PROTECT

Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium

Control for time-dependent confounding using Inverse Probability Weighting of Marginal Structural Models

18 February 2015

Objectives

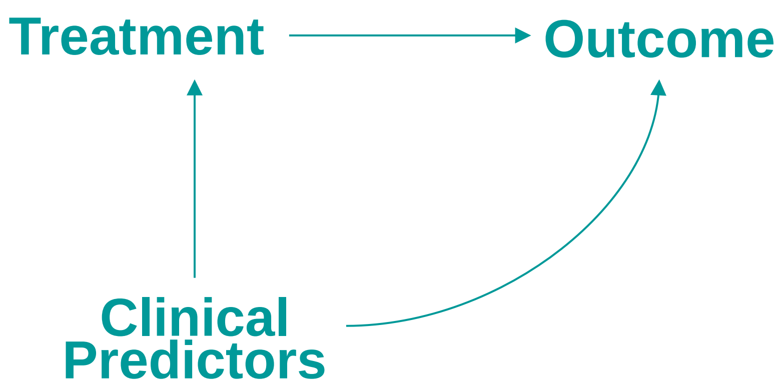
At the end of this session, you will be able to:

- Understand of time-dependent confounding in pharmacoepidemiologic research
- Limitations of conventional regression methods to handle time-dependent confounding
- Inverse probability weighting of marginal structural models
- Limitations of IPW and possible alternatives to address time-dependent confounding

Outline

- Confounding in causal diagrams terminology?
- How conventional methods adjust for confounding?
- How medication use in real life affect this assumption?
- What would this mean in terms of causal diagrams and adjustment for confounding: the Problem?
- Marginal structural models (MSMs): the solution?
- MSMs whose parameters are estimated using Inverse Probability of Treatment Weighting (IPTW)
- Assumptions and Limitations of MSMs?
- Example of antidepressant use and the risk of hip fracture
- Take home message

Confounding ?



- Terminologies: - Directed/back-door path
  - Collider : common effect
  - Confounder : common cause
  - Intermediate
- A path is blocked if it has one or more collider

Controlling for Confounding :

- **Stratification**

  For example, using Mantel-Haenszel methods,

  Propensity score Stratification

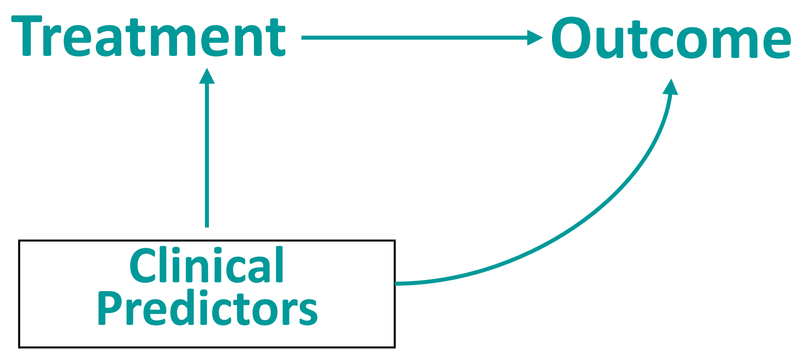
- **Regression models**

  Including treatment and confounders as covariates,

  Covariate adjustment using propensity score

Blocking all back-door paths!

Controlling for Confounding :



- Blocking all back-door paths!

Conditioning on:

- Collider will open a closed path
- Confounder/intermediate will close a path

Medication Use in Real Life ?

- Traditional methods make assumptions that treatment is fixed during follow-up.

