Causal inference with data subject to covariate-dependent observation times: An application to a cohort of new users of antidepressants

Janie Coulombe
Université de Montréal
janie.coulombe@umontreal.ca

Joint work with Dr. Erica E.M. Moodie, Dr. Susan M. Shortreed, and Dr. Christel Renoux

Department of Mathematics and Statistics
Dalhousie University
September 22, 2022

Plan

Motivation

Optimize the choice of antidepressant

- Electronic health records from the United Kingdom (*Clinical Practice Research Datalink* (CPRD)).
- Cohort of new users of antidepressants (ADs) with a diagnostic of depression in the year before.

Optimize the choice of antidepressant

- Electronic health records from the United Kingdom (*Clinical Practice Research Datalink* (CPRD)).
- Cohort of new users of antidepressants (ADs) with a diagnostic of depression in the year before.
- We focus on two selective serotonin reuptake inhibitors (SSRIs): **Citalopram** and **fluoxetine** (31120 patients).
- Adverse effects include change in weights and appetite.

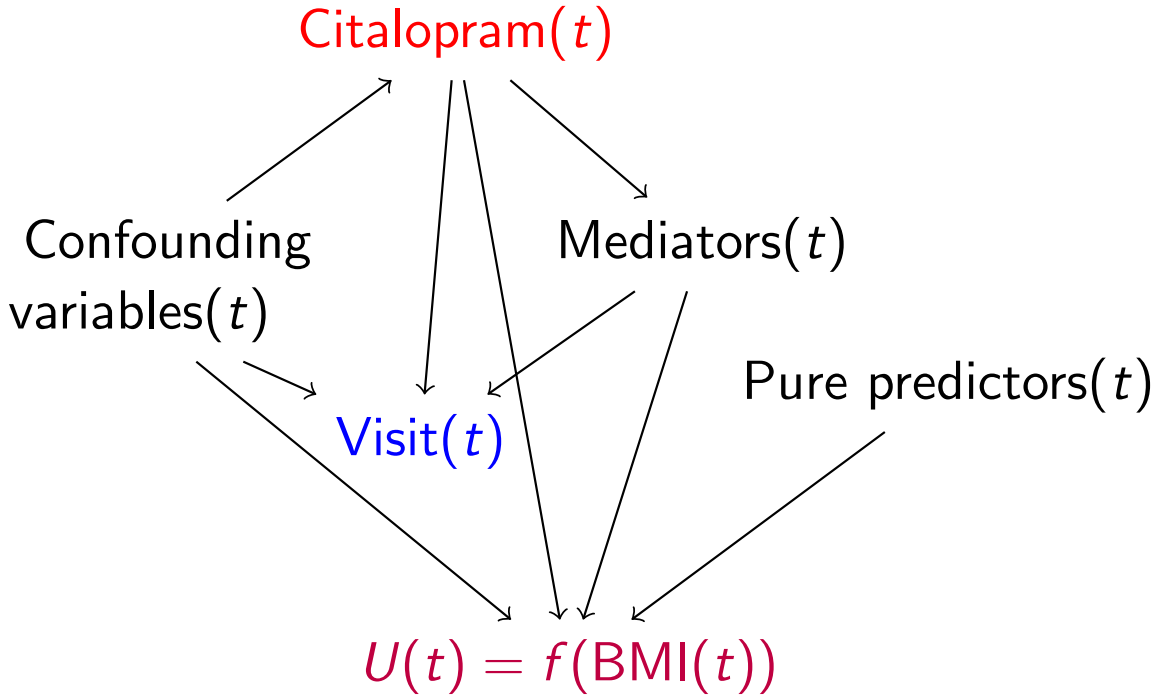
Optimize the choice of antidepressant

- Electronic health records from the United Kingdom (*Clinical Practice Research Datalink* (CPRD)).
- Cohort of new users of antidepressants (ADs) with a diagnostic of depression in the year before.
- We focus on two selective serotonin reuptake inhibitors (SSRIs): **Citalopram** and **fluoxetine** (31120 patients).
- Adverse effects include change in weights and appetite.
- We aim to develop an individualized treatment rule of the type

"**Treat with citalopram if detrimental changes in weight are minimized under citalopram, treat with fluoxetine otherwise.**"

Optimize the choice of antidepressant

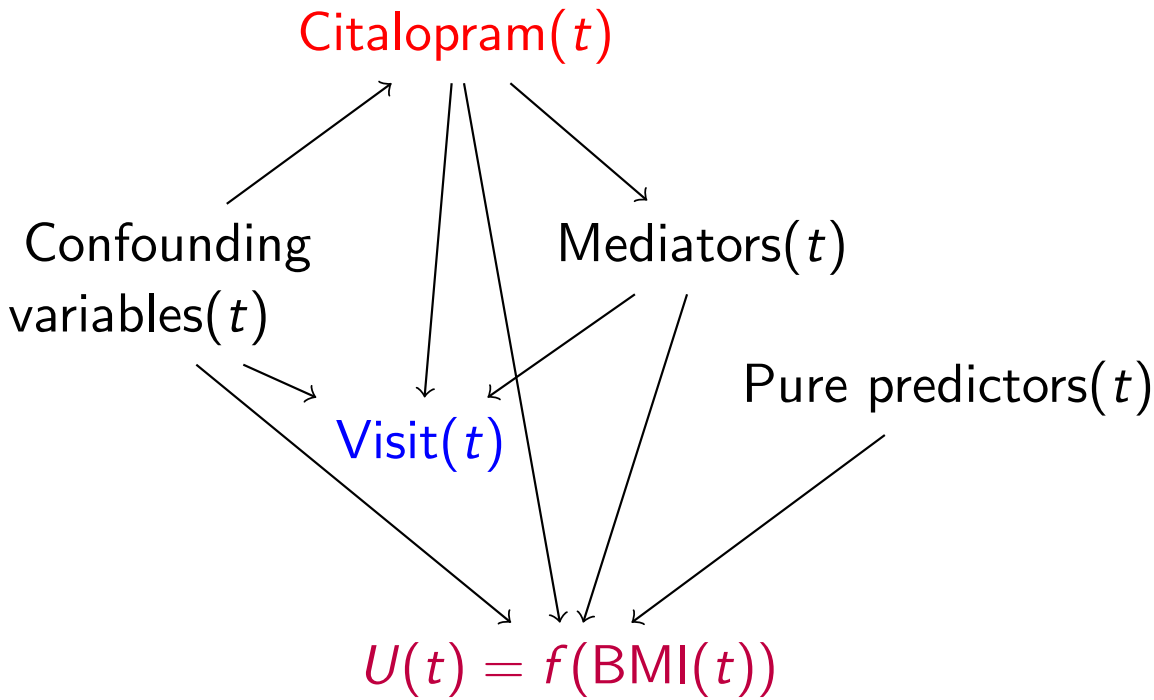
The causal diagram that is postulated at each time t:



with $\text{Visit}(t) = 1$ corresponding to the observation of $U(t) = f(\text{BMI})$, a function of body mass index (BMI).

Optimize the choice of antidepressant

The causal diagram that is postulated at each time t:

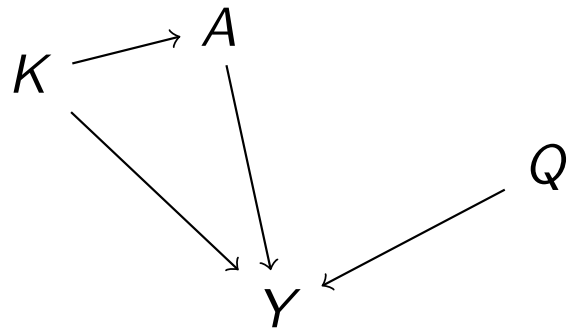


with $\text{Visit}(t) = 1$ corresponding to the observation of $U(t) = f(\text{BMI})$, a function of body mass index (BMI).

The inference on treatment effects is affected by:

- **Confounding** (variables that affect prescription mechanism and BMI)
- **Irregular** (and "informative") observation times for BMI

Note on effect modification



Some potential models for Y:

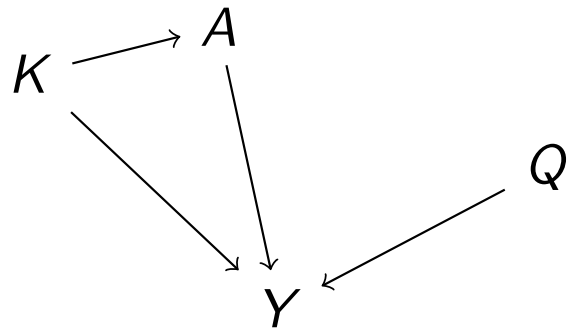
$$\mathbb{E}[Y|A,K,Q] = \beta_0 + \beta_a A + \beta_k K + \beta_q Q, \quad \text{(A)}$$

$$\mathbb{E}[Y|A,K,Q] = \beta_0 + \beta_a A + \beta_k K + \beta_q Q + \beta_{int} A \times K, \quad \text{(B)}$$

$$\mathbb{E}[Y|A,K,Q] = \beta_0 + \beta_a A + \beta_k K + \beta_q Q + \beta_{int} A \times Q, \quad \text{(C)}$$

$$\mathbb{E}[Y|A,K,Q] = \beta_0 + \beta_a A + \beta_k K + \beta_q Q + \beta_{int1} A \times K + \beta_{int2} A \times Q, \quad \text{(D)}$$

Note on effect modification



Some potential models for Y:

$$\mathbb{E}[Y|A,K,Q] = \beta_0 + \beta_a A + \beta_k K + \beta_q Q, \quad \text{(A)}$$

$$\mathbb{E}[Y|A,K,Q] = \beta_0 + \beta_a A + \beta_k K + \beta_q Q + \beta_{int} A \times K, \quad \text{(B)}$$

$$\mathbb{E}[Y|A,K,Q] = \beta_0 + \beta_a A + \beta_k K + \beta_q Q + \beta_{int} A \times Q, \quad \text{(C)}$$

$$\mathbb{E}[Y|A,K,Q] = \beta_0 + \beta_a A + \beta_k K + \beta_q Q + \beta_{int1} A \times K + \beta_{int2} A \times Q, \quad \text{(D)}$$

Models (B), (C) and (D) allow for the effect of A on Y to vary by K, by Q, or both. I.e.,

$$\mathbb{E}[Y|A=1,K,Q] - \mathbb{E}[Y|A=0,K,Q]$$

depends on the values of K and Q.

