WEB APPENDIX for “A tutorial on dealing with time-varying eligibility for treatment: Comparing the risk of major bleeding with DOACs versus warfarin”

AUTHORS: Mireille E. Schnitzer[1], Robert W. Platt[2], and Madeleine Durand[3]

AFFILIATIONS

[1]Faculty of Pharmacy and Department of Social and Preventative Medicine, Université de Montréal, Montreal, Canada
[2]Departments of Pediatrics and of Epidemiology, Biostatistics, and Occupational Health, McGill University, and the Research Institute of the McGill University Health Centre, Montreal, Canada
[3]Internal Medicine Service, Department of Medicine, Centre Hospitalier de l’Université de Montréal, and the Research Center of Centre Hospitalier de l’Université de Montréal, Montréal, Canada

A1. SUMMARY MEASURES FOR THE THREE SIMULATED SCENARIOS

	A_0=1 (DOAC)	**A_0=0 (warfarin)**
Scenario 1		
	n=15314	*n*=9686
Baseline variables		
W, *mean (sd)*	-0.17 (0.98)	0.30 (0.96)
U, *mean (sd)*	0.00 (0.99)	0.00 (1.00)
Post-initial exposure variables		
RF, *p*	0.07	0.10
A_1, *p*	0.84	0.15
A_1 amongst those with RF=0	0.91	0.17
A_1 amongst those with RF=1	0	0
Y, *p*	0.22	0.25
Scenario 2		
	n=15314	*n*=9686
Baseline variables		
W, *mean (sd)*	-0.17 (0.98)	0.30 (0.96)
U, *mean (sd)*	0.00 (0.99)	0.00 (1.00)
Post-initial exposure variables		
RF, *p*	0.07	0.10
A_1, *p*	0.84	0.15
A_1 amongst those with RF=0	0.91	0.17
A_1 amongst those with RF=1	0	0
Y, *p*	0.21	0.24
Scenario 3		
	n=15314	*n*=9686
Baseline variables		
W, *mean (sd)*	-0.17 (0.97)	0.30 (0.96)
U, *mean (sd)*	0.00 (0.99)	0.00 (1.00)
Post-initial exposure variables		
RF, *p*	0.14	0.07
A_1, *p*	0.78	0.16
A_1 amongst those with RF=0	0.91	0.17
A_1 amongst those with RF=1	0	0
Y, *p*	0.22	0.24

Web Table 1. Baseline and Post-initial Exposure Summary Measures for Both Initial Exposure Groups in the Simulated Data Example.

A2. CODE FOR GENERAL NONPARAMETRIC BOOTSTRAP AND APPLIED TO IPTW

When inverse probability weights are used, it is often recommended to estimate the standard error using the bootstrap. Here we provide a simple general nonparametric implementation that can be applied to an arbitrary estimator of a vector parameter.

Box A2.1: Nonparametric bootstrap function for a vector parameter

```
bootstrap.vect<-function( funct, DAT , K=500, n=dim(DAT)[1] ){
result.full<-funct(DAT)
mat.est<-matrix(0,ncol=length(result.full),nrow=K)
for (k in 1:K){
 #resample the original dataset
 resamp<-sample(1:n, n, replace=T)
 DATk<-as.data.frame(DAT[resamp,])
 estimate<-funct(DATk)
 mat.est[k,]<-estimate
 }
VAR=apply(mat.est,2,var)
CIlow=apply(mat.est,2,function(x) quantile(x,0.025))
CIhigh=apply(mat.est,2,function(x) quantile(x,0.975))

#Output: variance, 95% confidence limits,
#and the bootstrap estimates
return(list(VAR=VAR,CIlow=CIlow,CIhigh=CIhigh,
estList=mat.est))

}
```

In order to apply the above bootstrap, we must create a function that inputs the data and outputs the vector estimate. Here is such a function of the IPTW estimator described in this manuscript for both the risk ratio and odds ratio.