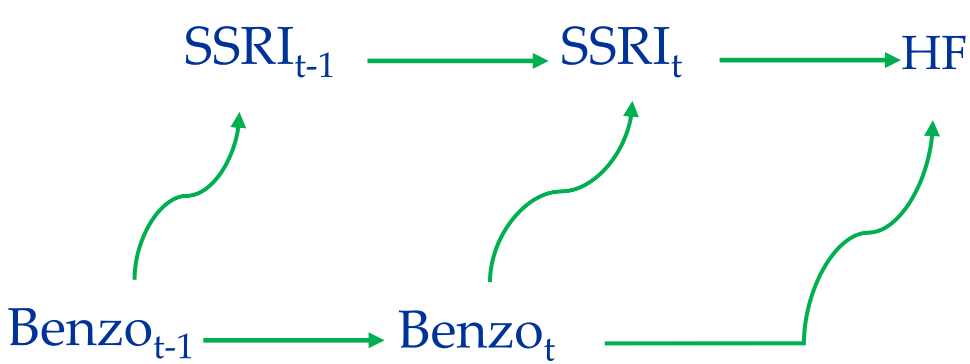
- In real life, patients could switch or stop treatment for several reasons:

  - Non-response
  - Side-effects
  - Recovery

- Can we capture it? Assumptions in Electronic Health care records

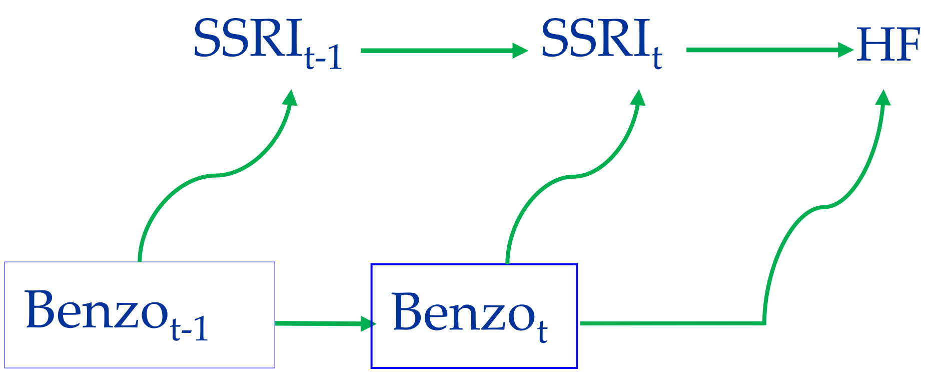
Medication Use in Real Life ?

- Selective Serotonin Reuptake Inhibitors (SSRI) use and the risk of hip fracture (HF)
- Benzodiazepine (Benzo) use as one of the confounding variables (time-fixed/time-varying)



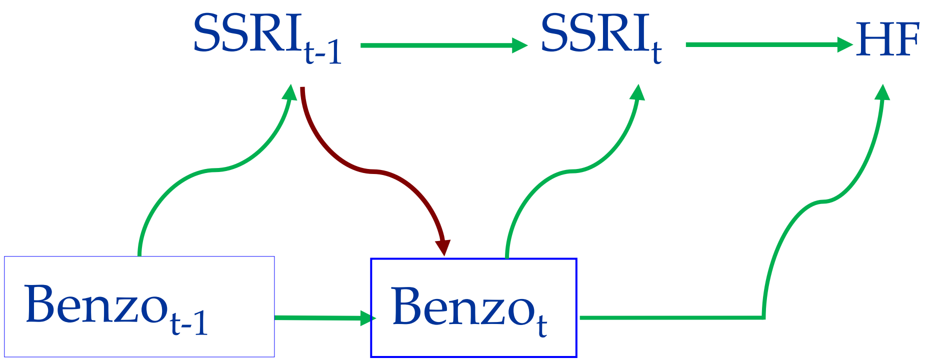
Medication Use in Real Life ?

- Cox proportional hazards models with time-varying coefficients to control for time-varying nature of covariates