Informative observation in causal inference

Informative observation

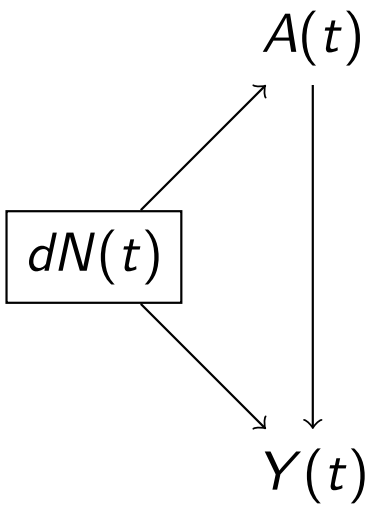
- Variables are observed at times that are not common across patients (and even across variables).
- Often, the observation of a variable **depends on patient characteristics** ("informative").
- Some exceptions: Routine visits, yearly mammography, etc. (considered as "visiting at random").

Informative observation

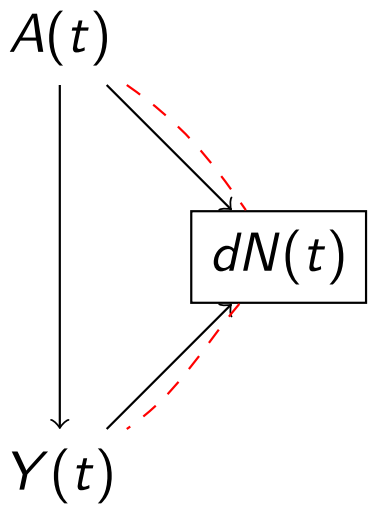
Bias caused by informative observation:

- Confounding by the visit process
- **Stratification on a collider** (Greenland, 2003)

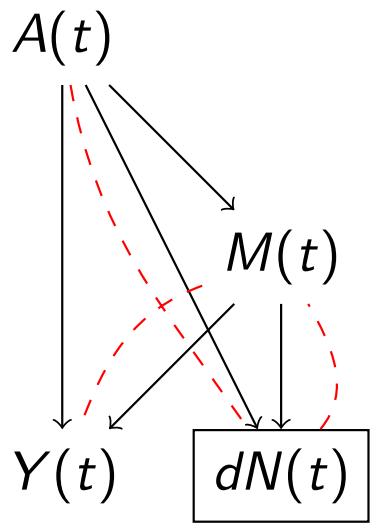
E.g., if $dN(t)$ is a visit indicator, $A(t)$ a treatment, $M(t)$ a mediator of the treatment effect and $Y(t)$ an outcome at time t:



Confounder



Collider

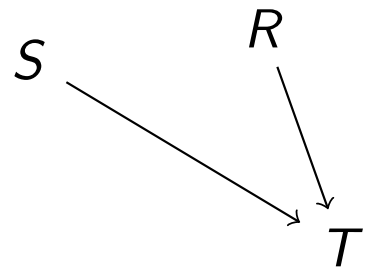


Collider

Parenthesis about collider-stratification bias

Suppose:

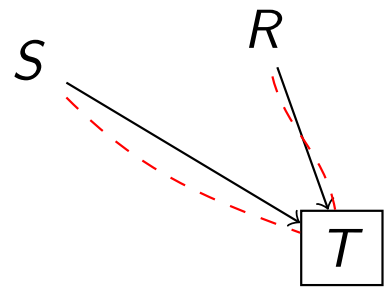
- going for a run (R) $\perp$ eating a salty meal for lunch (S)
- the risk of feeling thirsty (T) increases with sport and salt consumption



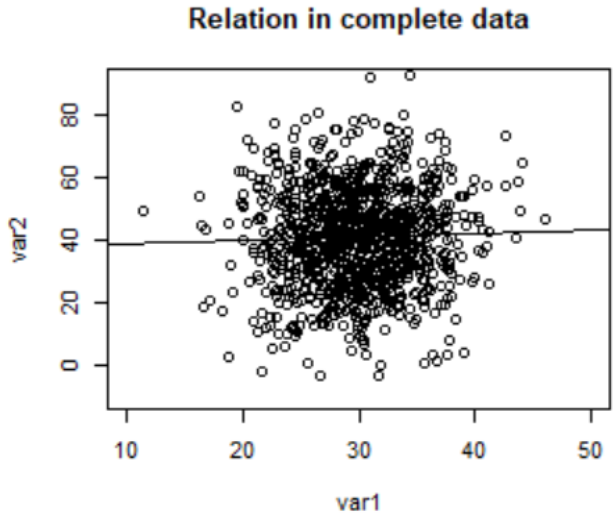
What can we infer

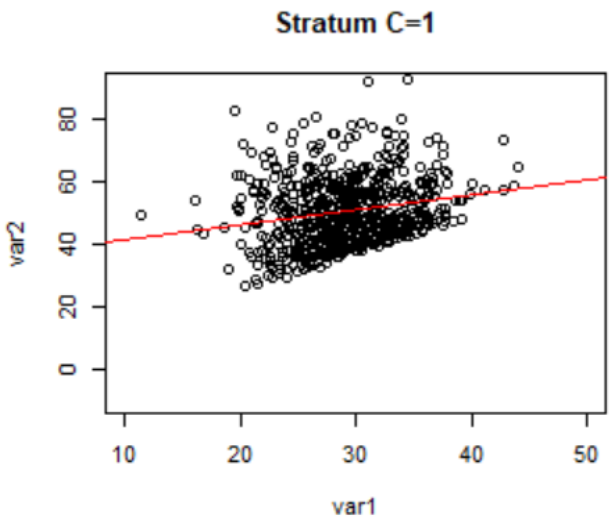
- on an individual who feels thirsty ($T = 1$) who did not run today?
- on an individual who does not feel thirsty ($T = 0$) who had a salty meal for lunch?

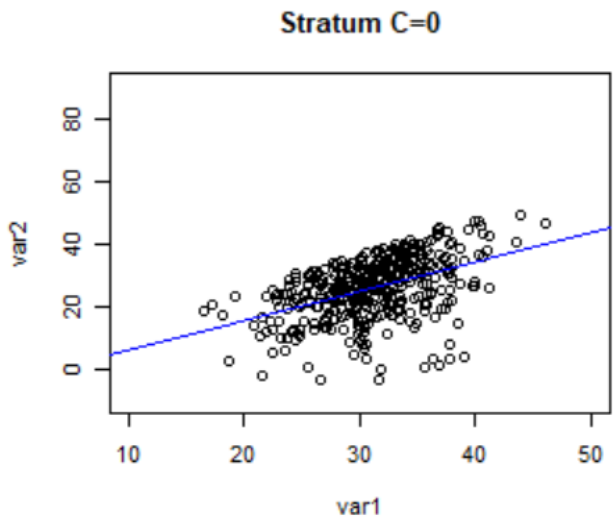
I.e., $R \not\perp S | T$.