Box A2.2: Function for IPTW

```
iptw_RF<-function(DAT){
#define regimen W as: stay on warfarin
#define regimen D as: stay on DOAC unless RF, then warfarin
#patients following reg W at first time point
RW0<-DAT$A0==0
#patients following reg W at first AND second time points
RW01<-(DAT$A0==0&DAT$A1==0)
#patients following reg D at first time point
RD0<-DAT$A0==1
#patients following reg D at first AND second time points
RD01<-(DAT$A0==1&DAT$A1==1) | (DAT$A0==1&DAT$RF==1)

A0mod<-glm(A0~W,family=binomial(),data=DAT)
A1mod<-glm(A1~A0+W,family=binomial(),subset=(RF==0),data=DAT)

P1<-predict(A0mod,type="response");  P0<-1-P1
P00<-((1-predict(A1mod,
 newdata=as.data.frame(cbind(W=DAT$W,A0=0)),type="response"))*
 (DAT$RF==0)+1*(DAT$RF==1))*P0
P11<-((DAT$RF==0)*
 predict(A1mod,newdata=as.data.frame(cbind(W=DAT$W,A0=1)),
 type="response") + (DAT$RF==1))*P1

w<-(1/P00*RW01 + 1/P11*RD01)

sub<-(RD01==1|RW01==1)

#5% weight truncation
#w[w>quantile(w[sub],0.05)]<-quantile(w[sub],0.05)

#Option 1: RISK DIFFERENCE
IPTWmod<-lm(Y~RD01,weights=w,subset=sub,data=DAT)
est_RD<-IPTWmod$coef[2]

#Option 2: RISK RATIO
IPTWmod<-glm(Y~RD01,weights=w,subset=sub,
 family=quasibinomial(link=log),data=DAT)
est_RR<-exp(IPTWmod$coef[2])

#Option 3: ODDS RATIO
IPTWmod<-
glm(Y~RD01,weights=w,subset=sub,family=quasibinomial(),data=DAT)
est_OR<-exp(IPTWmod$coef[2])

return(c(est_RD,est_RR,est_OR)) #output all three estimates
}
```

The next box shows how to use the above functions to obtain bootstrapped estimates of the standard errors of the risk difference, risk ratio, and odds ratio given data O=(W,A0,RF,A1,Y) as defined previously.

Box A2.3: Using the bootstrap and IPTW functions

```
#Both functions must first be run
#Save the data as a dataframe
DAT<-as.data.frame(cbind(W,A0,RF,A1,Y))

#K is the number of bootstrap resamples
bsres<-bootstrap.vect(iptw_RF,DAT,K=500)

#Estimated standard error for each parameter estimate
sqrt(bsres$VAR)

#Lower-limit of the 95% confidence interval for each estimate
bsres$CIlow

#Upper-limit of the 95% confidence interval for each estimate
bsres$CIhigh

#List of bootstrapped estimates (vector of size K)
#Can use to plot histogram and/or identify failed runs
Bsres$estList
```

A3. SIMULATION STUDY RESULTS

We compare the performance of the four adherence analyses (M1-M4) with the performance of the IPW application contrasting the two treatment strategies. Recall that, in our simulated data with the data-generating functions given in Box 2, there is no true difference between the effects of DOACs and warfarin on bleeding (i.e. no arrows leading from either exposure node to the outcome, Y). Thus, a correct result from this analysis would result in null estimates (OR=1).

We independently sample 1,000 datasets of size n=25,000 and run the models M1-M4 and IPW for treatment strategy on each of the 1,000 datasets. The parameters estimated by the first four methods are the exponential of the coefficients in the models M1-M4, i.e. the conditional odds ratios related to A_0 and A_1, respectively. The parameter estimated by the IPW method contrasting strategies R_D and R_W is the marginal odds ratio $(P[Y(R_D)=1]/\{1- P[Y(R_D)=1]\})/(P[Y(R_W)=1]/\{1- P[Y(R_W)=1]\})$. We then save the point estimates from each model run on each dataset and also note whether the resulting confidence interval includes the null. If not, this is considered a false positive, incorrectly concluding that there is an effect of anticoagulant choice on bleeding.

