eAppendix 1

Data-generating process

Data of the training and test datasets were generated using the same algorithm. First, j independent confounding variables Z were generated. Z_1 was sampled from a normal distribution with mean 3 and variance 1. The remaining Z_{j-1} variables were sampled from independent Bernoulli distributions, each with a success probability of 0.5. A subject's probability of treatment was given by the model:

$$logit(p_{i,treatment}) = \alpha_0 + \alpha_1 z_{i1} + \cdots + \alpha_j z_{ij} \quad [1]$$

The value of α_0 was set so that on average 50% of the subjects were exposed to each treatment. Please see Table 1 and the simulation scenarios section below for an overview parameter values used. For each ith individual the probability of experiencing an event was given by:

$$logit(p_{i,event}) = \delta_0 + \delta_1 x_i + \delta_2 z_{i1} + \cdots + \delta_j z_{ij} + \delta_{int} x_i z_{i1} \quad [2]$$

The intercept (δ_0) was chosen so that on average 50% of the subjects experienced an event. Depending on the value of δ_{int} there was an interaction between treatment and continuous confounder Z_1. The treatment and outcome states were sampled from Bernoulli distributions:

$$x_i \sim Bernoulli(p_{i,treatment})$$

$$y_i \sim Bernoulli(p_{i,event})$$

Data analyses

The test data contained post launch information on subjects receiving new drug A (indicated by $X = 1$) or drug C (indicated by $X = 0$). The association between treatment and the outcome was estimated using the following methods.

Crude logistic regression

$$logit(prob[y_i = 1|x_i]) = \hat{\theta}_0 + \hat{\theta}_1 x_i \quad [3]$$

With $\hat{\theta}_1$ an estimate of the unadjusted natural logarithm of the odds ratio of the association between treatment A (relative to treatment C) and outcome

Multivariable Logistic regression

To adjust for confounding the following LR model was used:

$$logit\left(prob[y_i = 1|x_i, z_{i1}, \ldots, z_{ij}]\right) = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_1 z_{i1} + \cdots + \hat{\beta}_j z_{ij} \quad [4]$$

Where $\hat{\beta}_1$ is an estimate of the ln(odds ratio) of the association between treatment and outcome adjusted for confounders Z.

Propensity score

An alternative to LR models is to first estimate the associations of the confounder with the treatment variable. In a second step, the (logit of the) predicted probability of treatment (i.e., the propensity score) can be used to control for confounding:

$$PS = logit\left(prob[x_i = 1 \mid x_i, z_{i1}, \ldots, z_{ij}]\right) = \hat{\varphi}_0 + \hat{\varphi}_1 z_{i1} + \cdots + \hat{\varphi}_{j1} z_{ij} \quad [5]$$

$$logit(prob[y_i = 1 | x_i, \widehat{ps}_i]) = \hat{\omega}_0 + \hat{\omega}_1 x_i + \hat{\omega}_2 \widehat{ps}_i \quad [6]$$

Here $\hat{\omega}_1$, can be interpreted as the ln(odds ratio) of the treatment-outcome association adjusted for the confounders included the PS model (equation 5).

Inverse probability weighting

Including the estimated PS in an outcome regression, equation 6, assumes that the PS is linearly related with the outcome which is not necessarily the case. Besides stratification and matching on the PS, [1] an alternative strategy that does not assume linearity is to weight a crude logistic regression (equation 3) by inverse probability weights (IPW). [2-4]

$$SW_i = \frac{prob(x_i = 1)^{x_i}[1 - prob(x_i = 1)]^{1-x_i}}{prob(x_i = 1 \mid z_{i1}, \ldots, z_{ij})^{x_i}\left[1 - prob(x_i = 1 \mid z_{i1}, \ldots, z_{ij})\right]^{1-x_i}}$$

In our simulations we evaluated stabilized weights which are less variant than the unstabilized alternative; however stabilized weights may still require adjustment of the variance estimator [5]. This was implemented by replacing the usual variance estimators by sandwich estimators [2,6] using the R sandwich package .[6]

Disease risk score adjustment

Another approach to control for confounding is outcome regression adjusted for a disease risk score (DRS). First, the associations between the confounders and outcome are estimated in a training dataset, using equation 4. In a second stage, these associations are used to calculate

the logit of the predicted probability of the outcome (the DRS) for the patients included in the test data.

$$logit(prob[y_i = 1|x_i, drs_i]) = \hat{\gamma}_0 + \hat{\gamma}_1 x_i + \hat{\gamma}_2 drs_i \quad [7]$$

Where $\hat{\gamma}_1$, can be interpreted as the ln(odds ratio) of the association between treatment and the outcome, adjusted for the confounders included in the first stage.

eTable 1 Simulation results for scenario I assessing performance of different confounding adjustment methods with different events per coefficient[a].

