Introduction to Causal Inference and Causal Data Science

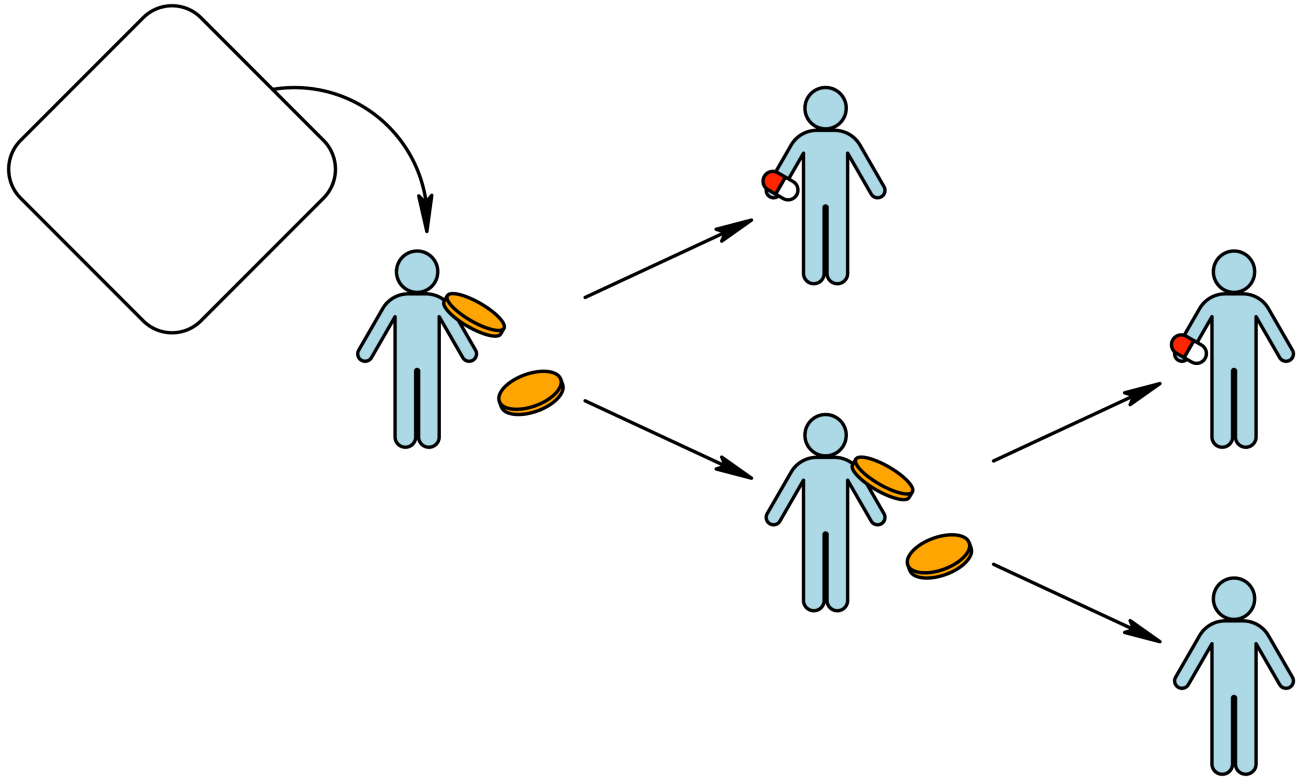
Complex longitudinal settings: When traditional methods fail

Bas Penning de Vries

Statin-cancer example revisited

Previous studies implicitly compared long-term statin users versus non-users – doesn't necessarily answer questions like …

- What would be my 10-year cancer risk if – possibly contrary to fact – I would start statin treatment now? And what if I wouldn't?
- What would be my 10-year cancer risk if – possibly contrary to fact – I would start statin treatment now and **adhered** to it? And what if I wouldn't start now *or* in the future?



Inference about time-varying treatments

If treatment/exposure is time-varying, there are many possible causal contrasts …

Single versus multiple-point interventions

Single-point (baseline) intervention

- Eg: *assign*/*initiate* versus withhold drug treatment at baseline
- Individuals are allowed to deviate (*intention-to-treat*)
- Randomisation at baseline only

Multiple-point (joint) intervention

- Eg: sustained/daily/weekly/monthly drug use versus continuous non-use (*per-protocol*)
- Randomisation at multiple points

