Using multiple imputations and DWSurv to develop an individualized treatment rule for the choice of antidepressant drug

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

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

SSC Annual Meeting 2023, Ottawa
May 29, 2023

- Choice of antidepressant drugs: Can we do more than prescribe the best treatment marginally?

- Choice of antidepressant drugs: Can we do more than prescribe the best treatment marginally?
- Effect modification by e.g., diagnosis of anxiety

- Choice of antidepressant drugs: Can we do more than prescribe the best treatment marginally?
- Effect modification by e.g., diagnosis of anxiety
- We may be interested in maximizing a survival outcome or a continuous outcome

DWOLS and DWSurv

Wallace and Moodie (2015) and Simoneau et al. (2020). Doubly-robust methods based on weighted generalized estimating equations.

DWOLS and DWSurv

Wallace and Moodie (2015) and Simoneau et al. (2020). Doubly-robust methods based on weighted generalized estimating equations.

- T_i the survival time for individual i and A_i a binary exposure,
- δ_i the indicator of experiencing the event before the end of study,
- $\mathbf{X}_i^{\psi}$ the potential effect modifiers,
- $\mathbf{X}_i^{\beta}$ some confounders or predictors of the outcome.

DWOLS and DWSurv

Wallace and Moodie (2015) and Simoneau et al. (2020). Doubly-robust methods based on weighted generalized estimating equations.

- T_i the survival time for individual i and A_i a binary exposure,
- δ_i the indicator of experiencing the event before the end of study,
- $\mathbf{X}_i^{\psi}$ the potential effect modifiers,
- $\mathbf{X}_i^{\beta}$ some confounders or predictors of the outcome.

Suppose the outcome model:

$$\mathbb{E}[\log T_i | A_i, \mathbf{X}_i] = f\left\{\mathbf{X}_i^{\beta}; \boldsymbol{\beta}\right\} + A_i \boldsymbol{\psi}' \mathbf{X}_i^{\psi}$$

with the "treatment-free model" $f\left\{\mathbf{X}_i^{\beta}; \boldsymbol{\beta}\right\}$ and the "blip" $\boldsymbol{\psi}' \mathbf{X}_i^{\psi}$.

In the case of a one-stage treatment rule, it corresponds to solving the following equations:

$$U(\boldsymbol{\beta}, \boldsymbol{\psi}) = \sum_{i=1}^{n} \int_0^{\tau} w_i \begin{bmatrix} \frac{\partial f\left\{\mathbf{X}_i^{\beta};\boldsymbol{\beta}\right\}}{\partial \boldsymbol{\beta}} \\ A_i(t)\mathbf{X}_i^{\psi} \end{bmatrix} \left[\log(T_i) - f\left\{\mathbf{X}_i^{\beta};\boldsymbol{\beta}\right\} - A_i(t)\boldsymbol{\psi}'\mathbf{X}_i^{\psi}\right] \delta_i = 0.$$

In the case of a one-stage treatment rule, it corresponds to solving the following equations:

$$U(\boldsymbol{\beta}, \boldsymbol{\psi}) = \sum_{i=1}^{n} \int_0^{\tau} w_i \begin{bmatrix} \frac{\partial f\left\{\mathbf{X}_i^{\beta}; \boldsymbol{\beta}\right\}}{\partial \boldsymbol{\beta}} \\ A_i(t)\mathbf{X}_i^{\psi} \end{bmatrix} \left[\log(T_i) - f\left\{\mathbf{X}_i^{\beta}; \boldsymbol{\beta}\right\} - A_i(t)\boldsymbol{\psi}'\mathbf{X}_i^{\psi}\right] \delta_i = 0.$$

The weights w_i must satisfy a balancing condition (Theorem 1, Wallace and Moodie).

In Simoneau et al., they proposed the use of inverse probability of censoring weights to account for informative censoring multiplied by overlap weights (Li et al., 2018).

In data from the Clinical Practice Research Datalink (CPRD), we looked for effect modification to:

- Maximize time to a severe depression-related outcome (Coulombe et al., 2021)
- Minimize "detrimental weight changes" (Coulombe et al., 2023)

We did not find any "good" tailoring variable for the choice of an antidepressant drug, or a class (neither in other research, see e.g., Iniesta et al. (2016), Green et al. (2017), Taliaz et al. (2021)).

