

Règle de traitement individuelle optimale en présence d'observation irrégulière

Janie Coulombe

janie.coulombe@mail.mcgill.ca

Travail conjoint avec Erica E.M. Moodie, Susan M. Shortreed et Christel Renoux

Département de mathématiques
Université du Québec à Montréal
17 mars 2022

Plan

Motivation

Observation informative en inférence causale

Méthodes

Application

Discussion

Motivation

Optimiser le choix d'un antidépresseur

Contexte:

- ▶ Dossiers médicaux électroniques du Royaume-Uni (*Clinical Practice Research Datalink* (CPRD)).
- ▶ Cohorte de nouveaux utilisateurs d'antidépresseurs avec un diagnostic de dépression dans l'année précédente.
- ▶ Inhibiteurs sélectifs de la recapture de la sérotonine (ISRS) ciblés: **Citalopram** et **fluoxétine** (31,120 patients).
- ▶ Effets indésirables incluent les changements dans l'appétit et le poids.

Optimiser le choix d'un antidépresseur

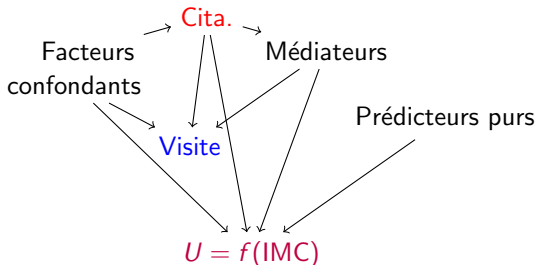
- ▶ On souhaite développer une règle de traitement individuelle optimale pour choisir un des deux antidépresseurs, de type

“Traiter avec citalopram si l'espérance de l'issue est maximisée sous citalopram, et sinon, avec fluoxétine”

- ▶ Le but: Minimiser les **changements de poids “détrimentaux”**.
- ▶ On utilise l'indice de masse corporelle (IMC) comme proxy pour le poids. L'IMC est mesuré **à répétition** et de façon **irrégulière**.
- ▶ On crée une fonction d'utilité à valeurs continues, $U(t)$, évaluée aux mêmes temps que l'IMC.

Optimiser le choix d'un antidépresseur

Le diagramme causal postulé à chaque temps t :

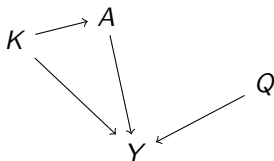


Avec $\text{Visite}(t) = 1$ correspondant à l'observation de $U(t) = f(\text{IMC})$.

L'estimation d'effets de traitement est affectée par:

- **Confusion** (variables affectant le choix de prescription et l'IMC)
- **Temps d'observation irréguliers** (et "informatifs") de l'IMC

Note sur la modification d'effet



Certains modèles possibles pour Y :

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

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

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

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

Modèles (B), (C) et (D) permettent à l'effet de A sur Y de varier selon K , ou Q , ou les deux. C.-à-d. que

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

dépend des valeurs de K et Q .

Observation informative en inférence causale

Observation informative

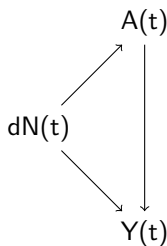
- ▶ Les variables sont observées à des temps qui ne sont pas communs entre les patients (et même entre les variables).
- ▶ Souvent, l'observation d'une variable (au moment d'une visite) **dépend des caractéristiques** du patient (visite "informative").
- ▶ Des exceptions: Examen de routine avec son médecin de famille, mammographie annuelle, etc. (considérées comme des visites aléatoires, "visit at random").

Observation informative

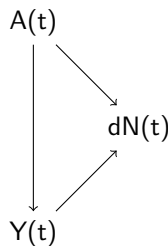
Biais causé par l'observation informative:

- Confusion par le processus de visite
- **Stratification sur un facteur de collision** (Greenland, 2003)

Par ex., si $dN(t)$ dénote un indicateur de visite, $A(t)$ un traitement et $Y(t)$ une issue au temps t :



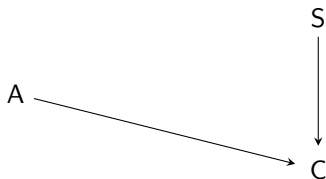
Confusion



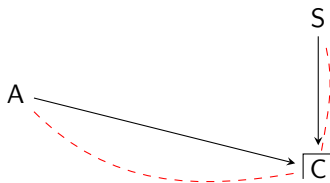
Collision

Facteur de collision C: Intérêt dans la relation A-S

A et S marginalement indépendants:



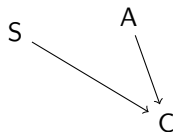
A et S conditionnellement dépendants:



