Marginal structural models

Laura Pazzagli, Ph. D.
Postdoc researcher in ”Methodologies in Pharmacoepidemiology”

Department of Medicine, Centre for pharmacoepidemiology,
Karolinska Institutet, S-171 76 Stockholm, Sweden
laura.pazzagli@ki.se

Dec 19, 2018

Outline

- Background
- Association versus causation
- Crash introduction to directed acyclic graphs
- Theoretical framework
- Estimation of causal effects in observational studies
- Cross-sectional and longitudinal studies
- Estimation of causal effects in cross-sectional studies
- Inverse probability weighting
- Marginal structural models

Outline

- Estimation of causal effects in longitudinal studies
- Why do we need different statistical methods?
- Time-varying treatments
- Time-varying confounders
- Treatment switching
- Marginal structural models
- Software
- Summary
- References

Background

- The increased availability of population based health registers gives an important opportunity for researchers to have access to a wealth of longitudinal information.
- The availability of various databases in the real world setting makes it possible to test different advanced methodologies proposed for time-varying data.
- At the state of art, in many observational studies, data are analyzed as in a clinical trials setting, using an intention to treat approach (ITT).

Background

- Data from longitudinal studies analyzed as data from cross-sectional studies.
- The ITT approach remains one of the most commonly applied.
- Time-varying data structures require particular attentions.
- Several advanced methods have been proposed but not so commonly applied.

Association versus causation

Association

Measure of the relationship between a treatment and an outcome defined in two different groups of the population and given by the treatment that individuals actually receive.

$$\textit{Treatment} \not\perp\!\!\!\perp \textit{Outcome}$$

Causation

Measure of the relationship between a treatment and an outcome defined in the entire population and given that the individuals are all potential receivers of the treatment.

$$\textit{Treatment} \not\perp\!\!\!\perp \textit{Outcome}$$

Association versus causation

Association

It is a symmetric measure

$$Treatment \rightarrow Outcome \quad or \quad Treatment \leftarrow Outcome$$

Causation

It is not a symmetric measure

$$Treatment \rightarrow Outcome$$

Directed acyclic graphs

Graphs are an intuitive tool to summarize knowledge and assumptions behind a causal structure.

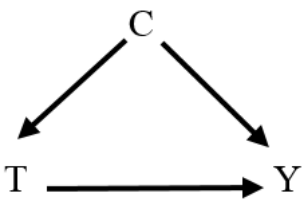
Causal direct acyclic graphs (DAGs)

Causal A DAG may be defined causal when for each pair of variables all confounders are represented in the graph

(D) Directed The edges imply a direction T ⟶ Y

(A) Acyclic There are not cycles T ⟲

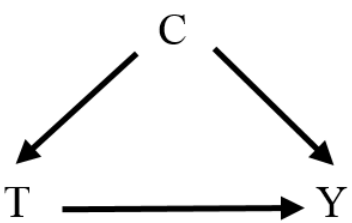
(G) Graphs



Directed acyclic graphs

Marginal randomized experiment

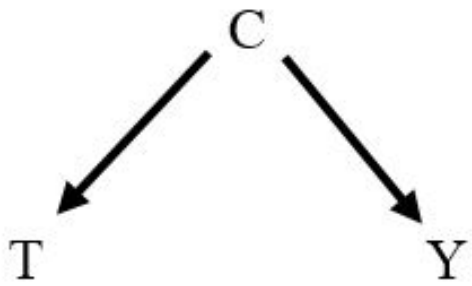
Conditional randomized experiment and observational study



Directed acyclic graphs

Confounder and intermediate

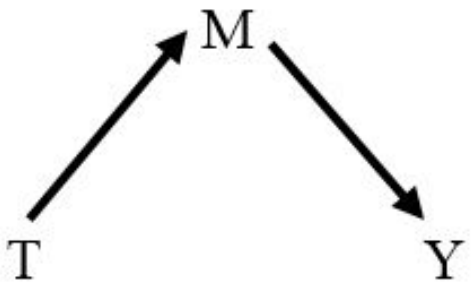
Confounder



C is a common cause

The relationship between
T and Y is confounded by C

Intermediate



T causes M and M causes Y

The relationship between
T and Y is mediated by D

Intermediate role

- The **intermediate role** that a variable can assume is a key concept in longitudinal studies.
- It drives to the need of special precautions when. researchers attempt to draw causal inference
- But let's begin with the theoretical framework for causal inference.

