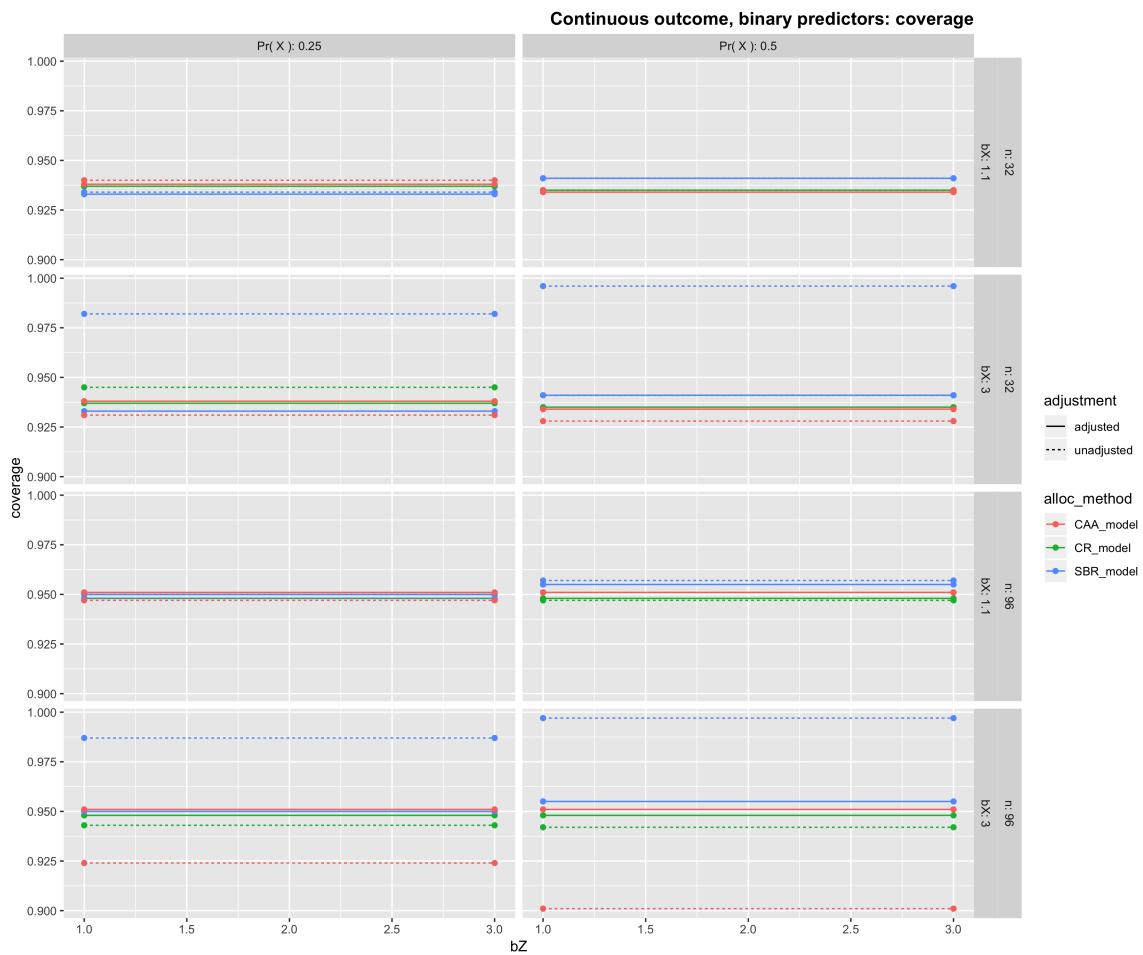


Figure 7.13: Batch 3: Coverage probability

Coverage probability remains near 95 percent in all methods and analyses, except for the unadjusted estimates when prognostic factor effect size is large.