- Do we find similar results in American data?
- Access to data from **Kaiser Permanente Washington** (KPW)
- Large nonprofit healthcare organization in Washington State that has a virtual data warehouse combining electronic health records data and insurance billing information

- Do we find similar results in American data?
- Access to data from **Kaiser Permanente Washington** (KPW)
- Large nonprofit healthcare organization in Washington State that has a virtual data warehouse combining electronic health records data and insurance billing information
- We gathered a cohort of patients (13 years or older) who received a diagnosis for depression and initiated antidepressant drugs between 2008-2018

- Do we find similar results in American data?
- Access to data from **Kaiser Permanente Washington** (KPW)
- Large nonprofit healthcare organization in Washington State that has a virtual data warehouse combining electronic health records data and insurance billing information
- We gathered a cohort of patients (13 years or older) who received a diagnosis for depression and initiated antidepressant drugs between 2008-2018
- In KPW data, as opposed to the CPRD, we have access to patient health questionnaires (PHQ)

PHQ-9 questionnaire (a validated tool)

- The DSM 5th version (Diagnostic and Statistical Manual) is a classification tool for common mental disorders
- The PHQ-9 scores each of the 9 depression-related items from the DSM from 0 (not at all) to 3 (nearly every day)

PHQ-9 questionnaire (a validated tool)

- The DSM 5th version (Diagnostic and Statistical Manual) is a classification tool for common mental disorders
- The PHQ-9 scores each of the 9 depression-related items from the DSM from 0 (not at all) to 3 (nearly every day)
  - PHQ first 8 items (**PHQ-8**) ranging from 0 to 24, and
  - **PHQ-9i**, the 9th item, which focuses on suicidal ideation or self-harm, ranging from 0 to 3

(Kroenke et al., 2001)

Challenge

The PHQ-9 score is supposed to be measured rather regularly (at cohort entry, day 30, 60, 90, 180, 270 and 365 as well as at second medication initiation, initiation of antipsychotics, or initiation of psychotherapy, if there is).

Challenge

The PHQ-9 score is supposed to be measured rather regularly (at cohort entry, day 30, 60, 90, 180, 270 and 365 as well as at second medication initiation, initiation of antipsychotics, or initiation of psychotherapy, if there is).

Inspired by Shortreed et al. (2014), we use a sequential imputation approach for PHQ-8, PHQ-9i and weight outcomes that will allow the study of outcomes such as "**time to 50% reduction in PHQ-8 score**".

Challenge

The PHQ-9 score is supposed to be measured rather regularly (at cohort entry, day 30, 60, 90, 180, 270 and 365 as well as at second medication initiation, initiation of antipsychotics, or initiation of psychotherapy, if there is).

Inspired by Shortreed et al. (2014), we use a sequential imputation approach for PHQ-8, PHQ-9i and weight outcomes that will allow the study of outcomes such as "**time to 50% reduction in PHQ-8 score**".

Idea: to borrow information from previous months and baseline to impute future PHQ-8, PHQ-9 and weight values. Shortreed et al.: "Time-ordered nested conditional imputation approach".

Data

- Cohort of around 90,000 patients
- We focused on the first year of follow-up (12 months)

Data

- Cohort of around 90,000 patients
- We focused on the first year of follow-up (12 months)
- Creation of monthly variables: $PHQ8_m$, $PHQ9i_m$, weight$_m$, $m = 1, ..., 12$

Data

- Cohort of around 90,000 patients
- We focused on the first year of follow-up (12 months)
- Creation of monthly variables: $PHQ8_m$, $PHQ9i_m$, weight$_m$, $m = 1, ..., 12$
- Other variables available:
  - demographics (age, sex, race and ethnicity, insurance type, etc.);
  - medication and treatment (psychotherapy, antipsychotics, etc.); and
  - outcomes (PHQ-8, PHQ-9i, suicide attempt, self-harm, weight, etc.).
- Other longitudinal variables can be transformed into indicators (such as monthly indicators of initiating psychotherapy during month $j, j = 1, ..., 12$)

Time-varying variables that were created:

- 1) Initiation and 2) continuation of SGA or FGA
- 1) Initiation and 2) continuation of psychotherapy
- End of the initiating treatment
- Adding a second medication during that month
- Ending the second medication
- PHQ-8 and PHQ-9i measurements closest in time, measured before the month day minus 10 days
- Indicator of self-harm diagnosis, death, death by suicide, hospitalization for depression
- Psychiatric diagnoses: Autism spectrum disorder, anxiety, PTSD, schizophrenia, other psychosis, bipolar disorder, OCD, opioid use disorder, personality disorder, sedative use disorder that occurred anytime before
- Indicator of at least one psychiatric contact on a given month

Imputation

Sequential approach with multiple imputations with chained equations (MICE):

- Impute baseline variables first
- Impute month 1 data using the baseline information and time-varying variables measured at month 1
- Impute month j data ($j = 2, ..., 12$) using the baseline information + PHQ-8, PHQ-9i and weight imputed at month $j - 1$ and time-varying variables from month j

We created 25 such imputed datasets.