	10 EPC	5 EPC	2.5 EPC	1 EPC^	0.5 EPC
Mean odds ratio					
Crude	1.18	1.19	1.19	1.21	1.25
LR	1.00	1.00	1.00	N/A#	N/A#
PS	1.00	1.00	1.00	1.09	N/A#
DRS1	1.01	1.03	1.05	1.10	1.15
DRS2	1.01	1.03	1.05	1.10	1.15
DRS3	1.01	1.03	1.05	1.10	1.15
DRS4	1.03	1.05	1.08	1.13	1.19
IPW	1.00	1.00	0.99	0.92	1.25
Relative bias (%)					
Crude	18.44	18.56	19.26	21.46	25.55
LR	-0.03	-0.28	-0.32	N/A#	-100
PS	-0.08	-0.28	-0.19	8.74	-100
DRS1	1.36	2.52	4.72	9.58	14.75
DRS2	1.36	2.52	4.72	9.59	14.75
DRS3	1.38	2.54	4.76	9.62	14.76
DRS4	2.55	4.53	7.77	13.48	18.83
IPW	0.09	-0.15	-0.51	-8.49	24.55
Coverage					
Crude	0.865	0.858	0.862	0.834	0.810
LR	0.950	0.941	0.917	0.656	1.000
PS	0.957	0.956	0.954	0.945	0.974
DRS1	0.952	0.945	0.946	0.924	0.897
DRS2	0.952	0.945	0.946	0.925	0.898
DRS3	0.951	0.945	0.944	0.925	0.897
DRS4	0.948	0.941	0.933	0.903	0.867
IPW	0.958	0.954	0.952	0.656	0.810
RMSE					
Crude	0.26	0.26	0.27	0.28	0.30
LR	0.22	0.25	0.30	$2.7*10^{14}$	$4.5*10^{4}$
PS	0.21	0.21	0.23	0.27	$4.8*10^{4}$
DRS1	0.21	0.21	0.21	0.23	0.25
DRS2	0.21	0.21	0.21	0.23	0.25
DRS3	0.21	0.21	0.21	0.23	0.25
DRS4	0.21	0.21	0.22	0.24	0.27
IPW	0.21	0.22	0.26	4.19	0.30
Mean SE					
Crude	0.20	0.20	0.20	0.20	0.20
LR	0.22	0.23	0.27	$2.6*10^{6}$	$10.81*10^{6}$
PS	0.21	0.22	0.23	0.26	61.30
DRS1	0.21	0.21	0.21	0.21	0.21
DRS2	0.21	0.21	0.21	0.21	0.21
DRS3	0.21	0.21	0.21	0.21	0.21
DRS4	0.21	0.21	0.21	0.21	0.20
IPW	0.21	0.22	0.26	0.49	0.20
Empirical SE					
Crude	0.20	0.20	0.20	0.20	0.20
LR	0.22	0.25	0.30	$2.7*10^{14}$	$4.5*10^{4}$
PS	0.21	0.21	0.23	0.25	$4.8*10^{4}$
DRS1	0.21	0.21	0.21	0.21	0.25
DRS2	0.21	0.21	0.21	0.21	0.25

DRS3	0.21	0.21	0.21	0.23	0.25
DRS4	0.21	0.21	0.21	0.24	0.27
IPW	0.21	0.22	4.18	4.19	0.30

* LR, logistic regression; PS, propensity score; DRS, disease risk score; IPW, inverse probability weights; RMSE indicates the square root of the mean squared error. # While all LR samples converged, the extreme estimate of the mean natural logarithm of the odds ratio resulted in an error when calculating the mean OR and relative bias.^ at an EPC of 1.0 the IPW model failed to converge 2,763 time out of 10,000 replications the other methods did converge. Defining extreme estimates as an absolute estimate above 5 resulted in 7,219 (EPC 1.0), and 9,415 (EPC 0.5) extreme results for the LR method, 4,706 (EPC 0.5) extreme replications for the PS method, and 3,627 (EPC 1.0), and 0 (EPC 0.5) extreme replications for the IPW method.

eTable 2 Simulation results for scenario II and III comparing different DRS models in the presence of an interaction effect in the training data*.