Time-dependent Confounding

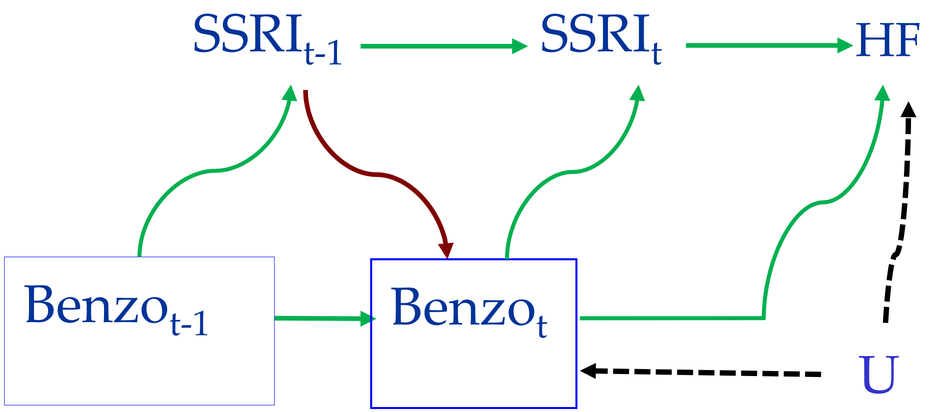
- When time-varying confounders are themselves affected by the previous treatment (SSRI use), conventional time-varying Cox model:
  - Can no longer provide unbiased estimates of the treatment effect:
  - It will adjust-away some effect of treatment



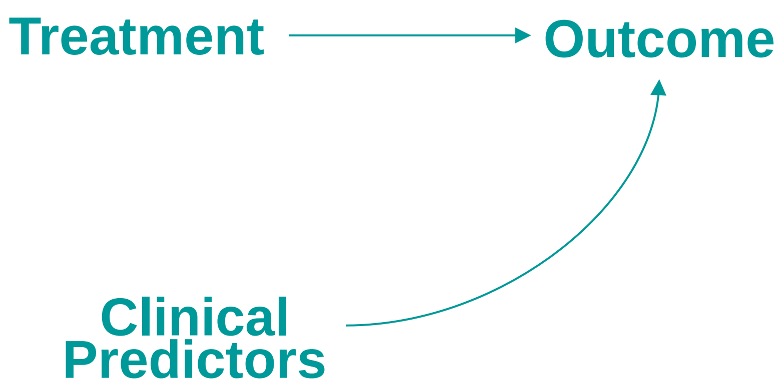
Time-dependent Confounding

- In the presence of unmeasured common causes of confounders and outcome (U):

  - It leads to collider-stratification bias



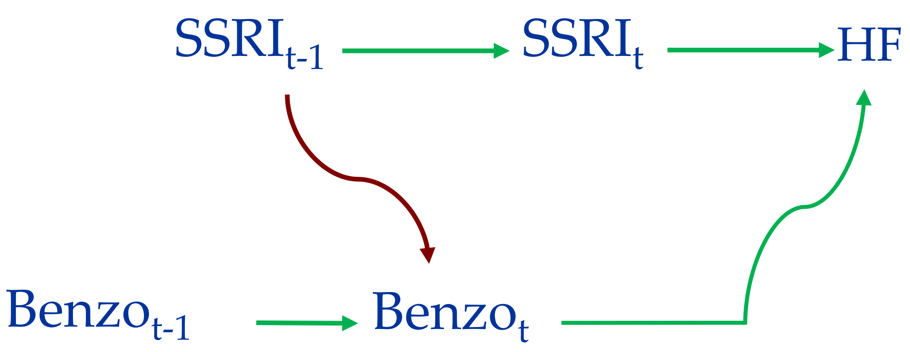
Inverse probability of Treatment Weighting, IPTW?



- Reweighting the population using the inverse of subject's probability to receive treatment given baseline covariates (Clinical predictors) = **Pseudopopulation**

Treatment and Clinical Predictors are not associated!

Inverse probability of Treatment Weighting, IPTW?

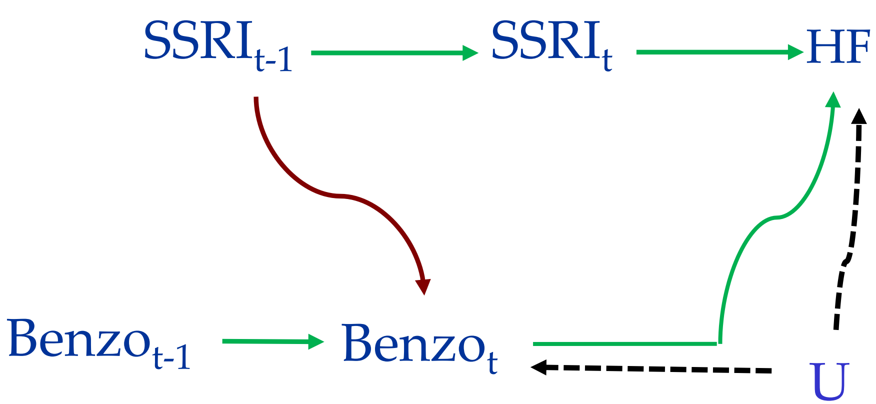


- Reweighting the population using the inverse of subject's probability to receive treatment given baseline confounders