Other important details

- Ultimate goal is causal inference

Other important details

- Ultimate goal is causal inference
- PHQ-8 and PHQ-9i provided different information and were treated separately

Other important details

- Ultimate goal is causal inference
- PHQ-8 and PHQ-9i provided different information and were treated separately
- Fully conditional specification - Testing to find the optimal distribution or imputation approach

Other important details

- Ultimate goal is causal inference
- PHQ-8 and PHQ-9i provided different information and were treated separately
- Fully conditional specification - Testing to find the optimal distribution or imputation approach
- Treatment of loss to follow-up

Other important details

- Ultimate goal is causal inference
- PHQ-8 and PHQ-9i provided different information and were treated separately
- Fully conditional specification - Testing to find the optimal distribution or imputation approach
- Treatment of loss to follow-up
- Had to create 31 imputed datasets to have 25 completed (and sensible) datasets

Other important details

- Ultimate goal is causal inference
- PHQ-8 and PHQ-9i provided different information and were treated separately
- Fully conditional specification - Testing to find the optimal distribution or imputation approach
- Treatment of loss to follow-up
- Had to create 31 imputed datasets to have 25 completed (and sensible) datasets
- Post-processing and censoring

Tailoring variables (and confounders)

- Age, sex, race and ethnicity, weight at cohort entry, tobacco use, Charlson comorbidity index[1]
- Psychotherapy (previous year)
- Anxiety or generalized anxiety disorder (GAD)
- Indicator of other psychiatric diagnosis at cohort entry (autism spectrum, obsessive-compulsive, bipolar, personality, sedative use, or alcohol use disorders, schizophrenia, PTSD[2])
- Number of hospit. for mental health diagnosis or suicide attempt or self-harm (previous 6 months)
- Number of antidepressant drugs in previous 5 years
- Had a baseline PHQ score
- PHQ-8 and PHQ-9i at baseline

[1] categorizes comorbidities based on the risk of mortality
[2] post-traumatic stress disorder

Missing data

- Baseline covariates: Between 0% and 24% missing values
- PHQ-8 and PHQ-9i at baseline: 58 % missing

Missing data

- Baseline covariates: Between 0% and 24% missing values
- PHQ-8 and PHQ-9i at baseline: 58 % missing
- PHQ-9: Between 86% and 99% missing between months 1 and 12 (roughly 3/4 of the patients have all PHQ missing from months 1 to 12)
- Naturally, months 4, 5, 7, 8, 10, and 11 not corresponding to the measurement schedule of PHQ-9 contained a lot of missing data.

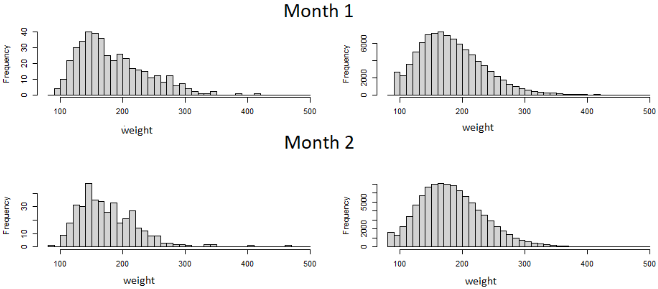
Keeping complete cases until month 3

Correlation matrix for PHQ-8 scores:

	baseline	month 1	month 2	month 3
baseline	1.0000000	0.5029500	0.4500437	0.3854434
month 1	0.5029500	1.0000000	0.6896819	0.6045574
month 2	0.4500437	0.6896819	1.0000000	0.7496655
month 3	0.3854434	0.6045574	0.7496655	1.0000000

Diagnostics for the imputation

Example with weight (left, before; right, after imputation)



Diagnostics for causal inference

SSRI: Selective serotonin reuptake inhibitors.

SNRI: Serotonin and norepinephrine reuptake inhibitors.

5337 patients using SNRI vs **68,721** patients using SSRI.