Parenthesis about collider-stratification bias

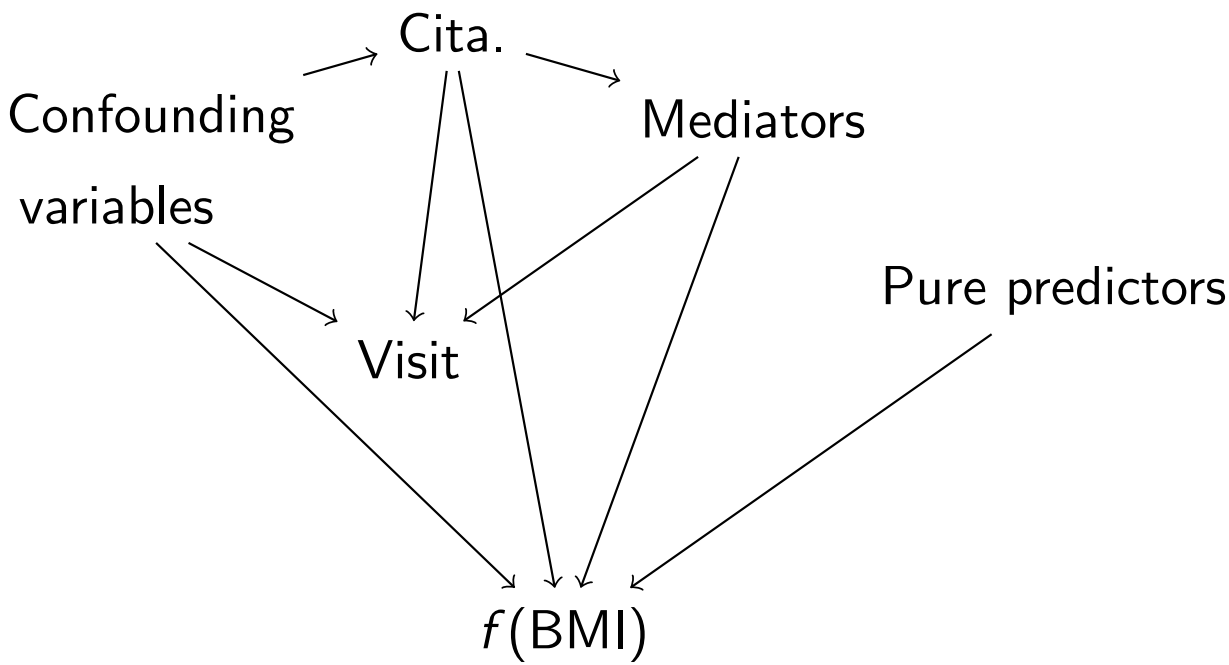






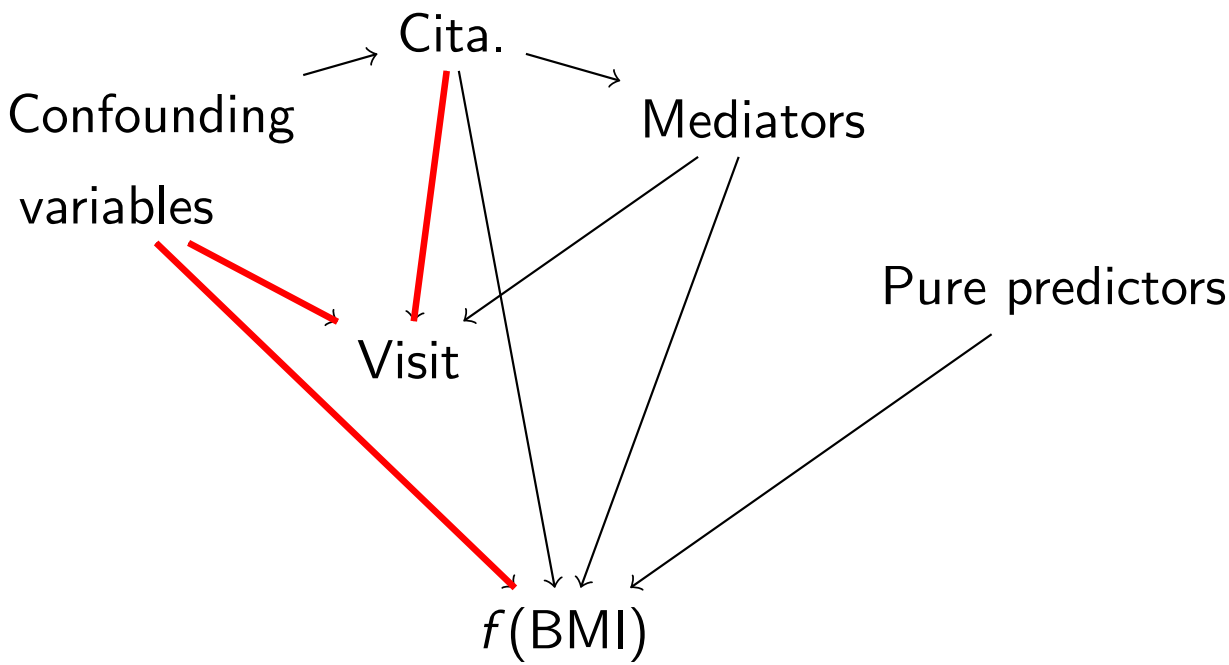
Motivation: Choice of antidepressant

Causal diagram at time t:



Motivation: Choice of antidepressant

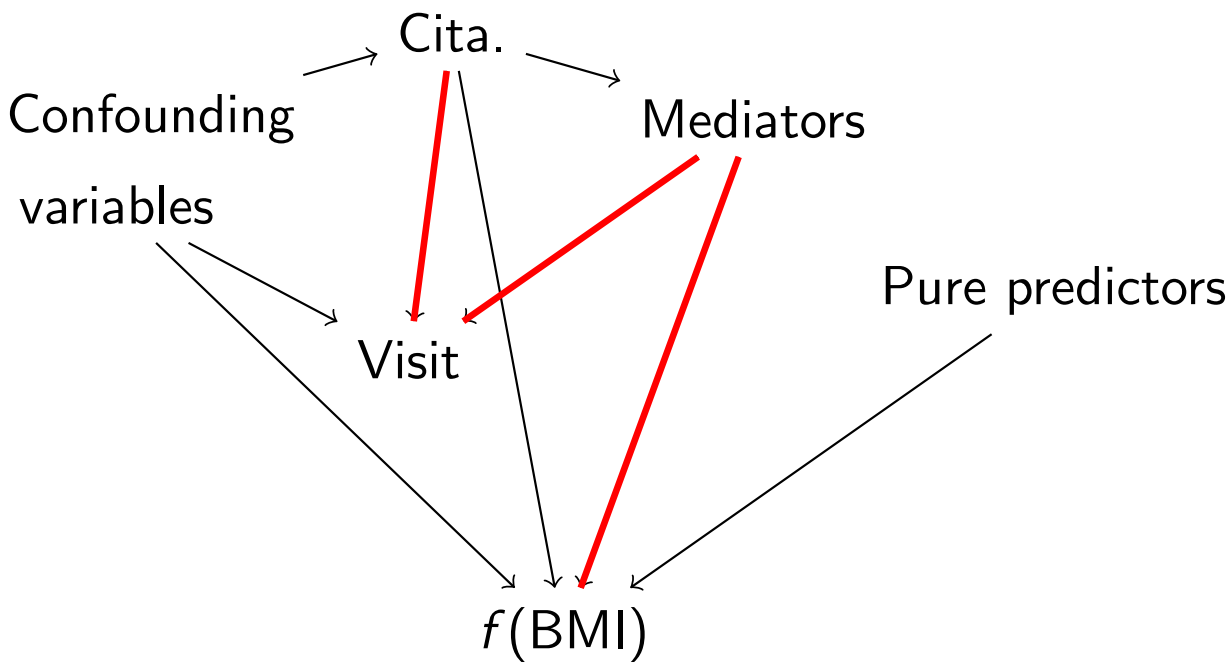
Causal diagram at time t:



When conditioning on observed data, we open a path (since the visit indicator acts as a collider).

Motivation: Choice of antidepressant

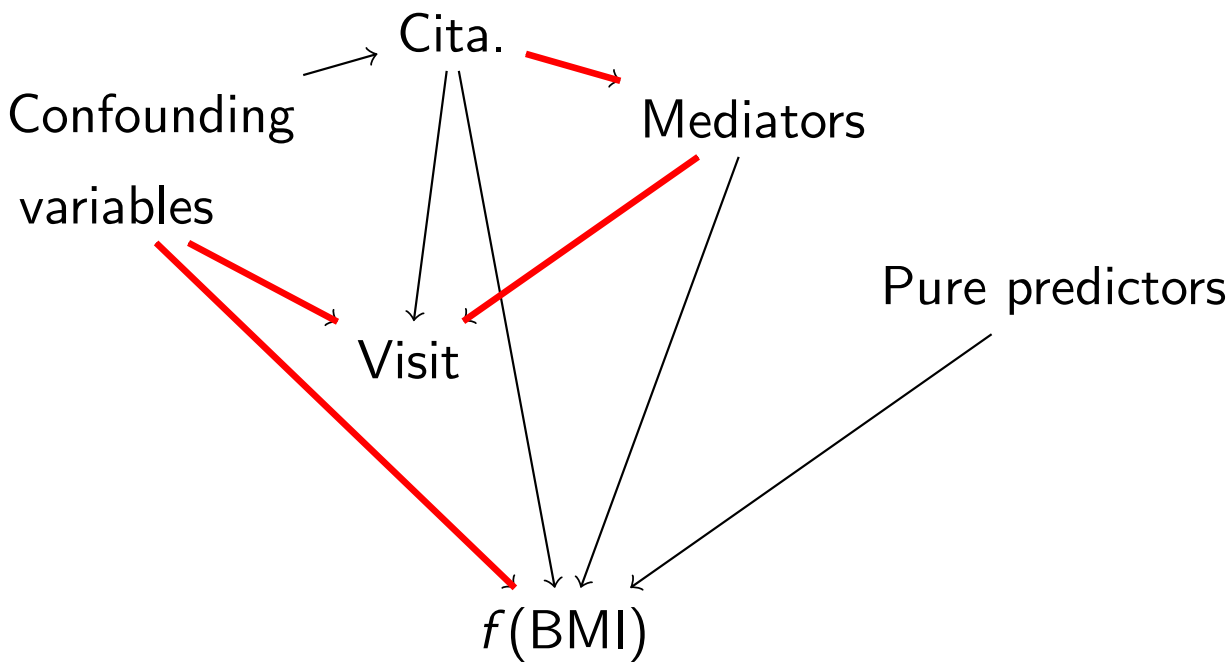
Causal diagram at time t:



When conditioning on observed data, we open a path (since the visit indicator acts as a collider).