Facteur de collision C: Intuition

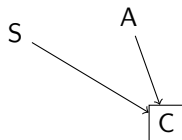
Soit une population où:

- ▶ l'âge (A) et le sexe (S) sont marginalement indépendants
- ▶ les risques de maladie du coeur (C) augmentent avec l'âge et sont plus élevés chez les personnes de sexe masculin



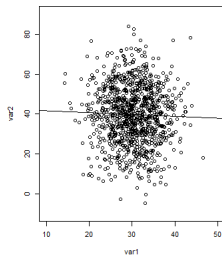
Que pouvons-nous inférer

- ▶ à propos d'un individu avec maladie du coeur ($C = 1$) de sexe féminin?
- ▶ à propos d'un individu sans maladie du coeur ($C = 0$) très âgé?

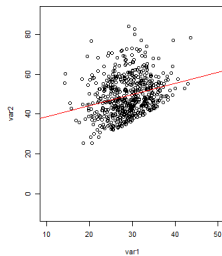


Facteur de collision C: Intuition

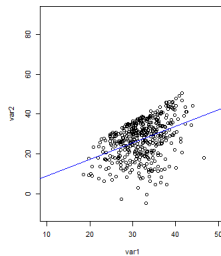
Relation dans les données complètes



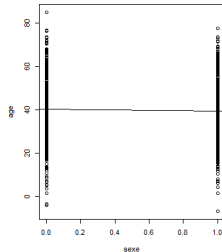
Strate C=1



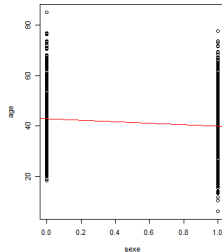
Strate C=0



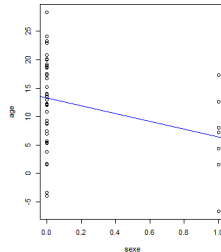
Relation dans les données complètes



Strate C=1

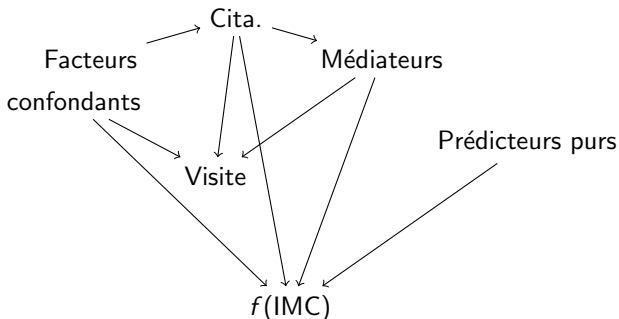


Strate C=0



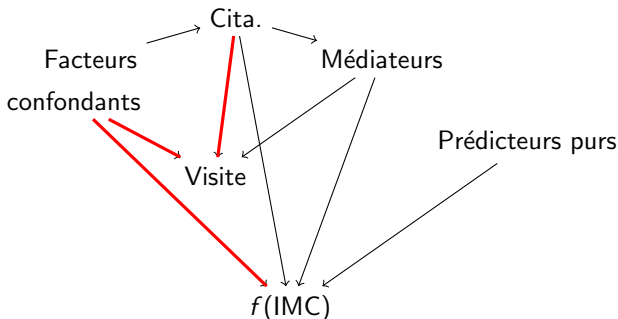
Motivation: Choix d'un antidépresseur

Diagramme causal au temps t :



Motivation: Choix d'un antidépresseur

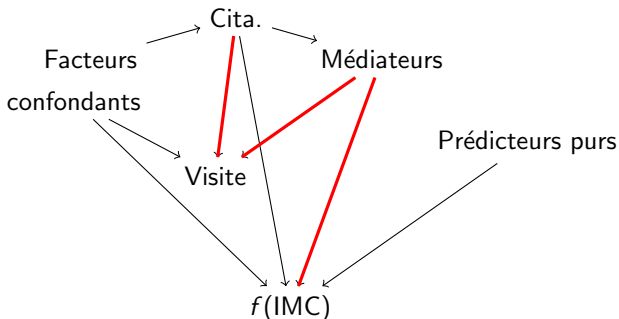
Diagramme causal au temps t :



En conditionnant sur les données observées, ouverture d'un trajet (puisque l'indicateur de visite agit comme facteur de collision).

Motivation: Choix d'un antidépresseur

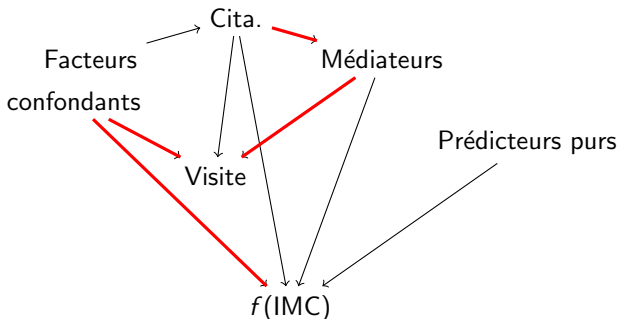
Diagramme causal au temps t :



En conditionnant sur les données observées, ouverture d'un trajet (puisque l'indicateur de visite agit comme facteur de collision).

