

# A Comparison of Parametric Propensity Score-Based Methods for Causal Inference with Multiple Treatments and a Binary Outcome Supplemental Materials

## Supplemental Section 1. Covariate Balance Checking

To check for balance in covariates for PEN-GAM in our case with four treatment groups, we fit a generalized linear model for each covariate  $X$  including splines of generalized propensity scores and treatment groups as predictors. Then we used a likelihood ratio test to conduct a 3-degree-of-freedom global test on all of the treatment coefficients being zero. Categorical covariates with more than two levels were represented with multiple indicator variables, and statistical tests were conducted on each indicator. P-values before and after adjusting for splines of propensity scores for each covariate are reported in Supplemental Table 1.1.

Let  $s_z^2$  be the variance of covariate  $X$  in treatment group  $z$  and  $\bar{X}_p$  be the mean of  $X$  in the target population. To conduct balancing checking for weighting-based and matching-based estimators, we followed Li and Li [1] and inspected the absolute standardized difference in means between each treatment group and the target population,

$$d = |\bar{X}_z - \bar{X}_p|/s,$$

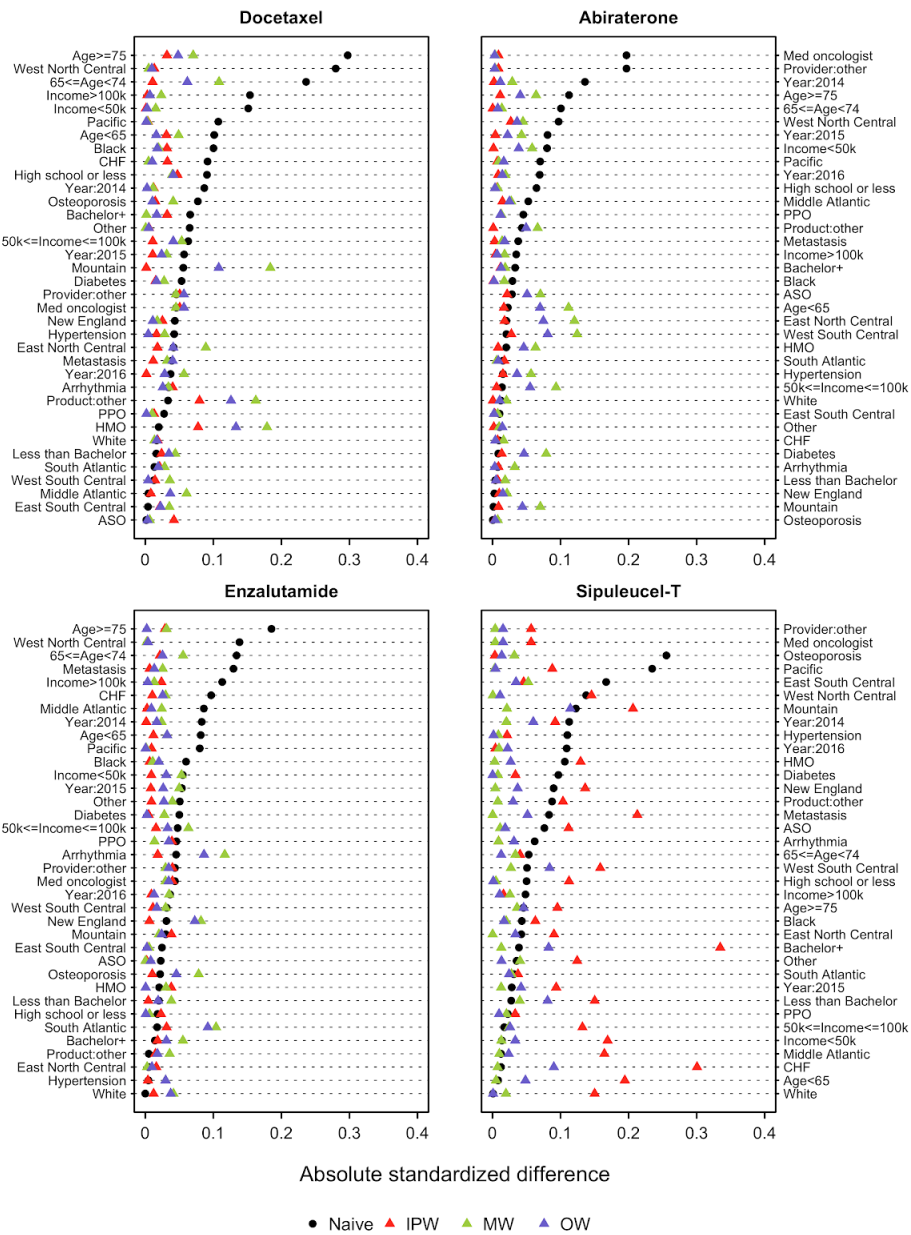
where  $s^2 = J^{-1} \sum_{z=1}^J s_z^2$ . For weighting-based estimators (IPW, MW, and OW),  $\bar{X}_z$  is the weighted average of  $X$  from the  $z$ th group, while for matching-based estimators (MCOV, MGPSV, and MGPS),  $\bar{X}_z$  is the group-specific mean of  $X$  after imputation. Supplemental Figures 1.1 and 1.2 present the absolute standardized differences for each covariate for weighting-based and matching-based estimators, respectively. The results for the naive estimator indicate the presence of large imbalance for some covariates across treatment groups (e.g. age, provider type, etc.). IPW balanced the covariates well for docetaxel, abiraterone, and enzalutamide, but not for sipuleucel-T. The absolute standardized differences for MW and OW remained small in general for all treatment groups. Improvement of balance was achieved using matching-based estimators for all treatment groups other than sipuleucel-T, with MGPSV and MGPS performing slightly better than MCOV.

**Table 1. P-values of likelihood ratio test (3-df) on all treatment coefficients being zero before and after adjusting for splines of propensity scores in a model of covariate.**

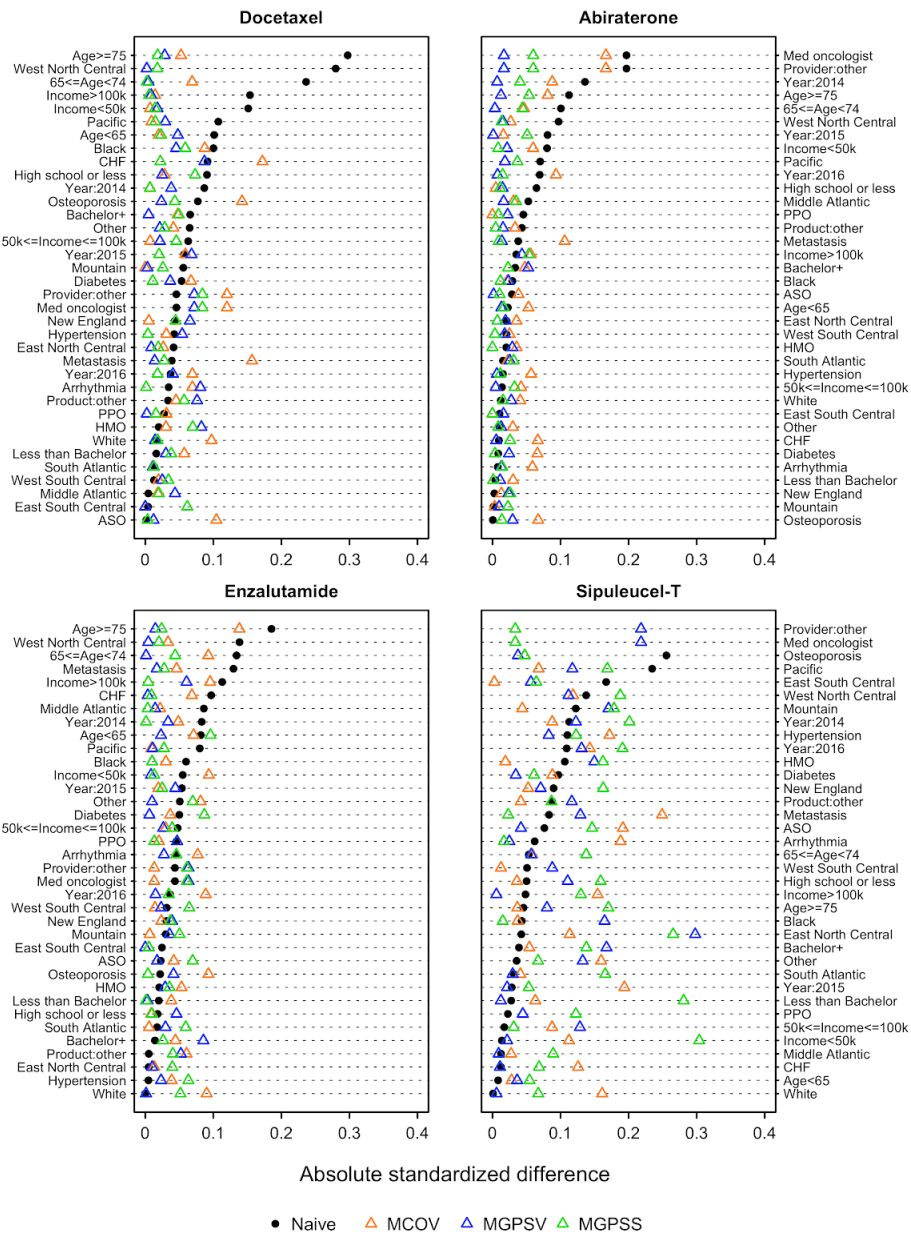
Covariates <sup>a</sup>	Before adjusting	After adjusting
<b>Age</b>		
<65	0.038	0.997
65-74	<0.001	0.997
≥75	<0.001	0.988
<b>Race</b>		
White	0.969	0.998
Black	0.043	0.990

Other	0.322	0.998
<b>Education level</b>		
High School Diploma or Less	0.055	0.987
High School Graduate and Less than Bachelor Degree	0.936	0.998
Bachelor Degree Plus	0.360	0.991
Unknown	0.001	0.983
<b>Household income range</b>		
<50k	<0.001	0.993
50k-100k	0.351	0.987
>100k	<0.001	0.956
Unknown	0.006	0.997
<b>Geographic Region</b>		
South Atlantic	0.915	0.998
New England	0.510	0.998
Middle Atlantic	0.155	0.953
East North Central	0.693	0.980
East South Central	0.314	0.996
West North Central	<0.001	0.916
West South Central	0.767	0.998
Mountain	0.266	0.984
Pacific	<0.001	0.994
<b>Product</b>		
HMO	0.562	0.993
PPO	0.422	0.970
Other	0.423	0.996
<b>Metastatic (Yes)</b>	0.024	0.968
<b>ASO (Yes)</b>	0.686	0.998
<b>Year of First Prescription</b>		
2014	<0.001	0.994
2015	0.053	0.999
2016	0.090	0.999
<b>Diabetes</b>	0.282	0.998
<b>Hypertension</b>	0.462	0.999
<b>Arrhythmia</b>	0.571	0.998
<b>CHF</b>	0.034	0.981
<b>Osteoporosis</b>	0.009	0.950
<b>Provider Type</b>		
Medical oncologist	<0.001	0.873
Others	<0.001	0.873

<sup>a</sup>Categorical covariates with more than two levels were represented with multiple indicator variables, and statistical tests were conducted on each indicator.



**Supplemental Figure 1.1.** Group-specific absolute standardized differences for weighting-based methods. Levels of covariates were sorted by the absolute standardized difference for the naive estimator.



**Supplemental Figure 1.2.** Group-specific absolute standardized differences for matching-based methods. Levels of covariates were sorted by the absolute standardized difference for the naive estimator.

## Reference

1. Li F, Li F. Propensity score weighting for causal inference with multiple treatments. *Ann Appl Stat.* 2019;13:2389–415.

## Supplemental Tables

Supplemental Table 1. Parameter settings for simulation studies					
	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
<b>Covariates in propensity model (<math>X_i</math>)</b>	$(1, X_{i1}, X_{i2}, X_{i3}, X_{i4}, X_{i5}, X_{i6})^T$	$(1, X_{i1}, X_{i2}, X_{i3}, X_{i4}, X_{i5}, X_{i6})^T$	$(1, X_{i1}, X_{i2}, X_{i3}, X_{i4}, X_{i5}, X_{i6})^T$	$(1, X_{i1}, X_{i2}, X_{i2}^2, X_{i3}, X_{i4}, X_{i5}, X_{i6}, X_{i1}X_{i3})^T$	$(1, X_{i1}, X_{i2}, X_{i3}, X_{i4}, X_{i5}, X_{i6})^T$
$\beta_1^T$	(0, 0, 0, 0, 0, 0, 0)	(0, 0, 0, 0, 0, 0, 0)	(0, 0, 0, 0, 0, 0, 0, 0)	(0, 0, 0, 0, 0, 0, 0, 0)	(0, 0, 0, 0, 0, 0, 0)
$\beta_2^T$	(-0.4, 0.1, 0.1, 0.2, 0.1, 0.2, 0.3)	(-0.9, 0.4, 0.4, 0.5, 0.4, 0.4, 0.3)	(-0.9, 0.4, 0.4, 0.5, 0.4, 0.4, 0.3)	(-1.2, 0.6, 0.6, 0.5, 0.5, 0.5, 0.2, 0.4, 0.3)	(-0.9, 0.4, 0.4, 0.5, 0.4, 0.4, 0.3)
$\beta_3^T$	(-0.2, 0.1, 0.2, 0.05, 0.1, 0.1, 0.1)	(-0.4, -0.4, -0.4, 0.5, 0.3, 0.4, -0.3)	(-0.4, -0.4, -0.4, 0.5, 0.3, 0.4, -0.3)	(-0.1, -0.3, -0.5, -0.5, 0.5, 0.5, 0.3, -0.2, 0.2)	(-0.4, -0.4, -0.4, 0.5, 0.3, 0.4, -0.3)
$\alpha_1^T$	(log(0.05), -0.2, -0.5, -1, 1, 0.5, 0.2)	(log(0.05), -0.2, -0.5, -1, 1, 0.5, 0.2)	(log(0.25), -0.8, 0.8, 0.8, 0.8, -1, 0.8)	(log(0.05), -0.2, -0.5, -1, 1, 0.5, 0.2)	(log(0.02), -0.2, -0.5, -1, 1, 0.5, 0.2)
$\alpha_2^T$	(log(0.2), 0.5, 0.3, -0.3, 0.8, -0.3, 0.8)	(log(0.2), 0.5, 0.3, -0.3, 0.8, -0.3, 0.8)	(log(0.15), 0.5, 0.8, -0.6, 0.8, 0.6, 0.8)	(log(0.2), 0.5, 0.3, -0.3, 0.8, -0.3, 0.8)	(log(0.01), 0.5, 0.3, -0.3, 0.8, -0.3, 0.8)
$\alpha_3^T$	(log(0.3), -0.5, -0.5, 0.1, 0.2, -0.3, -0.3)	(log(0.3), -0.5, -0.5, 0.1, 0.2, -0.3, -0.3)	(log(0.15), 0.6, 0.7, -1, 1, 0.5, 0.6)	(log(0.3), -0.5, -0.5, 0.1, 0.2, -0.3, -0.3)	(log(0.04), -0.5, -0.5, 0.1, 0.2, -0.3, -0.3)

<b>Supplemental Table 2. Root mean squared error (RMSE) <math>\times</math> 1000 for n=1500</b>															
<b>Methods*</b>	Scenario 1			Scenario 2			Scenario 3			Scenario 4			Scenario 5		
	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3
NAIVE	44	31	61	157	140	36	252	161	410	170	98	79	66	46	28
OREG (c)	26	25	28	26	24	29	30	32	29	28	23	30	15	15	12
OREG (m)	52	26	46	79	29	66	78	41	100	90	28	75	43	15	37
PEN-GAM (c, c, GLMPS)	25	26	28	30	29	34	33	37	35	32	33	37	17	18	16
PEN-GAM (c, m, GLMPS)	25	25	28	31	29	34	33	37	35	32	33	37	17	18	16
PEN-GAM (m, c, GLMPS)	25	26	28	30	27	32	32	35	34	29	30	35	17	17	15
PEN-GAM (m, m, GLMPS)	51	27	45	92	32	77	93	44	118	29	30	34	50	18	44
IPW (c, GLMPS)	26	26	29	34	31	40	39	53	52	36	45	52	17	17	14
IPW (m, GLMPS)	52	26	46	91	33	77	88	53	117	94	37	116	50	17	45
IPW (c, CBPS)	26	26	29	32	28	34	37	40	37	37	34	45	17	17	14
IPW (m, CBPS)	52	26	46	91	30	79	85	48	111	40	28	45	50	17	45
AIPW (c, c, GLMPS)	26	25	28	29	27	32	32	36	33	34	37	41	17	16	13
AIPW (c, m, GLMPS)	26	25	29	30	29	34	33	42	39	34	37	43	17	16	13
AIPW (m, c, GLMPS)	26	25	28	29	26	32	32	34	32	49	25	49	18	16	15
AIPW (m, m, GLMPS)	52	26	46	94	30	82	91	42	112	65	25	66	51	17	46
AIPW (c, c, CBPS)	26	25	28	29	28	33	32	36	33	34	35	43	17	17	14
AIPW (c, m, CBPS)	26	25	29	30	28	34	33	38	36	37	35	45	17	17	14
AIPW (m, c, CBPS)	26	25	28	29	26	32	32	34	32	31	25	32	18	16	15
AIPW (m, m, CBPS)	52	26	46	93	30	81	91	42	112	31	25	33	51	17	46
MW (c, GLMPS)	30	26	36	50	31	46	53	58	38	50	34	45	35	18	27
MW (m, GLMPS)	52	25	53	65	29	61	55	89	124	33	29	37	36	17	39
MW (c, CBPS)	31	26	38	57	33	51	60	57	43	59	39	52	37	19	28
MW (m, CBPS)	53	25	53	64	29	61	53	91	121	46	31	44	36	18	39
AMW (c, c, GLMPS)	30	26	36	50	30	45	52	57	36	49	34	44	35	18	27
AMW (c, m, GLMPS)	30	26	36	50	30	46	52	57	37	49	34	44	35	18	27
AMW (m, c, GLMPS)	26	27	30	42	29	38	58	65	34	32	37	41	22	19	15
AMW (m, m, GLMPS)	52	25	52	66	29	63	56	88	124	32	38	45	36	17	39
AMW (c, c, CBPS)	30	26	37	53	31	50	54	58	37	52	37	48	36	19	28
AMW (c, m, CBPS)	31	26	37	55	31	50	55	57	38	55	37	50	36	19	28
AMW (m, c, CBPS)	26	27	30	43	29	41	62	70	34	46	29	45	22	19	15
AMW (m, m, CBPS)	52	25	52	65	29	61	54	93	126	43	30	41	36	17	39
OW (c, GLMPS)	27	25	31	42	28	40	42	45	36	53	29	53	30	17	23
OW (m, GLMPS)	52	26	48	72	27	68	67	68	119	35	26	40	39	16	41
OW (c, CBPS)	27	26	32	49	31	43	49	44	44	64	33	61	33	18	25
OW (m, CBPS)	52	26	48	71	28	69	64	69	114	43	28	44	39	16	41
AOW (c, c, GLMPS)	27	25	31	42	29	39	41	43	34	52	29	52	31	18	24
AOW (c, m, GLMPS)	29	26	35	46	29	43	45	45	36	51	30	50	33	18	25
AOW (m, c, GLMPS)	26	25	29	35	28	33	43	47	33	32	31	35	20	18	14
AOW (m, m, GLMPS)	51	26	47	71	28	68	66	67	116	32	32	38	39	16	41
AOW (c, c, CBPS)	27	25	31	45	29	42	43	45	35	56	32	56	33	18	25
AOW (c, m, CBPS)	30	26	37	51	30	47	47	46	38	57	33	56	34	18	26
AOW (m, c, CBPS)	26	25	29	36	28	34	46	50	33	41	27	42	20	18	14
AOW (m, m, CBPS)	51	26	47	71	28	66	64	71	118	42	28	41	38	17	40