Cross-sectional studies

Theoretical framework

- $Y(1)$ and $Y(0)$ are the potential outcome random variables.
- T is the treatment variable.
- X is a vector of covariates not affected by the treatment.
- $Y = TY(1) + (1-T)Y(0)$ is the observed response variable.

Theoretical framework

The parameter of interest in causal inference is the

Individual Causal Effect $\rightarrow ICE = Y_i(1) - Y_i(0)$

$\Downarrow$

each i is a potential receiver of all possible treatment choices

$\Downarrow$

each i is an actual receiver of one treatment choice at the time

$\Downarrow$

the ICE cannot be identified from the observed data

$\Downarrow$

causal inference is a problem of missing data
(fundamental problem of causal inference).

Theoretical framework

- A parameter of interest that can be identified from the observed data is

$$\textbf{Average Causal Effect} \rightarrow ACE = E[Y(1) - Y(0)]$$

- There are different measures to quantify the magnitude of a causal effect:

$$\text{Risk difference} \rightarrow P[Y(1) = 1] - P[Y(0) = 1]$$

$$\text{Risk ratio} \rightarrow \frac{P[Y(1)=1]}{P[Y(0)=1]}$$

$$\text{Odds ratio} \rightarrow \frac{P[Y(1)=1]/P[Y(1)=0]}{P[Y(0)=1]/P[Y(0)=0]}$$

Theoretical framework

Assumptions

(1) **Unconfoundedness (exchangeability)**

$$(Y(1), Y(0)) \perp\!\!\!\perp T \mid X$$

(2) **Overlap (positivity)**

$$\eta < P(T = 1 \mid X) < 1 - \eta$$

The combination of ”Unconfoundedness” and ”Overlap” assumptions is referred by Rosenbaum and Rubin (1983a) as ”strongly ignorable treatment assignment”

Theoretical framework

Identification of the ACE

The identification of $ACE = E[Y(1) - Y(0)]$ involves the question if the observed data can be used to draw inference about it. A causal effect can be identified in two cases.

(1) There is not confounding.

(2) There are enough measured covariates in X in a such a way that there is not unmeasured confounding.

Theoretical framework

Identification of the ACE

(1) There is not confounding

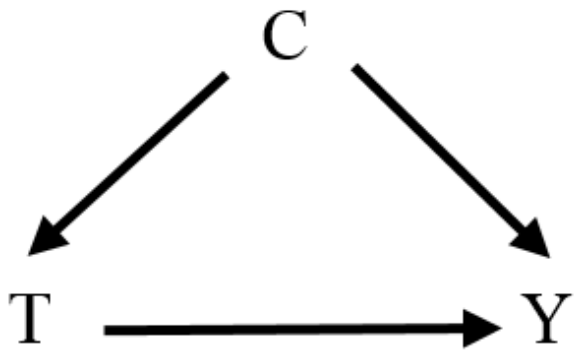
Marginal randomized experiment

Theoretical framework

Identification of the ACE

(2) There are enough measured covariates X in such a way that there is not unmeasured confounding.

Conditional randomized experiment and observational study



Estimation of causal effects in observational studies

- Observational studies are confounded by definition.
- The estimation of causal effects in this context is matter of **"Confounding adjustment"**.
- Several methods to adjust for confounding have been proposed.
- Each of them has strengths and limitations.
- The temporal window involved in the research question should be strongly taken into account.

Cross-sectional and longitudinal studies

- A **cross-sectional** study involves the comparison of different population groups at a single point in time.
- A **longitudinal** study involves several observations of the same subjects over a period of time.
- Accounting for the changes over time of the involved variables longitudinal studies may be a better tool to draw inference about cause-effect relationships.