Motivation: Choix d'un antidépresseur

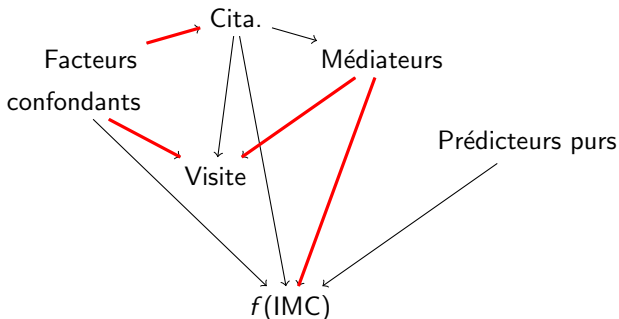
Diagramme causal au temps t :



En conditionnant sur les données observées, ouverture d'un trajet (puisque l'indicateur de visite agit comme facteur de collision).

Motivation: Choix d'un antidépresseur

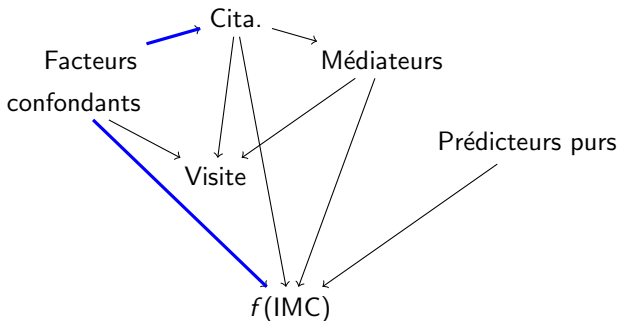
Diagramme causal au temps t :



En conditionnant sur les données observées, ouverture d'un trajet (puisque l'indicateur de visite agit comme facteur de collision).

Motivation: Choix d'un antidépresseur

Diagramme causal au temps t :



Confusion (chemin de porte arrière)

Méthodes

Notation (1)

On s'intéresse à un échantillon aléatoire de patients indexés par $i = 1, \dots, n$.

Soient:

- ▶ $A_i(t)$ et $Y_i(t)$ le traitement binaire et l'issue continue de l'individu i au temps t
- ▶ Les facteurs confondants dénotés par $\mathbf{K}_i(\mathbf{t})$
- ▶ Les prédicteurs purs de $Y_i(t)$, dénotés par $\mathbf{Q}_i(\mathbf{t})$
- ▶ L'ensemble des prédicteurs de $Y_i(t)$, $\mathbf{X}^\beta(\mathbf{t}) = [\mathbf{1} \ \mathbf{K}(\mathbf{t}) \ \mathbf{Q}(\mathbf{t})]$
- ▶ Les modificateurs d'effet $\mathbf{X}^\psi(\mathbf{t})$
- ▶ Soit $\mathbf{X}(\mathbf{t}) = [\mathbf{X}^\beta(\mathbf{t}) \ \mathbf{X}^\psi(\mathbf{t})]$

Notation (2)

Pour modéliser les temps d'observation de $Y_i(t)$:

- ▶ $dN_i(t)$ l'indicateur de visite et d'observation de $Y_i(t)$
- ▶ Par définition, $N_i(t) = \int_{s=0}^t \sum_{j=1}^{F_i} \mathbb{I}(s = T_{ij})$ pour T_{ij} les temps de visite du patient i
- ▶ C_i le temps passé dans l'étude, pour le patient i
- ▶ $\xi_i(t) = \mathbb{I}(C_i \geq t)$ l'indicateur d'être encore dans l'étude
- ▶ On assume que $dN_i(t) \perp Y_i(t) | \mathbf{V}_i(\mathbf{t})$ pour $\mathbf{V}_i(\mathbf{t})$ un ensemble de variables affectant les visites.

Règle de traitement

L'ensemble de **modificateurs d'effet** est dénoté par $\mathbf{X}^\psi(t)$ (voir p. ex., Bian et al., 2021 pour la sélection).

On postule un modèle pour l'issue qui incorpore un modèle “sans traitement” (*treatment-free*) et un blip $\psi' \mathbf{X}_i^\psi(t)$:

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

La règle suivante:

“Traiter avec citalopram si l'espérance de l'issue est maximisée sous citalopram, et sinon, avec fluoxétine”

devient donc

“Traiter avec citalopram si $\hat{\psi}' \mathbf{X}_i^\psi(t) \geq 0$, et sinon, avec fluoxétine.”