Pseudopopulation

SSRI use and benzo use are not associated!

Inverse probability of Treatment Weighting, IPTW?



- Reweighting the population using the inverse of subject's probability to receive treatment given baseline confounders

Pseudopopulation

SSRI use and benzo use are not associated!

Fitting MSMs (IPW)

- Given, treatment (A), time points k, censoring status (C), and time-varying covariates (L)
- Estimate Probability of Treatment conditional on covariates
- Estimate Treatment weights

$$STW_i(t) = \prod_{k=0}^{t} \frac{\Pr(A(k) = a_i(k) \mid \bar{A}(k-1) = \bar{a}_i(k-1))}{\Pr(A(k) = a_i(k) \mid \bar{A}(k-1) = \bar{a}_i(k-1), \bar{L}(k) = \bar{l}_i(k))}$$

- Estimate Censoring weights

$$SCW_i(t) = \prod_{k=0}^{t} \frac{\Pr(C(k) = 1 \mid \bar{C}(k-1) = 0, \bar{A}(k-1) = \bar{a}_i(k-1))}{\Pr(C(k) = 1 \mid \bar{C}(k-1) = 0, \bar{A}(k-1) = \bar{a}_i(k-1), \bar{L}(k) - \bar{l}_i(k)}$$

- Fit the outcome model

What can we estimate?

- Using IPTW, one can estimate
  - Cumulative treatment effect
  - Instantaneous treatment effect

Assumptions of MSMs

- **Exchangeability**
  - No unmeasured confounding and
  - Non-informative censoring

    Censoring Weights!

- **Positivity**
  - One must be able to estimate the average causal effect in each subset of the population defined by the confounders.
  - Estimated weights with the mean far from one or very extreme values are indicative of non- positivity or misspecification of the weight model.

Assumptions of MSMs

- **Consistency**
  - Which often is stated such that an individual's potential outcome under her observed exposure history is precisely her observed outcome
- **Correct model specification of models**
  - Weighted estimation of the parameters of MSMs requires fitting several models:
    1) the structural (i.e., weighted) model,
    2) the exposure model, and
    3) the censoring model.

Limitations of MSMs (IPTW)

- Can not be used to model interaction of treatment with time-dependant covariates
- Can not be used in studies in which all subjects with a particular covariate are certain to receive the same treatment

Non-positivity:

For example, occupational cohort study

- Can not be used in the presence of confounding by unmeasured factors (even if we have data on instrumental variable)

Alternatives to IPTW

- G- estimation of structural nested models (SNMs)
- G- computation algorithm formula estimator
- The iterative conditional expectations (ICE) estimator

MSMs (IPTW) in PROTECT

- Empirical Studies:
  - Beta2-agonist use and the risk of coronary heart disease (Eur J Epidemiol, 2013)
  - Antidepressant (SSRI) use and the risk of hip fracture

Antidepressant Use and the Risk of Hip Fracture

- Study using 2 EU databases (Mondriaan and BIFAP)
- A cohort of patients with a first prescription for Antidepressant (SSRI or tricyclic AD, TCA) period 2001-2009
- Data source(GP databases) :
  - the Dutch Mondriaan
  - the Spanish BIFAP
- Effects of SSRI use versus no SSRI use were estimated using time-varying Cox regression, PS stratification and regression analysis, as well as MSM.