Cross-sectional and longitudinal studies

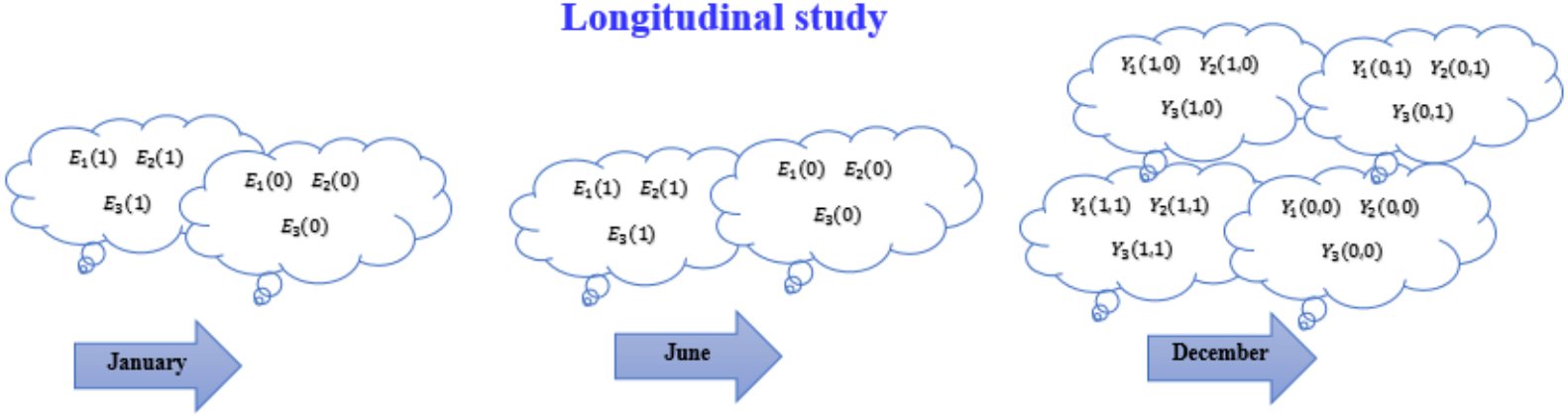
Cross-sectional study

3 individuals

1 binary exposure

1 point in time

Longitudinal study



3 individuals

1 binary exposure

2 points in time

Estimation of causal effects in cross-sectional studies

Methods to derive the causal effect of T on Y in one o more subsets of the population reproducing **"local conditional exchangeability"**:

- **Stratification/regression**: the population is stratified by the effect modifier variable (e.g sex). After stratification a regression model can be used to adjust for the confounder variable C.
- **Matching**: the subset of the population is derived matching for the confounder variable C.
- **Propensity score (PS)**: the population is subsetted by using one of the PS ($P(T = 1|X)$) approaches.

Estimation of causal effects in cross-sectional studies

- **Standardization**: the risk measure is standardized to the population $\sum_x E[Y \mid T = t, X = x]P(X = x)$
- **Inverse probability weighting (IPW)**: the population is weighted by the inverse of the PS.

Inverse probability weighting

- IPW is a technique based on an inverse function of the PS, the probability of receiving the treatment giving a set of covariates.
- IPW estimators were proposed in the sample theory by Horvits and Thompson (1952) for situations in which not all individuals have the same probability of being sampled.
- IPW aims to simulate a ***pseudo-population*** where each individual can be both treated and untreated.

Inverse probability weighting

- Via a weighting system IPW creates a pseudo-population. Each subject is weighted by the inverse of the conditional probability of receiving the treatment level actually received.
- IPW simulates what would have been observed if the set of covariates X would not have affected the treatment assignment.
- For discrete T and X a subject's i weight is equal to:

$$W_i^T = \frac{1}{P(T_i|X_i)}$$

where T_i and X_i are the individual treatment and and set of covariate levels.