Hypothèses causales

- Soient $Y_{i0}(t)$ et $Y_{i1}(t)$ deux issues potentielles sous les interventions 0 et 1 (Neyman, 1923; Rubin, 1974)
- Paramètre d'intérêt: effet conditionnel de traitement

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

- Données (si $\mathbf{X}^{\psi}(t) = X^{\psi}(t)$):

ID	$dN_i(t)$	$A_i(t)$	$X_i^{\psi}(t)$ (groupe âge)	$Y_i(t)$	$Y_{i0}(t)$	$Y_{i1}(t)$
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

Hypothèses causales

À partir des hypothèses causales suivantes, on peut estimer des fonctions de $Y_{i0}(t)$ et $Y_{i1}(t)$:

- Interchangeabilité conditionnelle

$$A_i(t) \perp \{Y_{i0}(t), Y_{i1}(t)\} | \mathbf{K}_i(\mathbf{t}), \mathbf{V}_i(\mathbf{t}), dN_i(t)$$

- Positivité de l'intervention et de l'observation

$$0 < \mathbb{P}(A_i(t) | \mathbf{K}_i(\mathbf{t}), \mathbf{X}_i^\psi) < 1$$

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

- Consistence de l'issue

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

Pourquoi?

Estimateur doublement pondéré

On utilise le concept de “pseudo-population”:

$$\begin{aligned} & \mathbb{E}[Y_{i1}(t) - Y_{i0}(t) | \mathbf{X}_i^\psi(\mathbf{t})] \\ &= \mathbb{E}_{\mathbf{V}}[\mathbb{E}_{\mathbf{K}}[\mathbb{E}[Y_{i1}(t) | \mathbf{K}_i(\mathbf{t}), \mathbf{V}_i(\mathbf{t}), \mathbf{X}_i^\psi(\mathbf{t})]]] - \mathbb{E}_{\mathbf{V}}[\mathbb{E}_{\mathbf{K}}[\mathbb{E}[Y_{i0}(t) | \mathbf{K}_i(\mathbf{t}), \mathbf{V}_i(\mathbf{t}), \mathbf{X}_i^\psi(\mathbf{t})]]] \\ &= \mathbb{E}_{\mathbf{V}}[\mathbb{E}_{\mathbf{K}}[\mathbb{E}[\textcolor{blue}{Y}_{i1}(t) | \mathbf{K}_i(\mathbf{t}), \textcolor{blue}{A}_i(t) = 1, \mathbf{V}_i(\mathbf{t}), \textcolor{red}{dN}_i(t) = 1, \mathbf{X}_i^\psi(\mathbf{t})]]] \\ &\quad - \mathbb{E}_{\mathbf{V}}[\mathbb{E}_{\mathbf{K}}[\mathbb{E}[\textcolor{red}{Y}_{i0}(t) | \mathbf{K}_i(\mathbf{t}), \textcolor{red}{A}_i(t) = 0, \mathbf{V}_i(\mathbf{t}), \textcolor{red}{dN}_i(t) = 1, \mathbf{X}_i^\psi(\mathbf{t})]]] \\ &\quad \text{par interchangeabilité conditionnelle et positivité} \\ &= \mathbb{E}_{\mathbf{V}}[\mathbb{E}_{\mathbf{K}}[\mathbb{E}[\textcolor{blue}{Y}_i(t) | \mathbf{K}_i(\mathbf{t}), A_i(t) = 1, \mathbf{V}_i(\mathbf{t}), dN_i(t) = 1, \mathbf{X}_i^\psi(\mathbf{t})]]] \\ &\quad - \mathbb{E}_{\mathbf{V}}[\mathbb{E}_{\mathbf{K}}[\mathbb{E}[\textcolor{red}{Y}_i(t) | \mathbf{K}_i(\mathbf{t}), A_i(t) = 0, \mathbf{V}_i(\mathbf{t}), dN_i(t) = 1, \mathbf{X}_i^\psi(\mathbf{t})]]] \\ &\quad \text{par consistance.} \end{aligned}$$