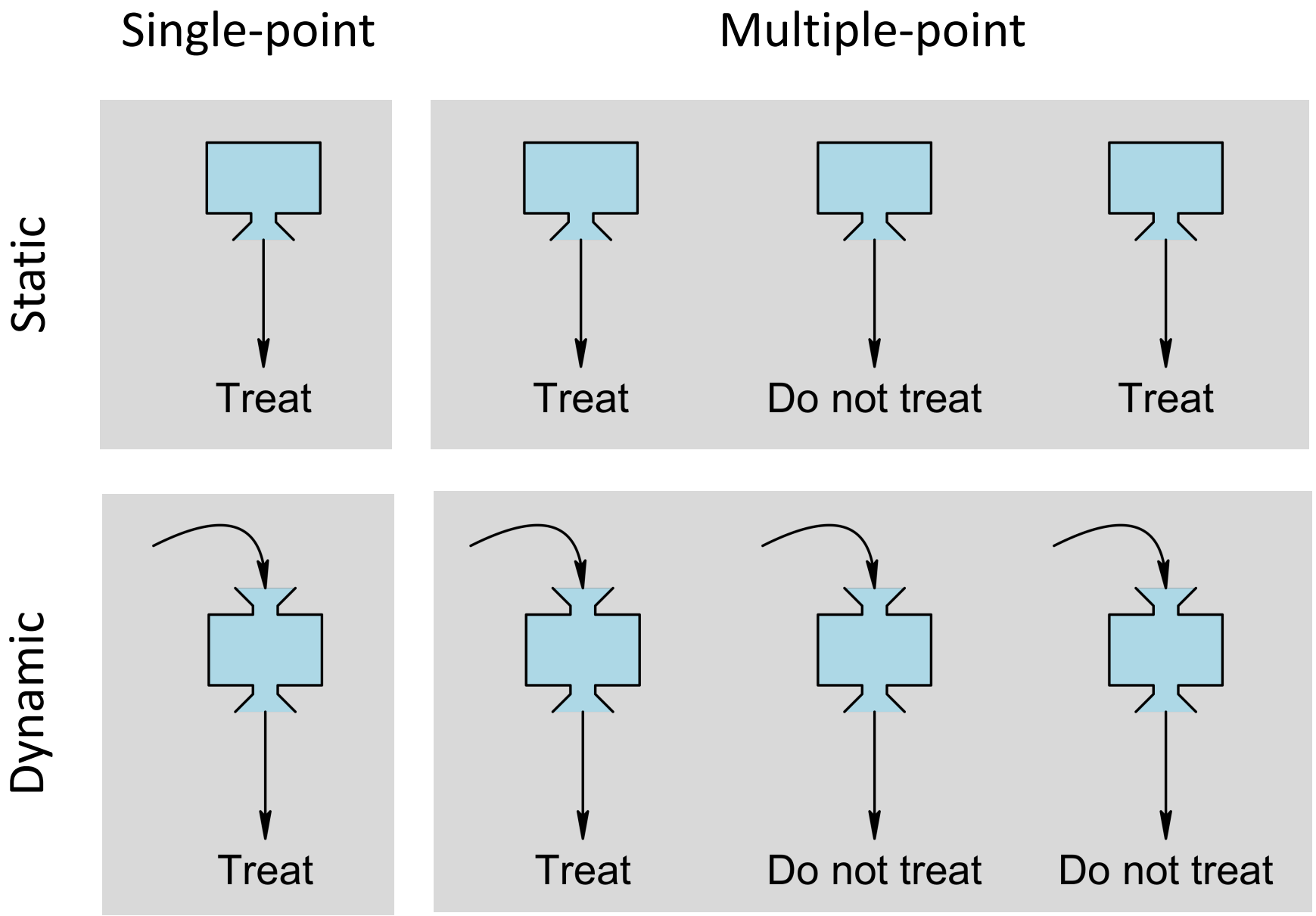
Static versus dynamic interventions

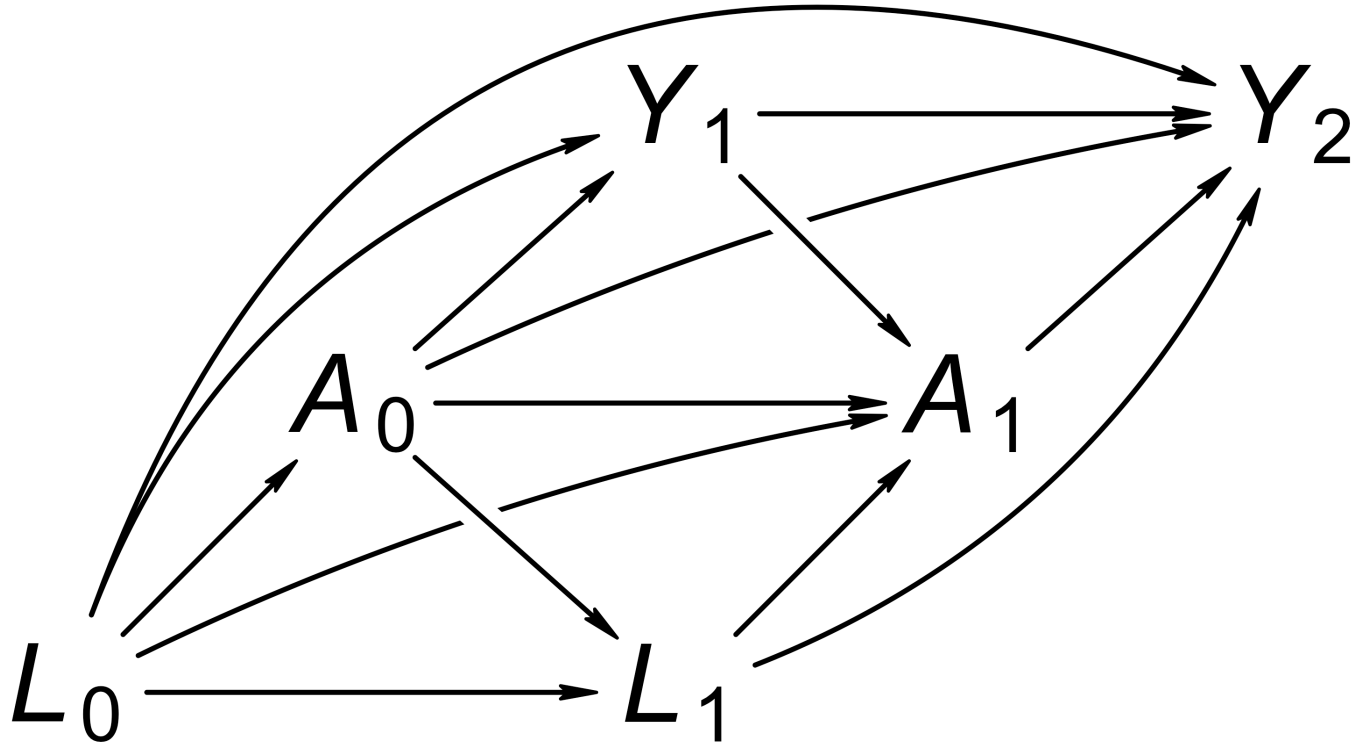
Static treatment rule/regime/protocol/…

- … assigns the same treatment option to everyone
- Eg: assign versus withhold treatment regime at baseline (intention-to-treat)
- Eg: always treat versus never treat (per-protocol)

Dynamic (individualised) treatment rule

- … assigns treatment based on the then-available information