	Crude	LR	PS	DRS1	DRS2	DRS3	DRS4	IPW
Scenario II#								
Mean odds ratio	1.18	1.00	1.00	1.05	1.00	1.02	1.04	1.00
Relative bias (%)	18.16	-0.29	-0.32	4.95	0.36	2.04	3.98	-0.15
Coverage	0.861	0.942	0.951	0.940	0.946	0.945	0.941	0.953
RMSE	0.26	0.23	0.21	0.22	0.21	0.21	0.21	0.21
Mean SE	0.20	0.22	0.21	0.21	0.21	0.21	0.21	0.21
Empirical SE	0.20	0.23	0.21	0.21	0.21	0.21	0.21	0.21
Scenario III^								
Mean odds ratio	1.18	1.00	1.00	1.09	1.17	1.01	1.03	1.00
Relative bias (%)	18.42	-0.04	-0.08	8.75	16.55	1.49	2.83	0.11
Coverage	0.863	0.946	0.953	0.931	0.879	0.948	0.948	0.955
RMSE	0.26	0.23	0.21	0.22	0.26	0.21	0.21	0.21
Mean SE	0.20	0.22	0.21	0.21	0.21	0.21	0.21	0.21
Empirical SE	0.20	0.23	0.21	0.21	0.21	0.21	0.21	0.21

* LR, logistic regression; PS, propensity score; DRS, disease risk score; IPW, inverse probability weights; RMSE indicates the square root of the mean squared error. # Treatment by confounder 1 interaction OR of 0.30.

^ Treatment by confounder 1 interaction OR of 3.0

eTable 3 Simulation results for scenario VI[#].

Treatment OR:	0.30	0.70	1.00	1.50	3.00
Mean odds ratio					
Crude	0.46	0.90	1.19	1.67	2.98
LR	0.27	0.68	1.00	1.57	3.52
PS	0.38	0.75	1.00	1.40	2.50
DRS1	0.43	0.86	1.15	1.61	2.88
DRS2	0.45	0.89	1.19	1.67	2.98
DRS3	0.39	0.78	1.05	1.49	2.69
DRS4	0.40	0.81	1.09	1.53	2.76
IPW	0.38	0.75	1.00	1.39	2.48
Relative bias (%)#					
Crude	51.70	28.03	19.45	11.28	-0.78
LR	-8.83	-2.56	-0.03	4.35	17.25
PS	25.78	7.12	0.01	-6.86	-16.77
DRS1	43.64	22.41	14.60	7.13	-3.95
DRS2	50.98	27.64	19.20	11.21	-0.69
DRS3	29.62	12.06	5.47	-0.89	-10.37
DRS4	34.61	15.83	8.81	1.96	-8.08
IPW	26.24	7.23	-0.06	-7.18	-17.39
Coverage#					
Crude	0.467	0.774	0.865	0.919	0.946
LR	0.908	0.917	0.916	0.912	0.891
PS	0.837	0.942	0.953	0.942	0.885
DRS1	0.581	0.832	0.899	0.938	0.940
DRS2	0.476	0.779	0.867	0.920	0.946
DRS3	0.763	0.914	0.940	0.947	0.918
DRS4	0.705	0.891	0.930	0.947	0.929
IPW	0.848	0.942	0.948	0.940	0.896
Rejection rate					
Crude	0.973	0.084	0.135	0.668	0.997
LR	0.994	0.306	0.084	0.368	0.966
PS	0.989	0.223	0.047	0.277	0.945
DRS1	0.984	0.116	0.101	0.591	0.995
DRS2	0.973	0.084	0.133	0.666	0.997
DRS3	0.994	0.206	0.060	0.439	0.988
DRS4	0.992	0.166	0.070	0.496	0.992
IPW	0.955	0.199	0.052	0.243	0.859
RMSE					
Crude	0.46	0.32	0.27	0.24	0.24
LR	0.34	0.31	0.31	0.33	0.41
PS	0.33	0.24	0.23	0.25	0.32
DRS1	0.42	0.29	0.25	0.23	0.24
DRS2	0.46	0.32	0.27	0.24	0.24
DRS3	0.34	0.24	0.22	0.22	0.26
DRS4	0.37	0.26	0.23	0.22	0.25
IPW	0.35	0.27	0.27	0.29	0.36
Mean SE					
Crude	0.20	0.20	0.21	0.21	0.23
LR	0.28	0.27	0.28	0.29	0.32
PS	0.24	0.23	0.24	0.24	0.26
DRS1	0.21	0.21	0.21	0.22	0.24
DRS2	0.21	0.20	0.21	0.21	0.23
DRS3	0.21	0.21	0.21	0.22	0.24