MCOV (c)	29	28	32	37	39	37	57	61	100	40	36	38	18	18	15
MCOV (m)	53	28	52	97	41	73	111	65	162	40	35	37	49	19	42
MGPSV (c, GLMPS)	32	31	35	36	33	40	39	45	42	40	43	50	20	20	16
MGPSV (m, GLMPS)	56	31	50	95	36	84	93	49	115	38	37	44	52	20	47
MGPSV (c, CBPS)	31	30	35	37	35	40	40	46	44	40	48	53	20	21	17
MGPSV (m, CBPS)	54	31	49	97	36	84	96	50	117	36	36	42	52	21	47
MGPSV (c, GLMPS)	30	29	32	35	32	38	39	45	44	39	39	45	19	19	15
MGPSV (m, GLMPS)	53	30	49	94	35	81	95	53	127	36	35	41	51	19	45
MGPSV (c, CBPS)	30	29	33	35	33	38	40	45	46	40	42	46	19	20	15
MGPSV (m, CBPS)	53	29	49	95	36	80	96	54	129	36	34	39	51	20	45

\*For methods that involve only the propensity score or the outcome model, (c) denotes correct specification and misspecification, respectively. For methods that involve both models, the first and second letter in the parentheses correspond to the treatment model and outcome model, respectively.

**Supplemental Table 3. Performance of different causal inference methods in scenario 1 (n=1500) of simulation studies.**

Methods*	Bias from ATE × 1000			Empirical SD × 1000			Average SE × 1000			95% Coverage Rate (%)		
	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3
NAIVE	35	-19	-54	27	25	29	27	24	28	75	88	53
OREG (c)	0	-1	-1	26	25	28	25	25	28	94	95	94
OREG (m)	45	9	-36	26	25	29	26	24	29	58	93	76
PEN-GAM (c, c, GLMPS)	-1	0	1	25	26	28	26	25	27	96	95	94
PEN-GAM (c, m, GLMPS)	-1	0	0	25	25	28	26	25	27	96	94	95
PEN-GAM (m, c, GLMPS)	-1	-1	0	25	25	28	26	25	27	96	94	95
PEN-GAM (m, m, GLMPS)	44	9	-35	26	25	29	26	25	28	62	92	76
IPW (c, GLMPS)	0	-1	-1	26	26	29	26	25	29	94	94	94
IPW (m, GLMPS)	45	9	-36	26	25	29	26	25	29	58	94	76
IPW (c, CBPS)	-1	-1	0	26	26	29	26	25	29	94	94	94
IPW (m, CBPS)	45	9	-36	26	25	29	26	25	29	58	93	76
AIPW (c, c, GLMPS)	0	-1	-1	26	25	28	26	25	28	95	95	94
AIPW (c, m, GLMPS)	0	-1	-1	26	25	29	26	25	28	95	94	94
AIPW (m, c, GLMPS)	0	-1	-1	26	25	28	25	25	28	94	95	94
AIPW (m, m, GLMPS)	45	9	-36	26	25	29	26	24	29	58	93	76
AIPW (c, c, CBPS)	0	-1	-1	26	25	28	25	25	28	95	94	94
AIPW (c, m, CBPS)	-1	0	0	26	25	29	26	25	28	94	94	94
AIPW (m, c, CBPS)	0	-1	-1	26	25	28	25	25	28	94	95	94
AIPW (m, m, CBPS)	45	9	-36	26	25	29	26	24	29	58	93	76
MW (c, GLMPS)	-15	5	20	26	26	30	26	25	31	95	94	95
MW (m, GLMPS)	45	3	-43	27	25	31	27	25	30	38	94	46
MW (c, CBPS)	-16	5	21	27	26	31	27	25	31	95	94	95
MW (m, CBPS)	45	3	-43	27	25	31	27	25	30	39	94	46
AMW (c, c, GLMPS)	-15	5	20	26	25	30	26	25	30	95	95	95
AMW (c, m, GLMPS)	-15	5	20	26	25	30	26	25	30	95	95	95
AMW (m, c, GLMPS)	0	-7	-7	26	26	30	26	25	29	91	91	85
AMW (m, m, GLMPS)	45	3	-43	27	25	31	27	25	30	39	94	45
AMW (c, c, CBPS)	-16	5	21	26	25	30	26	25	30	95	95	95
AMW (c, m, CBPS)	-16	5	21	26	25	31	26	25	31	95	95	95
AMW (m, c, CBPS)	0	-7	-7	26	26	30	26	25	29	91	91	84
AMW (m, m, CBPS)	45	3	-43	27	25	31	27	25	30	39	94	45
OW (c, GLMPS)	-8	3	11	26	25	29	26	25	29	95	94	94
OW (m, GLMPS)	45	7	-38	26	25	30	26	25	29	49	94	61
OW (c, CBPS)	-9	3	12	26	25	29	26	25	29	95	94	94
OW (m, CBPS)	45	7	-38	26	25	30	26	25	29	49	94	61
AOW (c, c, GLMPS)	-8	3	11	26	25	29	26	25	29	95	95	95
AOW (c, m, GLMPS)	-14	6	20	26	25	29	26	25	29	94	94	94
AOW (m, c, GLMPS)	0	-2	-2	26	25	29	26	25	28	93	94	92
AOW (m, m, GLMPS)	44	7	-37	26	25	29	26	24	29	49	94	61
AOW (c, c, CBPS)	-8	3	11	26	25	29	26	25	29	95	95	95
AOW (c, m, CBPS)	-15	7	22	26	25	30	26	25	30	94	94	93
AOW (m, c, CBPS)	0	-2	-2	26	25	29	26	25	28	93	94	92
AOW (m, m, CBPS)	44	7	-37	26	25	29	26	24	29	49	94	61
MCOV (c)	-2	-2	1	29	28	32	29	28	32	95	95	95
MCOV (m)	44	3	-40	30	28	33	30	28	33	69	94	77



<b>MGPSV (c, GLMPS)</b>	0	0	0	32	31	35	32	30	33	94	94	93
<b>MGPSV (m, GLMPS)</b>	46	9	-36	32	29	34	32	30	34	68	94	82
<b>MGPSV (c, CBPS)</b>	0	-2	-1	31	30	35	32	31	35	95	95	94
<b>MGPSV (m, CBPS)</b>	44	8	-36	31	30	34	32	30	35	72	93	83
<b>MGPSV (c, GLMPS)</b>	0	-1	-1	30	29	32	31	29	33	96	95	95
<b>MGPSV (m, GLMPS)</b>	44	8	-36	30	29	33	31	29	33	69	94	80
<b>MGPSV (c, CBPS)</b>	0	-1	-1	30	29	33	31	29	33	95	95	95
<b>MGPSV (m, CBPS)</b>	44	8	-36	30	28	33	31	29	33	70	94	81

\*For methods that involve only the propensity score or the outcome model, (c) denotes correct specification and misspecification, respectively. For methods that involve both models, the first and second letter in the parentheses correspond to the treatment model and outcome model, respectively.

**Supplemental Table 4. Performance of different causal inference methods in scenario 2 (n=1500) of simulation studies.**

Methods*	Bias from ATE × 1000			Empirical SD × 1000			Average SE × 1000			95% Coverage Rate (%)		
	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3
NAIVE	156	137	-19	62	63	69	62	60	69	28	39	94
OREG (c)	0	0	0	64	58	67	61	57	67	93	94	94
OREG (m)	74	14	-59	69	58	72	66	57	71	79	93	85
PEN-GAM (c, c, GLMPS)	3	5	2	30	28	34	32	30	35	96	96	95
PEN-GAM (c, m, GLMPS)	4	5	1	30	28	34	31	30	35	96	96	95
PEN-GAM (m, c, GLMPS)	6	3	-3	29	27	32	31	28	33	96	96	95
PEN-GAM (m, m, GLMPS)	86	17	-69	33	28	35	34	28	36	28	92	53
IPW (c, GLMPS)	0	2	2	34	31	40	32	30	37	94	94	94
IPW (m, GLMPS)	84	17	-67	35	28	39	35	28	38	32	91	55
IPW (c, CBPS)	-5	-1	4	31	28	34	30	28	33	93	95	94
IPW (m, CBPS)	84	14	-70	34	27	36	33	27	36	29	92	50
AIPW (c, c, GLMPS)	-1	0	1	29	27	32	29	27	31	94	95	94
AIPW (c, m, GLMPS)	0	1	1	30	29	34	30	28	33	94	95	95
AIPW (m, c, GLMPS)	-1	0	1	29	26	32	29	26	32	95	95	94
AIPW (m, m, GLMPS)	87	14	-73	36	27	38	34	26	36	26	92	48
AIPW (c, c, CBPS)	0	0	1	29	28	33	29	27	32	95	95	93
AIPW (c, m, CBPS)	-4	-1	4	30	28	33	30	28	32	94	95	94
AIPW (m, c, CBPS)	-1	0	1	29	26	32	29	26	31	94	95	94
AIPW (m, m, CBPS)	87	14	-72	34	27	36	33	27	35	25	92	47
MW (c, GLMPS)	-38	-10	28	33	29	37	33	29	37	95	95	94
MW (m, GLMPS)	56	6	-50	33	28	36	33	29	36	21	91	44
MW (c, CBPS)	-44	-12	32	36	30	39	35	31	39	94	96	94
MW (m, CBPS)	54	5	-49	34	29	36	34	30	37	25	92	47
AMW (c, c, GLMPS)	-38	-10	28	32	29	36	32	29	37	95	95	94
AMW (c, m, GLMPS)	-38	-10	28	33	29	36	33	29	37	95	95	94
AMW (m, c, GLMPS)	-28	-10	18	30	27	33	31	28	34	93	95	94
AMW (m, m, GLMPS)	57	6	-51	33	28	36	33	29	36	19	92	42
AMW (c, c, CBPS)	-41	-10	31	34	29	39	34	30	39	95	95	94
AMW (c, m, CBPS)	-42	-11	31	35	29	39	34	30	39	94	95	94
AMW (m, c, CBPS)	-31	-9	22	31	28	34	31	29	35	94	95	95
AMW (m, m, CBPS)	56	7	-49	33	28	36	33	29	37	21	91	46
OW (c, GLMPS)	-29	-9	20	31	27	34	31	27	35	95	95	94
OW (m, GLMPS)	65	6	-59	32	27	34	32	27	35	17	92	39
OW (c, CBPS)	-36	-12	23	33	28	36	33	29	36	94	95	94
OW (m, CBPS)	63	4	-60	32	27	35	32	28	35	21	92	40
AOW (c, c, GLMPS)	-28	-9	19	31	28	34	31	28	35	95	95	94
AOW (c, m, GLMPS)	-34	-9	24	32	28	35	32	28	36	94	95	94
AOW (m, c, GLMPS)	-19	-10	8	29	27	32	29	27	32	94	95	94
AOW (m, m, GLMPS)	64	5	-58	32	27	34	32	28	35	18	91	40
AOW (c, c, CBPS)	-31	-9	23	32	28	36	32	29	36	95	95	94
AOW (c, m, CBPS)	-38	-10	28	34	28	37	33	29	38	93	95	93
AOW (m, c, CBPS)	-21	-10	11	30	27	32	30	28	33	94	95	95
AOW (m, m, CBPS)	63	6	-56	32	27	35	32	28	35	20	92	43
MCOV (c)	18	24	6	32	30	36	32	30	36	92	88	95
MCOV (m)	91	28	-63	34	30	38	34	30	38	25	86	61

<b>MGPSV (c, GLMPS)</b>	2	3	1	35	32	38	35	33	38	95	95	95
<b>MGPSV (m, GLMPS)</b>	86	16	-70	39	31	41	38	32	40	38	93	59
<b>MGPSV (c, CBPS)</b>	3	5	2	35	32	37	35	33	38	96	95	95
<b>MGPSV (m, CBPS)</b>	87	18	-69	38	31	41	38	32	40	37	92	59

\*For methods that involve only the propensity score or the outcome model, (c) and (m) denote correct specification and misspecification, respectively. For methods that involve both models, the first and second letter in the parentheses correspond to the treatment model and outcome model, respectively.

**Supplemental Table 5. Performance of different causal inference methods in scenario 3 (n=1500) of simulation studies.**

Methods*	Bias from ATE × 1000			Empirical SD × 1000			Average SE × 1000			95% Coverage Rate (%)		
	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3
NAIVE	250	-158	-409	31	28	29	30	28	28	0	0	0
OREG (c)	-1	-1	0	30	32	29	29	32	29	94	95	94
OREG (m)	71	-24	-94	33	33	32	32	33	32	41	89	16
PEN-GAM (c, c, GLMPS)	5	-6	-12	32	36	34	33	37	37	95	95	97
PEN-GAM (c, m, GLMPS)	6	-6	-12	32	36	33	33	37	37	96	95	97
PEN-GAM (m, c, GLMPS)	7	-5	-12	31	35	32	32	35	36	95	95	96
PEN-GAM (m, m, GLMPS)	86	-26	-112	35	36	36	35	36	39	32	90	14
IPW (c, GLMPS)	0	-4	-4	39	53	52	36	48	46	93	92	93
IPW (m, GLMPS)	79	-29	-108	40	44	46	38	43	43	45	87	30
IPW (c, CBPS)	-8	0	8	36	40	36	34	40	35	92	94	92
IPW (m, CBPS)	76	-28	-104	39	39	39	37	39	37	46	88	19
AIPW (c, c, GLMPS)	-1	-1	0	32	36	33	31	35	32	95	94	94
AIPW (c, m, GLMPS)	-1	-1	-1	33	42	39	32	39	36	94	95	94
AIPW (m, c, GLMPS)	-1	-1	0	32	34	32	31	34	31	94	94	93
AIPW (m, m, GLMPS)	83	-23	-105	38	35	38	36	35	36	37	90	16
AIPW (c, c, CBPS)	-1	-1	0	32	36	33	31	35	32	95	94	94
AIPW (c, m, CBPS)	-3	4	7	33	38	35	32	36	34	94	93	92
AIPW (m, c, CBPS)	-1	-1	0	32	34	32	31	34	31	94	94	93
AIPW (m, m, CBPS)	83	-23	-105	37	35	37	35	35	36	36	90	15
MW (c, GLMPS)	-36	-42	-6	40	40	37	39	40	37	94	95	94
MW (m, GLMPS)	39	-80	-119	39	39	37	39	40	37	50	85	15
MW (c, CBPS)	-41	-35	6	44	44	42	43	44	41	94	94	93
MW (m, CBPS)	33	-81	-114	42	41	39	41	42	39	61	85	22
AMW (c, c, GLMPS)	-35	-42	-6	38	38	36	37	38	35	95	95	94
AMW (c, m, GLMPS)	-35	-42	-7	39	39	36	38	39	36	95	95	94
AMW (m, c, GLMPS)	-45	-53	-8	36	37	33	35	37	33	94	94	94
AMW (m, m, GLMPS)	40	-79	-119	38	38	36	38	38	36	48	84	13
AMW (c, c, CBPS)	-36	-42	-6	40	41	36	40	41	36	94	95	94
AMW (c, m, CBPS)	-37	-38	-1	41	43	38	40	43	37	95	94	93
AMW (m, c, CBPS)	-49	-58	-8	37	38	33	37	39	33	93	93	95
AMW (m, m, CBPS)	37	-84	-120	40	40	36	39	40	36	54	82	12
OW (c, GLMPS)	-21	-24	-3	36	37	36	35	38	36	94	95	94
OW (m, GLMPS)	56	-58	-114	37	37	36	36	38	36	44	86	14
OW (c, CBPS)	-28	-16	12	40	41	42	39	41	40	93	94	91
OW (m, CBPS)	51	-57	-108	39	39	39	38	39	38	53	87	23
AOW (c, c, GLMPS)	-20	-24	-4	35	36	34	35	36	34	94	95	94
AOW (c, m, GLMPS)	-25	-26	0	37	37	36	36	37	35	94	95	94
AOW (m, c, GLMPS)	-27	-31	-5	34	35	32	33	35	32	94	94	94
AOW (m, m, GLMPS)	55	-56	-111	36	36	35	36	36	35	45	86	14
AOW (c, c, CBPS)	-22	-25	-4	37	37	35	36	38	35	95	95	94
AOW (c, m, CBPS)	-27	-22	6	38	40	38	38	40	37	94	94	93
AOW (m, c, CBPS)	-30	-35	-5	35	36	32	34	36	33	94	94	94
AOW (m, m, CBPS)	52	-60	-112	37	37	35	37	38	35	49	84	14
MCOV (c)	44	-48	-93	35	37	37	35	37	37	76	74	29
MCOV (m)	104	-53	-157	38	38	40	38	39	39	20	72	2

<b>MGPSV (c, GLMPS)</b>	6	-11	-17	39	44	41	39	46	42	95	95	95
<b>MGPSV (m, GLMPS)</b>	86	-34	-119	42	41	44	41	43	44	46	90	22
<b>MGPSV (c, CBPS)</b>	9	-12	-21	38	44	41	39	46	42	94	95	94
<b>MGPSV (m, CBPS)</b>	87	-34	-121	42	42	44	41	43	44	45	89	20

\*For methods that involve only the propensity score or the outcome model, (c) and (m) denote correct specification and misspecification, respectively. For methods that involve both models, the first and second letter in the parentheses correspond to the treatment model and outcome model, respectively.