Motivation: Choice of antidepressant

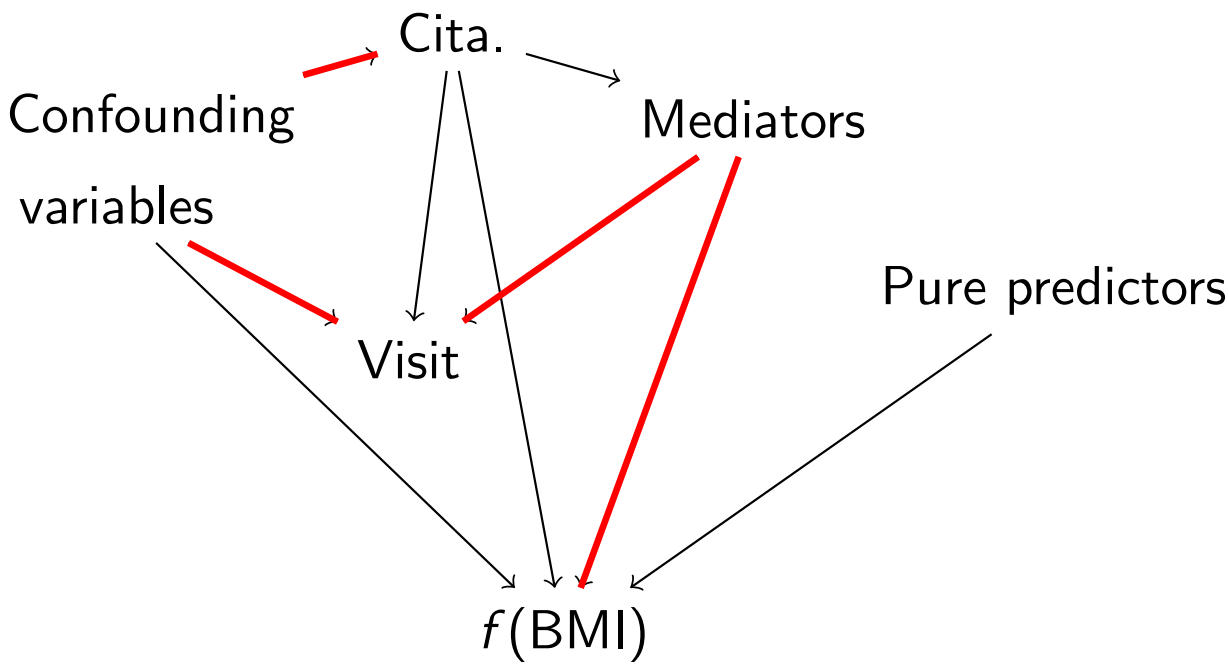
Causal diagram at time t:



When conditioning on observed data, we open a path (since the visit indicator acts as a collider).

Motivation: Choice of antidepressant

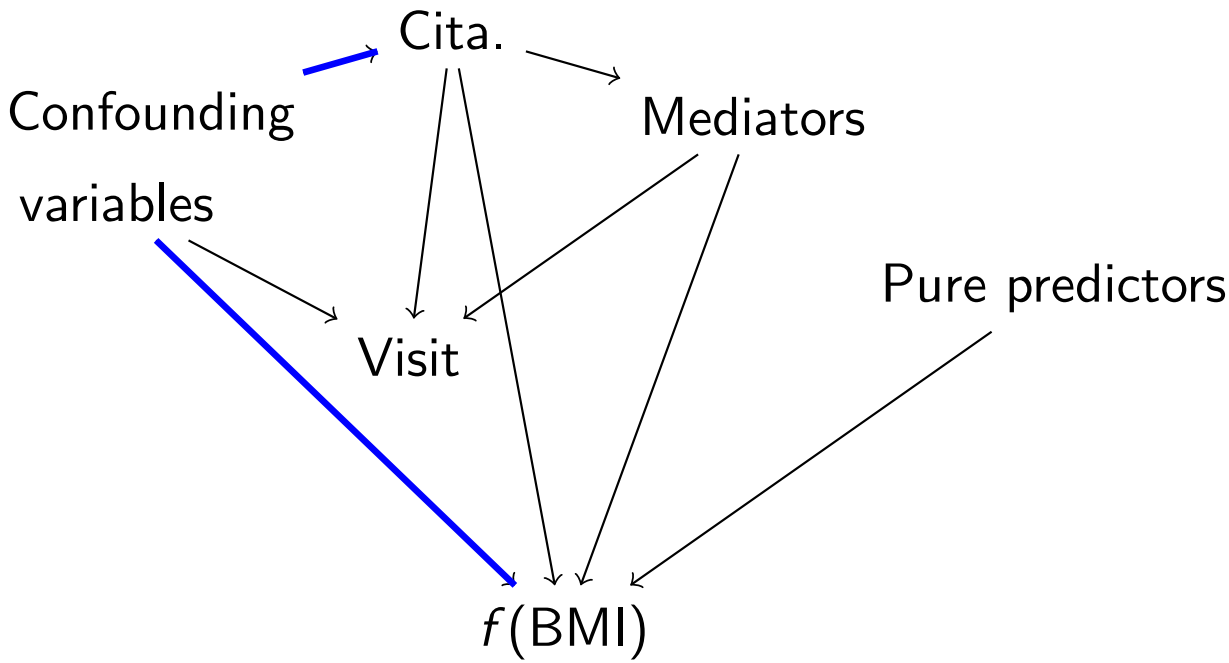
Causal diagram at time t:



When conditioning on observed data, we open a path (since the visit indicator acts as a collider).

Motivation: Choice of antidepressant

Causal diagram at time t:



Confounding (backdoor path)

Methods

Notation

Suppose a random sample of patients indexed by $i = 1, ..., n$. Let:

- $A_i(t)$ and $Y_i(t)$ the binary treatment and the continuous outcome, $\mathbf{K_i(t)}$ the confounding variables

Notation

Suppose a random sample of patients indexed by $i = 1, ..., n$. Let:

- $A_i(t)$ and $Y_i(t)$ the binary treatment and the continuous outcome, $\mathbf{K_i(t)}$ the confounding variables
- The set $\mathbf{X}^{\beta}(\mathbf{t}) = [\mathbf{1} \;\; \mathbf{K(t)} \;\; \mathbf{Q(t)}]$
- The effect modifiers $\mathbf{X}^{\psi}(\mathbf{t})$
- $\mathbf{X(t)} = \left[\mathbf{X}^{\beta}(\mathbf{t}) \;\; \mathbf{X}^{\psi}(\mathbf{t})\right]$

Notation

Suppose a random sample of patients indexed by $i = 1, ..., n$. Let:

- $A_i(t)$ and $Y_i(t)$ the binary treatment and the continuous outcome, $\mathbf{K_i(t)}$ the confounding variables
- The set $\mathbf{X}^{\beta}(\mathbf{t}) = [\mathbf{1} \;\; \mathbf{K(t)} \;\; \mathbf{Q(t)}]$
- The effect modifiers $\mathbf{X}^{\psi}(\mathbf{t})$
- $\mathbf{X(t)} = \left[\mathbf{X}^{\beta}(\mathbf{t}) \;\; \mathbf{X}^{\psi}(\mathbf{t})\right]$
- $dN_i(t)$ the visit indicator (= observation of $Y_i(t)$)
- C_i the follow-up time of patient i
- $\xi_i(t) = \mathbb{I}(C_i \geq t)$ the indicator for still being at risk
- We assume $dN_i(t) \perp Y_i(t) | \mathbf{V_i(t)}$ a set of variables

Outcome model

The outcome model includes a treatment-free model and a blip:

$$\mathbb{E}[Y_i(t)|A_i(t), \mathbf{X}_i(t)] = f\left\{\mathbf{X}_i^{\beta}(t); \boldsymbol{\beta}\right\} + A_i(t)\boldsymbol{\psi}'\mathbf{X}_i^{\psi}(t).$$

The following rule:

"Treat with citalopram if the outcome expectation is optimized under citalopram, and with fluoxetine otherwise"

becomes

"Treat with citalopram if $\hat{\boldsymbol{\psi}}'\mathbf{X}_i^{\psi}(t) \geq 0$, and with fluoxetine otherwise."

(See, e.g., Bian et al. 2021 for variable selection for $\mathbf{X}_i^{\psi}(t)$).

Causal assumptions

- Let $Y_{i0}(t)$ and $Y_{i1}(t)$ be two potential outcomes under treatments 0 and 1, respectively (Neyman, 1923; Rubin, 1974)
- Causal estimand: conditional treatment effect

$$\text{ECT} = \mathbb{E}[Y_{i1}(t) - Y_{i0}(t) | \mathbf{X}_\mathbf{i}^\psi(t)]$$

- Data (if $\mathbf{X}^\psi(t) = X^\psi(t)$):

ID	$dN_i(t)$	$A_i(t)$	$X_i^\psi(t)$ (age group)	$Y_i(t)$	$Y_{i0}(t)$	$Y_{i1}(t)$
1	1	1	18-25	25	-	25
2	1	0	26-35	22	22	-
			...			
$n-1$	1	0	18-25	23	23	-
n	1	1	26-35	15	-	15

Causal assumptions

Using the following causal assumptions, one can estimate functions of $Y_{i0}(t)$ and $Y_{i1}(t)$:

- Conditional exchangeability:

$$A_i(t) \perp \{Y_{i0}(t), Y_{i1}(t)\} \,|\mathbf{K_i(t)}, \mathbf{V_i(t)}, dN_i(t)$$

- Positivity:

$$0 < \mathbb{P}(A_i(t)|\mathbf{K_i(t)}, \mathbf{X_i^{\psi}}) < 1$$

$$0 < \mathbb{P}(dN_i(t)|\mathbf{V_i(t)}) < 1$$

- Consistency:

$$Y_i(t) = A_i(t)\,Y_{i1}(t) + (1 - A_i(t))\,Y_{i0}(t)$$

Why?