M Sanni Ali, et al. submitted

Associations Between SSRI Use and the Risk of Hip Fracture Using Time-Varying Cox Models

Adjusted for	Mondriaan		BIFAP	
	HR	95% CI	HR	95% CI
Crude	**1.75**	**1.12, 2.72**	**2.09**	**1.89, 2.32**
Gender	**1.73**	**1.10, 2.69**	**2.07**	**1.87, 2.30**
Gender +Age	**2.36**	**1.51, 3.68**	**1.51**	**1.37, 1.68**
Gender +Age + TCA_t	**2.59**	**1.63, 4.12**	**1.56**	**1.40, 1.73**
Gender +Age + TCA_t + $Benzo_t$	**2.60**	**1.63, 4.16**	**1.54**	**1.38, 1.71**
All Confounders*	**2.62**	**1.63, 4.19**	**1.52**	**1.37, 1.69**

*Age, Gender, TCA_t use, Benzodiazepine use ($Benzo_t$), Bone related medications, Anti-inflammatory medications, cardiovascular co-morbidities, Neurological co-morbidities, Respiratory co-morbidities, Previous history of fractures, and Gastrointestinal medications.

Associations Between SSRI Use and the Risk of Hip Fracture Using Propensity Score Based Cox Analyses

Adjusted for		Mondriaan		BIFAP	
		HR	95% CI	HR	95% CI
Crude		**1.75**	**1.12, 2.72**	**2.09**	**1.89, 1.71**
PS* Stratification	Quintiles	**2.64**	**1.63, 4.25**	**1.54**	**1.39, 1.71**
	Deciles	**2.72**	**1.63, 4.54**	**1.53**	**1.38, 1.70**
PS* Adjustment		**2.82**	**1.63, 4.25**	**1.61**	**1.45, 1.78**

* Included in the PS model: Age, Gender, TCA_t use, Benzodiazepine use ($Benzo_t$), Bone related medications, Anti-inflammatory medications, cardiovascular co-morbidities, Neurological co-morbidities, Respiratory co-morbidities, Previous history of fractures, and Gastrointestinal medications.

Association Between SSRI Use and the Risk of Hip Fracture Using Trimmed IPW estimation+ of Marginal Structural Models Without and With Accounting for Censoring

	Adjusted for	Mondriaan		BIFAP	
		HR	95% CI	HR	95% CI
Without Accounting for Censoring*	Crude	**1.69**	**1.05, 2.67**	**2.14**	**1.91, 2.39**
	Gender	**1.68**	**1.06, 2.67**	**2.11**	**1.89, 2.38**
	Gender +Age	**2.46**	**1.55, 3.99**	**1.54**	**1.37, 1.72**
Accounting for Censoring**	Crude	**1.73**	**1.08, 2.77**	**2.05**	**1.83, 2.30**
	Gender	**1.71**	**1.07, 2.74**	**2.03**	**1.81, 2.28**
	Gender +Age	**2.47**	**1.53, 3.98**	**1.51**	**1.35, 1.70**

*Only inverse probability of treatment weights were used

**Combined inverse probability of treatment and censoring weights were used, + Trimming at 1% and 99%

How to … data for IPTW studies

R Console

```
data3[1:15,]
```

id	start_date	stop_date	TCAt	TCA1	SSRIt	SSRI1	Cen	Gender	Benzo	Fract	Comed	Cardi	AInfla	Resp	Comorb	Event
3	2005-04-07	2005-05-31	0	1	0	0	1	0	1	0	0	0	0	0	0	0
3	2005-05-31	2005-07-30	1	0	0	0	1	0	1	0	0	0	0	0	0	0
3	2005-07-30	2006-01-28	0	0	0	0	1	0	1	0	0	0	0	0	0	0
3	2006-01-28	2006-07-29	0	0	0	0	1	0	1	0	0	0	1	0	0	0
3	2006-07-29	2007-01-27	0	0	0	0	1	0	1	0	0	0	0	0	0	0
3	2007-01-27	2007-07-28	0	0	0	0	1	0	1	0	0	0	1	0	0	0
3	2007-07-28	2008-01-26	0	0	0	0	1	0	1	0	1	0	1	0	0	0
3	2008-01-26	2008-07-26	0	0	0	0	1	0	1	0	0	0	0	0	0	0
3	2008-07-26	2009-01-24	0	0	0	0	1	0	0	0	0	0	0	0	0	0
3	2009-01-24	2009-02-11	0	0	0	0	0	0	1	1	1	1	1	0	0	0
4	2005-09-27	2005-12-22	0	0	0	1	1	0	0	0	0	1	0	0	1	0
4	2005-12-22	2006-02-20	0	0	1	0	1	0	1	0	0	1	0	0	1	0
4	2006-02-20	2006-08-21	0	0	0	0	1	0	1	0	0	1	0	0	1	0
4	2006-08-21	2007-02-19	0	0	0	0	1	0	0	0	0	1	0	0	1	0
4	2007-02-19	2007-08-20	0	0	0	0	1	0	0	0	0	1	0	0	1	0