DRS4	0.21	0.21	0.21	0.22	0.24
IPW	0.26	0.26	0.26	0.27	0.30
Empirical SE					
Crude	0.20	0.20	0.21	0.22	0.24
LR	0.32	0.31	0.31	0.33	0.38
PS	0.23	0.23	0.23	0.24	0.26
DRS1	0.21	0.21	0.21	0.22	0.24
DRS2	0.21	0.20	0.21	0.22	0.24
DRS3	0.21	0.21	0.22	0.22	0.24
DRS4	0.21	0.21	0.21	0.22	0.24
IPW	0.26	0.26	0.27	0.28	0.31

* LR, logistic regression; PS, propensity score; DRS, disease risk score; IPW, inverse probability weights; RMSE indicates the square root of the mean squared error. All models converged, and there were no extreme estimates; defined as an estimate above 5.
# relative bias and coverage indicate bias and coverage of the conditional treatment OR (as indicated above each column). Propensity score models, inverse probability weighted models, and crude models estimated marginal treatment ORs hence deviations from nominal values are expected when the true treatment OR is not 1.00.

eTable 4 Simulation results for scenario I repeated with firth penalized logistic regression models instead of logit binomial (logistic regression) model [#].

	10 EPC	5 EPC	2.5 EPC	1 EPC^	0.5 EPC
Mean odds ratio					
Crude	1.18	1.19	1.19	1.21	1.24
LR	1.00	1.00	1.00	1.00	0.97
PS	1.00	1.00	1.00	1.06	0.97
DRS1	1.01	1.03	1.05	1.10	1.14
DRS2	1.01	1.03	1.05	1.10	1.15
DRS3	1.20	1.20	1.20	1.22	1.24
DRS4	1.02	1.05	1.08	1.13	1.18
IPW	1.01	1.02	1.04	1.15	1.24
Relative bias (%)					
Crude	18.22	18.57	19.34	21.44	24.25
LR	-0.03	-0.07	0.24	0.09	-3.04
PS	-0.02	-0.06	0.35	5.57	-3.46
DRS1	1.30	2.67	4.96	9.65	14.44
DRS2	1.30	2.59	4.95	9.65	14.50
DRS3	20.29	20.42	20.43	21.76	23.81
DRS4	2.46	4.73	7.95	13.49	18.48
IPW	0.96	1.84	4.31	15.16	24.17
Coverage					
Crude	0.865	0.868	0.864	0.837	0.813
LR	0.951	0.958	0.962	0.927	1.00
PS	0.951	0.954	0.956	0.950	0.963
DRS1	0.946	0.950	0.945	0.927	0.900
DRS2	0.938	0.944	0.940	0.923	0.899
DRS3	0.848	0.851	0.858	0.840	0.819
DRS4	0.944	0.946	0.934	0.905	0.871
IPW	0.955	0.938	0.923	0.920	0.913
RMSE					
Crude	0.26	0.26	0.27	0.28	0.30
LR	0.22	0.22	0.24	0.45	8.81
PS	0.21	0.21	0.22	0.27	19.66
DRS1	0.21	0.21	0.21	0.23	0.25
DRS2	0.22	0.21	0.22	0.23	0.25
DRS3	0.29	0.28	0.28	0.29	0.30
DRS4	0.21	0.21	0.22	0.24	0.27
IPW	0.21	0.21	0.23	0.27	0.30
Mean SE					
Crude	0.20	0.20	0.20	0.20	0.20
LR	0.22	0.23	0.25	0.39	29.16
PS	0.21	0.22	0.23	0.27	12.94
DRS1	0.21	0.21	0.21	0.21	0.21
DRS2	0.21	0.21	0.21	0.21	0.21
DRS3	0.21	0.21	0.21	0.20	0.20
DRS4	0.21	0.21	0.21	0.20	0.20
IPW	0.20	0.20	0.21	0.23	0.25
Empirical SE					
Crude	0.20	0.20	0.20	0.20	0.20
LR	0.22	0.23	0.24	0.45	8.81
PS	0.21	0.22	0.22	0.27	19.66
DRS1	0.21	0.21	0.21	0.21	0.21
DRS2	0.22	0.21	0.21	0.21	0.21