Doubly weighted estimator

We use the concept of "pseudo-population":

$$\mathbb{E}[Y_{i1}(t) - Y_{i0}(t)|\mathbf{X}_{\mathbf{i}}^{\psi}(\mathbf{t})]$$

$$= \mathbb{E}_{\mathbf{V}}[\mathbb{E}_{\mathbf{K}}[\mathbb{E}[Y_{i1}(t)|\mathbf{K_i}(\mathbf{t}), \mathbf{V_i}(\mathbf{t}), \mathbf{X}_{\mathbf{i}}^{\psi}(\mathbf{t})]]] - \mathbb{E}_{\mathbf{V}}[\mathbb{E}_{\mathbf{K}}[\mathbb{E}[Y_{i0}(t)|\mathbf{K_i}(\mathbf{t}), \mathbf{V_i}(\mathbf{t}), \mathbf{X}_{\mathbf{i}}^{\psi}(\mathbf{t})]]]$$

$$= \mathbb{E}_{\mathbf{V}}[\mathbb{E}_{\mathbf{K}}[\mathbb{E}[Y_{i1}(t)|\mathbf{K_i}(\mathbf{t}), A_i(t) = 1, \mathbf{V_i}(\mathbf{t}), dN_i(t) = 1, \mathbf{X}_{\mathbf{i}}^{\psi}(\mathbf{t})]]]$$
$$- \mathbb{E}_{\mathbf{V}}[\mathbb{E}_{\mathbf{K}}[\mathbb{E}[Y_{i0}(t)|\mathbf{K_i}(\mathbf{t}), A_i(t) = 0, \mathbf{V_i}(\mathbf{t}), dN_i(t) = 1, \mathbf{X}_{\mathbf{i}}^{\psi}(\mathbf{t})]]]$$

by cond. exchangeability and positivity

$$= \mathbb{E}_{\mathbf{V}}[\mathbb{E}_{\mathbf{K}}[\mathbb{E}[Y_i(t)|\mathbf{K_i}(\mathbf{t}), A_i(t) = 1, \mathbf{V_i}(\mathbf{t}), dN_i(t) = 1, \mathbf{X}_{\mathbf{i}}^{\psi}(\mathbf{t})]]]$$
$$- \mathbb{E}_{\mathbf{V}}[\mathbb{E}_{\mathbf{K}}[\mathbb{E}[Y_i(t)|\mathbf{K_i}(\mathbf{t}), A_i(t) = 0, \mathbf{V_i}(\mathbf{t}), dN_i(t) = 1, \mathbf{X}_{\mathbf{i}}^{\psi}(\mathbf{t})]]]$$

by consistency.

Doubly weighted estimator

For each part of the equation, we can show:

$$\mathbb{E}_{\mathbf{V}}[\mathbb{E}_{\mathbf{K}}[\mathbb{E}[Y_i(t)|\mathbf{K_i(t)}, A_i(t)=1, \mathbf{V_i(t)}, dN_i(t)=1, \mathbf{X_i^{\psi}(t)}]]]$$

$$= \sum_{\mathbf{V}} \sum_{\mathbf{K}} \frac{\mathbb{P}(A_i(t)=1|\mathbf{K_i(t)})\mathbb{P}(dN_i(t)=1|\mathbf{V_i(t)})}{\mathbb{P}(A_i(t)=1|\mathbf{K_i(t)})\mathbb{P}(dN_i(t)=1|\mathbf{V_i(t)})}$$
$$\times \mathbb{E}[Y_i(t)|\mathbf{K_i(t)}, A_i(t)=1, \mathbf{V_i(t)}, dN_i(t)=1, \mathbf{X_i^{\psi}(t)}]\mathbb{P}(\mathbf{K}=\mathbf{K_i(t)})\mathbb{P}(\mathbf{V}=\mathbf{V_i(t)})$$

$$= \ldots$$

$$= \mathbb{E}\left[\frac{\mathbb{I}(A_i(t)=1, dN_i(t)=1)Y_i(t)}{\mathbb{P}(A_i(t)=1|\mathbf{K_i(t)})\mathbb{P}(dN_i(t)=1|\mathbf{V_i(t)})}|\mathbf{X_i^{\psi}(t)}\right],$$

and similarly for $A_i(t)=0$, leading to

$$\mathbb{E}\left[\frac{\mathbb{I}(A_i(t)=0, dN_i(t)=1)Y_i(t)}{\mathbb{P}(A_i(t)=0|\mathbf{K_i(t)})\mathbb{P}(dN_i(t)=1|\mathbf{V_i(t)})}|\mathbf{X_i^{\psi}(t)}\right]$$

Doubly weighted estimator

- We model observation times using the proportional rate model:

$$\mathbb{E}[dN_i(t)|\mathbf{V_i(t)}] = \xi_i(t)\exp\{\gamma'\mathbf{V_i(t)}\}\lambda_0(t)dt$$

where γ are estimated using the Andersen and Gill model (1982).

This is used to compute inverse intensity of visit (IIV) weights $\varphi_i(\gamma, \mathbf{V_i(t)})$ (Lin et al., 2004).

- IIV weights are combined with inverse probability of treatment (IPT) weights:

$$e_i(\omega, \mathbf{K_i(t)}) = \frac{\mathbb{I}(A_i(t)=1)}{\mathbb{P}(A_i(t)=1|\mathbf{K_i(t)};\omega)} + \frac{\mathbb{I}(A_i(t)=0)}{\mathbb{P}(A_i(t)=0|\mathbf{K_i(t)};\omega)}.$$

Doubly weighted estimator

- The estimating equation to solve (extension of the dWOLS proposed by Wallace and Moodie (2015)):

$$U(\beta, \psi; \hat{\gamma}, \hat{\omega}) = \sum_{i=1}^{n} \int_0^{\tau} \varphi_i(\hat{\gamma}, \mathbf{V_i(t)}) e_i(\hat{\omega}, \mathbf{K_i(t)})$$
$$\times \begin{bmatrix} \frac{\partial f\left\{\mathbf{X}_i^{\beta}(t); \beta\right\}}{\partial \beta} \\ A_i(t)\mathbf{X}_i^{\psi}(t) \end{bmatrix} \left[Y_i(t) - f\left\{\mathbf{X}_i^{\beta}(t); \beta\right\} - A_i(t)\psi' \mathbf{X}_i^{\psi}(t) \right] dN_i(t) = \mathbf{0}.$$

- We can compute the asymptotic variance using theory on two-step estimators (Newey et McFadden, 1994)
- Note: It requires an acute treatment effect, no carryover effect of subsequent treatments (Dong et al., 2021)

Doubly robust - intuition

As mentioned in Wallace and Moodie (2015), the facts of

1. using weights that satisfy the "*balancing condition*" (like IPT or overlaping weights) and
2. to correctly specify the blip $\psi' \mathbf{X}_i^{\psi}(t)$,

in addition to meeting all other assumptions mentioned previously, lead to a doubly robust estimator.

Doubly robust - intuition

Using correctly specified IPT weights make $A_i(t)$ and $\mathbf{X}_i^{\beta}(t)$ independent in

$$\mathbb{E}[Y_i(t)|A_i(t), \mathbf{X}_i(t)] = \underbrace{f\left\{\mathbf{X}_i^{\beta}(t); \boldsymbol{\beta}\right\}}_{\text{contains confounders}} + \ A_i(t)\boldsymbol{\psi}'\mathbf{X}_i^{\psi}(t).$$

Hence, the coefficients $\boldsymbol{\psi}$ can be estimated consistently even if $f\left\{\mathbf{X}_i^{\beta}(t); \boldsymbol{\beta}\right\}$ is wrongly specified (and vice versa).

Of note, the observation model must be correctly specified, at least w.r.t. a set of important variables ("**partial**" double robustness).