Parameter	**M1. Adjust for W**	**M2. Subset on RF**	**M3. Subset on RF with weights**	**M4. Adjust for RF**	**IPW for strategy**
Scenario 1					
OR of					
A_0	1.07 (31.9)	1.00 (4.2)	1.00 (5.4)	1.00 (4.8)	
A_1	0.90 (61.0)	1.00 (4.6)	1.00 (5.2)	1.00 (4.6)	
R_D vs R_W					1.00 (5.4)
Scenario 2					
OR of					
A_0	1.11 (69.4)	1.08 (29.2)	1.07 (27.7)	1.02 (7.7)	
A_1	0.82 (98.3)	0.88 (63.4)	0.89 (61.8)	0.94 (20.3)	
R_D vs R_W					1.00 (6.6)
Scenario 3					
OR of					
A_0	1.61 (100.0)	0.82 (95.4)	0.82 (97.6)	0.81 (99.7)	
A_1	0.41 (100.0)	1.00 (3.6)	1.01 (5.7)	1.00 (3.6)	
R_D vs R_W					1.00 (5.0)

Web Table 2. Mean of the Point Estimates and % False Positives for Regression Methods M1-M4 and the IPW Treatment Strategy Approach over 1,000 Simulated Draws.

Web Table 2 presents the mean of the odds ratio point estimates and the % of false positives over the 1,000 simulated datasets in each scenario. For scenario 1, there is a small average bias for both exposure effects, leading to 32% and 61% false positives for A_0 and A_1 respectively. For the models M2-M4 there is no bias with roughly 5% false positives in each case as desired. There is similarly no bias for the IPW approach and 5% false positives, indicating that the standard errors were well-estimated by bootstrap. For scenario 2, we see that for all regression models, there is an upwards mean bias in the odds ratio estimates of A_0 and a downwards bias for A_1. The biases are particularly large for M1, which ignored the contraindication entirely, and lowest for M4, which adjusted for RF in the regression model. The high level of false positives in all scenarios suggests that an investigator is likely to conclude that there is indeed an effect. The IPW method had no bias on average and 6.6% type 1 error. In scenario 3, where RF is a collider, models M1-M4 produced the largest average bias overall resulting in very high levels of type 1 errors. In particular, the direction of bias of the effect of A_0 for models M2-M4 was reversed compared to the previous scenario, suggesting a protective effect of early exposure. In M2-M4, there was no average bias in the estimation of the effect of A_1, which makes sense as these models correctly adjusted for the confounder RF of the relationship A_1 – Y. The IPW estimator was unbiased on average with optimal type 1 error.

In order to evaluate the IPW estimator for a true parameter off the null, we repeated the same simulation as scenario 3, except that we changed the outcome generation to

```
Y0<-rbinom(size=1,n=ssize,p=plogis(-2+0.5*W+2*U+0.2*A0+0.4*A1))
```

so that both early and later treatments had positive effects on the probability of an outcome. The results are in Web Table 3.

Parameter	M1. Adjust for W	M2. Subset on RF	M3. Subset on RF with weights	M4. Adjust for RF	IPW for strategy
OR of					
A_0	1.78 (1.00)	0.96 (0.15)	0.94 (0.23)	0.93 (0.35)	
A_1	0.56 (1.00)	1.31 (1.00)	1.32 (1.00)	1.31 (1.00)	
R_D vs R_W					1.38 (1.00)

Web Table 3. Mean of the Point Estimates and % True Positives (power) for Regression Methods M1-M4 and the IPW Treatment Strategy Approach over 1,000 Simulated Draws. In this scenario, both early and later exposures have small positive effects (conditional OR=1.22 and 1.49, resp.). True effect of treatment strategy OR = 1.38.

A4. CODE FOR STABILIZED WEIGHTS

Better finite-sample properties can be obtained by computing stabilized weights that include a numerator probability that does not adjust for W. Weight stabilization has a negligible impact on the current finite sample simulation results.