Estimateur doublement pondéré

On peut ensuite montrer pour chaque partie de l'expression:

$$\begin{aligned} & \mathbb{E}_{\mathbf{V}}[\mathbb{E}_{\mathbf{K}}[\mathbb{E}[Y_i(t)|\mathbf{K}_i(\mathbf{t}), A_i(t) = 1, \mathbf{V}_i(\mathbf{t}), dN_i(t) = 1, \mathbf{X}_i^\psi(\mathbf{t})]]] \\ &= \sum_{\mathbf{V}} \sum_{\mathbf{K}} \frac{\mathbb{P}(A_i(t) = 1|\mathbf{K}_i(\mathbf{t}))\mathbb{P}(dN_i(t) = 1|\mathbf{V}_i(\mathbf{t}))}{\mathbb{P}(A_i(t) = 1|\mathbf{K}_i(\mathbf{t}))\mathbb{P}(dN_i(t) = 1|\mathbf{V}_i(\mathbf{t}))} \\ & \quad \times \mathbb{E}[Y_i(t)|\mathbf{K}_i(\mathbf{t}), A_i(t) = 1, \mathbf{V}_i(\mathbf{t}), dN_i(t) = 1, \mathbf{X}_i^\psi(\mathbf{t})]\mathbb{P}(\mathbf{K} = \mathbf{K}_i(\mathbf{t}))\mathbb{P}(\mathbf{V} = \mathbf{V}_i(\mathbf{t})) \\ &= \dots \\ &= \mathbb{E} \left[\frac{\mathbb{I}(A_i(t) = 1, dN_i(t) = 1) Y_i(t)}{\mathbb{P}(A_i(t) = 1|\mathbf{K}_i(\mathbf{t}))\mathbb{P}(dN_i(t) = 1|\mathbf{V}_i(\mathbf{t}))} | \mathbf{X}_i^\psi(\mathbf{t}) \right], \end{aligned}$$

et similairement pour $A_i(t) = 0$, menant à

$$\mathbb{E} \left[\frac{\mathbb{I}(A_i(t) = 0, dN_i(t) = 1) Y_i(t)}{\mathbb{P}(A_i(t) = 0|\mathbf{K}_i(\mathbf{t}))\mathbb{P}(dN_i(t) = 1|\mathbf{V}_i(\mathbf{t}))} | \mathbf{X}_i^\psi(\mathbf{t}) \right]$$

Estimateur doublement pondéré

- ▶ On modélise les temps d'observation à partir du modèle à taux proportionnel:

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

avec γ estimé avec Andersen et Gill (1982). Le modèle ci-haut sert à calculer un poids inverse à l'intensité de visite (IIV) dénoté par $\varphi_i(\gamma, \mathbf{V}_i(\mathbf{t}))$ (Lin et al., 2004).

- ▶ Ce poids est combiné à un poids inverse à la probabilité de traitement (IPT):

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

Estimateur doublement pondéré

- L'équation d'estimation à résoudre (extension des dWOLS proposés par Wallace et Moodie (2015)):

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

- La variance asymptotique peut être calculée (Newey et McFadden, 1994):

$$\sqrt{n}(\hat{\psi}_{\text{DW}} - \psi_0) \xrightarrow{d} N(\mathbf{0}, \Sigma_{\mathbf{A}}),$$

pour $\Sigma_{\mathbf{A}}$ un bloc dans

$$\Sigma = \mathbf{G}_{\beta, \psi}^{-1} \mathbb{E} \left[\left\{ \mathbf{g}(\mathbf{o}; \beta_0, \psi_0, \omega_0, \gamma_0) - \mathbf{G}_{\omega, \gamma} \mathbf{M}^{-1} \mathbf{m}(\mathbf{o}; \omega_0, \gamma_0) \right\}^{\otimes 2} \right] \mathbf{G}_{\beta, \psi}^{-1}.$$

Autres hypothèses

Note 1: Les hypothèses suivantes doivent être faites pour considérer les mesures répétées (Dong et al., 2021):

1. Effet de traitement aigu, non différé
2. Pas d'effets rémanents de traitements consécutifs ou d'interactions entre des épisodes de traitement.

Note 2: Tel que décrit dans Wallace et Moodie (2015), le fait

1. d'utiliser un poids qui satisfait la "*balancing condition*" (comme IPT ou les poids de chevauchement) et
2. de bien spécifier le blip $\psi' \mathbf{X}_i^\psi(t)$,

en plus de satisfaire les autres hypothèses mentionnées plus tôt, mène à un estimateur doublement robuste.

Doublement robuste - intuition

Le fait d'utiliser un poids IPT correctement spécifié rend $A_i(t)$ et $\mathbf{X}_i^\beta(t)$ indépendants dans:

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

Ainsi, les coefficients ψ peuvent être estimés de façon consistente même si $f\left\{\mathbf{X}_i^\beta(t); \beta\right\}$ est mal spécifié (et vice-versa si $f\left\{\mathbf{X}_i^\beta(t); \beta\right\}$ est bien spécifié, mais pas le IPT).

À noter, le modèle pour l'observation doit être correctement spécifié, au moins par rapport à un ensemble de variables importantes (d'où le “**partiellement**” doublement robuste).