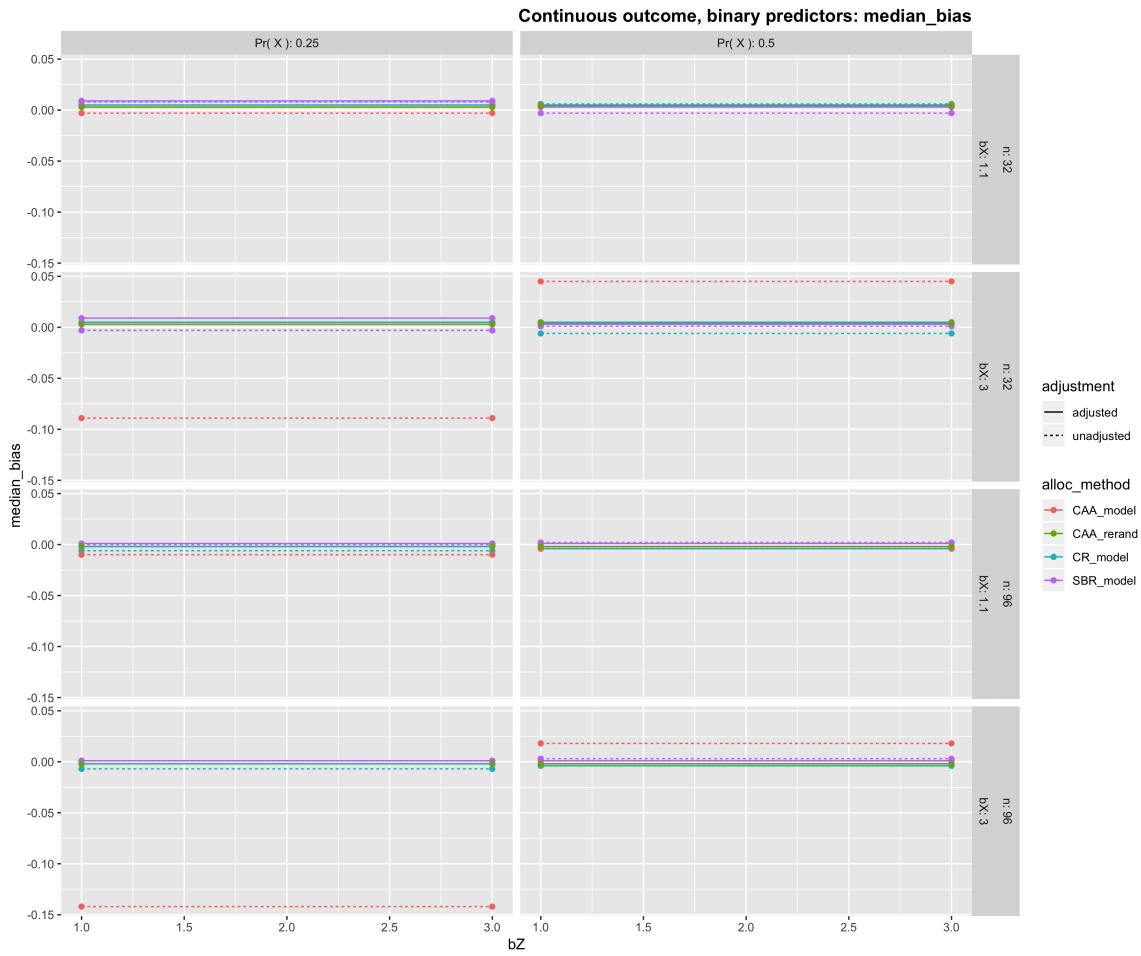


Figure 7.14: Batch 3: Median bias

For adjusted estimates, median bias is controlled under all considered settings.