Inverse probability weighting

Equivalence of IPW and standardization

For treatment level t the standardized mean and the IPW mean

$$\sum_x E[Y \mid T = t, X = x]P(X = x) \text{ and } E\frac{[I(T=t)Y]}{f[T|X]}$$

are mathematically equivalent given the definition of an expectation and under positivity:

$$\begin{aligned} E\frac{[I(T=t)Y]}{f[T|X]} &= \sum_x \frac{1}{f[t|x]} E[Y \mid T = t, X = x] f[t \mid x] P(X = x) \\ &= \sum_x E[Y \mid T = t, X = x] P(X = x). \end{aligned}$$

Adding exchangeability can be proofed that both the standardized and the IPW are equal to the counterfactual mean $E[Y^t]$.

Inverse probability weighting

IPW weights estimation

- ***Nonparametric estimation***: $P(T = 1 \mid X)$ is equal to the number of treated in ech stratum defined by X divided by the number of individuals in each stratum.
- It is out of the question with high-dimensional data with many confounders.
- Increasing the number of strata, and having a relatively small number of individuals, the stratum-specific estimates become meaningful.

Inverse probability weighting

IPW weights estimation

- ***Parametric estimation***: $P(T = 1 \mid X)$ can be estimated with a logistic regression model including as regressors all the covariates included in X. The estimated weights for the treated are:

$$\frac{1}{\widehat{P}(T=1|X)}.$$

The estimated weights for the untreated are:

$$\frac{1}{1-\widehat{P}(T=1|X)}.$$

Inverse probability weighting

Pseudopopulation

- The weights will be assigned to treated and untreated to reproduce the ***pseudo-population***.
- The satured outcome model can be fitted

$$E[Y \mid T] = \theta_0 + \theta_1 T$$

with individuals in the population weighted.

- The ***average treatment effect*** (ATE) can be estimated as the contrast

$$\widehat{E}[Y \mid T = 1] - \widehat{E}[Y \mid T = 0].$$

- The estimator is consistent for the ***causal*** difference defined in term of potential outcomes

$$E[Y(1)] - E[Y(0)].$$

Inverse probability weighting

Confidence interval

- In estimating the variance we need to account for the IPW technique and there are 3 options.

1. Derive analytically the ***variance estimator*** and program it.
2. Approximate the variance by nonparametric ***boostrapping***.
3. Use the ***robust variance estimator*** with an independent working correlation matrix (the same used for gee models). It provides valid estimates and it is conservative.

Inverse probability weighting

Stabilized weights

- ${W_i}^T = \frac{1}{P(T=1|X)}$ are ***nonstabilized weights*** but we could use any $\frac{p}{P(T=1|X)}$ with $0 < p \leq 1$ to obtain the same effect estimates.
- ***Stabilized weights*** are a common approach to assign to treated and untreated the same probability of being treated and untreated in the original population

$$ {SW_i}^T = \frac{P(T)}{P(T|X)} $$

- SW_i weights are usually more stable resulting in a ***narrower confidence interval*** when the treatment model is not satured (time-varying or continuos treatments).

Inverse probability weighting

Checking positivity

- ***Structural violations of positivity***: individuals with certain values in the set of covariates X cannot possibly be treated (for example survivor bias). Causal inference for all population cannot be made.
- ***Random violations of positivity***: given that the sample is finite we may have zero cells in specific strata defined by X. The probability of being treated in the strata with random zeroes can be estimated using data from individuals in the other strata.

Marginal structural models

- ***Marginal structural mean model***:

$$E[Y(t)] = \beta_0 + \beta_1 t$$

where β_1 is the ATE so that is:

$$\beta_1 = E[Y(1)] - E[Y(0)]$$

because $E[Y(0)] = \beta_0$ and $E[Y(1)] = \beta_0 + \beta_1$.

- ***Marginal structural logistic model*** for a binary Y:

$$logitP[Y = 1 \mid T] = \theta_0 + \theta_1 T$$

where $\hat{\theta}_1$ is a consistent estimator of the ATE β_1.

Marginal structural models and effect modification

- One may include covariates in the marginal structural model (MSM) to assess effect modification:

$$E[Y \mid T, V] = \theta_0 + \theta_1 T + \theta_2 VT + \theta_3 V$$

  where the set of covariates X needs to include V.