**Supplemental Table 6. Performance of different causal inference methods in scenario 4 (n=1500) of simulation studies.**

Methods*	Bias from ATE × 1000			Empirical SD × 1000			Average SE × 1000			95% Coverage Rate (%)		
	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3
NAIVE	167	95	-73	28	26	31	27	27	31	0	6	34
OREG (c)	0	1	1	28	23	30	27	24	29	94	95	94
OREG (m)	85	16	-68	30	23	32	29	24	31	19	89	41
PEN-GAM (c, c, GLMPS)	1	9	8	32	31	36	33	35	39	96	97	96
PEN-GAM (c, m, GLMPS)	2	9	6	32	31	36	33	35	40	96	97	97
PEN-GAM (m, c, GLMPS)	0	8	8	29	29	34	30	33	36	96	97	96
PEN-GAM (m, m, GLMPS)	-1	7	8	29	29	34	30	32	35	96	97	96
IPW (c, GLMPS)	2	10	8	36	44	51	35	35	43	94	93	92
IPW (m, GLMPS)	-78	25	104	51	27	53	45	28	47	55	85	39
IPW (c, CBPS)	-12	2	15	35	34	43	34	32	40	91	95	92
IPW (m, CBPS)	-25	6	31	31	27	33	30	28	32	82	94	81
AIPW (c, c, GLMPS)	1	1	0	34	37	41	32	33	38	94	96	94
AIPW (c, m, GLMPS)	1	2	0	34	37	43	33	33	40	94	96	94
AIPW (m, c, GLMPS)	3	1	-1	49	25	49	37	25	38	93	95	93
AIPW (m, m, GLMPS)	-38	2	40	53	25	53	48	26	49	86	95	85
AIPW (c, c, CBPS)	1	1	1	34	35	43	33	33	40	94	95	94
AIPW (c, m, CBPS)	-12	0	12	35	35	43	34	33	41	92	95	94
AIPW (m, c, CBPS)	1	1	0	31	25	32	30	25	31	94	95	94
AIPW (m, m, CBPS)	3	-1	-4	31	25	32	29	26	31	94	95	94
MW (c, GLMPS)	-34	-16	18	37	30	41	36	31	40	94	95	94
MW (m, GLMPS)	-4	-9	-5	33	28	37	32	29	36	85	94	90
MW (c, CBPS)	-43	-17	26	41	35	45	41	36	44	94	95	93
MW (m, CBPS)	-30	-8	22	35	30	39	35	30	38	94	94	94
AMW (c, c, GLMPS)	-34	-16	17	36	30	40	35	31	39	94	95	94
AMW (c, m, GLMPS)	-33	-16	17	37	30	41	36	31	40	93	95	95
AMW (m, c, GLMPS)	-7	-26	-19	32	27	36	32	28	35	86	94	81
AMW (m, m, GLMPS)	-1	-27	-26	32	27	36	32	28	36	82	94	76
AMW (c, c, CBPS)	-36	-17	19	38	33	44	38	34	43	94	95	94
AMW (c, m, CBPS)	-39	-17	22	39	33	45	39	34	44	94	96	94
AMW (m, c, CBPS)	-32	-9	23	33	28	38	33	28	37	94	94	94
AMW (m, m, CBPS)	-27	-11	16	34	28	38	34	29	37	94	95	94
OW (c, GLMPS)	-40	-4	35	35	28	39	34	29	38	93	95	94
OW (m, GLMPS)	-16	2	18	31	26	35	31	27	34	88	94	92
OW (c, CBPS)	-51	-8	43	39	32	43	38	33	42	93	95	93
OW (m, CBPS)	-28	-3	25	33	28	36	32	28	35	93	95	93
AOW (c, c, GLMPS)	-39	-4	35	35	29	39	34	30	38	93	95	95
AOW (c, m, GLMPS)	-36	-7	29	36	29	40	35	30	39	93	95	94
AOW (m, c, GLMPS)	-9	-17	-9	31	26	34	30	27	33	82	92	73
AOW (m, m, GLMPS)	-5	-19	-14	32	26	35	31	27	35	80	92	69
AOW (c, c, CBPS)	-42	-5	37	37	32	43	37	33	42	94	96	94
AOW (c, m, CBPS)	-42	-9	33	39	32	45	39	33	43	94	96	94
AOW (m, c, CBPS)	-27	-4	22	31	26	35	31	27	35	93	95	93
AOW (m, m, CBPS)	-26	-7	19	33	27	37	33	27	36	92	95	92
MCOV (c)	20	20	0	34	29	38	34	31	38	91	92	95
MCOV (m)	22	20	-1	34	29	37	33	30	37	90	91	95

<b>MGPSV (c, GLMPS)</b>	6	5	-1	39	39	45	38	38	44	94	95	94
<b>MGPSV (m, GLMPS)</b>	5	2	-3	36	35	41	36	35	40	94	96	94
<b>MGPSV (c, CBPS)</b>	11	14	4	38	39	46	39	39	45	94	95	95
<b>MGPSV (m, CBPS)</b>	-8	-4	3	36	33	39	35	34	38	94	95	95

\*For methods that involve only the propensity score or the outcome model, (c) and (m) denote correct specification and misspecification, respectively. For methods that involve both models, the first and second letter in the parentheses correspond to the treatment model and outcome model, respectively.

**Supplemental Table 7. Performance of different causal inference methods in scenario 5 (n=1500) of simulation studies.**

Methods*	Bias from ATE × 1000			Empirical SD × 1000			Average SE × 1000			95% Coverage Rate (%)		
	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3
<b>NAIVE</b>	63	42	-21	19	17	20	19	18	20	9	33	84
<b>OREG (c)</b>	0	0	0	15	15	12	15	15	13	95	94	97
<b>OREG (m)</b>	38	5	-34	19	15	17	19	15	17	46	93	49
<b>PEN-GAM (c, c, GLMPS)</b>	1	3	3	17	18	16	18	20	18	97	97	98
<b>PEN-GAM (c, m, GLMPS)</b>	1	3	2	17	18	16	18	19	18	97	97	98
<b>PEN-GAM (m, c, GLMPS)</b>	2	2	0	17	17	15	18	18	17	97	97	98
<b>PEN-GAM (m, m, GLMPS)</b>	45	7	-39	22	17	21	23	18	23	50	96	58
<b>IPW (c, GLMPS)</b>	1	1	0	17	17	14	17	16	14	95	95	97
<b>IPW (m, GLMPS)</b>	45	5	-39	23	16	22	22	16	21	49	93	50
<b>IPW (c, CBPS)</b>	-1	0	1	17	17	14	17	16	14	95	95	97
<b>IPW (m, CBPS)</b>	45	4	-40	23	16	21	22	16	20	47	93	46
<b>AIPW (c, c, GLMPS)</b>	1	1	0	17	16	13	16	16	13	95	95	97
<b>AIPW (c, m, GLMPS)</b>	1	1	0	17	16	13	16	16	14	95	95	97
<b>AIPW (m, c, GLMPS)</b>	0	0	0	18	16	15	17	16	15	95	95	97
<b>AIPW (m, m, GLMPS)</b>	45	5	-41	24	16	22	22	16	21	48	93	47
<b>AIPW (c, c, CBPS)</b>	1	1	0	17	17	14	16	16	14	95	95	98
<b>AIPW (c, m, CBPS)</b>	0	0	1	17	17	14	16	16	14	95	95	97
<b>AIPW (m, c, CBPS)</b>	0	0	0	18	16	15	17	16	14	94	95	97
<b>AIPW (m, m, CBPS)</b>	45	5	-41	23	16	21	22	16	20	46	93	46
<b>MW (c, GLMPS)</b>	-31	-8	23	17	17	14	17	17	14	94	95	95
<b>MW (m, GLMPS)</b>	29	-4	-34	21	17	19	21	17	19	19	93	15
<b>MW (c, CBPS)</b>	-32	-8	24	17	17	14	17	17	14	95	95	95
<b>MW (m, CBPS)</b>	29	-5	-34	21	17	19	22	17	20	20	94	16
<b>AMW (c, c, GLMPS)</b>	-31	-8	23	17	17	14	16	17	14	95	95	95
<b>AMW (c, m, GLMPS)</b>	-30	-8	23	17	17	14	17	17	14	95	95	95
<b>AMW (m, c, GLMPS)</b>	-14	-9	5	17	17	14	17	17	14	81	94	74
<b>AMW (m, m, GLMPS)</b>	30	-4	-34	21	17	19	21	17	19	18	93	14
<b>AMW (c, c, CBPS)</b>	-32	-8	24	17	17	14	17	17	14	95	95	95
<b>AMW (c, m, CBPS)</b>	-32	-8	24	17	17	14	17	17	15	95	95	95
<b>AMW (m, c, CBPS)</b>	-15	-9	6	17	17	14	17	17	14	83	94	77
<b>AMW (m, m, CBPS)</b>	29	-4	-33	21	17	19	21	17	20	20	94	16
<b>OW (c, GLMPS)</b>	-26	-7	19	16	16	13	15	16	13	94	95	95
<b>OW (m, GLMPS)</b>	33	-3	-36	20	16	19	20	16	19	16	93	13
<b>OW (c, CBPS)</b>	-28	-7	21	17	16	14	16	16	14	94	95	95
<b>OW (m, CBPS)</b>	33	-4	-36	21	16	19	21	16	19	18	93	14
<b>AOW (c, c, GLMPS)</b>	-26	-7	19	16	16	14	16	16	14	95	95	96
<b>AOW (c, m, GLMPS)</b>	-28	-7	21	16	16	14	16	16	14	95	95	95
<b>AOW (m, c, GLMPS)</b>	-10	-8	3	16	16	13	16	16	14	82	95	76
<b>AOW (m, m, GLMPS)</b>	33	-3	-36	21	16	19	21	16	19	18	93	14
<b>AOW (c, c, CBPS)</b>	-28	-7	21	17	17	14	17	17	14	95	95	96
<b>AOW (c, m, CBPS)</b>	-30	-7	22	17	17	14	17	17	14	94	95	95
<b>AOW (m, c, CBPS)</b>	-11	-8	3	17	16	14	17	16	14	84	95	79
<b>AOW (m, m, CBPS)</b>	32	-3	-36	21	16	19	21	16	19	19	94	16
<b>MCOV (c)</b>	3	6	3	17	17	15	18	18	16	95	94	97
<b>MCOV (m)</b>	44	7	-37	22	17	21	22	18	21	50	93	58



<b>MGPSV (c, GLMPS)</b>	1	1	0	20	20	16	19	19	16	95	95	97
<b>MGPSV (m, GLMPS)</b>	44	5	-39	25	19	23	25	19	22	57	95	57
<b>MGPSV (c, CBPS)</b>	1	2	1	19	20	15	19	19	16	94	94	97
<b>MGPSV (m, CBPS)</b>	44	5	-39	25	19	23	25	19	22	58	93	59

\*For methods that involve only the propensity score or the outcome model, (c) and (m) denote correct specification and misspecification, respectively. For methods that involve both models, the first and second letter in the parentheses correspond to the treatment model and outcome model, respectively.

**Supplemental Table 8.  $100 \times$  Ratio of 95% confidence interval width to 95% confidence interval width of GLMPS based IPW(c) for n=1500**

Methods*	Scenario 1			Scenario 2			Scenario 3			Scenario 4			Scenario 5		
	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3
NAIVE	104	96	99	85	91	83	83	57	62	79	76	72	117	107	144
OREG (c)	98	97	98	82	82	77	80	67	63	77	67	67	91	90	90
OREG (m)	99	96	101	90	83	83	89	69	69	84	67	73	114	90	122
PEN-GAM (c, c, GLMPS)	101	100	96	98	100	93	90	77	82	95	100	92	111	122	129
PEN-GAM (c, m, GLMPS)	101	100	96	98	100	93	90	77	82	97	100	93	110	119	126
PEN-GAM (m, c, GLMPS)	100	99	96	96	94	89	89	73	78	87	93	84	111	112	123
PEN-GAM (m, m, GLMPS)	101	97	99	105	96	98	98	75	85	87	92	84	140	113	163
IPW (c, GLMPS)	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
IPW (m, GLMPS)	100	97	102	107	94	101	104	91	95	128	78	111	134	97	147
IPW (c, CBPS)	99	99	100	93	94	88	95	83	76	97	92	94	100	100	97
IPW (m, CBPS)	100	96	101	103	91	95	102	81	81	85	79	75	132	98	143
AIPW (c, c, GLMPS)	98	98	98	88	91	84	84	72	69	91	87	84	97	98	96
AIPW (c, m, GLMPS)	99	99	99	92	94	89	88	81	78	95	89	89	100	98	98
AIPW (m, c, GLMPS)	98	97	98	90	88	85	85	70	69	106	71	89	103	96	105
AIPW (m, m, GLMPS)	99	96	101	106	89	98	99	73	79	138	73	115	135	96	147
AIPW (c, c, CBPS)	98	98	98	90	92	85	86	73	69	96	94	94	99	100	98
AIPW (c, m, CBPS)	98	98	99	91	93	87	89	76	73	99	95	96	99	100	98
AIPW (m, c, CBPS)	98	97	98	89	89	84	85	70	68	85	72	73	102	97	103
AIPW (m, m, CBPS)	99	96	101	102	89	95	97	73	78	85	73	73	132	97	144
MW (c, GLMPS)	102	100	107	102	99	99	107	84	81	105	89	94	100	102	100
MW (m, GLMPS)	103	98	106	103	98	98	107	83	80	93	82	86	128	102	138
MW (c, CBPS)	103	100	108	109	105	106	119	93	90	117	101	104	104	107	104
MW (m, CBPS)	103	98	106	106	101	100	114	87	85	100	86	89	131	105	141
AMW (c, c, GLMPS)	101	98	105	100	98	98	103	80	77	102	88	92	99	102	100
AMW (c, m, GLMPS)	101	99	106	101	98	99	105	81	78	103	88	93	100	102	100
AMW (m, c, GLMPS)	101	99	103	94	95	91	98	78	73	91	79	83	101	101	100
AMW (m, m, GLMPS)	102	97	105	102	97	97	106	80	78	92	79	84	128	102	138
AMW (c, c, CBPS)	101	99	106	104	101	104	109	85	78	109	96	101	102	104	104
AMW (c, m, CBPS)	102	99	107	106	102	105	111	89	82	112	97	103	103	104	104
AMW (m, c, CBPS)	101	99	103	96	97	94	102	81	73	96	81	88	102	103	103
AMW (m, m, CBPS)	102	97	105	103	99	100	109	83	79	97	81	88	130	103	141
OW (c, GLMPS)	99	98	101	95	92	93	97	79	78	98	82	89	94	95	95
OW (m, GLMPS)	100	97	102	97	92	93	99	79	79	88	77	81	122	96	134
OW (c, CBPS)	100	98	102	101	98	98	107	86	87	109	94	98	97	99	98
OW (m, CBPS)	100	96	102	100	95	95	106	82	83	93	80	82	125	98	136
AOW (c, c, GLMPS)	99	97	100	96	94	93	96	76	74	98	85	89	98	99	99
AOW (c, m, GLMPS)	100	97	102	99	95	96	100	77	77	102	85	92	98	99	99
AOW (m, c, GLMPS)	99	98	99	91	92	87	91	74	71	88	76	79	98	99	98
AOW (m, m, GLMPS)	100	96	101	99	93	93	100	76	76	91	77	82	125	99	136
AOW (c, c, CBPS)	99	97	101	99	97	98	100	79	76	106	93	98	100	102	103
AOW (c, m, CBPS)	101	97	104	103	98	101	105	83	80	111	94	102	100	102	102

AOW (m, c, CBPS)	99	98	99	93	93	89	94	76	71	91	77	81	100	100	101
AOW (m, m, CBPS)	100	96	101	100	95	95	102	78	77	95	78	84	127	100	138
MCOV (c)	111	109	111	99	102	96	97	78	80	98	87	89	107	109	114
MCOV (m)	115	109	114	106	102	101	104	81	85	96	86	87	136	108	150
MGPS (c, GLMPS)	123	119	116	110	109	102	105	91	90	111	108	104	116	116	111
MGPS (m, GLMPS)	122	116	119	123	108	112	117	90	99	105	97	96	155	115	168
MGPS (c, CBPS)	124	120	121	111	114	106	109	98	96	114	124	118	118	119	116
MGPS (m, CBPS)	123	117	121	121	108	112	118	93	99	100	98	94	155	115	168
MGPSV (c, GLMPS)	118	116	114	108	110	101	107	95	92	111	108	103	115	116	110
MGPSV (m, GLMPS)	118	113	116	118	108	108	114	91	97	105	99	93	149	114	160
MGPSV (c, CBPS)	118	116	114	108	110	101	107	95	93	111	110	105	115	115	111
MGPSV (m, CBPS)	118	113	116	117	108	108	113	91	97	101	96	90	149	114	161

\*For methods that involve only the propensity score or the outcome model, (c) denotes correct specification and misspecification, respectively. For methods that involve both models, the first and second letter in the parentheses correspond to the treatment model and outcome model, respectively.