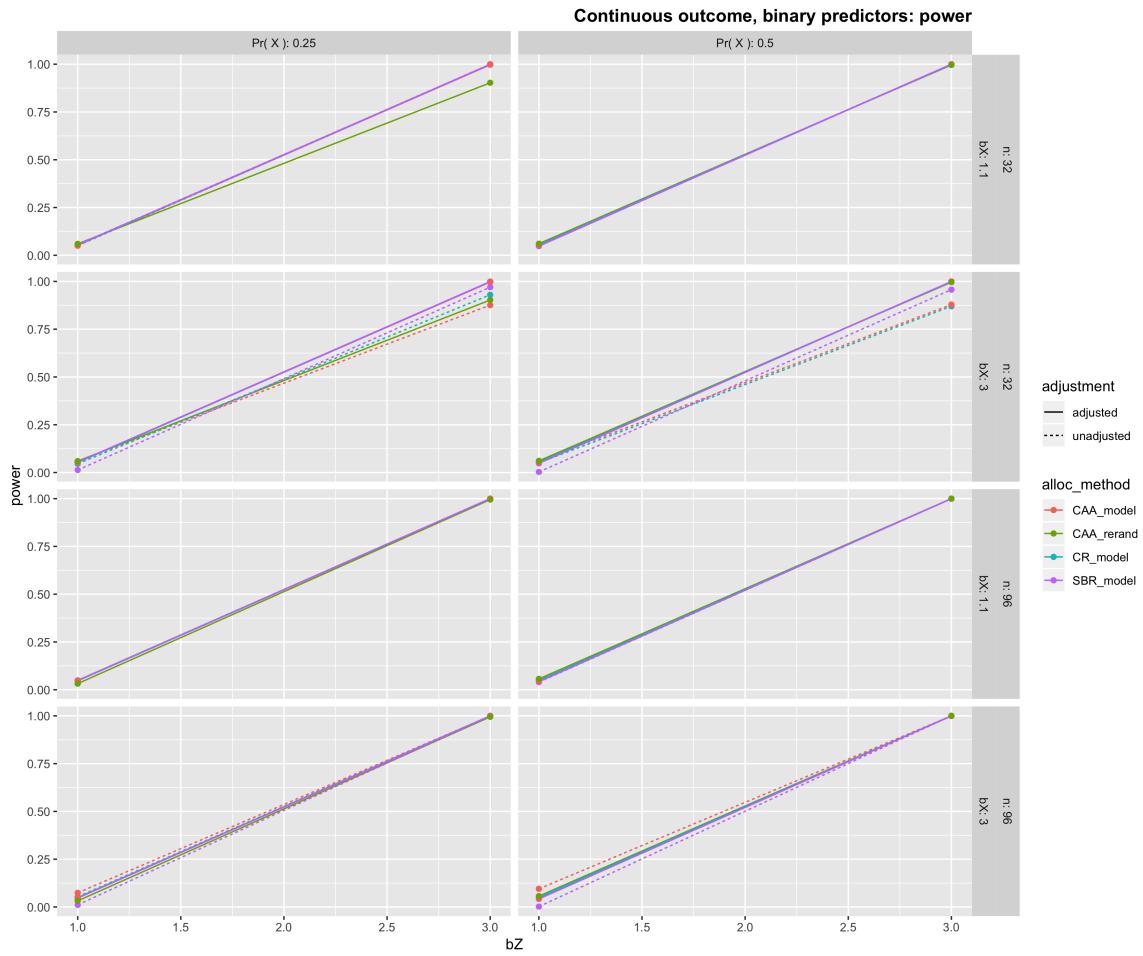


Figure 7.15: Batch 3: Power

More simulations between low and high treatment assignment effect size are necessary to view the trend