- Then the weights can be estimated as

$$SW_i^T(V) = \frac{P(T|V)}{P(T|X)}.$$

Marginal structural models and effect modification

- If the investigator believes that all covariates are effect modifiers then no IPW is necessary and all the covariates can be included in the outcome model.
- ***Confounding adjustment***: for example IPW technique.
- ***Effect modification assessement***: adding treatment covariates interaction term to a MSM. The model for the effect of treatment is estimated within levels of the covariates.

Censoring and missing data

- Selecting only individuals with nonmissing outcome values (censoring) may introduce selection bias.
- A ***censoring indicator*** C is equal 1 for censored individuals and equal 0 for uncensored individuals.
- We can estimate the ATE as if nobody had been censored

$$E[Y \mid T = 1, C = 0] - E[Y \mid T = 0, C = 0].$$

- We can estimated the weights for uncensored individuals to recreate a pseudo-population of the same size as the original population, after censoring:

$$SW_i{}^C = \frac{P(C=0|T)}{P(C=0|T,X)}.$$

Censoring and missing data

- We need a further assumption of conditional exchangeability of censored and uncensored individuals.
- Uncensored individuals are weighted to represent the censored part of the population. For example uncensored individuals with same covariates values as censored will get larger weight to reconstruct the missing information.
- the total individual weight is given by:

$$SW_i^{T,C} = SW^T{}_i \times SW^C{}_i.$$

Longitudinal studies

Cross-sectional and longitudinal studies

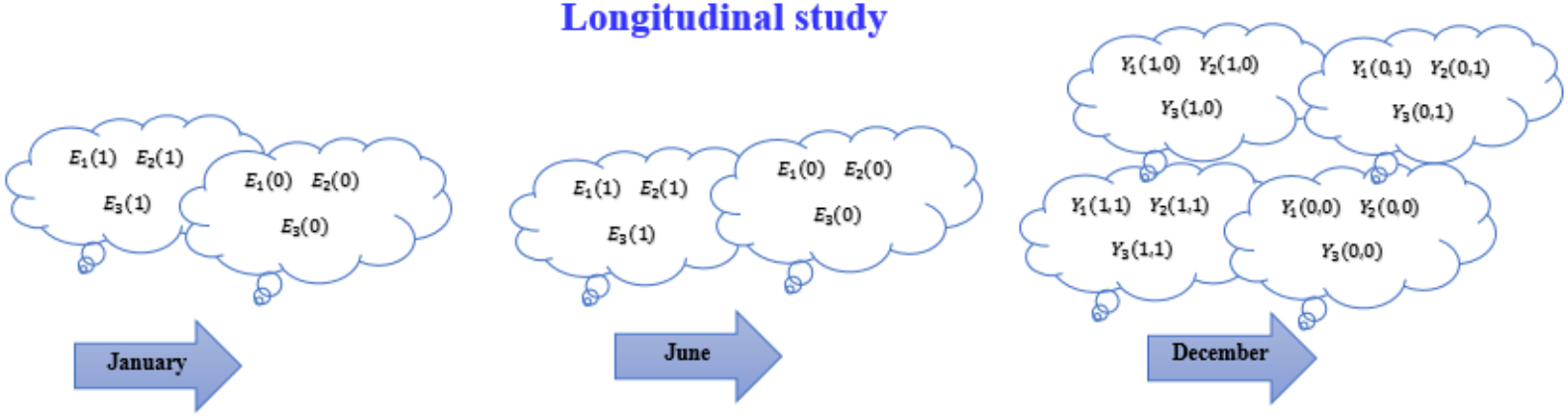
Cross-sectional study

3 individuals

1 binary exposure

1 point in time

Longitudinal study



3 individuals

1 binary exposure

2 points in time

Estimation of causal effects in longitudinal studies

Methods to derive the causal effect of T on Y in the population reproducing **"global conditional exchangeability"**.

- **Standardization/g-formula**: let C_1 be a confounder level at time 1, using the law of total expectation $E[Y(t_0,t_1)]$ (for a discrete confounder) can be estimated under sequential exchangeability as

$$\sum_{X_1} E[Y|T_0 = t_0, T_1 = t_1, C_1 = c_1]P(c_1|t_0)$$

parametrically (g-formula) or non, and where t_0 and t_1 are the levels of treatment at time 0 and 1.