**Supplemental Table 9. Root mean squared error (RMSE)  $\times$  1000 for n=300**

Methods*	Scenario 1			Scenario 2			Scenario 3			Scenario 4			Scenario 5		
	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3
NAIVE	69	58	83	168	151	72	261	171	416	179	113	101	75	57	48
OREG (c)	57	57	64	64	58	67	68	75	70	62	55	67	35	35	28
OREG (m)	73	56	74	101	60	94	105	81	123	110	57	99	57	35	52
PEN-GAM (c, c, GLMPS)	57	53	62	65	64	76	75	90	99	68	65	80	38	39	36
PEN-GAM (c, m, GLMPS)	57	53	62	65	64	76	75	88	97	68	65	80	38	39	36
PEN-GAM (m, c, GLMPS)	57	53	61	64	62	74	75	87	98	64	66	78	38	39	37
PEN-GAM (m, m, GLMPS)	65	52	70	102	67	91	119	95	163	64	67	79	63	40	58
IPW (c, GLMPS)	58	58	65	76	70	88	86	113	110	82	78	96	39	39	32
IPW (m, GLMPS)	73	57	75	120	69	113	122	108	156	120	70	141	72	38	66
IPW (c, CBPS)	58	58	65	75	70	82	89	98	86	89	79	104	40	39	32
IPW (m, CBPS)	73	56	75	116	67	112	115	98	138	79	68	87	68	38	63
AIPW (c, c, GLMPS)	58	57	64	73	64	79	76	83	81	76	70	89	38	38	31
AIPW (c, m, GLMPS)	57	57	64	75	65	83	79	94	92	80	74	94	40	38	33
AIPW (m, c, GLMPS)	57	57	64	72	62	77	75	81	80	113	59	115	41	37	35
AIPW (m, m, GLMPS)	73	56	74	127	64	121	127	86	145	169	63	171	194	37	192
AIPW (c, c, CBPS)	58	57	64	74	70	80	81	88	82	92	86	108	41	40	34
AIPW (c, m, CBPS)	57	57	64	77	71	83	84	92	88	102	88	116	41	40	34
AIPW (m, c, CBPS)	57	57	64	71	65	77	75	82	78	74	63	78	42	38	35
AIPW (m, m, CBPS)	73	56	74	119	66	113	120	88	139	75	64	80	69	38	64
MW (c, GLMPS)	60	58	70	86	71	91	95	99	86	90	74	95	49	39	39
MW (m, GLMPS)	75	56	79	96	69	99	99	118	147	74	66	84	56	39	56
MW (c, CBPS)	62	58	73	101	78	101	116	105	99	113	89	113	53	42	41
MW (m, CBPS)	75	56	79	96	72	102	103	121	137	90	73	95	57	41	57
AMW (c, c, GLMPS)	60	57	70	85	71	90	92	95	83	88	73	93	50	40	39
AMW (c, m, GLMPS)	59	57	69	85	71	90	93	96	83	89	72	94	49	39	39
AMW (m, c, GLMPS)	59	57	66	78	69	82	92	101	79	73	67	84	42	40	32
AMW (m, m, GLMPS)	75	56	78	96	68	99	99	115	146	73	68	85	56	39	56
AMW (c, c, CBPS)	60	57	71	92	74	99	101	104	86	100	84	107	52	41	42
AMW (c, m, CBPS)	61	57	72	94	75	98	103	104	88	103	85	109	52	41	41
AMW (m, c, CBPS)	59	57	66	81	71	87	100	109	80	85	68	92	43	42	34
AMW (m, m, CBPS)	75	56	78	97	71	101	100	124	150	84	69	90	57	40	57
OW (c, GLMPS)	58	57	67	78	66	84	85	90	83	88	68	95	45	37	36
OW (m, GLMPS)	74	56	76	99	65	101	102	103	143	73	62	83	57	36	56
OW (c, CBPS)	59	57	69	95	74	95	106	97	100	113	84	114	50	39	39
OW (m, CBPS)	74	56	76	97	68	104	102	105	130	86	68	91	57	38	58
AOW (c, c, GLMPS)	58	57	66	79	68	84	84	87	80	87	69	95	47	39	37
AOW (c, m, GLMPS)	58	57	68	82	69	87	87	87	82	89	70	96	47	38	38
AOW (m, c, GLMPS)	58	57	64	73	66	77	82	89	77	71	63	79	40	39	32
AOW (m, m, GLMPS)	73	56	75	99	66	100	101	100	140	72	64	82	57	38	56
AOW (c, c, CBPS)	58	57	67	87	72	93	92	94	83	99	81	108	50	40	40
AOW (c, m, CBPS)	61	57	71	92	72	95	96	95	87	103	82	111	50	40	40
AOW (m, c, CBPS)	58	57	64	76	69	82	89	96	78	80	65	88	41	40	33
AOW (m, m, CBPS)	73	56	75	99	68	101	102	109	144	82	66	89	57	39	57

MCOV (c)	66	64	73	83	84	82	116	114	181	86	76	84	40	41	36
MCOV (m)	81	64	87	127	84	102	155	119	230	87	76	83	65	41	58
MGPSV (c, GLMPS)	71	71	80	85	80	90	93	111	109	89	90	107	46	46	37
MGPSV (m, GLMPS)	85	71	88	128	79	121	134	111	159	81	80	95	74	44	68
MGPSV (c, CBPS)	71	68	78	86	82	96	97	113	114	93	96	111	45	44	40
MGPSV (m, CBPS)	83	69	86	129	80	124	141	111	166	81	83	95	75	44	70
MGPSV (c, GLMPS)	65	66	75	80	75	86	90	103	106	88	82	96	44	44	35
MGPSV (m, GLMPS)	80	65	84	122	76	112	132	108	174	82	77	92	71	43	64
MGPSV (c, CBPS)	66	65	75	83	76	88	95	102	116	90	87	97	43	43	36
MGPSV (m, CBPS)	80	65	84	124	78	112	136	111	180	79	78	90	71	43	64

\*For methods that involve only the propensity score or the outcome model, (c) denotes correct specification and misspecification, respectively. For methods that involve both models, the first and second letter in the parentheses correspond to the treatment model and outcome model, respectively.

**Supplemental Table 10. Performance of different causal inference methods in scenario 1 (n=300) of simulation studies.**

Methods*	Bias from ATE × 1000			Empirical SD × 1000			Average SE × 1000			95% Coverage Rate (%)		
	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3
NAIVE	34	-20	-54	60	54	62	60	54	63	91	93	85
OREG (c)	-1	-1	0	57	57	64	58	57	64	95	94	94
OREG (m)	44	8	-36	58	55	65	58	55	65	87	94	91
PEN-GAM (c, c, GLMPS)	-11	-5	6	56	52	62	63	63	66	97	98	97
PEN-GAM (c, m, GLMPS)	-10	-5	6	56	53	62	63	62	66	97	98	96
PEN-GAM (m, c, GLMPS)	-13	-7	5	55	52	61	64	63	67	97	98	97
PEN-GAM (m, m, GLMPS)	31	1	-29	57	52	63	64	62	68	95	98	95
IPW (c, GLMPS)	0	0	-1	58	58	65	59	59	66	95	95	95
IPW (m, GLMPS)	45	8	-37	58	56	66	59	56	66	88	94	91
IPW (c, CBPS)	-1	0	2	58	58	65	58	57	65	95	94	94
IPW (m, CBPS)	45	8	-37	58	56	65	59	56	66	87	94	91
AIPW (c, c, GLMPS)	0	-1	-1	58	57	64	58	58	64	95	95	94
AIPW (c, m, GLMPS)	1	0	-1	57	57	64	59	59	65	95	95	94
AIPW (m, c, GLMPS)	0	-1	0	57	57	64	58	57	64	95	94	94
AIPW (m, m, GLMPS)	44	8	-36	58	55	65	59	56	65	87	94	91
AIPW (c, c, CBPS)	0	-1	-1	58	57	64	58	57	64	95	94	94
AIPW (c, m, CBPS)	-1	1	2	57	57	64	58	57	65	95	94	94
AIPW (m, c, CBPS)	0	-1	0	57	57	64	58	57	64	94	94	94
AIPW (m, m, CBPS)	44	8	-36	58	55	65	59	56	65	87	94	91
MW (c, GLMPS)	-15	4	19	58	57	67	61	59	70	96	95	95
MW (m, GLMPS)	45	4	-42	60	56	67	61	57	69	83	94	85
MW (c, CBPS)	-18	6	24	59	58	69	63	60	72	96	95	95
MW (m, CBPS)	46	4	-41	60	56	67	62	58	70	83	95	85
AMW (c, c, GLMPS)	-14	5	19	58	57	67	60	58	69	95	95	95
AMW (c, m, GLMPS)	-14	5	18	58	57	67	61	58	69	96	95	95
AMW (m, c, GLMPS)	0	-5	-5	59	57	66	60	58	67	95	94	93
AMW (m, m, GLMPS)	45	4	-41	60	56	67	61	57	68	83	95	85
AMW (c, c, CBPS)	-16	6	22	58	57	68	61	58	70	96	95	95
AMW (c, m, CBPS)	-18	5	23	59	57	68	62	59	70	96	95	95
AMW (m, c, CBPS)	0	-6	-5	59	57	66	60	58	67	95	94	93
AMW (m, m, CBPS)	45	4	-41	60	55	67	61	57	69	83	95	85
OW (c, GLMPS)	-9	3	12	57	57	65	59	57	66	95	94	95
OW (m, GLMPS)	45	6	-39	58	56	65	59	56	66	85	94	88
OW (c, CBPS)	-13	4	17	58	57	67	60	57	68	95	94	95
OW (m, CBPS)	45	7	-39	58	55	66	59	56	66	85	95	88
AOW (c, c, GLMPS)	-9	3	12	57	56	65	59	57	66	95	95	94
AOW (c, m, GLMPS)	-13	6	19	57	56	66	60	57	67	95	95	95
AOW (m, c, GLMPS)	-1	-3	-2	58	57	64	59	57	65	95	95	94
AOW (m, m, GLMPS)	44	6	-38	59	55	65	60	56	66	86	95	88
AOW (c, c, CBPS)	-10	4	14	57	56	66	59	57	67	95	95	94
AOW (c, m, CBPS)	-18	6	24	58	57	67	61	58	69	95	95	95
AOW (m, c, CBPS)	-1	-3	-2	58	57	64	59	57	65	95	95	94
AOW (m, m, CBPS)	44	6	-38	59	55	65	60	56	66	86	95	88
MCOV (c)	-4	-5	-1	66	63	73	66	63	72	95	94	94
MCOV (m)	42	-4	-46	69	64	74	68	63	74	90	95	89

<b>MGPSV (c, GLMPS)</b>	0	-1	-1	65	66	75	69	66	75	96	95	94
<b>MGPSV (m, GLMPS)</b>	44	5	-38	67	65	75	69	65	76	91	95	92
<b>MGPSV (c, CBPS)</b>	0	-3	-2	66	65	74	69	66	75	96	95	95
<b>MGPSV (m, CBPS)</b>	44	5	-39	67	65	75	69	65	76	90	95	92

\*For methods that involve only the propensity score or the outcome model, (c) and (m) denote correct specification and misspecification, respectively. For methods that involve both models, the first and second letter in the parentheses correspond to the treatment model and outcome model, respectively.

**Supplemental Table 11. Performance of different causal inference methods in scenario 2 (n=300) of simulation studies.**

Methods*	Bias from ATE × 1000			Empirical SD × 1000			Average SE × 1000			95% Coverage Rate (%)		
	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3
NAIVE	156	137	-19	62	63	69	62	60	69	28	39	94
OREG (c)	0	0	0	64	58	67	61	57	67	93	94	94
OREG (m)	74	14	-59	69	58	72	66	57	71	79	93	85
PEN-GAM (c, c, GLMPS)	1	18	16	65	62	74	78	80	91	98	99	98
PEN-GAM (c, m, GLMPS)	3	15	12	65	62	75	78	79	90	97	98	98
PEN-GAM (m, c, GLMPS)	6	13	7	64	61	73	78	76	88	98	99	97
PEN-GAM (m, m, GLMPS)	75	25	-50	68	62	76	79	76	89	88	98	94
IPW (c, GLMPS)	6	7	1	76	70	88	73	67	82	93	94	93
IPW (m, GLMPS)	87	18	-68	82	67	90	77	64	84	76	93	83
IPW (c, CBPS)	-11	-4	7	74	70	82	73	68	80	93	94	94
IPW (m, CBPS)	83	12	-71	81	66	86	77	65	82	79	94	83
AIPW (c, c, GLMPS)	2	0	-2	73	64	79	68	66	77	93	95	95
AIPW (c, m, GLMPS)	4	1	-3	75	65	83	76	68	85	94	95	95
AIPW (m, c, GLMPS)	0	1	0	72	62	77	69	62	75	93	95	94
AIPW (m, m, GLMPS)	88	14	-75	91	62	95	80	61	86	80	94	84
AIPW (c, c, CBPS)	2	0	-1	74	70	80	76	72	82	95	96	96
AIPW (c, m, CBPS)	-12	-3	9	76	71	82	79	74	86	94	96	96
AIPW (m, c, CBPS)	0	1	1	71	65	77	70	65	76	93	95	94
AIPW (m, m, CBPS)	87	14	-73	81	65	86	78	65	83	79	95	83
MW (c, GLMPS)	-37	-10	27	78	70	87	77	69	86	94	94	95
MW (m, GLMPS)	57	5	-52	77	68	84	77	68	84	76	94	83
MW (c, CBPS)	-54	-17	37	86	76	95	90	80	97	95	96	95
MW (m, CBPS)	52	0	-52	81	72	88	85	74	92	81	96	85
AMW (c, c, GLMPS)	-37	-10	27	77	70	86	76	69	85	93	94	94
AMW (c, m, GLMPS)	-35	-10	26	77	70	86	76	69	85	94	94	94
AMW (m, c, GLMPS)	-27	-10	18	73	68	80	71	67	80	93	94	94
AMW (m, m, GLMPS)	58	5	-53	77	68	84	76	67	84	75	94	82
AMW (c, c, CBPS)	-44	-10	34	81	73	93	84	76	94	95	96	95
AMW (c, m, CBPS)	-47	-13	33	82	74	92	85	77	95	95	96	95
AMW (m, c, CBPS)	-31	-9	22	75	71	84	76	72	86	95	95	95
AMW (m, m, CBPS)	56	6	-51	79	71	87	81	72	89	77	95	84
OW (c, GLMPS)	-28	-8	20	73	66	82	71	64	80	94	94	94
OW (m, GLMPS)	66	6	-60	74	65	81	72	63	80	74	94	81
OW (c, CBPS)	-49	-19	30	81	71	90	83	74	91	94	95	94
OW (m, CBPS)	59	-2	-61	78	68	85	79	69	86	79	95	84
AOW (c, c, GLMPS)	-28	-9	20	74	68	82	73	66	81	94	95	95
AOW (c, m, GLMPS)	-31	-8	23	76	68	84	75	67	83	94	94	95
AOW (m, c, GLMPS)	-18	-9	9	70	66	77	69	65	77	93	94	95
AOW (m, m, GLMPS)	64	5	-59	75	66	81	74	65	81	75	94	83
AOW (c, c, CBPS)	-36	-9	27	79	71	89	81	74	91	95	96	95
AOW (c, m, CBPS)	-44	-12	31	81	71	90	83	75	92	94	96	95
AOW (m, c, CBPS)	-22	-9	13	73	68	81	74	69	82	94	95	95
AOW (m, m, CBPS)	62	5	-57	77	68	84	79	69	86	78	95	85
MCOV (c)	37	44	7	74	71	81	73	69	82	92	90	95
MCOV (m)	100	43	-56	79	72	85	76	70	84	71	89	88



<b>MGPSV (c, GLMPS)</b>	5	3	-2	85	80	90	77	69	83	92	92	93
<b>MGPSV (m, GLMPS)</b>	89	17	-72	92	77	98	84	71	90	79	93	83
<b>MGPSV (c, CBPS)</b>	14	9	-4	85	82	95	87	84	100	95	95	96
<b>MGPSV (m, CBPS)</b>	91	18	-73	91	78	101	90	78	99	81	94	86
<b>MGPSV (c, GLMPS)</b>	9	9	-1	79	75	86	80	73	87	94	94	95
<b>MGPSV (m, GLMPS)</b>	87	22	-66	85	73	91	84	73	90	80	94	87
<b>MGPSV (c, CBPS)</b>	15	15	0	82	75	88	79	74	87	93	94	95
<b>MGPSV (m, CBPS)</b>	89	24	-65	86	74	91	84	73	90	80	94	88

\*For methods that involve only the propensity score or the outcome model, (c) and (m) denote correct specification and misspecification, respectively. For methods that involve both models, the first and second letter in the parentheses correspond to the treatment model and outcome model, respectively.