- Eg: choose dose depending on baseline covariates
- Eg: start when blood marker first drops below threshold
- Eg: stop when toxicity occurs


Single-point
Multiple-point
Static
Dynamic
Treat
Treat
Do not treat
Treat
Treat
Treat
Do not treat
Do not treat



Treatment-covariate feedback

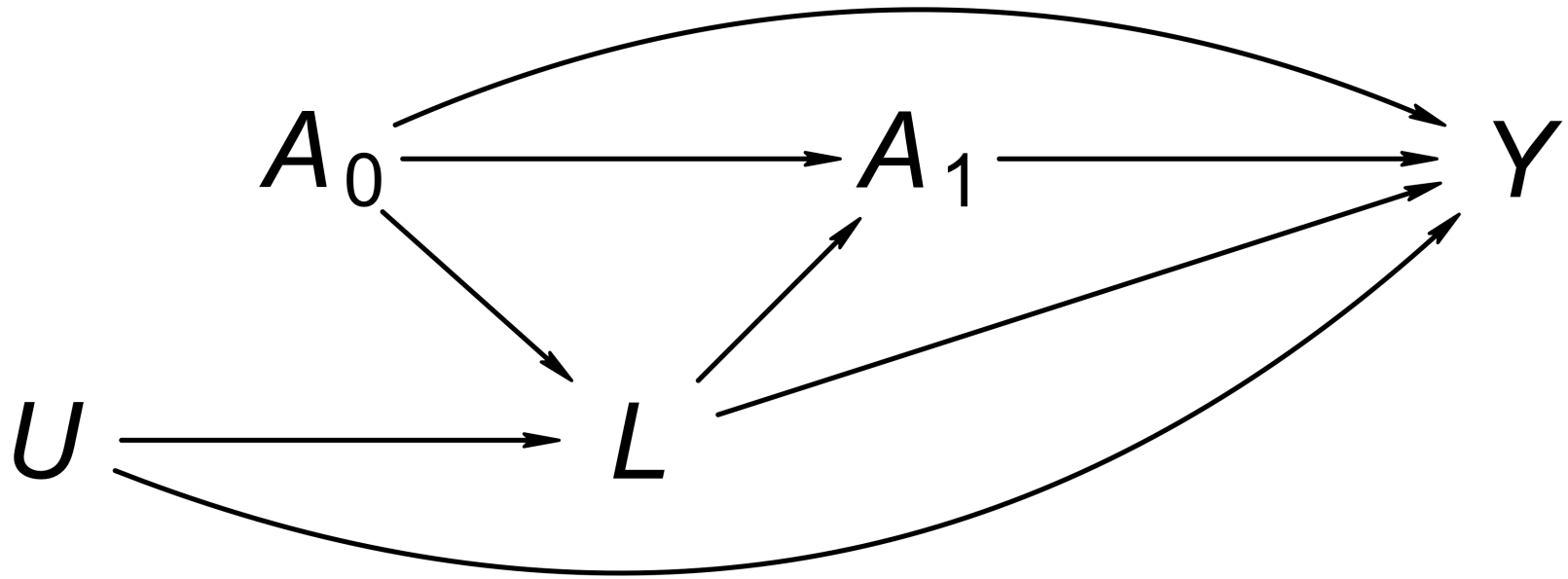
Adjust for L_1 if we want to know the effect of always- versus never treatment?