Simulation study

$$\mathbf{K_i} = \{K_{1i}, K_{2i}, K_{3i}\} \sim \{N(1,1), Bern(0,55), N(0,1)\}$$

$$A_i \sim Bern(p_i) \text{ avec } p_i = \text{expit}\{0{,}5 + 0{,}55K_{1i} - 0{,}2K_{2i} - 1K_{3i}\}$$

$$Z_i(t) \sim \begin{cases} N(2,1) \text{ if } A_i = 1 \\ N(4,2) \text{ if } A_i = 0 \end{cases}$$

$$Q_i(t) \sim Bern(0.5)$$

$$Y_i(t) = \sqrt{(t/100)} - 2\, A_i + 2.5\,\{Z_i(t) - \mathbb{E}[Z_i(t)|A_i]\} +$$
$$0{,}4K_{i1i} + 0{,}05K_{2i} - 0{,}6K_{3i} + 0.5\,\{A_i \times Q_i(t)\} - 1\,\{A_i \times K_{1i}\} + \epsilon_i(t)$$

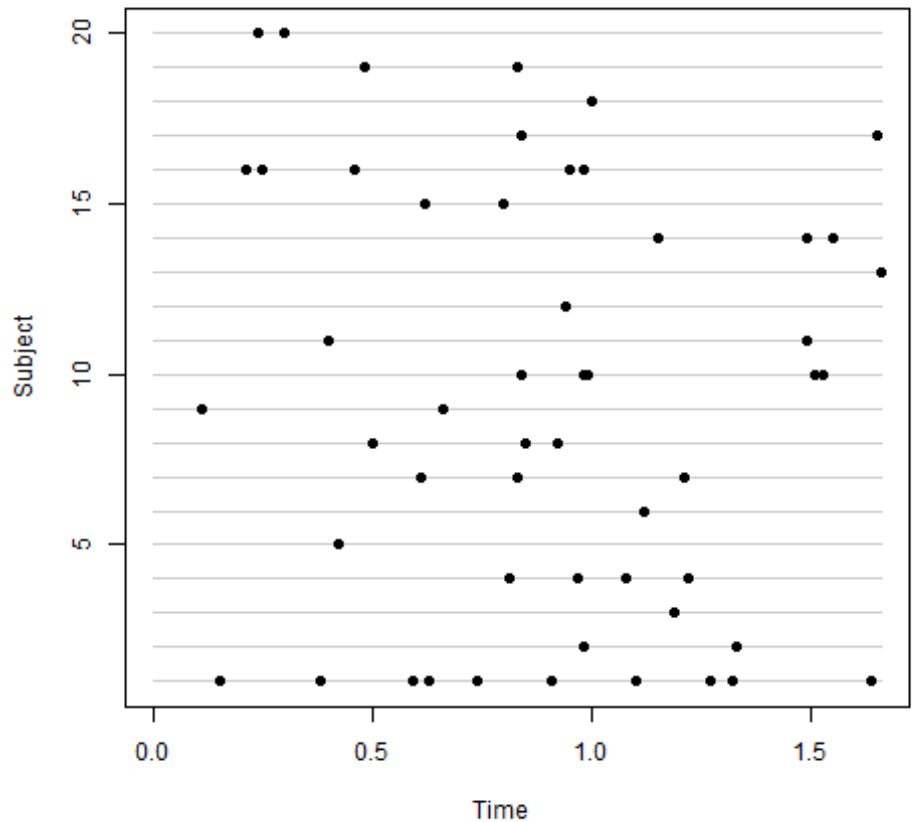
$$\text{where } \epsilon_i(t)|\phi_i \sim N(\phi_i, 0{,}01), \phi_i \sim N(0, 0{,}04)$$

$$dN_i(t) \sim Poisson(\lambda_i(t)) \text{ avec } \lambda_i(t) = \eta_i \exp\{\gamma_A A_i + \gamma_Z Z_i(t) + \gamma_{K2} K_2 + \gamma_{K3} K_3\}$$

where η_i is a random effect, $(\gamma_A, \gamma_Z, \gamma_{K2}, \gamma_{K3})$ the dependence parameters

Simulation study

Visits are simulated according to a **proportional rate model**:



Abacus plot produced with the `IrregLong` package in R (Pullenayegum, 2022)

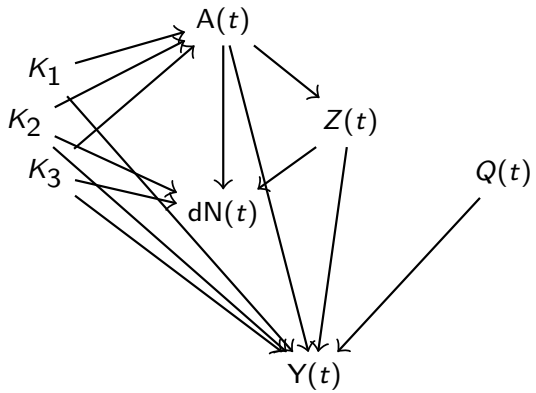
Estimators that were compared

To assess the method's robustness, we compare:

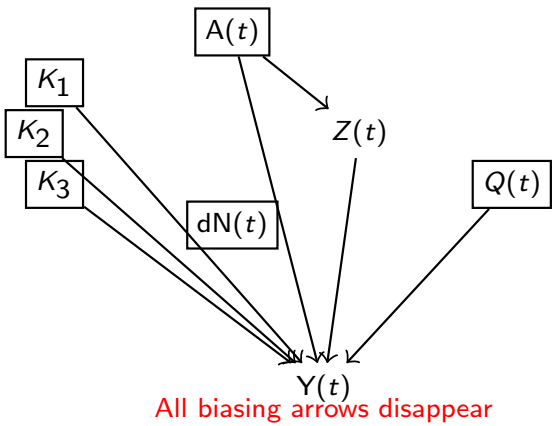
	Misspecified model (x)			
Estimator	Visit (partial[1])	Visit (full[2])	Treatment[3]	Outcome[4]
$\hat{\psi}_{DW1}$				
$\hat{\psi}_{DW2}$	x			x
$\hat{\psi}_{DW3}$	x		x	
$\hat{\psi}_{DW4}$		x		
$\hat{\psi}_{OLS}$		x	x	
$\hat{\psi}_{IPT}$		x		

1. Adjusted for the important variables (treatment and mediator)
2. Not adjusted for the mediator, adjusted for the treatment and K_2
3. Wrong functional form (squared terms) for the linear terms of K_1 and K_3
4. Misses the adjustment for K_2

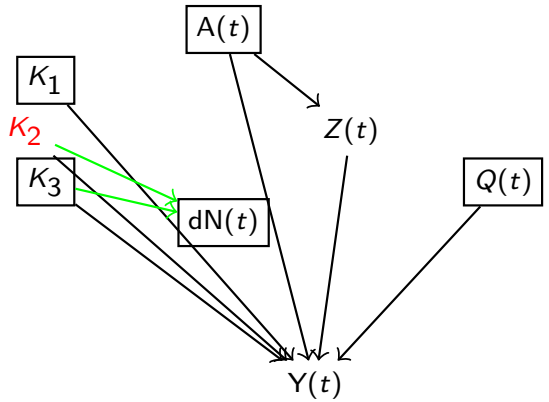
Corresponding causal diagrams



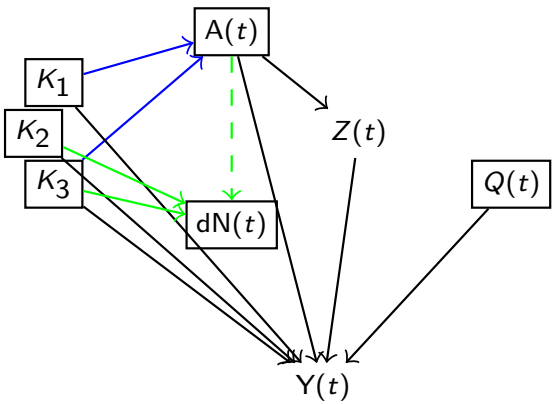
(a) Data generating mechanism



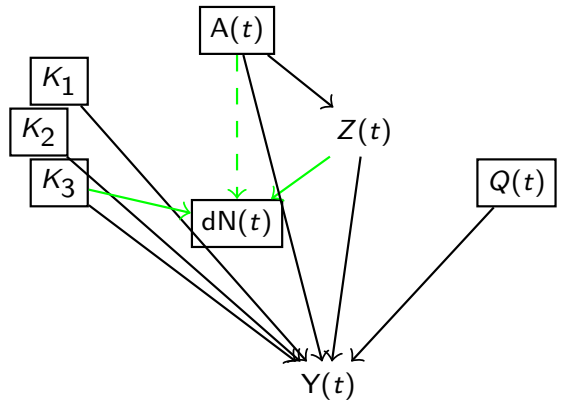
(b) Proposed estimator



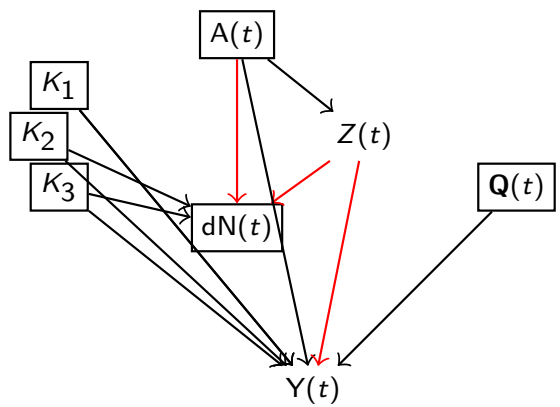
(c) DW2: Wrong observation and outcome models



(d) DW3: Wrong observation and treatment models



(e) DW4: Fully wrong observation model ($dN_i(t) \sim A(t) + K_2$)



(f) IPT estimator (prone to collider stratification bias)

Note: A box means that we condition upon that variable.

Results

Table 1: Results of the simulation study ($M = 1000$ simulations) - error rate for the optimal treatment decision

Sample size	Parameters $\boldsymbol{\gamma}^{\upsilon}$	No. obs. times mean (IQR)	Error rate $\hat{\psi}_{DW1}$	$\hat{\psi}_{DW2}$	$\hat{\psi}_{DW3}$	$\hat{\psi}_{DW4}$	$\hat{\psi}_{OLS}$	$\hat{\psi}_{IPT}$
250	1	3 (1-3)	0.02	0.01	0.01	0.04	0.03	0.04
	2	3 (2-5)	0.05	0.06	0.05	0.16	0.15	0.16
	3	6 (3-9)	0.06	0.03	0.03	0.26	0.25	0.26
	4	10 (8-12)	0.01	0.01	0.00	0.01	0.00	0.01
500	1	3 (1-3)	0.01	0.01	0.01	0.03	0.03	0.03
	2	3 (1-5)	0.02	0.03	0.02	0.14	0.13	0.14
	3	6 (3-9)	0.04	0.02	0.02	0.25	0.25	0.25
	4	10 (8-12)	0.00	0.00	0.00	0.00	0.00	0.00

υ.1. (-2, -0.3, 0.2, -1.2); 2. (0.3, -0.6, -0.4, -0.3); 3. (0.4, -0.8, 1, 0.6); 4. (0, 0, 0, 0). Abbrev.: IQR, interquartile range.