Estimation of causal effects in longitudinal studies

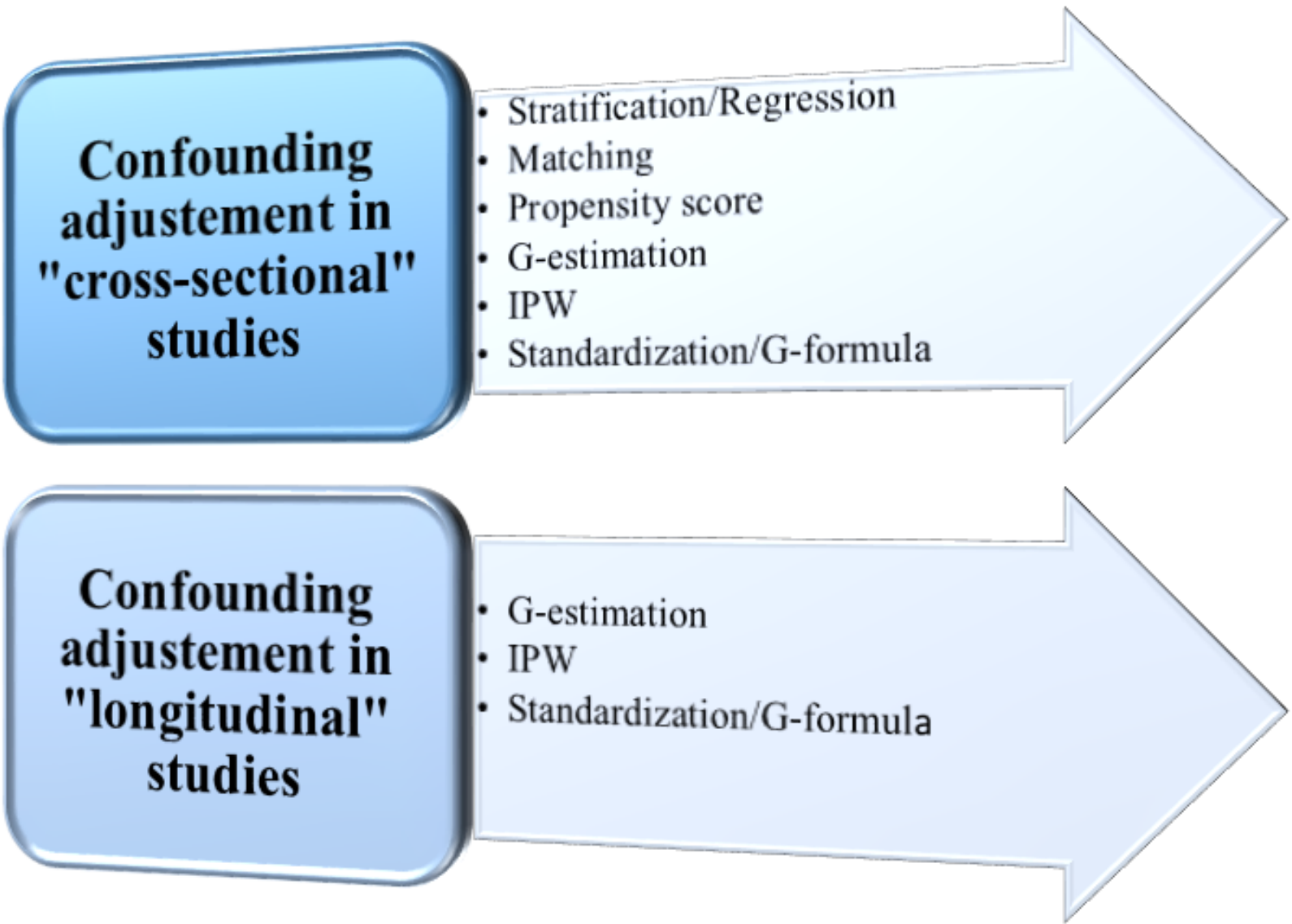
Methods to derive the causal effect of T on Y in the population reproducing **"global conditional exchangeability"**.