Étude de simulation

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

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

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

$$Q_i(t) \sim \text{Bern}(0.5)$$

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

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

$$dN_i(t) \sim \text{Poisson}(\lambda_i(t)) \text{ avec } \lambda_i(t) = \eta_i \exp \{\gamma_A A_i + \gamma_Z Z_i(t) + \gamma_{K2} K_2 + \gamma_{K3} K_3\} \\ \text{où } \eta_i \text{ un effet aléatoire, } (\gamma_A, \gamma_Z, \gamma_{K2}, \gamma_{K3}) \text{ les param. de dépendance}$$

Étude de simulation

Les visites simulées selon le **modèle à taux proportionnel**:

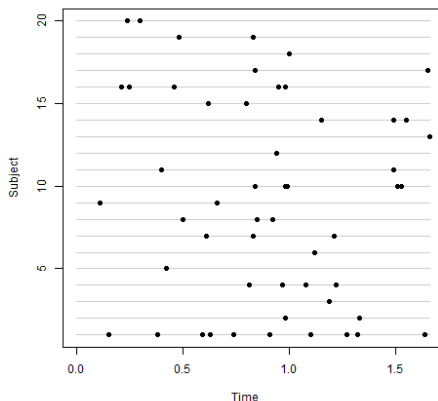


Diagramme abacus produit à partir du logiciel IrregLong (Pullenayegum, 2022)

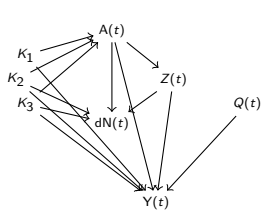
Estimateurs comparés

Pour évaluer la robustesse de la méthode, on compare 6 estimateurs:

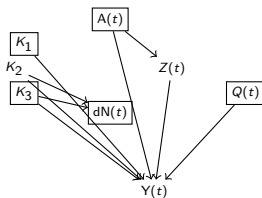
Estimateur	Modèle mal spécifié (x)			
	Visite (partiel ¹)	Visite (complet ²)	Traitement ³	Issue ⁴
$\hat{\psi}_{DW1}$				
$\hat{\psi}_{DW2}$	x			x
$\hat{\psi}_{DW3}$	x		x	
$\hat{\psi}_{DW4}$		x		
$\hat{\psi}_{OLS}$		x	x	
$\hat{\psi}_{IPT}$		x		

1. Ajusté pour les variables importantes (traitement et médiateur)
2. Non ajusté pour le médiateur, ajusté pour le traitement et K_2
3. Mauvais format (au carré) pour termes linéaires de K_1 et K_3
4. Manque l'ajustement pour K_2

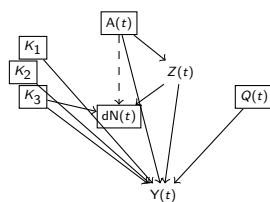
Diagrammes causaux correspondants



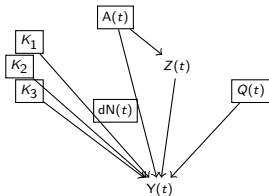
(a) Mécanisme génération données



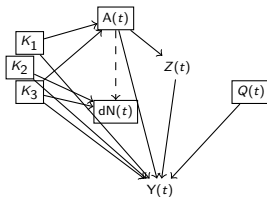
(c) Modèles observation, issue mal spécifiés



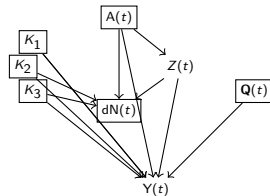
(e) Modèle observation mal spécifié au complet ($dN_i(t) \sim A(t) + K_2$)



(b) Estimateur proposé



(d) Modèles observation, traitement mal spécifiés



(f) Juste ajusté pour IPT

Note: Une boîte signifie qu'on conditionne sur la variable.

Résultats

Tableau 1: Résultats de l'étude de simulation ($M = 1000$ simulations) - taux d'erreur de la décision optimale pour le traitement

Taille éch.	Paramètres γ^v	No. temps obs. moyenne (EIQ)	Taux erreur					
			$\hat{\psi}_{DW1}$	$\hat{\psi}_{DW2}$	$\hat{\psi}_{DW3}$	$\hat{\psi}_{DW4}$	$\hat{\psi}_{OLS}$	$\hat{\psi}_{IPT}$
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

v .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.: EQM, erreur quadratique moyenne; EIQ, étendue interquartile.