7.4 Batch 4: Continuous Outcome, Continuous Predictors

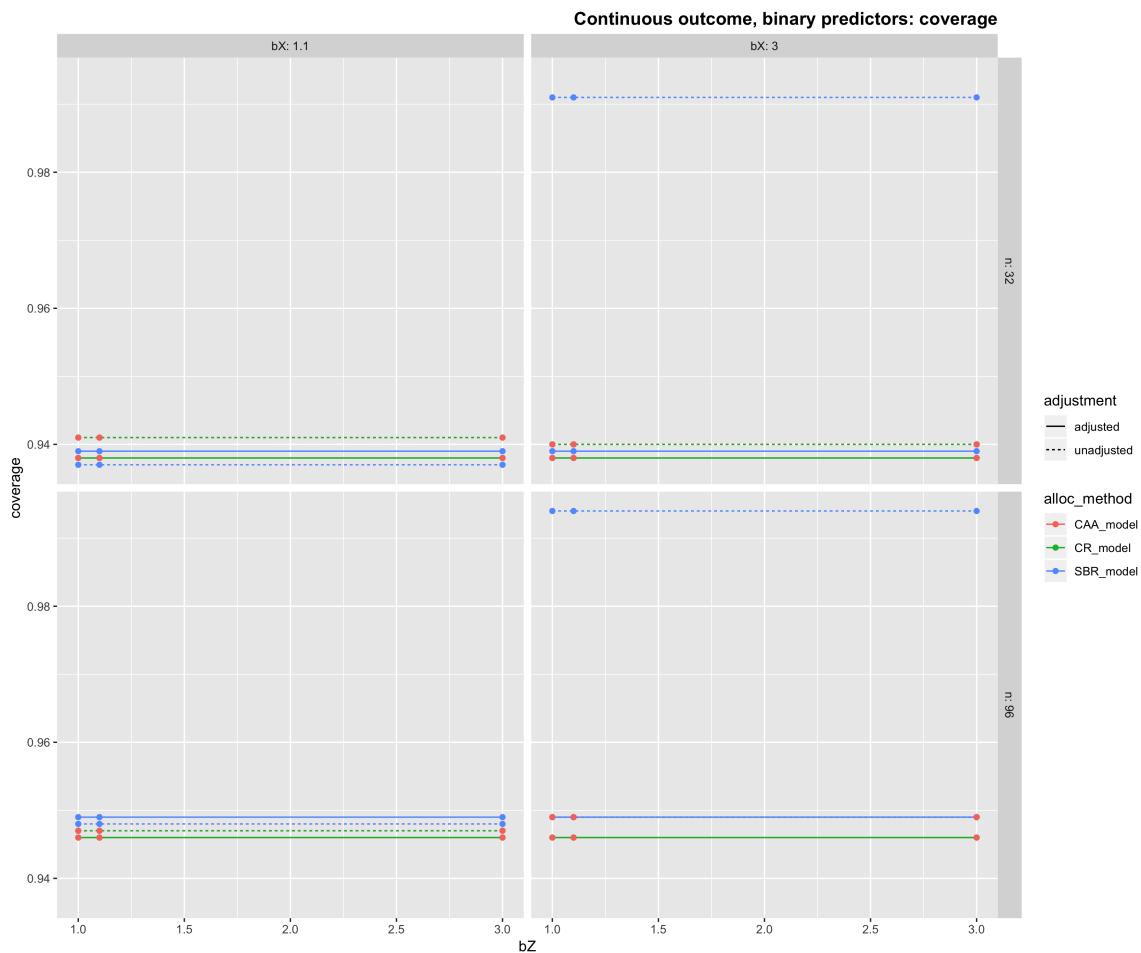


Figure 7.16: Batch 4: Coverage probability