**Supplemental Table 12. Performance of different causal inference methods in scenario 3 (n=300) of simulation studies.**

Methods*	Bias from ATE × 1000			Empirical SD × 1000			Average SE × 1000			95% Coverage Rate (%)		
	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3
<b>NAIVE</b>	251	-159	-410	67	61	63	67	61	63	4	27	0
<b>OREG (c)</b>	1	-5	-6	67	74	68	67	78	71	95	96	95
<b>OREG (m)</b>	73	-24	-98	73	76	73	74	78	75	83	95	78
<b>PEN-GAM (c, c, GLMPS)</b>	23	-41	-64	71	80	76	83	97	101	97	96	97
<b>PEN-GAM (c, m, GLMPS)</b>	23	-38	-61	71	80	75	82	96	100	97	97	97
<b>PEN-GAM (m, c, GLMPS)</b>	26	-39	-65	70	78	73	81	92	97	96	96	97
<b>PEN-GAM (m, m, GLMPS)</b>	92	-51	-143	76	80	78	84	91	97	84	94	75
<b>IPW (c, GLMPS)</b>	6	-13	-19	84	110	106	85	104	101	95	93	93
<b>IPW (m, GLMPS)</b>	84	-31	-115	88	100	103	86	98	98	83	92	75
<b>IPW (c, CBPS)</b>	-17	-5	12	83	94	86	84	95	84	94	94	93
<b>IPW (m, CBPS)</b>	74	-30	-105	85	89	88	86	90	86	87	93	79
<b>AIPW (c, c, GLMPS)</b>	1	-3	-4	71	81	77	74	89	83	95	97	95
<b>AIPW (c, m, GLMPS)</b>	3	-6	-9	74	90	86	84	107	100	96	96	96
<b>AIPW (m, c, GLMPS)</b>	1	-4	-6	72	79	76	75	83	81	95	96	95
<b>AIPW (m, m, GLMPS)</b>	89	-23	-112	210	81	213	96	84	101	85	95	81
<b>AIPW (c, c, CBPS)</b>	2	-3	-5	77	86	80	82	92	84	96	97	95
<b>AIPW (c, m, CBPS)</b>	-10	5	15	79	89	87	87	96	91	96	96	95
<b>AIPW (m, c, CBPS)</b>	2	-4	-6	73	80	76	75	84	79	95	96	95
<b>AIPW (m, m, CBPS)</b>	85	-23	-108	83	82	85	84	85	86	84	95	79
<b>MW (c, GLMPS)</b>	-32	-37	-5	90	92	85	91	94	87	95	94	95
<b>MW (m, GLMPS)</b>	43	-76	-119	89	90	83	90	91	86	86	93	75
<b>MW (c, CBPS)</b>	-46	-20	26	102	103	97	109	111	102	96	96	94
<b>MW (m, CBPS)</b>	30	-72	-103	96	97	90	102	102	95	91	94	84
<b>AMW (c, c, GLMPS)</b>	-31	-37	-6	88	89	80	87	90	83	95	94	95
<b>AMW (c, m, GLMPS)</b>	-30	-38	-9	88	89	81	89	91	84	96	94	95
<b>AMW (m, c, GLMPS)</b>	-42	-53	-11	83	86	76	83	88	79	95	95	95
<b>AMW (m, m, GLMPS)</b>	44	-75	-119	87	87	81	89	89	83	85	94	73
<b>AMW (c, c, CBPS)</b>	-33	-38	-5	95	96	83	99	101	89	96	96	96
<b>AMW (c, m, CBPS)</b>	-34	-28	6	94	99	86	101	105	91	97	96	96
<b>AMW (m, c, CBPS)</b>	-47	-59	-12	88	90	78	90	94	82	95	95	96
<b>AMW (m, m, CBPS)</b>	41	-82	-123	91	92	83	96	95	86	88	94	74
<b>OW (c, GLMPS)</b>	-20	-23	-4	82	87	82	83	88	83	95	94	94
<b>OW (m, GLMPS)</b>	58	-57	-115	82	86	82	83	86	83	85	93	74
<b>OW (c, CBPS)</b>	-40	-3	37	94	97	95	99	103	96	95	95	93
<b>OW (m, CBPS)</b>	43	-52	-95	90	91	88	94	95	91	90	94	84
<b>AOW (c, c, GLMPS)</b>	-18	-24	-5	82	84	78	82	86	80	95	94	95
<b>AOW (c, m, GLMPS)</b>	-21	-25	-4	84	84	80	85	86	82	95	95	95
<b>AOW (m, c, GLMPS)</b>	-25	-34	-9	78	82	75	78	84	78	95	95	95
<b>AOW (m, m, GLMPS)</b>	57	-56	-113	82	83	79	84	85	82	85	93	74
<b>AOW (c, c, CBPS)</b>	-22	-27	-5	88	90	81	93	95	86	96	96	96
<b>AOW (c, m, CBPS)</b>	-27	-18	10	89	93	85	97	100	90	97	96	96
<b>AOW (m, c, CBPS)</b>	-31	-41	-10	81	86	76	84	89	80	95	95	95
<b>AOW (m, m, CBPS)</b>	53	-63	-117	85	87	81	91	91	85	88	94	75
<b>MCOV (c)</b>	83	-78	-161	80	82	81	80	79	81	81	82	50
<b>MCOV (m)</b>	131	-81	-212	85	84	86	84	82	84	65	82	30

<b>MGPSV (c, GLMPS)</b>	22	-25	-47	87	100	94	90	101	96	94	93	92
<b>MGPSV (m, GLMPS)</b>	96	-47	-143	91	95	96	93	96	98	83	92	71
<b>MGPSV (c, CBPS)</b>	36	-29	-65	88	98	97	89	100	97	93	94	90
<b>MGPSV (m, CBPS)</b>	102	-49	-150	92	92	96	92	96	98	80	92	68

\*For methods that involve only the propensity score or the outcome model, (c) and (m) denote correct specification and misspecification, respectively. For methods that involve both models, the first and second letter in the parentheses correspond to the treatment model and outcome model, respectively.

**Supplemental Table 13. Performance of different causal inference methods in scenario 4 (n=300) of simulation studies.**

Methods*	Bias from ATE × 1000			Empirical SD × 1000			Average SE × 1000			95% Coverage Rate (%)		
	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3
NAIVE	167	93	-74	61	61	70	61	60	68	23	64	80
OREG (c)	-2	1	3	61	55	66	62	56	67	95	95	95
OREG (m)	83	16	-67	66	55	71	67	55	71	77	94	84
PEN-GAM (c, c, GLMPS)	-1	17	18	68	63	78	82	81	95	98	99	98
PEN-GAM (c, m, GLMPS)	0	16	15	68	63	78	83	80	94	98	99	98
PEN-GAM (m, c, GLMPS)	-3	20	24	64	63	75	76	80	90	97	99	98
PEN-GAM (m, m, GLMPS)	-3	19	22	64	64	76	76	80	90	97	98	98
IPW (c, GLMPS)	3	19	16	81	77	99	82	74	94	95	94	94
IPW (m, GLMPS)	-65	27	93	104	67	109	94	64	99	86	92	81
IPW (c, CBPS)	-25	6	31	84	80	100	83	75	93	92	94	91
IPW (m, CBPS)	-30	2	32	72	66	80	73	68	80	92	96	92
AIPW (c, c, GLMPS)	-2	1	3	73	73	89	88	92	102	96	97	96
AIPW (c, m, GLMPS)	0	3	3	79	77	96	101	96	119	97	97	97
AIPW (m, c, GLMPS)	1	3	2	116	65	120	115	65	120	97	96	96
AIPW (m, m, GLMPS)	-45	5	49	161	77	159	205	72	208	99	96	97
AIPW (c, c, CBPS)	-1	2	3	91	89	109	91	83	103	95	95	95
AIPW (c, m, CBPS)	-34	-1	32	94	90	113	96	85	109	93	95	94
AIPW (m, c, CBPS)	1	3	2	72	63	78	77	67	80	96	97	96
AIPW (m, m, CBPS)	-8	0	8	73	64	79	80	69	84	97	97	96
MW (c, GLMPS)	-39	-16	23	83	72	93	87	75	94	95	96	95
MW (m, GLMPS)	-9	-9	0	73	66	85	76	68	85	94	95	94
MW (c, CBPS)	-61	-20	41	97	86	106	107	95	113	96	97	96
MW (m, CBPS)	-42	-14	28	81	71	90	88	77	94	97	96	96
AMW (c, c, GLMPS)	-37	-16	22	81	71	92	84	73	92	95	96	95
AMW (c, m, GLMPS)	-37	-15	22	83	71	93	85	73	93	95	96	95
AMW (m, c, GLMPS)	-10	-24	-14	73	64	83	75	66	83	94	95	93
AMW (m, m, GLMPS)	-4	-25	-21	73	64	83	75	66	83	93	95	92
AMW (c, c, CBPS)	-45	-17	28	91	82	106	98	88	109	96	96	95
AMW (c, m, CBPS)	-49	-18	31	92	83	107	100	89	111	96	96	96
AMW (m, c, CBPS)	-36	-11	24	77	66	89	83	72	92	96	96	95
AMW (m, m, CBPS)	-33	-14	18	78	67	88	84	73	91	96	96	96
OW (c, GLMPS)	-43	-6	38	78	67	89	80	68	88	95	95	95
OW (m, GLMPS)	-19	1	20	70	62	80	71	63	80	95	95	95
OW (c, CBPS)	-70	-17	53	92	81	103	99	88	106	95	96	95
OW (m, CBPS)	-44	-11	33	76	66	85	81	71	88	96	96	96
AOW (c, c, GLMPS)	-42	-6	36	79	69	89	81	71	89	95	95	95
AOW (c, m, GLMPS)	-39	-7	32	82	69	91	84	71	92	95	95	95
AOW (m, c, GLMPS)	-12	-17	-5	70	62	79	72	63	79	94	95	92
AOW (m, m, GLMPS)	-8	-18	-10	72	62	81	74	64	81	93	95	91
AOW (c, c, CBPS)	-49	-9	40	89	80	103	95	86	106	96	96	95
AOW (c, m, CBPS)	-51	-12	39	92	82	106	99	87	109	96	96	95
AOW (m, c, CBPS)	-32	-7	25	74	64	84	80	69	87	96	96	95
AOW (m, m, CBPS)	-32	-10	22	76	65	86	82	70	89	96	96	95
MCOV (c)	39	31	-7	75	68	83	76	68	84	92	92	95
MCOV (m)	41	33	-8	73	69	83	76	68	83	92	92	94

<b>MGPSV (c, GLMPS)</b>	4	8	4	91	92	109	80	72	88	91	90	90
<b>MGPSV (m, GLMPS)</b>	8	4	-4	82	84	96	77	72	87	93	92	92
<b>MGPSV (c, CBPS)</b>	15	18	3	93	97	116	100	100	122	96	96	96
<b>MGPSV (m, CBPS)</b>	3	3	0	80	83	97	85	85	100	96	96	96
<b>MGPSV (c, GLMPS)</b>	16	12	-5	85	82	97	86	80	96	94	94	94
<b>MGPSV (m, GLMPS)</b>	10	5	-4	79	77	88	82	75	90	95	94	95
<b>MGPSV (c, CBPS)</b>	24	27	2	83	83	97	86	81	97	94	94	95
<b>MGPSV (m, CBPS)</b>	4	5	0	77	75	86	80	74	88	95	95	95

\*For methods that involve only the propensity score or the outcome model, (c) and (m) denote correct specification and misspecification, respectively. For methods that involve both models, the first and second letter in the parentheses correspond to the treatment model and outcome model, respectively.

**Supplemental Table 14. Performance of different causal inference methods in scenario 5 (n=300) of simulation studies.**

Methods*	Bias from ATE × 1000			Empirical SD × 1000			Average SE × 1000			95% Coverage Rate (%)		
	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3
NAIVE	62	41	-21	43	39	44	43	39	45	71	81	94
OREG (c)	-4	-3	1	35	35	28	38	39	32	97	97	98
OREG (m)	37	2	-34	43	35	39	45	37	41	90	96	90
PEN-GAM (c, c, GLMPS)	-1	2	4	37	39	36	55	58	59	100	100	100
PEN-GAM (c, m, GLMPS)	-1	5	7	38	38	35	54	58	58	100	100	100
PEN-GAM (m, c, GLMPS)	-3	2	5	38	39	37	53	57	57	99	100	100
PEN-GAM (m, m, GLMPS)	42	7	-34	48	40	47	60	57	64	96	100	98
IPW (c, GLMPS)	0	1	1	39	39	32	37	37	31	95	94	96
IPW (m, GLMPS)	44	5	-39	57	38	53	48	36	44	87	94	89
IPW (c, CBPS)	-4	-2	2	40	39	32	39	39	31	96	95	97
IPW (m, CBPS)	43	3	-40	53	38	49	49	37	44	87	95	89
AIPW (c, c, GLMPS)	-2	-2	1	38	38	31	40	42	34	97	97	98
AIPW (c, m, GLMPS)	0	-1	-1	40	38	33	46	42	38	97	97	98
AIPW (m, c, GLMPS)	-4	-2	2	41	37	35	43	40	37	96	97	98
AIPW (m, m, GLMPS)	47	3	-44	188	37	187	61	38	57	90	96	91
AIPW (c, c, CBPS)	-2	-1	0	41	40	34	43	44	35	97	98	99
AIPW (c, m, CBPS)	-5	-1	4	41	40	34	44	43	37	97	98	98
AIPW (m, c, CBPS)	-3	-2	2	41	38	35	42	41	36	96	97	98
AIPW (m, m, CBPS)	43	3	-40	54	38	49	51	39	46	89	97	90
MW (c, GLMPS)	-31	-9	22	38	38	32	38	39	32	95	95	98
MW (m, GLMPS)	29	-5	-34	49	38	44	49	38	44	77	95	78
MW (c, CBPS)	-34	-10	24	41	40	33	44	45	35	97	97	99
MW (m, CBPS)	27	-7	-34	50	40	46	53	42	48	81	96	83
AMW (c, c, GLMPS)	-32	-9	22	38	39	32	40	41	34	96	96	98
AMW (c, m, GLMPS)	-30	-9	21	38	38	33	41	40	34	96	96	98
AMW (m, c, GLMPS)	-17	-11	6	39	39	32	41	42	35	94	96	94
AMW (m, m, GLMPS)	28	-6	-34	49	39	44	50	40	45	78	96	78
AMW (c, c, CBPS)	-34	-10	24	40	40	34	44	45	36	97	97	98
AMW (c, m, CBPS)	-33	-10	23	40	40	34	44	44	37	97	97	99
AMW (m, c, CBPS)	-18	-11	7	40	40	33	43	44	37	95	97	95
AMW (m, m, CBPS)	28	-6	-34	50	40	46	52	42	47	81	96	81
OW (c, GLMPS)	-27	-8	19	36	36	30	36	36	30	95	94	97
OW (m, GLMPS)	32	-4	-36	47	36	43	46	36	42	75	94	76
OW (c, CBPS)	-32	-10	22	38	38	32	40	40	33	97	96	98
OW (m, CBPS)	30	-6	-37	48	37	45	50	38	45	80	95	80
AOW (c, c, GLMPS)	-28	-9	19	37	38	32	39	40	33	96	97	98
AOW (c, m, GLMPS)	-28	-8	20	37	37	32	40	39	34	97	96	98
AOW (m, c, GLMPS)	-14	-10	4	38	38	31	40	41	34	94	97	94
AOW (m, m, GLMPS)	31	-5	-36	48	37	43	48	39	44	79	96	79
AOW (c, c, CBPS)	-32	-9	22	39	39	33	43	43	36	97	97	98
AOW (c, m, CBPS)	-31	-9	22	39	39	33	43	43	36	97	97	99
AOW (m, c, CBPS)	-15	-10	5	39	39	33	42	43	36	95	97	95
AOW (m, m, CBPS)	30	-5	-36	49	39	45	51	41	47	82	96	82
MCOV (c)	5	9	4	40	40	35	40	40	37	94	93	96
MCOV (m)	43	9	-34	49	41	46	49	40	46	87	93	89

<b>MGPSV (c, GLMPS)</b>	0	0	0	46	46	37	40	39	32	92	91	96
<b>MGPSV (m, GLMPS)</b>	43	3	-39	57	43	51	53	41	48	87	94	90
<b>MGPSV (c, CBPS)</b>	2	3	0	43	43	36	42	41	35	94	94	96
<b>MGPSV (m, CBPS)</b>	44	5	-39	56	43	51	53	41	48	87	93	89

\*For methods that involve only the propensity score or the outcome model, (c) and (m) denote correct specification and misspecification, respectively. For methods that involve both models, the first and second letter in the parentheses correspond to the treatment model and outcome model, respectively.