How to … Fitting Cox Model

- Conventional regression with time-varying coefficients

F1 <- coxph(Surv(start,stop,Event) ~ SSRI1 + Gender + age +… + cluster(data$id), data=data)

Summary (F1)

	exp(coef)	exp(-coef)	lower .95	upper .95
SSRI1	1.5232	0.6565	1.3701	1.6932
Gender	0.8464	1.1814	0.7409	0.9669
age	1.1183	0.8942	1.1128	1.1238
TCAt	1.3855	0.7218	1.1013	1.7430
Benzo	1.1329	0.8827	1.0197	1.2586
Fract	1.4435	0.6928	1.2599	1.6537
Cardi	0.9916	1.0085	0.8842	1.1120
Comed	1.2268	0.8152	1.0920	1.3781
Comorb	1.0867	0.9202	0.9806	1.2043
Resp	0.8691	1.1506	0.7467	1.0115
AInfla	1.0349	0.9663	0.9159	1.1693

How to … Fitting MSM Models

- MSM with IPTWs

Generating Treatment weights

```
Psq.1 <- glm(SSRI1 ~ SSRIt + Gender + age +..  cluster(data$id),data=data,family =binomial)$fitted
Psq.0 <- glm(SSRI1 ~ SSRIt + Gender + age + cluster(data$id), data=data,family = binomial)$fitted
```

```
iptws[data$SSRI1==1] <-   data$Psq.0[data$SSRI1==1]/data$Psq.1[data$SSRI1==1]
iptws[data$SSRI1==0] <-   (1-data$Psq.0)[data$SSRI1==0]/(1-data$Psq.1)[data$SSRI1==0]
```

How to … data for IPTW studies

R Console

```
data3[1:15,]
 id start_date  stop_date TCAt TCA1 SSRIt SSRI1       Psq.1     Psq.0     iptws
  3 2005-04-07 2005-05-31    0    1     0     0 0.349921088 0.2792739 1.1086748
  3 2005-05-31 2005-07-30    1    0     0     0 0.005456679 0.2792739 0.7246805
  3 2005-07-30 2006-01-28    0    0     0     0 0.349921088 0.2792739 1.1086748
  3 2006-01-28 2006-07-29    0    0     0     0 0.332045237 0.2792739 1.0790044
  3 2006-07-29 2007-01-27    0    0     0     0 0.350971313 0.2806462 1.1083543
  3 2007-01-27 2007-07-28    0    0     0     0 0.333069298 0.2806462 1.0786035
  3 2007-07-28 2008-01-26    0    0     0     0 0.360467877 0.2820227 1.1226602
  3 2008-01-26 2008-07-26    0    0     0     0 0.352022982 0.2820227 1.1080289
  3 2008-07-26 2009-01-24    0    0     0     0 0.246568071 0.2834033 0.9511101
  3 2009-01-24 2009-02-11    0    0     0     0 0.365737811 0.2834033 1.1298115
  4 2005-09-27 2005-12-22    0    0     0     1 0.308560566 0.3206831 1.0392875
  4 2005-12-22 2006-02-20    0    0     1     0 0.527094605 0.4677512 1.1254868
  4 2006-02-20 2006-08-21    0    0     0     0 0.426681899 0.3206831 1.1848865
  4 2006-08-21 2007-02-19    0    0     0     0 0.309545764 0.3221680 0.9817189
  4 2007-02-19 2007-08-20    0    0     0     0 0.309545764 0.3221680 0.9817189
```