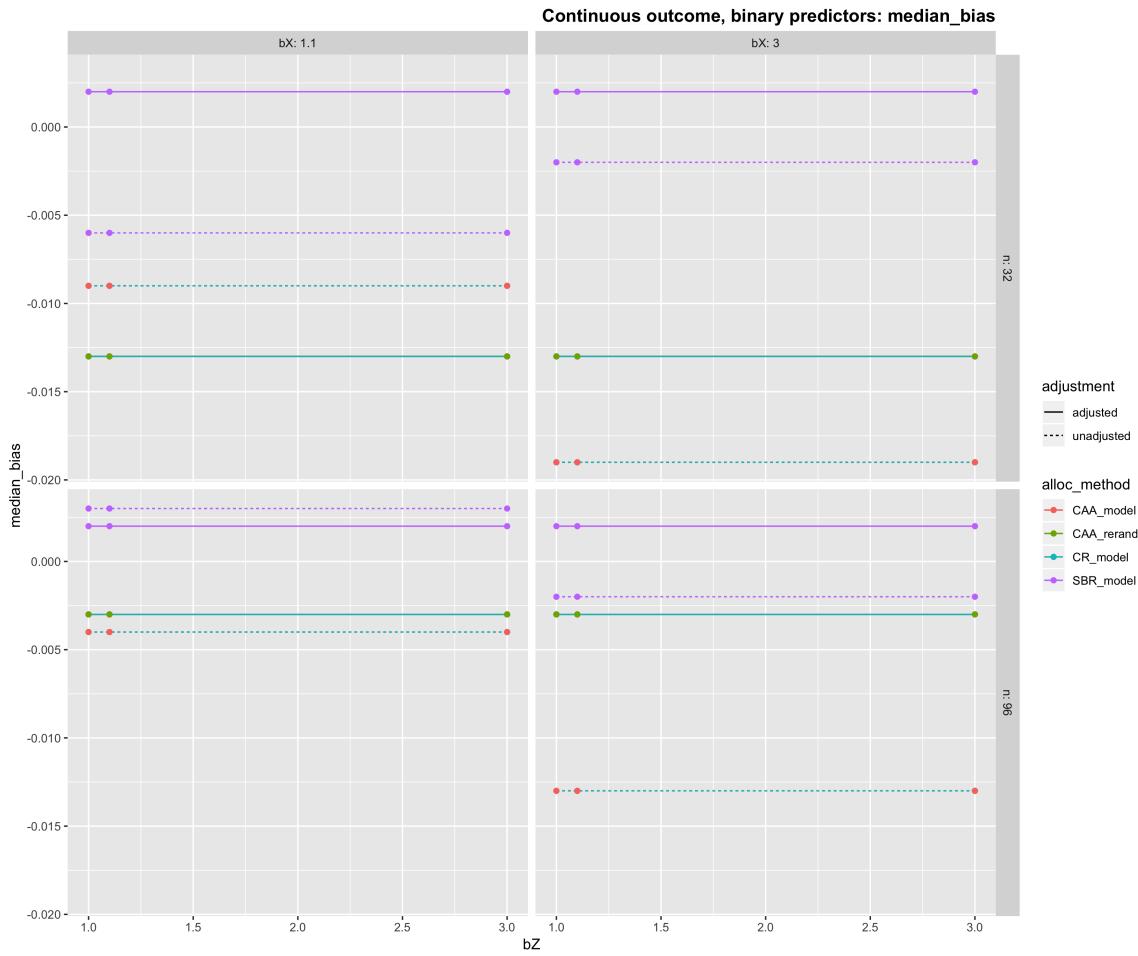


Figure 7.17: Batch 4: Median bias

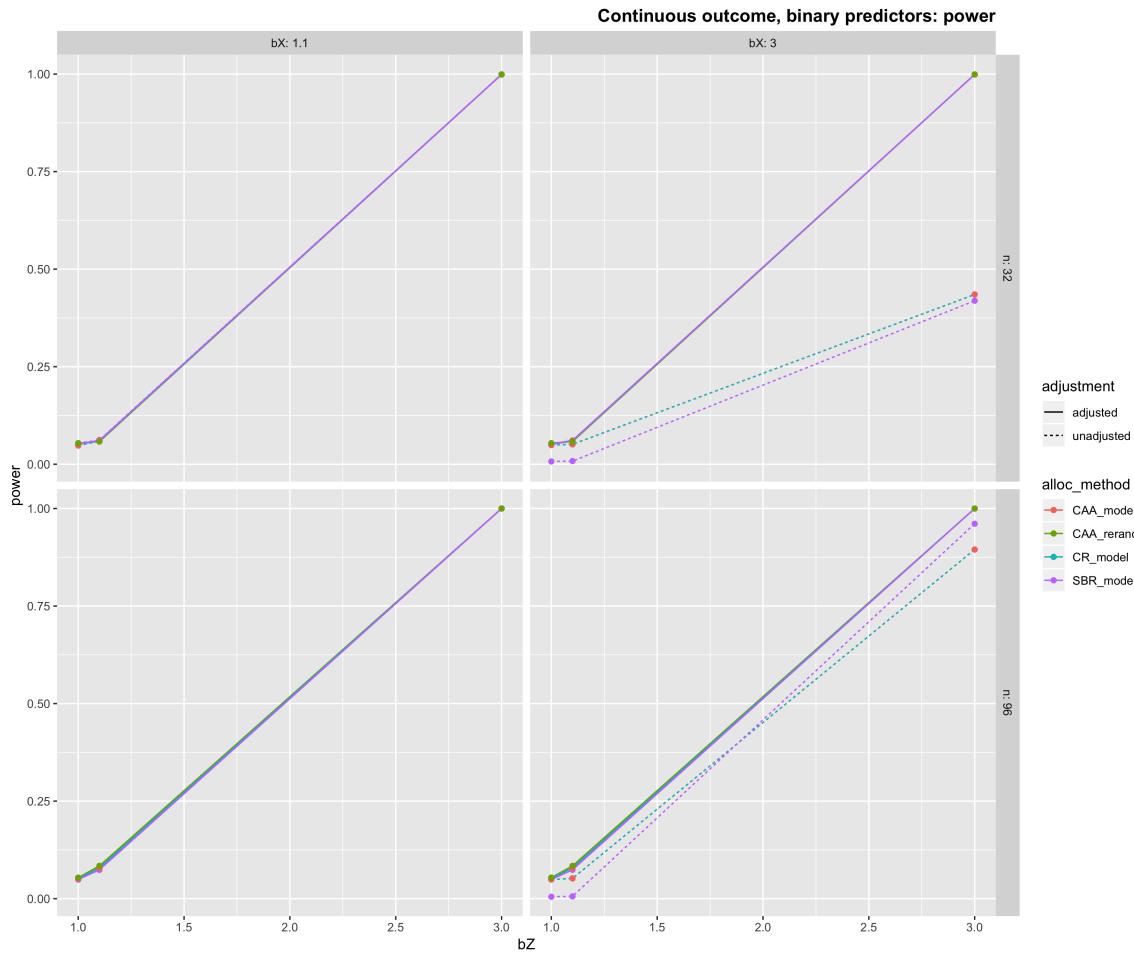


Figure 7.18: Batch 4: Power