Treatment-covariate feedback

- Traditional methods (multivariable regression modelling) are not suited to deal with time-varying confounding when it is affected by past treatment (treatment-covariate feedback)
- Methods that can handle treatment-covariate feedback and adjusting for time-varying confounding:
  - G-computation
  - Inverse probability weighting (IPW)

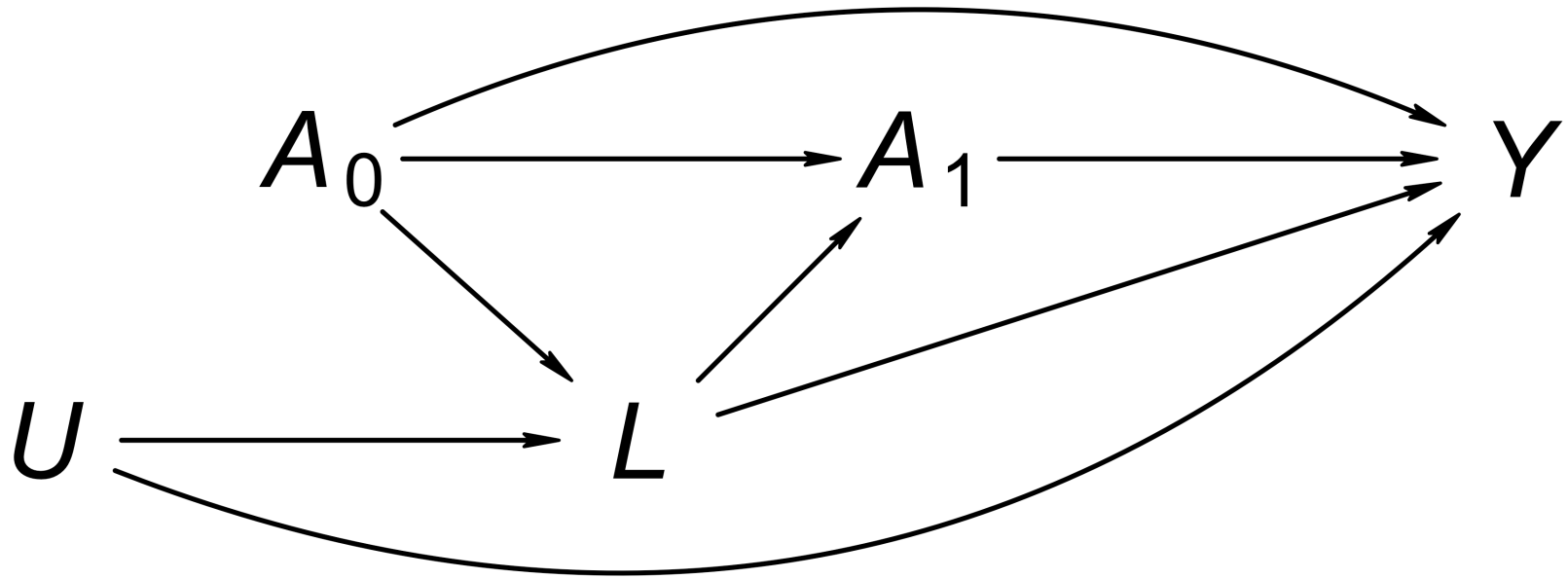
IPW for sustained treatments

- Let $Y_1^{a_0}, Y_2^{a_0,a_1}$ be counterfactual Y_1, Y_2 if all A_0, A_1 were a_0, a_1
- Target: $\Pr(Y_1^{a_0} = Y_2^{a_0,a_1} = 0)$ – survival probability if …?
- Same assumptions but slightly different form – eg, sequential conditional **exchangeability**:
  $Y_1^{a_0}, Y_2^{a_0,a_1}$ independent of
  - independent of A_0 given L_0
  - independent of A_1 given L_0, L_1, $A_0=a_0$, $Y_1=0$
- Idea: *patients who do not deviate from protocol compensate for "similar" patients who do*
- Weights are time-varying



Example

- Two time points at which we might intervene, $t = 0$ (baseline) and $t = 1$ (e.g., one month after baseline)
- At each time t, there are two options: issuing ($A_t = 1$) or withholding ($A_t = 0$) a one-month drug prescription



Example

Should we condition on ('adjust for') L to identify the effect of 'always treat' ($A_0 = A_1 = 1$) versus 'never treat' ($A_0 = A_1 = 0$)?

Example: IPW

- Under a version of the three identifiability conditions (**positivity, consistency and exchangeability**), we can use IPW to identify the effect of the time-varying treatment, expressed as a contrast between the average counterfactual outcomes under ‘treat always’ and ‘treat never’ regimes:

$$\mathrm{E}[Y^{1,1}] \text{ versus } \mathrm{E}[Y^{0,0}]$$

Example: IPW

- IPW for time-varying confounding starts with time-varying propensity scores. For the example with two time points, every individual will have two propensity scores:
  - $PS_0 = Pr(A_0 = 1)$
  - $PS_1 = Pr(A_1 = 1 \mid A_0, L)$

Example: IPW