Résultats

Tableau 2: Résultats de l'étude de simulation ($M = 1000$ simulations) - biais absolu du blip évalué aux variables $Q(t)$ et K_1

Taille éch.	Paramètres γ^v	No. temps obs. moyenne (EIQ)	Biais absolu					
			$\hat{\psi}_{DW1}$	$\hat{\psi}_{DW2}$	$\hat{\psi}_{DW3}$	$\hat{\psi}_{DW4}$	$\hat{\psi}_{OLS}$	$\hat{\psi}_{IPT}$
250	1	3 (1-3)	0.58	0.50	0.42	0.76	0.74	0.76
	2	3 (1-5)	1.00	1.07	0.97	1.57	1.53	1.57
	3	6 (3-9)	1.14	0.87	0.83	2.04	2.06	2.06
	4	10 (8-12)	0.24	0.24	0.21	0.24	0.19	0.24
1000	1	3 (1-3)	0.33	0.27	0.21	0.68	0.72	0.67
	2	3 (1-5)	0.55	0.60	0.53	1.50	1.49	1.50
	3	6 (3-9)	0.69	0.48	0.46	2.03	2.04	2.04
	4	10 (8-12)	0.12	0.12	0.10	0.12	0.09	0.12
2500	1	3 (1-3)	0.23	0.18	0.14	0.66	0.71	0.66
	2	3 (2-5)	0.36	0.40	0.35	1.51	1.51	1.50
	3	6 (3-9)	0.52	0.34	0.32	2.04	2.05	2.05
	4	10 (8-12)	0.08	0.08	0.07	0.08	0.06	0.08

$v.1.$ (-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). Abbrév.: EQM, erreur quadratique moyenne; EIQ, étendue interquartile.

Note. Le biais du blip évalué aux variables $Q(t)$ et K_1 dépend des valeurs de $Q(t)$ et K_1 . L'étude du biais des coefficients peut être plus pertinente.

Application

Question

- ▶ Règle de traitement pour choisir entre citalopram (1) et fluoxétine (0) pour réduire les variations de l'IMC
- ▶ Définition de l'issue:

$$\begin{aligned} U(t) = & 100 - 5 \times \mathbb{I}[\text{Changement détrimental catégorie IMC}(t)] \\ & + \mathbb{I}[\text{IMC}(0) < 18.5 \cup (18.5 \leq \text{IMC}(0) \leq 24.9 \cap \text{IMC}(t) < 20)] \times \{\% \text{ augment. IMC}(t)\} \\ & - \mathbb{I}[\text{IMC}(0) \geq 25 \cup (18.5 \leq \text{IMC}(0) \leq 24.9 \cap \text{IMC}(t) > 23.5)] \times \{\% \text{ augment. IMC}(t)\}, \end{aligned}$$

pour \mathbb{I} la fonction indicatrice. L'IMC est catégorisé comme suit:

< 18.5	Poids insuffisant
$18.5 - 24.9$	Poids normal
$25 - 29.9$	Excès de poids
≥ 30	Obésité

- ▶ 31,120 patients et 48,388 mesures pour $U(t)$

Caractéristiques des patients à l'entrée dans la cohorte

Tableau 3: Caractéristiques des 31,120 patients, n (%), CPRD, 1998-2017

Variable	Traitement	
	Citalopram (n=18, 671)	Fluoxétine (n=12, 449)
Âge, moyenne (ES)	48.5 (18.1)	45.1 (16.5)
Sexe masculin	5965 (32)	3609 (29)
Indice de déprivation multiple, moyenne (ES)	3.0 (1.4)	3.1 (1.4)
Année d'entrée dans la cohorte		
1998-2005	3751 (20)	4896 (39)
2006-2011	10,279 (55)	5703 (46)
2012-2017	4641 (25)	1850 (15)
A été fumeur	11,586 (62)	8017 (64)
Abus d'alcool	1478 (8)	869 (7)
Maladie psychiatrique [†]	521 (3)	321 (3)
Anxiété	5956 (32)	2987 (24)
Médication		
Antipsychotiques	2836 (15)	1675 (13)
Autres psychotropiques [‡]	4476 (24)	2546 (20)
Hypolipidémiants	3360 (18)	1614 (13)
Nb. d'hospitalisations psychiatriques dans 6 mois précédents, moyenne (ES)	0.04 (0.24)	0.03 (0.34)

Abbrev.: CPRD, Clinical Practice Research Datalink; ES, erreur standard.

[†]. Indicateur d'un diagnostic d'autisme, OCD, trouble bipolaire ou schizophrénie.

[‡]. Inclut les benzodiazépines, anxiolytiques, barbituriques et hypnotiques.