**Supplemental Table 15.  $100 \times$  Ratio of 95% confidence interval width to 95% confidence interval width of GLMPS based IPW(c) for n=300**

Methods*	Scenario 1			Scenario 2			Scenario 3			Scenario 4			Scenario 5		
	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3	1v2	1v3	2v3
NAIVE	101	93	96	84	90	84	80	59	64	75	83	74	115	106	144
OREG (c)	97	97	97	84	86	81	80	74	71	76	77	72	101	105	104
OREG (m)	98	94	99	91	85	86	87	75	75	82	76	77	120	99	131
PEN-GAM (c, c, GLMPS)	107	108	101	107	120	111	99	95	103	102	113	103	147	160	191
PEN-GAM (c, m, GLMPS)	106	107	100	106	118	109	99	94	102	103	112	102	145	159	189
PEN-GAM (m, c, GLMPS)	108	108	102	106	114	107	98	90	99	94	112	98	144	157	184
PEN-GAM (m, m, GLMPS)	108	106	103	109	114	109	100	88	99	94	111	98	162	155	207
IPW (c, GLMPS)	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
IPW (m, GLMPS)	99	96	101	104	96	102	103	94	98	113	88	105	128	96	142
IPW (c, CBPS)	98	98	99	100	102	97	100	92	85	102	103	100	104	105	99
IPW (m, CBPS)	99	95	100	105	97	99	102	87	87	90	94	85	131	100	141
AIPW (c, c, GLMPS)	98	98	98	93	99	93	87	83	80	99	109	103	107	113	109
AIPW (c, m, GLMPS)	100	100	99	104	102	103	97	95	94	115	117	117	122	114	123
AIPW (m, c, GLMPS)	97	97	97	93	92	91	87	79	79	132	85	119	114	108	119
AIPW (m, m, GLMPS)	98	95	99	110	92	104	100	80	88	211	89	186	163	103	182
AIPW (c, c, CBPS)	97	97	97	104	108	99	98	88	84	111	113	110	115	119	114
AIPW (c, m, CBPS)	98	97	98	108	110	105	104	93	91	118	116	116	118	117	118
AIPW (m, c, CBPS)	97	97	97	95	98	92	89	81	79	93	93	85	113	112	116
AIPW (m, m, CBPS)	98	95	99	106	97	101	100	82	86	98	96	90	137	107	149
MW (c, GLMPS)	103	100	106	105	103	104	108	91	88	106	103	102	103	105	102
MW (m, GLMPS)	103	98	105	105	101	102	107	88	86	93	93	91	130	104	142
MW (c, CBPS)	106	102	110	122	120	118	130	108	102	130	131	122	119	121	113
MW (m, CBPS)	104	99	106	116	111	111	121	98	95	108	107	102	142	114	153
AMW (c, c, GLMPS)	101	99	104	103	103	103	104	87	83	102	100	99	108	111	108
AMW (c, m, GLMPS)	102	98	105	104	103	104	106	87	84	104	101	100	108	108	109
AMW (m, c, GLMPS)	101	99	102	97	101	97	98	85	80	91	91	89	109	113	112
AMW (m, m, GLMPS)	103	97	104	104	100	102	105	85	83	92	91	90	133	108	144
AMW (c, c, CBPS)	103	99	106	114	114	115	118	98	89	120	121	117	118	121	116
AMW (c, m, CBPS)	104	100	107	116	115	115	121	102	92	123	123	119	119	120	119
AMW (m, c, CBPS)	101	99	102	104	107	104	107	90	82	101	99	99	116	119	119
AMW (m, m, CBPS)	103	97	104	111	107	108	114	92	87	102	100	99	140	114	152
OW (c, GLMPS)	99	97	101	97	96	97	98	85	84	98	94	95	95	97	96
OW (m, GLMPS)	100	95	101	99	94	97	99	83	84	87	87	86	123	97	135
OW (c, CBPS)	101	98	103	113	110	110	118	100	97	121	120	114	108	110	105
OW (m, CBPS)	100	95	101	108	103	104	111	92	91	99	98	95	132	104	144
AOW (c, c, GLMPS)	99	97	100	99	99	98	97	82	81	99	97	96	105	108	108
AOW (c, m, GLMPS)	100	97	102	102	99	101	100	83	83	102	98	99	106	106	109
AOW (m, c, GLMPS)	99	97	98	94	97	93	92	81	78	88	87	85	106	110	110
AOW (m, m, GLMPS)	100	95	101	101	97	98	100	81	82	90	88	88	130	105	142
AOW (c, c, CBPS)	100	97	102	110	110	110	111	92	86	116	118	114	114	117	115
AOW (c, m, CBPS)	102	98	104	113	112	112	115	96	90	121	120	117	116	116	117
AOW (m, c, CBPS)	99	97	99	100	104	100	100	86	80	97	95	94	112	116	117



AOW (m, m, CBPS)	101	95	101	107	103	105	108	87	85	100	97	96	137	110	150
MCOV (c)	111	107	110	100	103	99	96	77	82	94	94	91	106	108	118
MCOV (m)	114	107	112	103	104	102	100	80	85	93	94	90	132	109	150
MGPSV (c, GLMPS)	117	113	111	105	103	101	100	93	95	97	100	95	107	105	104
MGPSV (m, GLMPS)	118	112	115	115	106	110	108	94	102	94	98	93	140	107	153
MGPSV (c, CBPS)	123	120	120	119	126	121	116	116	120	122	136	130	123	123	127
MGPSV (m, CBPS)	122	117	121	123	117	120	118	107	115	103	117	107	153	117	170
MGPSV (c, GLMPS)	116	113	114	109	110	105	107	99	98	106	111	105	111	111	109
MGPSV (m, GLMPS)	117	111	115	114	109	109	110	93	100	100	103	95	142	111	154
MGPSV (c, CBPS)	116	113	114	108	111	105	106	97	97	105	112	105	112	112	112
MGPSV (m, CBPS)	117	111	115	114	109	109	109	93	99	97	102	94	141	111	154

\*For methods that involve only the propensity score or the outcome model, (c) denotes correct specification and misspecification, respectively. For methods that involve both models, the first and second letter in the parentheses correspond to the treatment model and outcome model, respectively.

<b>Supplemental Table 16. Characteristics of censored vs. uncensored subjects.</b>		
Variable	Uncensored Subjects (N=1,955) Count (%)	Censored Subjects (N=669) Count (%)
<b>Treatment</b>		
Docetaxel	565 (28.9)	161 (24.1)
Abiraterone	783 (40.1)	254 (38.0)
Enzalutamide	476 (24.3)	163 (24.4)
Sipuleucel-T	131 (6.7)	91 (13.6)
<b>Age</b>		
<65	255 (13.0)	121 (18.1)
65-74	657 (33.6)	231 (34.5)
≥75	1,043 (53.4)	317 (47.4)
<b>Race</b>		
White	1,310 (67.0)	471 (70.4)
Black	249 (12.7)	71 (10.6)
Other	396 (20.3)	127 (19.0)
<b>Education level</b>		
High School Diploma or Less	581 (29.7)	187 (28.0)
High School Graduate and Less than Bachelor Degree	970 (49.6)	340 (50.8)
Bachelor Degree Plus	274 (14.0)	102 (15.2)
Unknown	130 (6.6)	40 (6.0)
<b>Household income range</b>		
<50k	669 (34.2)	202 (30.2)
50k-100k	613 (31.4)	226 (33.8)
>100k	376 (19.2)	153 (22.9)
Unknown	297 (15.2)	188 (13.2)
<b>Geographic Region</b>		
South Atlantic	357 (18.3)	125 (18.7)
New England	100 (4.9)	29 (4.3)
Middle Atlantic	197 (10.3)	74 (11.1)
East North Central	317 (16.2)	119 (17.8)
East South Central	71 (3.6)	26 (3.9)
West North Central	181 (9.3)	55 (8.2)
West South Central	196 (10.0)	63 (9.4)
Mountain	241 (12.3)	76 (11.4)
Pacific	295 (15.1)	102 (15.2)
<b>Product</b>		
HMO	599 (30.6)	171 (25.6)
PPO	132 (6.8)	42 (6.3)
Other	1,224 (62.6)	456 (68.2)
<b>Metastatic (Yes)</b>	1,607 (82.2)	515 (77.0)
<b>ASO (Yes)</b>	238 (12.2)	86 (12.9)
<b>Year of First Prescription</b>		
2014	812 (41.5)	235 (35.1)
2015	846 (43.3)	199 (29.7)
2016	297 (15.2)	235 (35.1)
<b>Diabetes</b>	574 (29.4)	191 (28.6)
<b>Hypertension</b>	1,432 (73.2)	475 (71.0)
<b>Arrhythmia</b>	489 (25.0)	148 (22.1)
<b>CHF</b>	234 (12.0)	86 (12.9)
<b>Osteoporosis</b>	160 (8.2)	63 (9.4)
<b>Provider Type</b>		
Medical oncologist	1,196 (61.2)	409 (61.1)
Others	759 (38.8)	260 (38.9)

---

**Supplemental Table 17. Characteristics of uncensored subjects in the four treatment groups of interest**

---

Total (N=1,955)	Docetaxel (N=565)	Abiraterone (N=783)	Enzalutamide (N=476)	Sipuleucel-T1. Zhang M, Tsiatis AA, Davidian M. Improving efficiency of inferences in randomized clinical trials using auxiliary covariates. <i>Biometrics</i> . 2008;64:707–15.
				2. Moore KL, van der Laan MJ. Covariate adjustment in randomized trials with binary outcomes: targeted maximum likelihood estimation. <i>Stat Med</i> . 2009;28:39–64.
				3. Mittlböck M, Schemper M. Explained Variation for Logistic Regression. <i>Stat Med</i> . 1996;15:1987–97.
				4. Zhou T, Elliott MR, Little RJA. Penalized Spline of Propensity Methods for Treatment Comparison. <i>J Am Stat Assoc</i> . 2019;114:1–19.
				5. Li F, Li F. Propensity score weighting for causal inference with multiple treatments. <i>Ann Appl Stat</i> . 2019;13:2389–415.
				6. Hu L, Gu C, Lopez M, Ji J, Wisnivesky J. Estimation of causal effects of multiple treatments in observational studies with a binary outcome: <i>Stat Methods Med Res</i> [Internet]. SAGE PublicationsSage UK: London, England; 2020 [cited 2020 Sep 20]; Available from: <a href="https://journals.sagepub.com/doi/10.1177/0962280220921909">https://journals.sagepub.com/doi/10.1177/0962280220921909</a>
				7. Abadie A, Imbens GW. Large Sample Properties of Matching Estimators for Average Treatment Effects. <i>Econometrica</i> . 2006;74:235–67.
				8. Yang S, Imbens GW, Cui Z, Faries DE, Kadziola Z. Propensity score matching and subclassification in observational studies with multi-level treatments. <i>Biometrics</i> . 2016;72:1055–65.
				9. Lunceford JK, Davidian M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. <i>Stat Med</i> . 2004;23:2937–60.
				10. Li L, Greene T. A weighting analogue to pair matching in propensity score analysis. <i>Int J Biostat</i> . 2013;9:215–34.
				11. Yoshida K, Hernández-Díaz S, Solomon DH, Jackson JW, Gagne JJ, Glynn RJ, et al. Matching Weights to Simultaneously Compare Three Treatment Groups: Comparison to Three-way Matching. <i>Epidemiol Camb Mass</i> . 2017;28:387–95.

---

(N=131)

---

Variable	Count (%)	Count (%)	Count (%)	Count (%)	Count (%)
<b>Age</b>					
<65	255 (13.0)	96 (17.0)	92 (11.7)	49 (10.3)	18 (13.7)
65-74	657 (33.6)	267 (47.3)	216 (27.6)	126 (26.5)	48 (36.6)
≥75	1,043 (53.4)	202 (35.8)	475 (60.7)	301 (63.2)	65 (49.6)
<b>Race</b>					
White	1,310 (67.0)	390 (60.0)	516 (65.9)	318 (66.8)	86 (65.6)
Black	249 (12.7)	49 (8.7)	106 (13.5)	74 (15.5)	20 (15.3)
Other	396 (20.3)	126 (22.3)	161 (20.6)	84 (17.6)	25 (19.1)
<b>Education level</b>					
High School Diploma or Less	581 (29.7)	144 (25.5)	260 (33.2)	142 (29.8)	35 (26.7)
High School Graduate and Less than Bachelor Degree	970 (49.6)	277 (49.0)	389 (49.7)	241 (50.6)	63 (48.1)
Bachelor Degree Plus	274 (14.0)	87 (15.4)	101 (12.9)	69 (14.5)	17 (13.0)
Unknown	130 (6.6)	57 (10.1)	33 (4.2)	24 (5.0)	16 (12.2)
<b>Household income range</b>					
<50k	669 (34.2)	145 (25.7)	308 (39.3)	174 (36.6)	42 (32.1)
50k-100k	613 (31.4)	161 (28.5)	244 (31.2)	164 (34.5)	44 (33.6)
>100k	376 (19.2)	150 (26.5)	137 (17.5)	66 (13.9)	23 (17.6)
Unknown	297 (15.2)	109 (19.3)	94 (12.0)	72 (15.1)	22 (16.8)
<b>Geographic Region</b>					
South Atlantic	357 (18.3)	101 (17.9)	144 (18.4)	87 (18.3)	25 (19.1)
New England	100 (4.9)	33 (5.8)	42 (5.4)	21 (4.4)	4 (3.1)
Middle Atlantic	197 (10.3)	56 (9.9)	65 (8.3)	61 (12.8)	15 (11.5)
East North Central	317 (16.2)	82 (14.5)	131 (16.7)	79 (16.6)	25 (19.1)
East South Central	71 (3.6)	21 (3.7)	26 (3.3)	15 (3.2)	9 (6.9)
West North Central	181 (9.3)	112 (19.8)	42 (5.4)	20 (4.2)	7 (5.3)
West South Central	196 (10.0)	55 (9.7)	83 (10.6)	43 (9.0)	15 (11.5)

Mountain	241 (12.3)	58 (10.3)	95 (12.1)	64 (13.4)	24 (18.3)
Pacific	295 (15.1)	47 (8.3)	155 (19.8)	86 (18.1)	7 (5.3)
<b>Product</b>					
HMO	599 (30.6)	162 (28.7)	256 (32.7)	148 (31.1)	33 (25.2)
PPO	132 (6.8)	35 (6.2)	62 (7.9)	26 (5.5)	9 (6.9)
Other	1,224 (62.6)	368 (65.1)	465 (59.4)	302 (63.4)	89 (67.9)
<b>Metastatic (Yes)</b>	1,607 (82.2)	483 (85.5)	644 (82.2)	365 (76.7)	115 (87.8)
<b>ASO (Yes)</b>	238 (12.2)	66 (11.7)	102 (13.0)	56 (11.8)	14 (10.7)
<b>Year of First Prescription</b>					
2014	812 (41.5)	208 (36.8)	383 (48.9)	174 (36.6)	47 (35.9)
2015	846 (43.3)	262 (46.4)	303 (38.7)	222 (46.6)	59 (45.0)
2016	297 (15.2)	95 (16.8)	97 (12.4)	80 (16.8)	25 (19.1)
<b>Diabetes</b>	574 (29.4)	147 (26.0)	228 (29.1)	154 (32.4)	45 (34.4)
<b>Hypertension</b>	1,432 (73.2)	402 (71.2)	577 (73.7)	350 (73.5)	103 (78.6)
<b>Arrhythmia</b>	489 (25.0)	128 (22.7)	203 (25.9)	130 (27.3)	28 (21.4)
<b>CHF</b>	234 (12.0)	42 (7.4)	103 (13.2)	75 (15.8)	14 (10.7)
<b>Osteoporosis</b>	160 (8.2)	30 (5.3)	66 (8.4)	43 (9.0)	21 (16.0)
<b>Provider Type</b>					
Medical oncologist	1,196 (61.2)	321 (56.8)	565 (72.2)	280 (58.8)	30 (22.9)
Others	759 (38.8)	244 (43.2)	218 (27.8)	196 (41.2)	101 (77.1)

**Supplemental Table 18. Difference (95% confidence interval) in probability of at least one emergency room visit within 180 days of first prescription across four treatment groups (N=1776)**

Method	A – D	E – D	S – D	E – A	S – A	S – E
NAIVE	-0.130 <i>(-0.186, -0.073)</i>	-0.177 <i>(-0.239, -0.115)</i>	-0.099 <i>(-0.197, -0.001)</i>	-0.047 <i>(-0.105, 0.010)</i>	0.031 <i>(-0.064, 0.126)</i>	0.078 <i>(-0.020, 0.177)</i>
OREG	-0.128 <i>(-0.187, 0.068)</i>	-0.183 <i>(-0.248, -0.118)</i>	-0.108 <i>(-0.231, 0.036)</i>	-0.055 <i>(-0.115, 0.004)</i>	0.020 <i>(-0.121, 0.161)</i>	0.075 <i>(-0.069, 0.219)</i>
PEN-GAM	-0.120 <i>(-0.182, -0.058)</i>	-0.174 <i>(-0.241, -0.107)</i>	-0.169 <i>(-0.396, 0.058)</i>	-0.054 <i>(-0.115, 0.007)</i>	-0.049 <i>(-0.277, 0.178)</i>	0.005 <i>(-0.225, 0.235)</i>
IPW	-0.126 <i>(-0.187, -0.066)</i>	-0.176 <i>(-0.243, -0.108)</i>	-0.058 <i>(-0.259, 0.144)</i>	-0.050 <i>(-0.110, 0.011)</i>	0.068 <i>(-0.129, 0.266)</i>	0.118 <i>(-0.083, 0.319)</i>
AIPW	-0.125 <i>(-0.185, -0.064)</i>	-0.179 <i>(-0.246, -0.113)</i>	-0.095 <i>(-0.263, 0.074)</i>	-0.055 <i>(-0.115, 0.005)</i>	0.030 <i>(-0.137, 0.196)</i>	0.085 <i>(-0.084, 0.253)</i>
MW	-0.103 <i>(-0.199, -0.007)</i>	-0.179 <i>(-0.272, -0.085)</i>	-0.091 <i>(-0.206, 0.024)</i>	-0.076 <i>(-0.166, 0.014)</i>	0.011 <i>(-0.101, 0.124)</i>	0.087 <i>(-0.024, 0.199)</i>
AMW	-0.085 <i>(-0.181, 0.012)</i>	-0.161 <i>(-0.256, -0.065)</i>	-0.079 <i>(-0.189, 0.032)</i>	-0.076 <i>(-0.166, 0.014)</i>	0.006 <i>(-0.099, 0.111)</i>	0.082 <i>(-0.024, 0.188)</i>
OW	-0.111 <i>(-0.189, -0.033)</i>	-0.174 <i>(-0.252, -0.095)</i>	-0.076 <i>(-0.187, 0.035)</i>	-0.063 <i>(-0.137, 0.011)</i>	0.035 <i>(-0.073, 0.143)</i>	0.098 <i>(-0.012, 0.207)</i>
AOW	-0.099 <i>(-0.178, -0.020)</i>	-0.163 <i>(-0.244, -0.082)</i>	-0.084 <i>(-0.191, 0.024)</i>	-0.064 <i>(-0.139, 0.011)</i>	0.015 <i>(-0.087, 0.118)</i>	0.080 <i>(-0.026, 0.185)</i>
MCOV	-0.144 <i>(-0.211, -0.076)</i>	-0.202 <i>(-0.272, -0.131)</i>	-0.149 <i>(-0.291, -0.008)</i>	-0.058 <i>(-0.121, 0.005)</i>	-0.006 <i>(-0.143, 0.132)</i>	0.052 <i>(-0.087, 0.192)</i>
MGPS	-0.131 <i>(-0.197, -0.064)</i>	-0.195 <i>(-0.272, -0.119)</i>	-0.019 <i>(-0.232, 0.195)</i>	-0.065 <i>(-0.135, 0.005)</i>	0.112 <i>(-0.100, 0.324)</i>	0.177 <i>(-0.037, 0.391)</i>
MGPSV	-0.125 <i>(-0.197, -0.053)</i>	-0.164 <i>(-0.240, -0.087)</i>	-0.125 <i>(-0.301, 0.051)</i>	-0.039 <i>(-0.107, 0.029)</i>	0.000 <i>(-0.173, 0.173)</i>	0.039 <i>(-0.136, 0.040)</i>