Balance in baseline characteristics:

Variable	Before imputation			After imputation		
(% missing)	SNRI	SSRI	SMD[3]	SNRI	SSRI	SMD[3]
Weight (24)	190.58	179.73	0.20	192.06	180.32	0.22
PHQ-8 (58)	14.10	14.57	0.09	14.18	14.56	0.07
PHQ-9i (58)	0.52	0.58	0.07	0.52	0.57	0.06

[3] standardized mean difference

Diagnostics for causal inference (2)

Before (left) and after imputation (right):

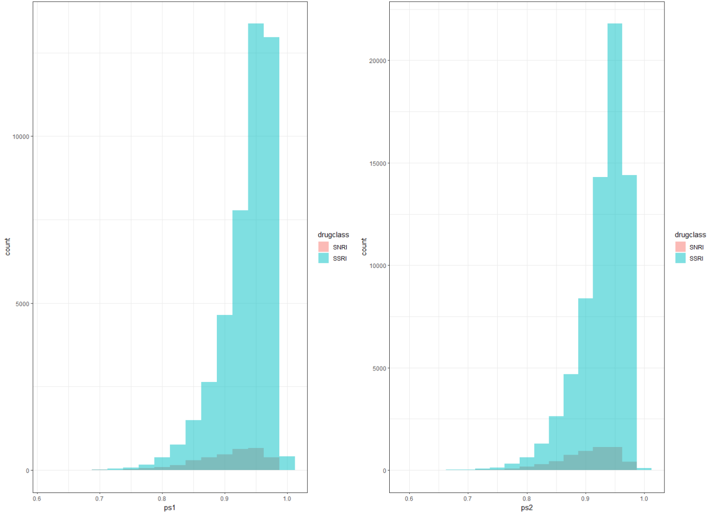


Figure 1: Overlap in the propensity score distributions. Blue: SSRI, Pink: SNRI

Time to 50% PHQ reduction - SSRI vs SNRI

Out of 25 imputed datasets:

- Sex (3)
- Anxiety or GAD (2)
- Diagnosis for a psychiatric diagnosis other than anxiety, GAD (2)
- No. of mental health inpatient stays in previous 6 months (2)
- No. mental health visits in previous 5 years (2)
- Had baseline PHQ (2)
- Psychotherapy (1)
- No. prior AD in previous 5 years (1)

Discussion and what's next?

- The approach relies on the imputation models we choose (MAR assumption, confounding, visit predictors)

Discussion and what's next?

- The approach relies on the imputation models we choose (MAR assumption, confounding, visit predictors)
- Smoothness in mean and interactions for future work (congeniality)

Discussion and what's next?

- The approach relies on the imputation models we choose (MAR assumption, confounding, visit predictors)
- Smoothness in mean and interactions for future work (congeniality)
- Investigate the causal inference assumptions (and properties, such as imputed values' variance)
- Other work...

Acknowledgements


Drs Erica E. M. Moodie
and Susan M. Shortreed

and the organizers of this session.

Thanks for your attention!

Funding and support:

NSERC CRSNG

compute canada | calcul canada

NIH National Institute of Mental Health

Contact:
janie.coulombe@umontreal.ca

References

Coulombe, J., Moodie, E. E. M., Shortreed, et al. (2021) Can the risk of severe depression-related outcomes be reduced by tailoring the antidepressant therapy to patient characteristics?. *American Journal of Epidemiology*, 190(7), pp. 1210-1219.

Coulombe, J., Moodie, E. E. M., Renoux, C., et al. (2023) Estimating individualized treatment rules in longitudinal studies with covariate-driven observation times. *Statistical Methods in Medical Research*.

Green E., Goldstein-Piekarski, A., Schatzberg, A. F., et al. (2017) Personalizing antidepressant choice by sex, body mass index, and symptom profile: An iSPOT-D report. *Personalized medicine in psychiatry*, 1, pp. 65-73.

Iniesta, R., Malki, K., Maier, W., et al. (2016) Combining clinical variables to optimize prediction of antidepressant treatment outcomes. *Journal of psychiatric research*, 78, pp. 94-102.

Kroenke K., Spitzer, R. L., and Williams, J. BW. (2001) The PHQ-9: validity of a brief depression severity measure. *Journal of general internal medicine*, 16(9), pp. 606-613.

Li, F., Morgan, K. L., and Zaslavsky, A. M. (2018) Balancing covariates via propensity score weighting. *Journal of the American Statistical Association*, 113(521), pp. 390-400.

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.

Shortreed, S. M., Laber, E., Scott Stroup, T., et al. (2014) A multiple imputation strategy for sequential multiple assignment randomized trials. *Statistics in medicine*, 33(24), pp. 4202-4214.

Simoneau, G., Moodie, E. E. M., Nijjar, J. S., et al. (2020) Estimating optimal dynamic treatment regimes with survival outcomes. *Journal of the American Statistical Association*, 115(531), pp. 1531-1539.

Taliaz, D., Spinrad, A., Barzilay, R., et al. (2021) Optimizing prediction of response to antidepressant medications using machine learning and integrated genetic, clinical, and demographic data. *Translational psychiatry*, 11(1), pp. 381.

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