Box A4.1: Stabilized weights for IPW for treatment strategy

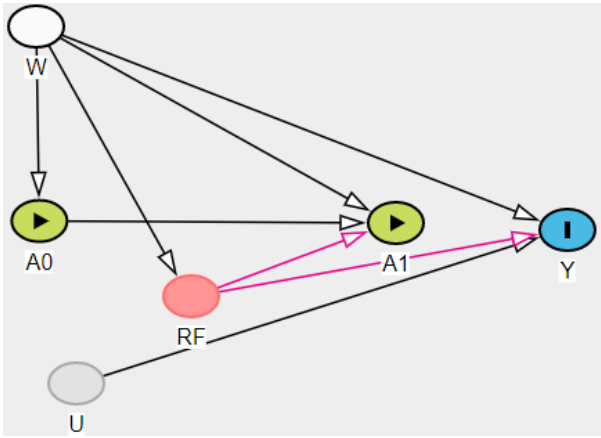
```
#denominator (same as before)
A0mod<-glm(A0~W,family=binomial())
A1mod<-glm(A1~A0+W,family=binomial(),subset=(RF==0))

P1<-predict(A0mod,type="response")
P0<-1-P1
P00<-((1-predict(A1mod,
newdata=as.data.frame(cbind(W,A0=0)),type="response"))*(RF==0)+1*
(RF==1))*P0
P11<-
((RF==0)*predict(A1mod,newdata=as.data.frame(cbind(W,A0=1)),type=
"response")+ (RF==1))*P1

#numerator
P1n<-mean(A0)
P0n<-1-P1n
P00n<-((1-mean(A1[RF==0&A0==0]))*(RF==0)+1*(RF==1))*P0n
P11n<-((RF==0)*mean(A1[RF==0&A0==1])+ (RF==1))*P1n

#stabilized weights
ws<-(P00n/P00*RW01 + P11n/P11*RD01)
```

A5. DAGITTY CODE FOR THE DAGS PRESENTED IN FIGURE 2

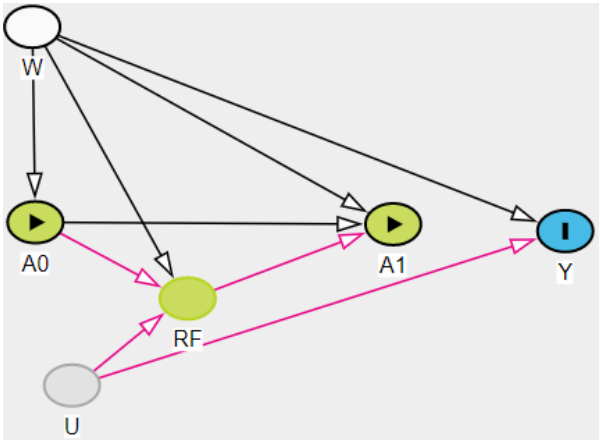
An anonymous reviewer generously provided DAGitty (http://www.dagitty.net/dags.html) code to reproduce the DAGs presented in Figure 2, which we relay to the reader.

<table>
<tr>
<td>

```
dag {
bb="0,0,1,1"
A0
[exposure,pos="0.170,0.488"]
A1
[exposure,pos="0.405,0.490"]
RF [pos="0.270,0.566"]
U [latent,pos="0.194,0.655"]
W [adjusted,pos="0.168,0.288"]
Y [outcome,pos="0.518,0.497"]
A0 -> A1
RF -> A1
RF -> Y
U -> Y
W -> A0
W -> A1
W -> RF
W -> Y
}
```

</td>
<td> 
</td>
</tr>
<tr>
<td>

```
dag {
bb="0,0,1,1"
A0
[exposure,pos="0.170,0.488"]
A1
[exposure,pos="0.405,0.490"]
RF [pos="0.270,0.566"]
U [latent,pos="0.194,0.655"]
W [adjusted,pos="0.168,0.288"]
Y [outcome,pos="0.518,0.497"]
A0 -> A1
A0 -> RF
RF -> A1
U -> RF
U -> Y
W -> A0
W -> A1
W -> RF
W -> Y
}
```

</td>
<td> 
</td>
</tr>
</table>