Ratios de taux de visites

Tableau 4: Ratios de taux estimés (ICs 95% bootstrap) pour l'observation de l'issue $U(t)$, CPRD, 1998-2017

Variable	Ratio de taux (CI 95%)
Traitement au citalopram	0.9 (0.9, 0.9)*
Âge	1.0 (1.0, 1.0)
Sexe masculin	0.9 (0.9, 0.9)*
Indice de déprivation multiple	1.0 (1.0, 1.0)*
Année d'entrée dans la cohorte (réf. = <2006)	
2006-2011	0.9 (0.9, 1.0)*
2012-2017	0.9 (0.9, 0.9)*
A été fumeur	1.7 (1.6, 1.7)*
Abus d'alcool	1.0 (0.9, 1.1)
Maladie psychiatrique [†]	1.0 (0.9, 1.2)
Anxiété	1.0 (1.0, 1.0)
Médication	
Antipsychotiques	1.1 (1.0, 1.2)*
Autres psychotropiques [‡]	1.2 (1.2, 1.3)*
Hypolipidémiants	1.2 (1.2, 1.3)*
Nb. d'hospitalisations psychiatriques 6 mois précédents	1.0 (0.9, 1.0)

Abbrev.: IC, intervalle de confiance; CPRD, Clinical Practice Research Datalink.

*. IC significatif.

†. Indicateur d'un diagnostic d'autisme, OCD, trouble bipolaire ou schizophrénie.

‡. Inclut les benzodiazépines, anxiolytiques, barbituriques et hypnotiques.

Règle de traitement

Traiter avec citalopram si $-1.45 + 0.16 \times \mathbb{I}[\text{Sexe masculin}]$
 $+0.13 \times [\text{Indice déprivation multiple}] + 0.08 \times \mathbb{I}[\text{A été fumeur}]$
 $+0.42 \times \mathbb{I}[\text{Abus d'alcool}] + 1.31 \times \mathbb{I}[\text{Maladie psychiatrique}]$
 $+0.35 \times \mathbb{I}[\text{Anxiété}] - 0.91 \times \mathbb{I}[\text{Utilisation d'antipsychotiques}]$
 $+0.30 \times \mathbb{I}[\text{Autres psychotropiques}] + 0.21 \times \mathbb{I}[\text{Hypolipidémiants}] > 0$

Tableau 5: Comparaison des issues ajustées, CPRD, 1998-2017

Traitement	Issue ajustée moyenne (ES^\dagger)			
	$\hat{\psi}_{OLS}$	$\hat{\psi}_{IPT}$	$\hat{\psi}_{IIV}$	$\hat{\psi}_{DW1}$
Reçu	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.: ES, erreur standard.

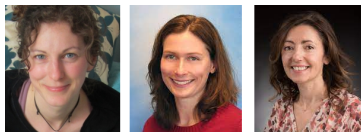
†. Basée sur les ES sommées et normalisées pour la variance de la moyenne des issues.

Discussion

Discussion

- ▶ Les règles de traitement individuelles permettent d'optimiser un choix de traitement selon les caractéristiques du patient
- ▶ Dans une analyse typique, on conditionne sur l'indicateur d'observation des variables. On doit étudier le diagramme causal!
- ▶ Les méthodes discutées s'appuient sur des hypothèses
 - causales
 - sur le mécanisme de génération des données
 - fenêtre temporelle pour l'effet de traitement
 - sur les modèles
- ▶ Extension à venir: Règle à plusieurs étapes et optimisation sur plusieurs points

Remerciements



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

ainsi que Marie-Hélène Descary

Merci de votre attention!

Financement et support:



Contact:

janie.coulombe@mail.mcgill.ca

Références

- 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 Thiébaud, 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.

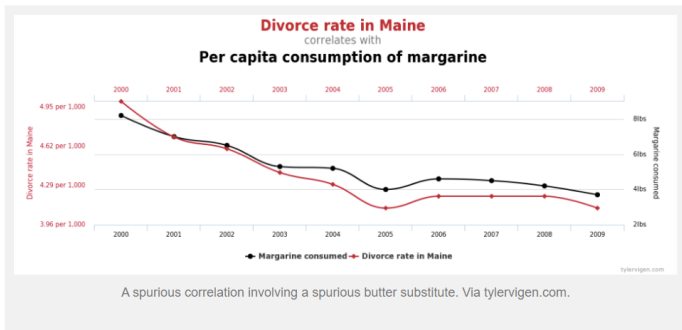
Blip évalué chez différents patients

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

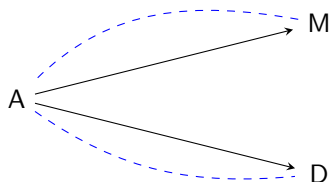
Posted on [April 10, 2017](#) by [annahaensch](#)



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):

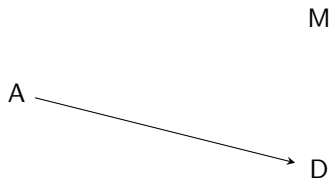


Tableau 6: Coefficients of the blip function (95% bootstrap CIs) for the optimal treatment rules as estimated by four alternative models: OLS which does not adjust for confounding or observation process, IPW which accounts only for confounding, IIV which accounts only for the observation process, and 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.