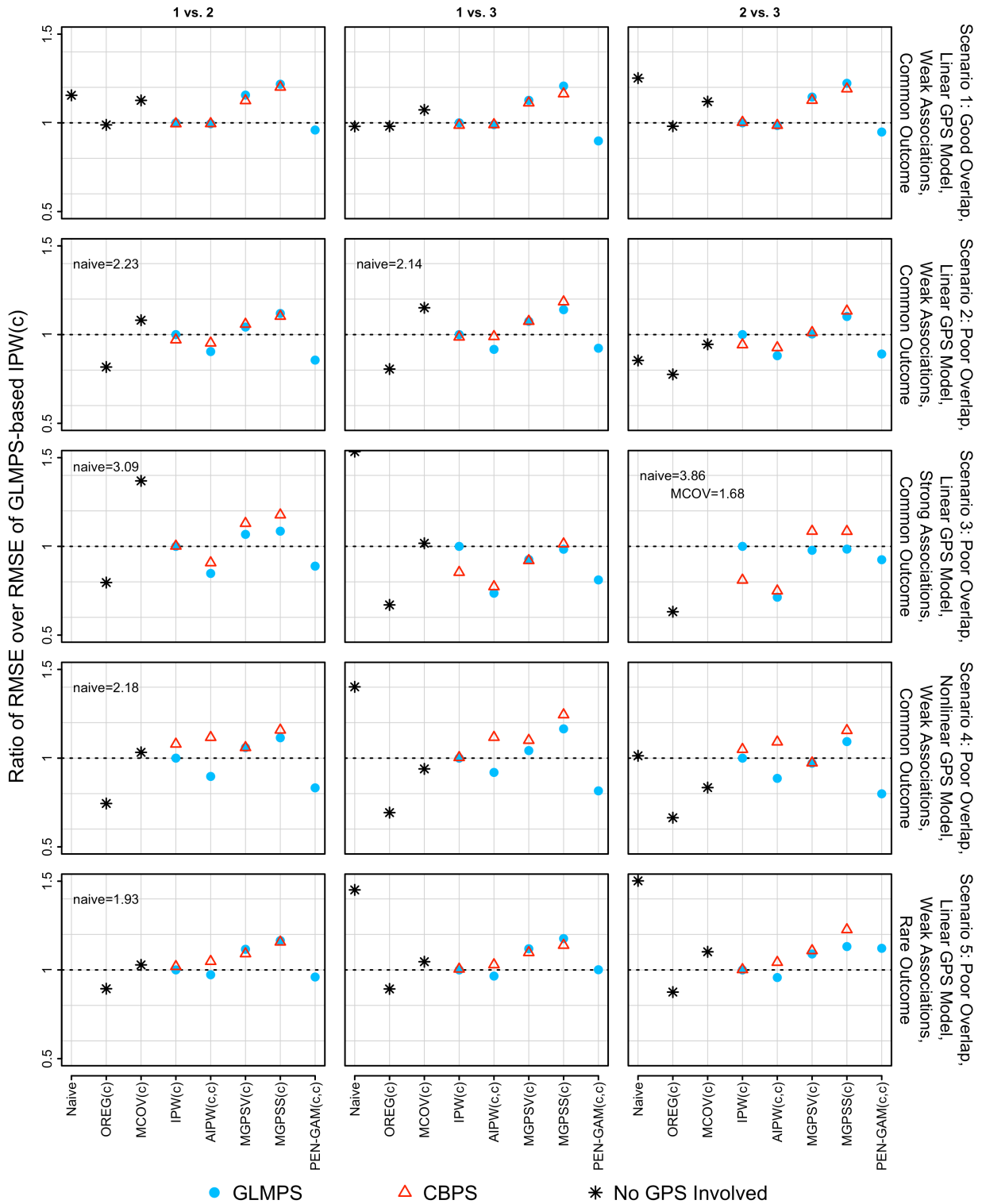
Abbreviations: A, abiraterone; D, docetaxel; E, enzalutamide; S, Sipuleucel-T.

95% confidence intervals that exclude 0 are italicized. The 95% confidence intervals were calculated using: (1) bootstrapped standard errors from 50 bootstrap samples for OREG, IPW, AIPW, MW, AMW, OW, AOW, and CBPS-based MGPS; (2) Wald-type confidence interval based on original data for NAIVE; (3) Abadie and Imbens (2006) confidence interval for MCOV and both GLMPS- and CBPS-based MGPSV; (4) Abadie and Imbens (2016) confidence interval for GLMPS-based MGPS; (5) Rubin's imputation rule for PEN-GAM.

**Supplemental Table 19. Average computational time across 100 simulated datasets for the methods under comparison.**

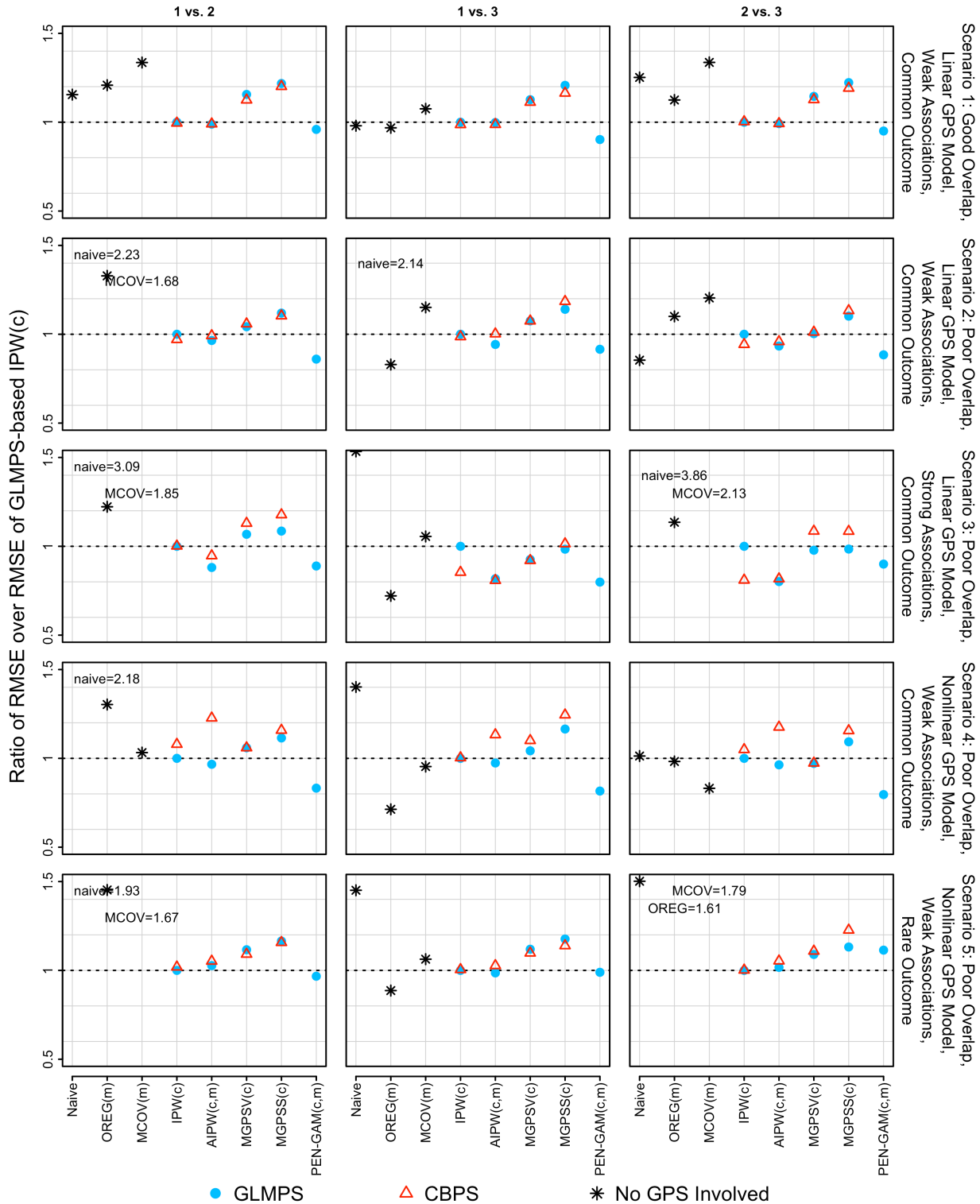
<b>Methods</b>	<b>Computational time (milliseconds)</b>	
<i>Estimation of GPS</i>		
	GLMPS	19.2
	Just-identified CBPS	1423.3
	Over-identified CBPS	2460.9
<i>Estimation of ATE<sup>a</sup></i>		
	OREG	8.7
	IPW	19.9
	AIPW	29.3
	MW	22.8
	OW	23.8
	MCOV	218.5
	MGPS	514.6
	MGPSV	171.9
	PENCOMP	1858.2
<sup>a</sup> GLMPS was used for propensity score-based methods		

## Supplemental Figures

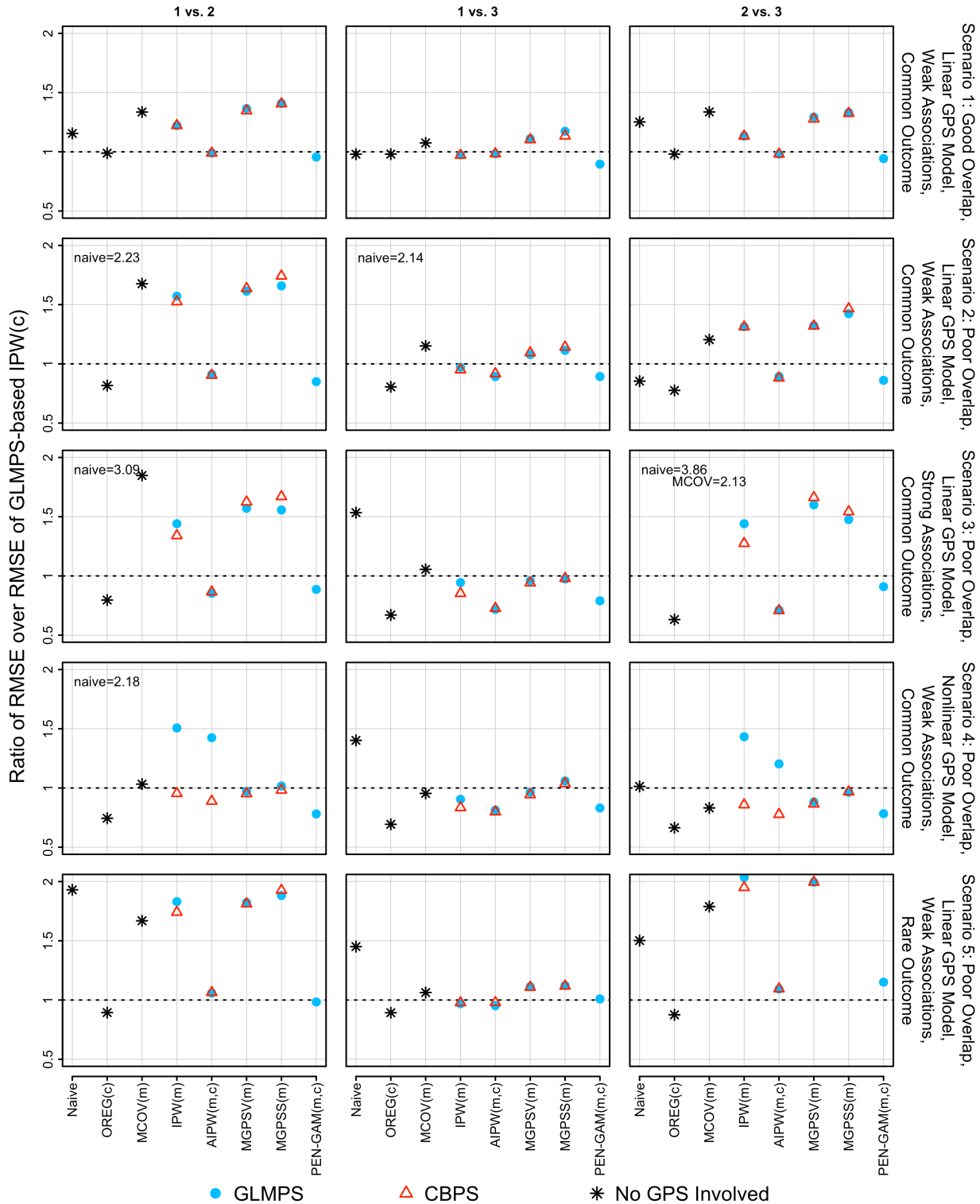


**Supplemental Figure 1.** Ratio of RMSE over RMSE of GLMPS-based IPW(c) for sample size 300 across methods based on correctly specified outcome and propensity models. The rows represent scenarios and columns represent pairs of comparison. Results were obtained using 2000 simulated datasets.