How to … Fitting MSM Models

- MSM with IPCWs

Generating Censoring Weights:

data$

PCen.1 <- glm(Cen ~ SSRIt + Gender + age+.. cluster(data$id), data=data,family = binomial)$fitted

PCen.0 <- glm(Cen ~ SSRIt + Gender + age + cluster(data$id), data=data,family = binomial)$fitted

ipcws[data$Cen==1] <-   data$PCen.0[data$Cen==1] / data$PCen.1 [data$Cen==1]

ipcws[data$Cen==0] <-   (1-data$PCen.0)[data$Cen==0] / (1-data$PCen.1)[data$Cen==0]

PROTECT

How to … data for IPTW studies

R Console

```
data3[1:15,]
 id start_date  stop_date TCAt TCA1 SSRIt SSRI1       Psq.1     Psq.0     iptws    PCen.1    PCen.0     ipcws
  3 2005-04-07 2005-05-31    0    1     0     0 0.349921088 0.2792739 1.1086748 0.8998709 0.8825604 0.9807634
  3 2005-05-31 2005-07-30    1    0     0     0 0.005456679 0.2792739 0.7246805 0.9502811 0.8825604 0.9287361
  3 2005-07-30 2006-01-28    0    0     0     0 0.349921088 0.2792739 1.1086748 0.8998709 0.8825604 0.9807634
  3 2006-01-28 2006-07-29    0    0     0     0 0.332045237 0.2792739 1.0790044 0.8969601 0.8825604 0.9839461
  3 2006-07-29 2007-01-27    0    0     0     0 0.350971313 0.2806462 1.1083543 0.8995490 0.8820682 0.9805672
  3 2007-01-27 2007-07-28    0    0     0     0 0.333069298 0.2806462 1.0786035 0.8966299 0.8820682 0.9837595
  3 2007-07-28 2008-01-26    0    0     0     0 0.360467877 0.2820227 1.1226602 0.8926524 0.8815742 0.9875895
  3 2008-01-26 2008-07-26    0    0     0     0 0.352022982 0.2820227 1.1080289 0.8992261 0.8815742 0.9803699
  3 2008-07-26 2009-01-24    0    0     0     0 0.246568071 0.2834033 0.9511101 0.8837166 0.8810784 0.9970146
  3 2009-01-24 2009-02-11    0    0     0     0 0.365737811 0.2834033 1.1298115 0.8676120 0.8810784 0.8982808
  4 2005-09-27 2005-12-22    0    0     0     1 0.308560566 0.3206831 1.0392875 0.8441833 0.8675457 1.0276746
  4 2005-12-22 2006-02-20    0    0     1     0 0.527094605 0.4677512 1.1254868 0.9221191 0.9223386 1.0002380
  4 2006-02-20 2006-08-21    0    0     0     0 0.426681899 0.3206831 1.1848865 0.8637357 0.8675457 1.0044111
  4 2006-08-21 2007-02-19    0    0     0     0 0.309545764 0.3221680 0.9817189 0.8437134 0.8670000 1.0276002
  4 2007-02-19 2007-08-20    0    0     0     0 0.309545764 0.3221680 0.9817189 0.8437134 0.8670000 1.0276002
```

How to … Fitting MSM Models

- MSM with IPCWs

Cumulative Weights:

```
c.iptws   <-   unsplit(sapply(split(data$iptws, data$id),cumprod), data$id)
c.ipcws   <-   unsplit(sapply(split(data$ipcws, data$id),cumprod), data$id)

data   <-   cbind(data,c.iptws,c.ipcws)
data$Cum.Wts   <-   data$c.ipcws*data$c.iptws
```