DRS3	0.22	0.22	0.21	0.21	0.21
DRS4	0.21	0.21	0.21	0.20	0.20
IPW	0.21	0.21	0.22	0.23	0.25

* LR, logistic regression; PS, propensity score; DRS, disease risk score; IPW, inverse probability weights; RMSE indicates the square root of the mean squared error. # All models converged. Defining extreme estimates as an absolute estimate above 5 resulted in 2,442 extreme coefficients (EPC 0.5) for the LR method and 4,262 extreme replications (EPC 0.5) for the PS method.

eTable 5 Simulation results for scenario I repeated including an instrumental variable in the adjustment models[#].

	10 EPC	5 EPC	2.5 EPC	1 EPC^	0.5 EPC
Mean odds ratio					
Crude	1.17	1.17	1.18	1.19	1.22
LR	1.00	1.00	1.00	NA#	-100
PS	1.00	1.00	1.00	1.09	NA#
DRS1	1.01	1.03	1.05	1.09	1.13
DRS2	1.01	1.03	1.05	1.09	1.13
DRS3	1.01	1.03	1.05	1.09	1.13
DRS4	1.03	1.05	1.07	1.12	1.17
IPW	1.00	1.00	1.00	1.19	1.22
Relative bias (%)					
Crude	16.58	17.18	17.57	19.06	22.09
LR	-0.06	0.37	0.02	NA#	NA#
PS	-0.29	0.13	0.08	8.92	NA#
DRS1	1.35	2.90	4.65	8.60	13.30
DRS2	1.36	2.91	4.66	8.60	13.30
DRS3	1.38	2.94	4.69	8.62	13.31
DRS4	2.60	4.93	7.43	12.02	16.19
IPW	0.04	0.47	-0.23	19.12	22.09
Coverage					
Crude	0.880	0.879	0.874	0.860	0.830
LR	0.943	0.935	0.919	0.672	1.000
PS	0.952	0.953	0.956	0.943	0.950
DRS1	0.947	0.945	0.943	0.932	0.902
DRS2	0.947	0.945	0.944	0.932	0.902
DRS3	0.947	0.945	0.944	0.931	0.903
DRS4	0.944	0.941	0.935	0.914	0.876
IPW	0.950	0.949	0.946	0.842	0.830
RMSE					
Crude	0.25	0.26	0.26	0.27	0.29
LR	0.24	0.26	0.32	NA#	NA#
PS	0.22	0.23	0.24	0.28	NA#
DRS1	0.21	0.21	0.21	0.22	0.24
DRS2	0.21	0.21	0.21	0.22	0.24
DRS3	0.21	0.21	0.21	0.22	0.24
DRS4	0.21	0.21	0.22	0.24	0.26
IPW	0.23	0.25	0.34	1.52	0.29
Mean SE					
Crude	0.20	0.20	0.20	0.20	0.20
LR	0.23	0.25	0.28	NA#	NA#
PS	0.22	0.23	0.24	0.26	NA#
DRS1	0.21	0.21	0.21	0.21	0.21
DRS2	0.21	0.21	0.21	0.21	0.21
DRS3	0.21	0.21	0.21	0.21	0.21
DRS4	0.21	0.21	0.21	0.20	0.20
IPW	0.23	0.25	0.31	0.22	0.20
Empirical SE					
Crude	0.20	0.20	0.20	0.20	0.20
LR	0.24	0.26	0.32	NA#	NA#
PS	0.22	0.23	0.24	0.26	NA#
DRS1	0.21	0.21	0.21	0.21	0.21
DRS2	0.21	0.21	0.21	0.21	0.21
DRS3	0.21	0.21	0.21	0.21	0.21

DRS4	0.21	0.21	0.21	0.21	0.21
IPW	0.23	0.25	0.34	1.51	0.20

* LR, logistic regression; PS, propensity score; DRS, disease risk score; IPW, inverse probability weights; RMSE indicates the square root of the mean squared error. # At an EPC of 1.0 1,078 IPW models failed to converge. Defining extreme estimates as an absolute estimate above 5 resulted in 7,737 (EPC 1.0) and 9,413 (EPC 0.5) extreme results LR method, 1,0000 extreme estimates (EPC 0.5) for the PS method and 1,177 (EPC 1.0) extreme results for the IPW model. # While all LR samples converged, the extreme estimate of the mean natural logarithm of the odds ratio resulted in an error when calculating the mean OR and relative bias

eTable 6 Simulation results repeating scenario 1 with extreme confounding bias*.