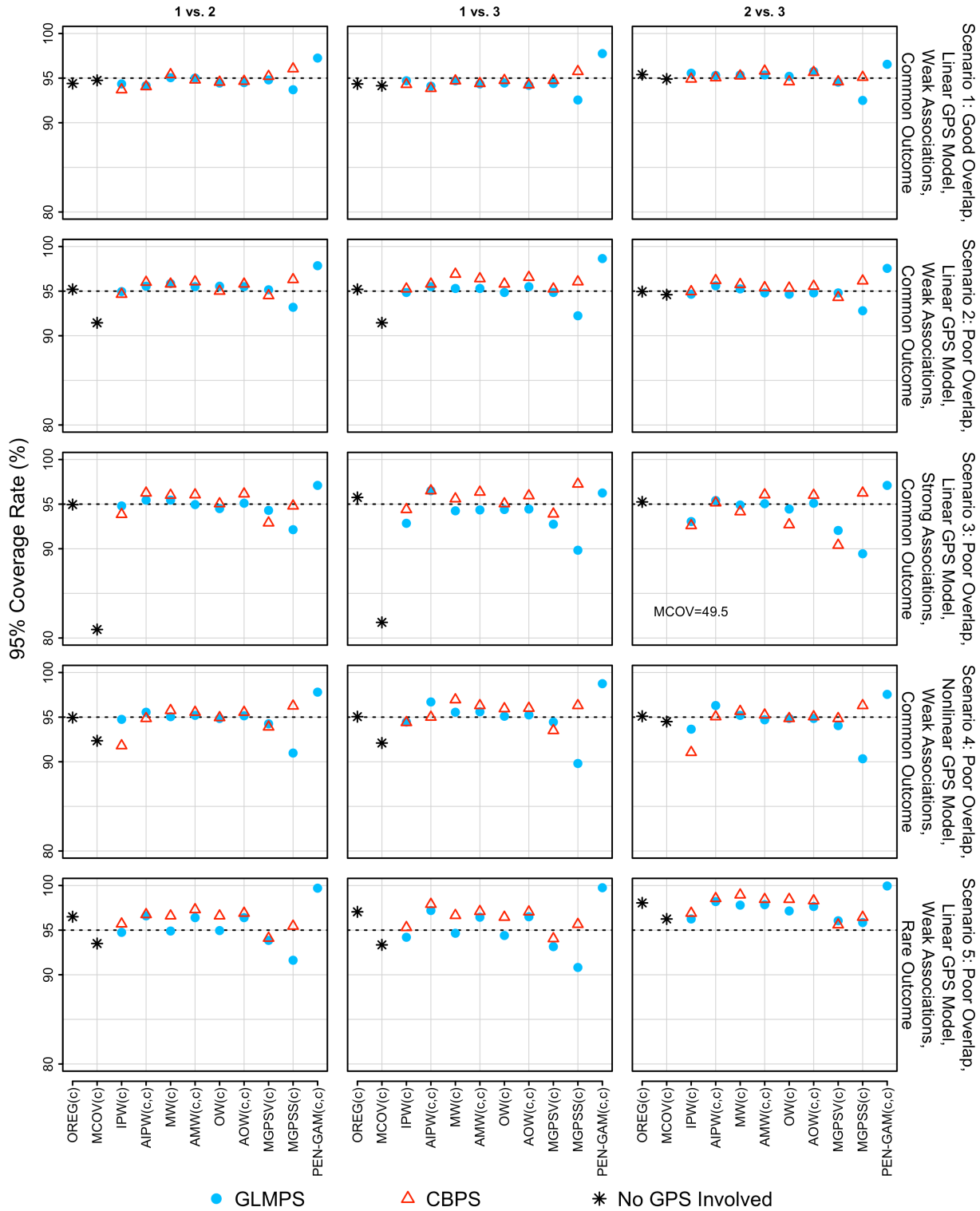




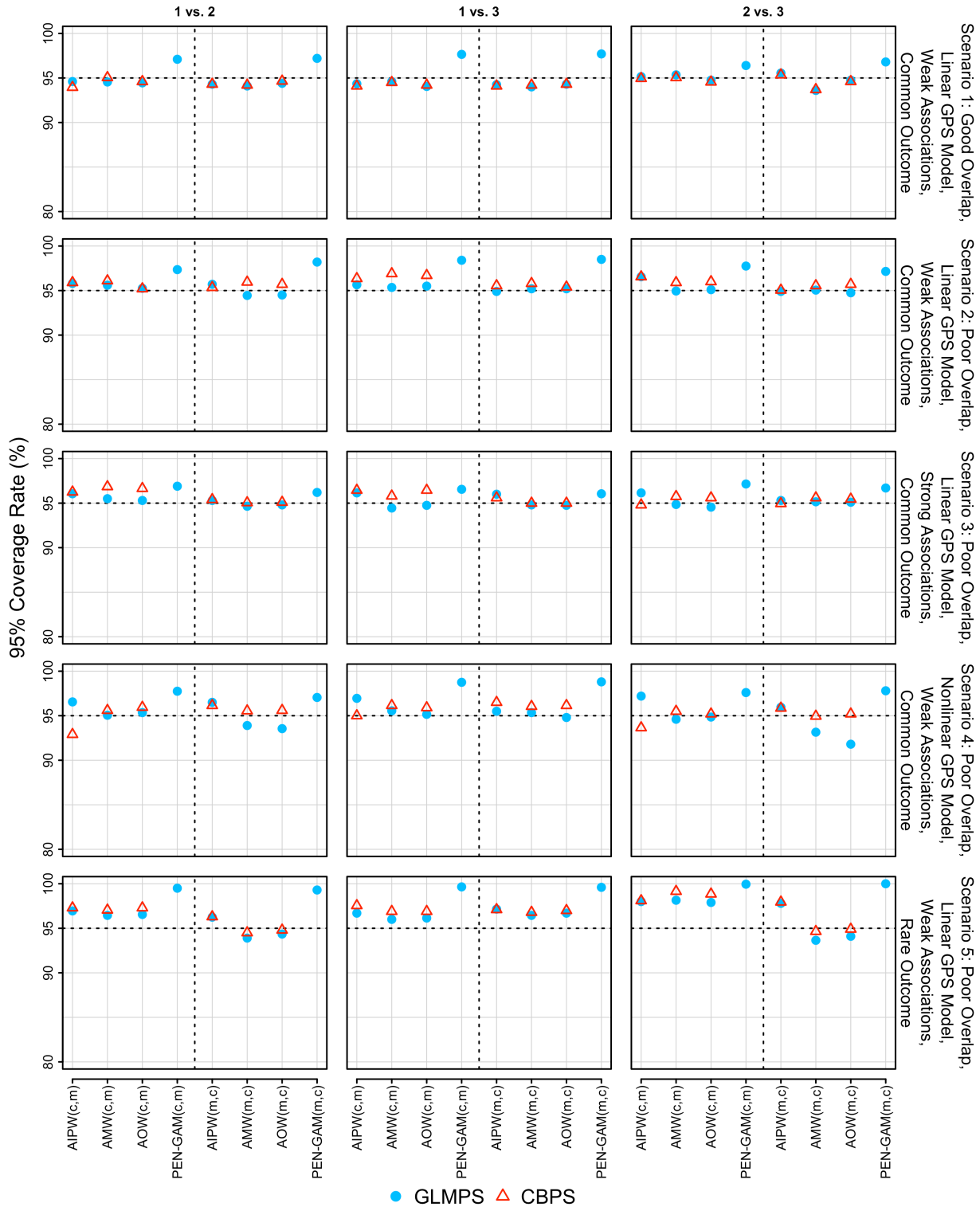
**Supplemental Figure 2.** Ratio of RMSE over RMSE of GLMPS-based IPW(c) for sample size 300 across methods based on a correctly specified propensity model only. The rows represent scenarios and columns represent pairs of comparison. Results were obtained using 2000 simulated datasets.



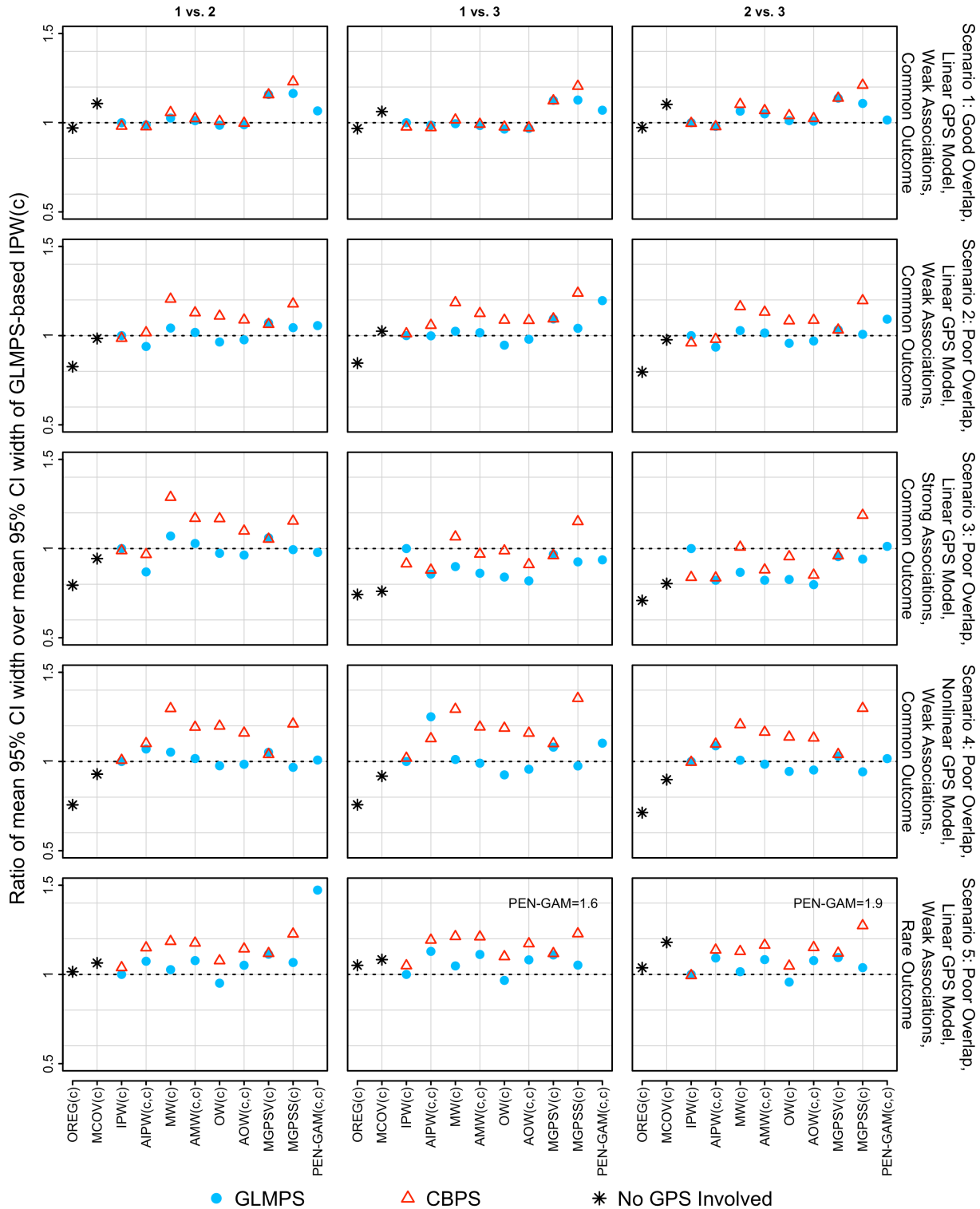
**Supplemental Figure 3.** Ratio of RMSE over RMSE of GLMPS-based IPW(c) for sample size 300 across methods based on a correctly specified outcome model only. The rows represent scenarios and columns represent pairs of comparison. Results were obtained using 2000 simulated datasets.



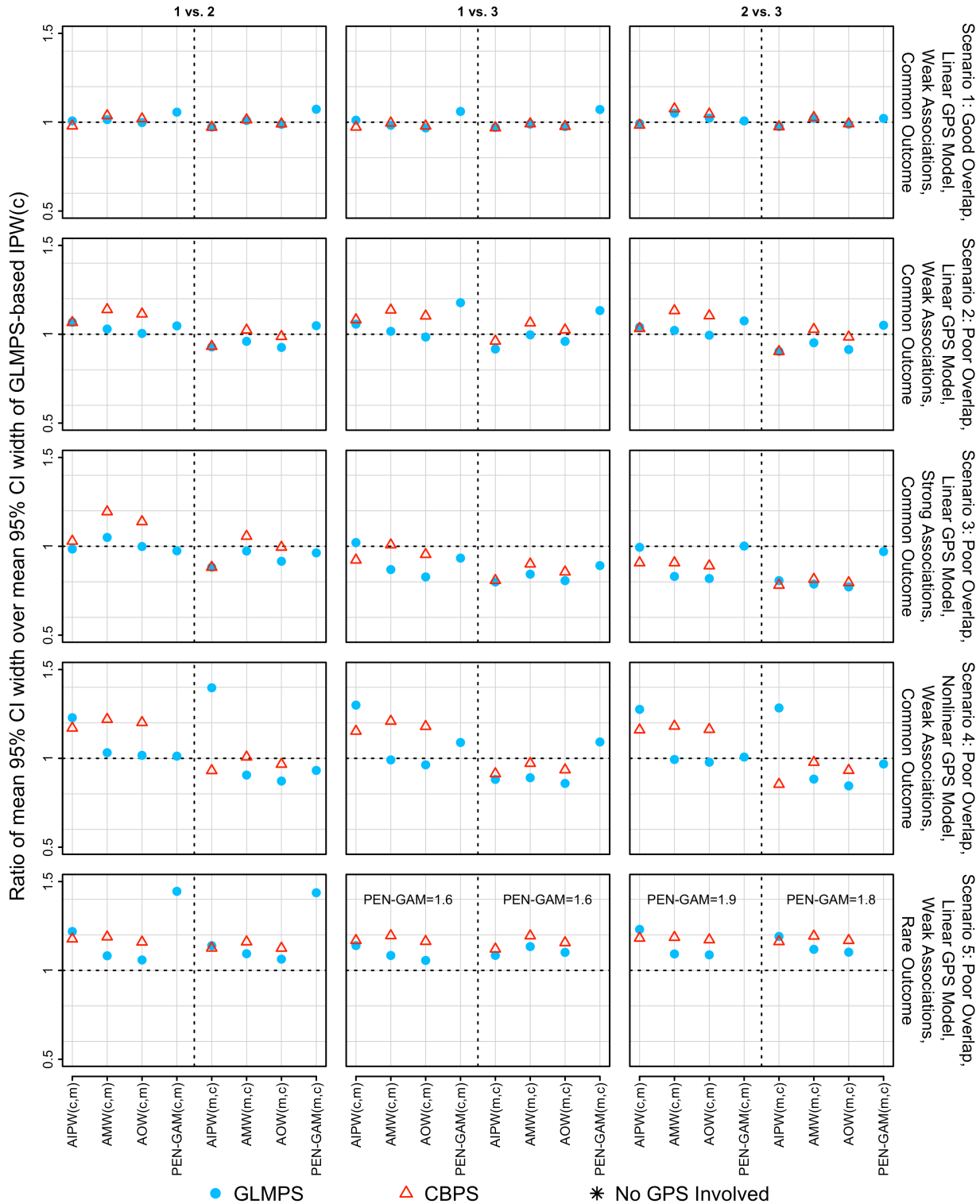
**Supplemental Figure 4.** 95% Coverage probability for sample size 300 across methods based on correctly specified outcome and propensity models. The rows represent scenarios and columns represent pairs of comparison. Results were obtained using 2000 simulated datasets.



**Supplemental Figure 5.** 95% Coverage probability for sample size 300 across methods based on a correctly specified propensity score or outcome model. For methods that involve both models, the first and second letter in the parentheses correspond to the propensity and outcome model, respectively. The rows represent scenarios and columns represent pairs of comparison. Results were obtained using 2000 simulated datasets.

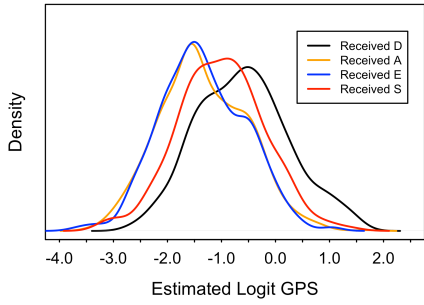


**Supplemental Figure 6.** Ratio of mean 95% CI width over mean 95% CI width of GLMPS-based IPW(c) for sample size 300 across methods based on correctly specified outcome and propensity models. The rows represent scenarios and columns represent pairs of comparison. Results were obtained using 2000 simulated datasets.

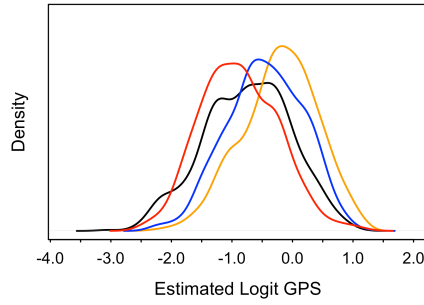


**Supplemental Figure 7.** Ratio of mean 95% CI width over mean 95% CI width of GLMPS-based IPW(c) for sample size 300 across methods based on a correctly specified propensity score or outcome model. For methods that involve both models, the first and second letter in the parentheses correspond to the propensity and outcome model, respectively. The rows represent scenarios and columns represent pairs of comparison. Results were obtained using 2000 simulated datasets.

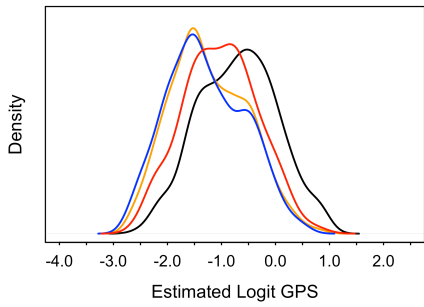
**A** Logit Propensity to Docetaxel (Untrimmed)



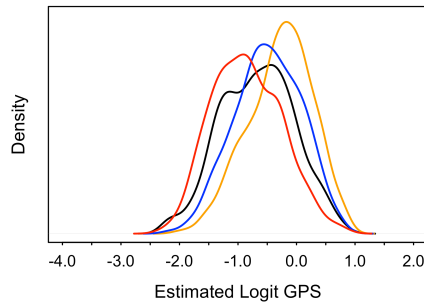
**C** Logit Propensity to Abiraterone (Untrimmed)



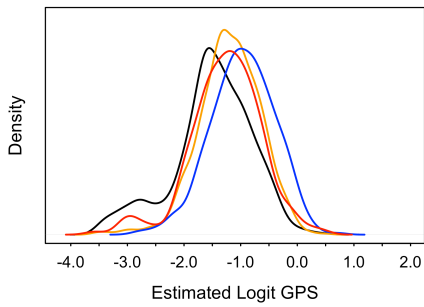
**B** Logit Propensity to Docetaxel (Trimmed)



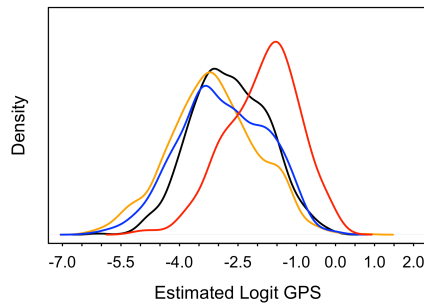
**D** Logit Propensity to Abiraterone (Trimmed)



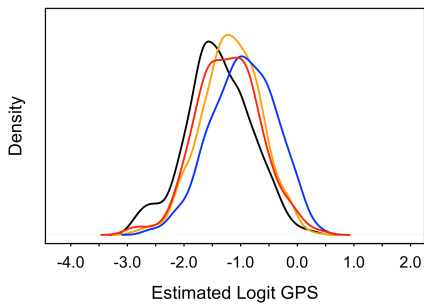
**E** Logit Propensity to Enzalutamide (Untrimmed)



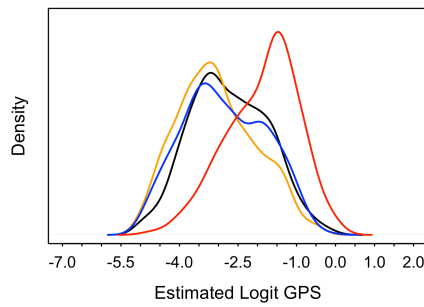
**G** Logit Propensity to Sipuleucel-T (Untrimmed)



**F** Logit Propensity to Enzalutamide (Trimmed)



**H** Logit Propensity to Sipuleucel-T (Trimmed)



**Supplemental Figure 8.** Distribution of the estimated generalized propensity scores in logit scale for the original (N=1,955) and trimmed samples (N=1,777).