Similar results for the comparison of the blip coefficients and the blip evaluated at patients' characteristics.

Illustration

Question

- Treatment rule to choose between citalopram (1) and fluoxetine (0) to reduce BMI variations
- Outcome definition:

$$
\begin{aligned}
U(t) = 100 &- 5 \times \mathbb{I}[\,\text{Detrimental change in BMI}(t)\text{ category}\,] \\
&+ \mathbb{I}[\,\text{BMI}(0) < 18.5 \;\cup\; (18.5 \leq \text{BMI}(0) \leq 24.9 \;\cap\; \text{BMI}(t) < 20)\,] \times \{\%\text{ increase BMI}(t)\} \\
&- \mathbb{I}[\,\text{BMI}(0) \geq 25 \;\cup\; (18.5 \leq \text{BMI}(0) \leq 24.9 \;\cap\; \text{BMI}(t) > 23.5)\,] \times \{\%\text{ increase BMI}(t)\}\,,
\end{aligned}
$$

where $\mathbb{I}$ is the indicator function. BMI is categorized as follows:

< 18.5	Underweight
$18.5 - 24.9$	Normal weight
$25 - 29.9$	Overweight
≥ 30	Obese

- 31,120 patients and 48,388 measures of $U(t)$

Characteristics of patients at cohort entry

Table 2: Characteristics of the 31,120 patients, n (%), CPRD, 1998-2017

Variable	Treatment	
	Citalopram (n=18,671)	Fluoxetine (n=12,449)
Age, mean (SE)	48.5 (18.1)	45.1 (16.5)
Sex (M)	5965 (32)	3609 (29)
Index of Multiple Deprivation, mean (SE)	3.0 (1.4)	3.1 (1.4)
Year of cohort entry		
1998-2005	3751 (20)	4896 (39)
2006-2011	10,279 (55)	5703 (46)
2012-2017	4641 (25)	1850 (15)
Has been a smoker	11,586 (62)	8017 (64)
Alcohol abuse	1478 (8)	869 (7)
Psychiatric disease†	521 (3)	321 (3)
Anxiety	5956 (32)	2987 (24)
Medication		
Antipsychotics	2836 (15)	1675 (13)
Other psychotropics‡	4476 (24)	2546 (20)
Lipid-lowering drugs	3360 (18)	1614 (13)
No. psychiatric hospitalisations in previous 6 months, mean (SE)	0.04 (0.24)	0.03 (0.34)

Abbrev.: CPRD, Clinical Practice Research Datalink; SE, standard error.

†. Indicator of a diagnostic of autism spectrum disorder, OCD, bipolar disorder, or schizophrenia.

‡. Includes benzodiazepines, anxiolytics, barbiturates, and hypnotics.

Visit rate ratios

Table 3: Rate ratios (95% bootstrap CIs) for the observation of $U(t)$, CPRD, 1998-2017

Variable	Rate ratio (95% CI)
Citalopram treatment	0.9 (0.9, 0.9)*
Age	1.0 (1.0, 1.0)
Sex (M)	0.9 (0.9, 0.9)*
Index of Multiple Deprivation	1.0 (1.0, 1.0)*
Year of cohort entry (ref.= <2006)	
2006-2011	0.9 (0.9, 1.0)*
2012-2017	0.9 (0.9, 0.9)*
Has been a smoker	1.7 (1.6, 1.7)*
Alcohol abuse	1.0 (0.9, 1.1)
Psychiatric disease †	1.0 (0.9, 1.2)
Anxiety	1.0 (1.0, 1.0)
Medication	
Antipsychotics	1.1 (1.0, 1.2)*
Other psychotropics‡	1.2 (1.2, 1.3)*
Lipid-lowering drugs	1.2 (1.2, 1.3)*
No. psychiatric hospitalisations in previous 6 months	1.0 (0.9, 1.0)

Abbrev.: CI, confidence interval; CPRD, Clinical Practice Research Datalink.

∗. Significant CI.

†. Indicator of a diagnostic of autism spectrum disorder, OCD, bipolar disorder, or schizophrenia.

‡. Includes benzodiazepines, anxiolytics, barbiturates, and hypnotics.

Treatment rule

$$\text{Treat with citalopram if } -1.45 + 0.16 \times \mathbb{I}\,[\text{Male sex}]$$

$$+0.13 \times [\text{Index Multiple Deprivation}] + 0.08 \times \mathbb{I}\,[\text{Has been smoker}]$$

$$+0.42 \times \mathbb{I}\,[\text{Alcohol abuse}] + 1.31 \times \mathbb{I}\,[\text{Psychiatric disease}]$$

$$+0.35 \times \mathbb{I}\,[\text{Anxiety}] - 0.91 \times \mathbb{I}\,[\text{Use of antipsychotics}]$$

$$+0.30 \times \mathbb{I}\,[\text{Other psychotropics}] + 0.21 \times \mathbb{I}\,[\text{Lipid lowering drugs}] > 0$$

Table 4: Comparison of fitted outcomes, CPRD, 1998-2017

	Mean fitted outcome (SE†)			
Treatment	$\hat{\psi}_{OLS}$	$\hat{\psi}_{IPT}$	$\hat{\psi}_{IIV}$	$\hat{\psi}_{DW1}$
Received	98.2 (0.001)	98.2 (0.001)	98.3 (0.001)	98.3 (0.001)
Optimal	98.9 (0.001)	98.9 (0.001)	99.1 (0.001)	99.0 (0.001)

Abbrev.: SE, standard error.
†. Based on individual SEs that were summed and normalized.

Discussion

Discussion

- Individualized treatment rules can be used to optimize a treatment choice according to patient characteristics
- (but!) in a standard analysis, we condition on the indicator of being observed. It is important to look at the causal diagram!

Discussion

- Individualized treatment rules can be used to optimize a treatment choice according to patient characteristics
- (but!) in a standard analysis, we condition on the indicator of being observed. It is important to look at the causal diagram!
- The methods discussed rely on the following assumptions:
  - causal assumptions
  - on the data generating mechanism
  - time window for the treatment effect
  - on the several models

Discussion

- Individualized treatment rules can be used to optimize a treatment choice according to patient characteristics
- (but!) in a standard analysis, we condition on the indicator of being observed. It is important to look at the causal diagram!
- The methods discussed rely on the following assumptions:
  - causal assumptions
  - on the data generating mechanism
  - time window for the treatment effect
  - on the several models
- Extension: Rule for a sequence of treatments (multiple time points)

Acknowledgments



Drs
Erica E. M. Moodie
Susan M. Shortreed
Christel Renoux

and Dr. Orla Murphy and Dr. Théo Michelot for the invitation and organization.

Thank you for your attention.

Funding and support:

NSERC CRSNG

compute canada | calcul canada

HEALTHY BRAINS HEALTHY LIVES

NIH National Institute of Mental Health

McGill UNIVERSITY

Contact:
janie.coulombe@umontreal.ca

References

Andersen, P. K., et Gill, R. D. (1982) Cox's regression model for counting processes: a large sample study. *The Annals of Statistics*, 10(4), pp. 1100-1120.

Bian, Z., Moodie, E. E. M., Shortreed, S. M., et Bhatnagar, S. (2021) Variable selection in regression-based estimation of dynamic treatment regimes. *Biometrics*, à venir.

Coulombe, J., Moodie, E. E. M., et Platt, R. W. (2021) Weighted regression analysis to correct for informative monitoring times and confounders in longitudinal studies. *Biometrics*, 77(1), pp. 162-174.

Coulombe, J., Moodie, E. E. M., Shortreed, S. M., et Renoux, C. (2022) Estimating Individualized Treatment Rules in Longitudinal Studies with Covariate-Driven Observation Times. *arXiv:2202.09611v1*, pp. 1- 62.

Dong, L., Moodie, E. E. M., Villain, L., et Thiébaut, R. Evaluating the use of generalized dynamic weighted ordinary least squares for individualized HIV treatment strategies. *arXiv:2109.01218v1*, pp. 1-38.

Greenland, S. (2003) Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology*, 14(3), pp. 300-306.

Lin, H., Scharfstein, D. O., et Rosenheck, R. A. (2004). Analysis of longitudinal data with irregular, outcome-dependent follow-up. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(3), pp. 791-813.

Newey, W. K., et McFadden, D. (1994) Large sample estimation and hypothesis testing. *Handbook of econometrics*, 4, pp. 2111-2245.

Neyman, J. S. (1923) On the application of probability theory to agricultural experiments. Essay on principles, section 9, *Statistical Science*, 5(14), pp. 465-472.

Pullenayegum, E. Package 'IrregLong'. 2022.

Wallace, M., et Moodie, E. E. M. (2015) Doubly-robust dynamic treatment regimen estimation via weighted least squares. *Biometrics*, 71(3), pp. 636-644.