PROTECT

How to … data for IPTW studies

R Console

```
data3[1:15,]
```

id	start_date	stop_date	TCAt	TCA1	SSRIt	SSRI1	Psq.1	Psq.0	iptws	PCen.1	PCen.0	ipcws	c.iptws	c.ipcws	Cum.Wts
3	2005-04-07	2005-05-31	0	1	0	0	0.349921088	0.2792739	1.1086748	0.8998709	0.8825604	0.9807634	1.1086748	0.9807634	1.0873477
3	2005-05-31	2005-07-30	1	0	0	0	0.005456679	0.2792739	0.7246805	0.9502811	0.8825604	0.9287361	0.8034350	0.9108704	0.7318251
3	2005-07-30	2006-01-28	0	0	0	0	0.349921088	0.2792739	1.1086748	0.8998709	0.8825604	0.9807634	0.8907482	0.8933483	0.7957484
3	2006-01-28	2006-07-29	0	0	0	0	0.332045237	0.2792739	1.0790044	0.8969601	0.8825604	0.9839461	0.9611212	0.8790066	0.8448319
3	2006-07-29	2007-01-27	0	0	0	0	0.350971313	0.2806462	1.1083543	0.8995490	0.8820682	0.9805672	1.0652629	0.8619250	0.9181767
3	2007-01-27	2007-07-28	0	0	0	0	0.333069298	0.2806462	1.0786035	0.8966299	0.8820682	0.9837595	1.1489962	0.8479269	0.9742648
3	2007-07-28	2008-01-26	0	0	0	0	0.360467877	0.2820227	1.1226602	0.8926524	0.8815742	0.9875895	1.2899324	0.8374037	1.0801942
3	2008-01-26	2008-07-26	0	0	0	0	0.352022982	0.2820227	1.1080289	0.8992261	0.8815742	0.9803699	1.4292824	0.8209654	1.1733914
3	2008-07-26	2009-01-24	0	0	0	0	0.246568071	0.2834033	0.9511101	0.8837166	0.8810784	0.9970146	1.3594049	0.8185145	1.1126926
3	2009-01-24	2009-02-11	0	0	0	0	0.365737811	0.2834033	1.1298115	0.8676120	0.8810784	0.8982808	1.5358713	0.7352559	1.1292584
4	2005-09-27	2005-12-22	0	0	0	1	0.308560566	0.3206831	1.0392875	0.8441833	0.8675457	1.0276746	1.0392875	1.0276746	1.0680494
4	2005-12-22	2006-02-20	0	0	1	0	0.527094605	0.4677512	1.1254868	0.9221191	0.9223386	1.0002380	1.1697043	1.0279192	1.2023616
4	2006-02-20	2006-08-21	0	0	0	0	0.426681899	0.3206831	1.1848865	0.8637357	0.8675457	1.0044111	1.3859668	1.0324535	1.4309463
4	2006-08-21	2007-02-19	0	0	0	0	0.309545764	0.3221680	0.9817189	0.8437134	0.8670000	1.0276002	1.3606298	1.0609494	1.4435594
4	2007-02-19	2007-08-20	0	0	0	0	0.309545764	0.3221680	0.9817189	0.8437134	0.8670000	1.0276002	1.3357560	1.0902318	1.4562837

How to … Fitting MSM Models

- Fitting Outcome Model

```
F <- coxph(Surv(start1,stop1,Event) ~ SSRI1 + cluster(data$id), data=data, weights= c.iptws)
summary(F)

F1 <- coxph(Surv(start1,stop1,Event) ~ SSRI1 + Gender + age + cluster(data$id), data=data,
                                         weights= c.iptws)
summary(F1)

F2 <- coxph(Surv(start1,stop1,Event) ~ SSRI1 + Gender + age + cluster(data$id),   data=data,
                                         weights= Cum.Wts )
summary(F2)
```

PROTECT

References:

- Robins JM, Hernán MÁ, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000; **11**: 550-560.
- Hernán MÁ, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. Epidemiology 2000; **11**: 561-570.

Take Home Messages:

- Time-dependent confounding could be substantial in pharmacoepidemiologic studies
- Several alternatives are outhere to conrol for time-dependent confounding
- When reading a study that applied IPTW analysis:
  - Range of weights?
  - How the treatment and outcome models are fitted?
  - Any motivation with respect to undelying assumptions?