- What we 'include' in the propensity scores (i.e., wat we 'condition on') depends on the setting – here, we condition on nothing for $t = 0$, because we're assuming that Y^{a_0,a_1} is (marginally) independent of A_0 (**no baseline confounding**). For $t = 1$, we condition on A_0 and L because we're assuming that Y^{a_0,a_1} is independent of A_1 given $A_0 = a_0$ and L (**no uncontrolled confounding at $t = 1$ given A_0 and L**).

A_0	L	A_1	**Stratum probability**	PS_0	PS_1	W
0	0	0	0.20	0.50	?	
0	0	1	0.05	0.50	?	
0	1	0	0.15	0.50	0.40	
0	1	1	0.10	0.50	0.40	
1	0	0	0.03	0.50	0.88	
1	0	1	0.22	0.50	0.88	
1	1	0	0.02	0.50	0.92	
1	1	1	0.23	0.50	0.92	

Example: IPW

- The next step is computing the weights. These look much like the weights for time-fixed treatments:
  - $W_0 = 1/\text{PS}_0$ if $A_0 = 1$ (treated at $t = 0$) and $W_0 = 1/(1 - \text{PS}_0)$ if $A_0 = 0$ (untreated)
  - $W_1 = 1/\text{PS}_1$ if $A_1 = 1$ (treated at $t = 1$) and $W_1 = 1/(1 - \text{PS}_1)$ if $A_1 = 0$ (untreated)
- The final weights W are obtained by taking the product of these time-varying weights: $W = W_0 W_1$

A_0	L	A_1	**Stratum probability**	PS_0	PS_1	W
0	0	0	0.20	0.50	0.20	2.50
0	0	1	0.05	0.50	0.20	?
0	1	0	0.15	0.50	0.40	3.33
0	1	1	0.10	0.50	0.40	5.00
1	0	0	0.03	0.50	0.88	16.67
1	0	1	0.22	0.50	0.88	2.27
1	1	0	0.02	0.50	0.92	25.00
1	1	1	0.23	0.50	0.92	2.17

Example: IPW

- Finally, having computed the weights *W*, an estimate of the always-versus-never treatment effects is obtained by taking the difference in mean outcome between the always treated ($A_0 = A_1 = 1$) and never treated ($A_0 = A_1 = 0$) individuals, weighted by *W.*

Concluding remarks

- IPW and g-computation, but not traditional methods, are suited to handle time-varying confouding affected by past treatment (feedback)
- As with IPW for time-fixed confounding, default standard error estimators of many software packages are not appropriate for weighted regressions, because the weights are falsely assumed to reflect actual observation frequencies
- IPW can (and need sometimes) be combined with marginal structural modelling (Robins et al., Epidemiology, 2000;11:550-560)