Rubin, D. B. (1974) Estimating causal effects of treatments in randomized and nonrandomized studies, *Journal of Educational Psychology*, 66(5), pp. 688-701.

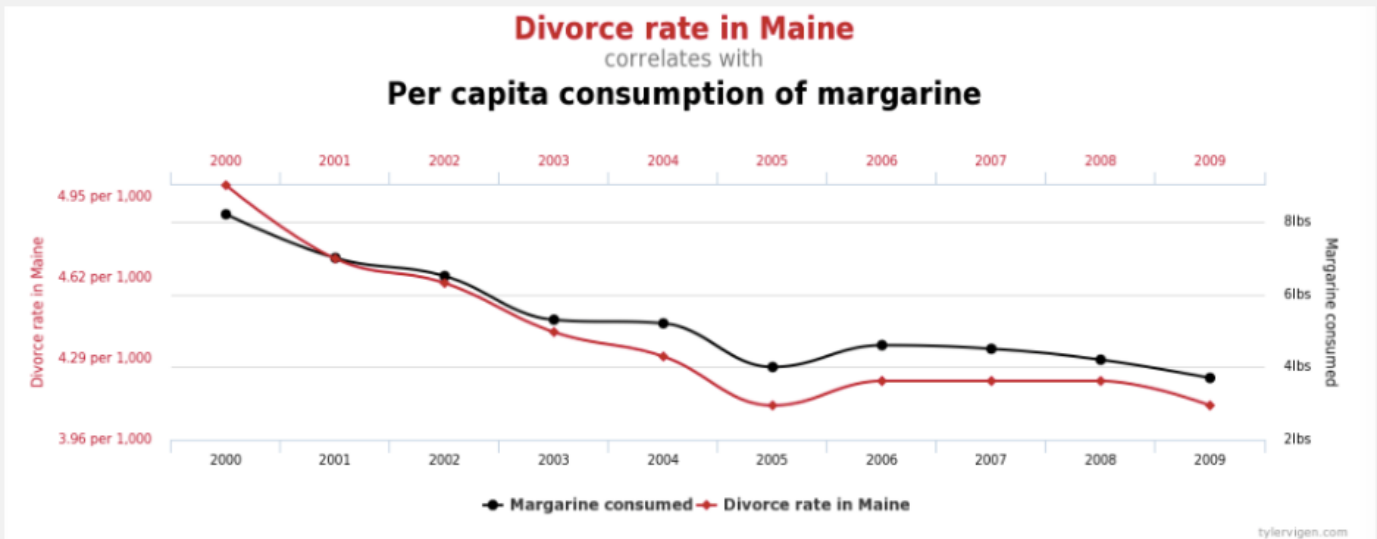
Blip évalué chez différents patients

Male sex (yes)	IMD (1 to 5)	Ever smoker (yes)	Alcohol abuse (yes)	Psychiatric diagnosis (yes)	Anxiety (yes)	Antipsy. drug (yes)	Psychotro. drug (yes)	Lipid lowering drug (yes)	Value blip function
0	1	0	0	0	0	1	0	0	-2.23
1	1	0	0	0	0	1	0	0	-2.07
0	3	0	0	0	0	1	1	1	-1.46
0	1	0	0	0	0	0	0	0	-1.32
1	3	0	0	0	0	1	1	1	-1.30
1	1	0	0	0	0	0	0	0	-1.16
0	3	1	0	0	0	0	0	0	-0.98
1	3	1	0	0	0	0	0	0	-0.82
0	5	0	0	0	0	0	0	0	-0.80
0	3	0	0	0	1	0	0	0	-0.71
1	5	0	0	0	0	0	0	0	-0.64
0	3	1	1	0	0	0	0	0	-0.56
1	3	0	0	0	1	0	0	0	-0.55
1	3	1	1	0	0	0	0	0	-0.40
0	5	1	1	0	0	0	0	0	-0.30
1	5	1	1	0	0	0	0	0	-0.14
0	3	0	1	1	0	0	0	0	0.67
0	3	1	1	1	0	0	0	0	0.75
1	3	0	1	1	0	0	0	0	0.83
0	3	1	0	1	1	0	0	1	0.89
1	3	1	1	1	0	0	0	0	0.91
0	3	0	1	1	1	0	0	0	1.02
0	5	1	1	1	0	0	0	0	1.01
1	3	1	0	1	1	0	0	1	1.05
1	5	1	1	1	0	0	0	0	1.17
1	3	0	1	1	1	0	0	0	1.18
1	5	1	1	1	1	0	1	1	2.03

Rappel: Facteur confondant

Divorce And Margarine

Posted on April 10, 2017 by annahaensch

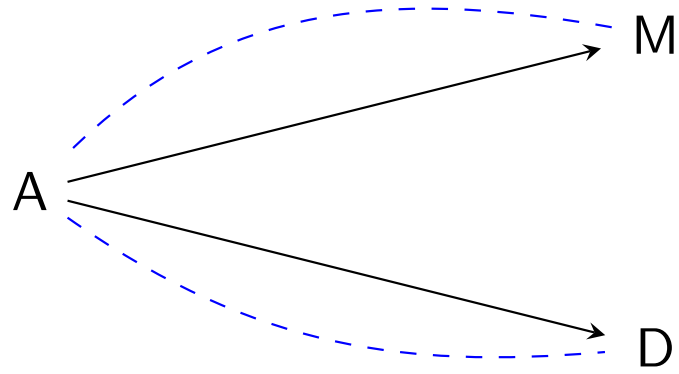


A spurious correlation involving a spurious butter substitute. Via tylervigen.com.

Source: blogs.ams.org

Facteur confondant A: Intérêt dans la relation M-D

Sans ajustement: M et D dépendants



Si modélisation de $M|A$ et ajustement (p. ex., poids IPT):

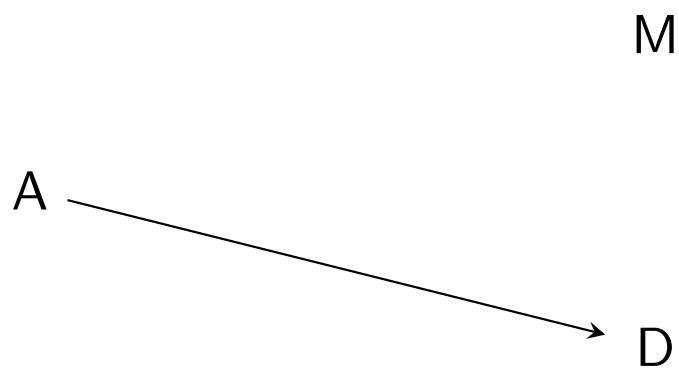


Table 5: Coefficients of the blip function (95% bootstrap CIs) for the optimal treatment rules as estimated by four alternative models: OLS which does not adjust for confounding or observation process, IPW which accounts only for confounding, IIV which accounts only for the observation process, and the proposed doubly-weighted estimator which accounts for both processes, CPRD, UK, 1998-2017, n=31,120 individuals.

Variable	$\hat{\psi}_{OLS}$	$\hat{\psi}_{IPT}$	$\hat{\psi}_{IIV}$	$\hat{\psi}_{DW}$
Intercept	-1.66 (-2.69, -0.46)	-1.38 (-2.62, -0.11)	-1.68 (-2.84, -0.58)	-1.45 (-2.66, -0.22)
Age	0.01 (-0.01, 0.03)	0.00 (-0.01, 0.03)	0.01 (-0.01, 0.03)	0.00 (-0.02, 0.03)
Male sex	-0.08 (-0.67, 0.55)	0.03 (-0.59, 0.62)	0.03 (-0.54, 0.65)	0.16 (-0.48, 0.76)
IMD	0.14 (-0.09, 0.31)	0.14 (-0.13, 0.32)	0.12 (-0.10, 0.31)	0.13 (-0.13, 0.31)
Ever smoker	0.23 (-0.38, 0.66)	0.13 (-0.47, 0.65)	0.21 (-0.41, 0.66)	0.08 (-0.50, 0.60)
Alcohol abuse	1.03 (-0.11, 2.24)	0.68 (-0.48, 1.88)	0.78 (-0.26, 1.99)	0.42 (-0.70, 1.60)
Psychiatric disease†	0.44 (-1.83, 2.10)	1.02 (-1.23, 2.84)	0.58 (-1.73, 1.93)	1.31 (-0.88, 3.05)
Anxiety	0.29 (-0.07, 1.12)	0.31 (-0.02, 1.21)	0.32 (-0.05, 1.15)	0.35 (0.00, 1.26)
Medication				
Antipsychotics	-0.73 (-1.56, 0.17)	-0.82 (-1.75, 0.10)	-0.78 (-1.61, 0.10)	-0.91 (-1.91, 0.03)
Other psychotropic drugs‡	0.03 (-0.81, 0.66)	0.07 (-0.73, 0.64)	0.22 (-0.49, 0.86)	0.30 (-0.47, 0.93)
Lipid lowering drugs	-0.16 (-0.73, 0.76)	0.04 (-0.70, 0.97)	-0.02 (-0.57, 0.95)	0.21 (-0.49, 1.23)

Abbreviations: CI, confidence interval; CPRD, Clinical Practice Research Datalink; UK, United Kingdom; IMD, Index of Multiple Deprivation.

†. An indicator for a diagnosis of either autism spectrum disorder, obsessive compulsive disorder, bipolar disorder, or schizophrenia.

‡. Which include benzodiazepine drugs, anxiolytics, barbiturates and hypnotics.