	10 EPC	5 EPC	2.5 EPC	1 EPC^	0.5 EPC
Mean odds ratio					
Crude	7.00	7.05	7.10	7.42	7.96
LR	0.99	0.99	0.98	126.98	0.00
PS	0.98	0.98	1.04	2.29	0.01
DRS1	1.01	1.04	1.09	1.27	1.62
DRS2	1.01	1.04	1.09	1.27	1.62
DRS3	1.01	1.04	1.09	1.27	1.62
DRS4	1.03	1.09	1.20	1.59	2.35
IPW	1.12	1.11	0.99	7.42	8.16
Relative bias (%)					
Crude	599.59	605.17	610.36	641.65	695.53
LR	-1.10	-1.44	-1.74	$1.26*10^4$	-100
PS	-1.93	-2.15	3.60	128.78	-98.79
DRS1	1.07	3.65	8.87	27.36	61.77
DRS2	1.06	3.64	8.86	27.34	61.78
DRS3	1.07	3.66	8.89	27.41	61.96
DRS4	3.47	8.99	20.22	58.73	135.18
IPW	12.19	10.80	-1.08	641.65	695.53
Coverage					
Crude	0.000	0.000	0.000	0.000	0.000
LR	0.942	0.929	0.893	1.000	1.000
PS	0.956	0.960	0.963	0.353	0.954
DRS1	0.952	0.945	0.935	0.863	0.617
DRS2	0.951	0.945	0.935	0.864	0.618
DRS3	0.952	0.945	0.935	0.863	0.615
DRS4	0.948	0.935	0.896	0.637	0.139
IPW	0.949	0.926	0.839	0.000	0.000
RMSE					
Crude	1.96	1.97	1.97	2.02	2.09
LR	0.35	0.40	0.61	17.05	$5.24*10^2$
PS	0.31	0.30	0.31	0.90	509.27
DRS1	0.31	0.32	0.33	0.39	0.57
DRS2	0.31	0.32	0.33	0.39	0.57
DRS3	0.31	0.32	0.33	0.39	0.57
DRS4	0.31	0.33	0.36	0.55	0.90
IPW	0.39	0.47	0.92	2.02	2.09
Mean SE					
Crude	0.23	0.23	0.23	0.23	0.23
LR	0.33	0.37	0.48	$2.16*10^5$	$7.73*10^6$
PS	0.31	0.32	0.32	0.36	50.63
DRS1	0.31	0.31	0.31	0.30	0.29
DRS2	0.31	0.31	0.31	0.30	0.29
DRS3	0.31	0.31	0.31	0.30	0.29
DRS4	0.31	0.31	0.30	0.29	0.28
IPW	0.39	0.43	0.58	0.23	0.23
Empirical SE					
Crude	0.23	0.23	0.23	0.23	0.23
LR	0.35	0.40	0.61	16.34	$5.24*10^2$
PS	0.31	0.30	0.30	0.35	509.25
DRS1	0.31	0.32	0.30	0.31	0.30
DRS2	0.31	0.32	0.31	0.31	0.30
DRS3	0.31	0.32	0.31	0.31	0.30

DRS4	0.31	0.31	0.31	0.30	0.30
IPW	0.38	0.43	0.92	0.23	0.23

* LR, logistic regression; PS, propensity score; DRS, disease risk score; IPW, inverse probability weights; RMSE indicates the square root of the mean squared error. All models converged.

Reference List

(1) Lunt M, Solomon D, Rothman K et al. Different methods of balancing covariates leading to different effect estimates in the presence of effect modification. *Am J Epidemiol.* 2009;169(7):909-917.

(2) Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology.* 2000;11(5):550-560.

(3) Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol.* 2008;168(6):656-664.

(4) van der Wal WM, Noordzij M, Dekker FW et al. Comparing mortality in renal patients on hemodialysis versus peritoneal dialysis using a marginal structural model. *Int J Biostat.* 2010;6(1):Article.

(5) Xu S, Ross C, Raebel MA, Shetterly S, Blanchette C, Smith D. Use of stabilized inverse propensity scores as weights to directly estimate relative risk and its confidence intervals. *Value Health.* 2010;13(2):273-277.

(6) Zeileis A. Object-Oriented Computation of Sandwich Estimators. *Journal of Statistical Software.* 2006;16(9):1-16.