- **g-estimation**: used to estimate the parameters of structural nested models (SNMs). A SNM is a model for the effect on Y of a last blip of treatment magnitude as a function of past treatment and covariate histories.
- **IPW**: it creates a pseudo-population via a weighting system. Used to estimate MSMs.

Why do we need different statistical methods?

- Standard methods may provide biased estimates in longitudinal studies.
- All the variables involved in the research question can change observed values over time.
- We can have time-varying treatments, baseline and **time-varying confounders**.
- Standard appraches usually treat a time-varing confounder as a baseline confounder.
- Patients can switch treatment during the follow-up introducing **censoring**.

Why do we need different statistical methods?



Time-varying treatments

Some treatments or exposures can vary over time as for example medical treatments, lifestyle habits, employment status, marital status, etc.

$\Downarrow$

A **treatment history** from time 0 to time k can be represented as $\overline{T}_k = (T_0, T_1, ..., T_k)$.

$\Downarrow$

The ATE of a time-varying treatment cannot be defined as a contrast at a single time k.

$\Downarrow$

$E[Y(T_k = 1)] - E[Y(T_k = 0)]$ reflects only the effect of T at a single time k.

Time-varying treatments

Treatment strategies

- **ATE** = contrast between the conterfactual mean outcomes under 2 treatment strategies that involve treatment at all times between the start and the end of the follow-up.
- Two possible treatment stategies are
  "always treat" $\overline{T}_a = (1,1,...,1)$
  or "never treat" $\overline{T}_n = (0,0,...,0)$.
- Then the ATE can be defined as

$$E[Y(1,1,...,1)] - E[Y(0,0,...,0)] = E[Y(\overline{T}_a)] - E[Y(\overline{T}_n)].$$

Time-varying treatments

Treatment strategies

- There are many other possible causal effects that can be defined contrasting potential outcomes under 2 treatment strategies of interest.
- The number of possible ATEs is 2^K like the number of possible combinations of treatment values from time 0 to time k for a binary treatment.

Time-varying treatments

Dinamic and static treatment strategies

- A **dinamic** treatment strategy is defined when the next treatment level for an individual depends on the individual's time-varying covariates.
- A **static** treatment strategy is defined when the next treatment level for an individual does not depends on the individual's covariates.
- The ATE of a time-varying treatment is correctly defined when the treatment strategies of interest are clearly specified.

Time-varying confounders

In a time-varying data structure confounding factors may also vary over time.

Often future levels of a time-varying confounder are affected by previous levels of the time-varying exposure.

The confounder is also an intermediate in the causal pathway between the treatment and the outcome.

Investigations which do not adjust properly for such factors may lead to bias.

Time-varying confounders

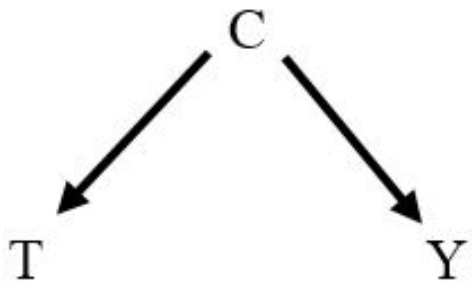
Treatment-confounder feedback

- The confounder affects the treatment and the treatment affects the confounder.
- Time-varying confounding can occur without treatment-confounder feedback.
- In presence of treatment-confounder feedback causal effects cannot be correctly estimated with traditional methods to adjust for confounding like stratification, regressions, and matching.
- This limitation applies also to settings in which the time-varying confounders share causes with prior treatment.

Directed acyclic graphs

Confounder and intermediate

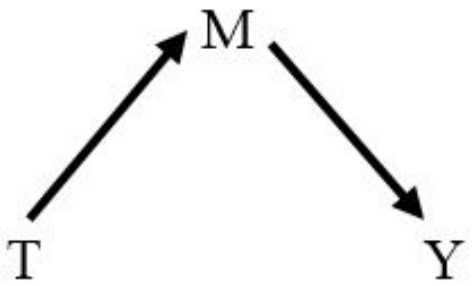
Confounder



C is a common cause

The relationship between
T and Y is confounded by C

Intermediate

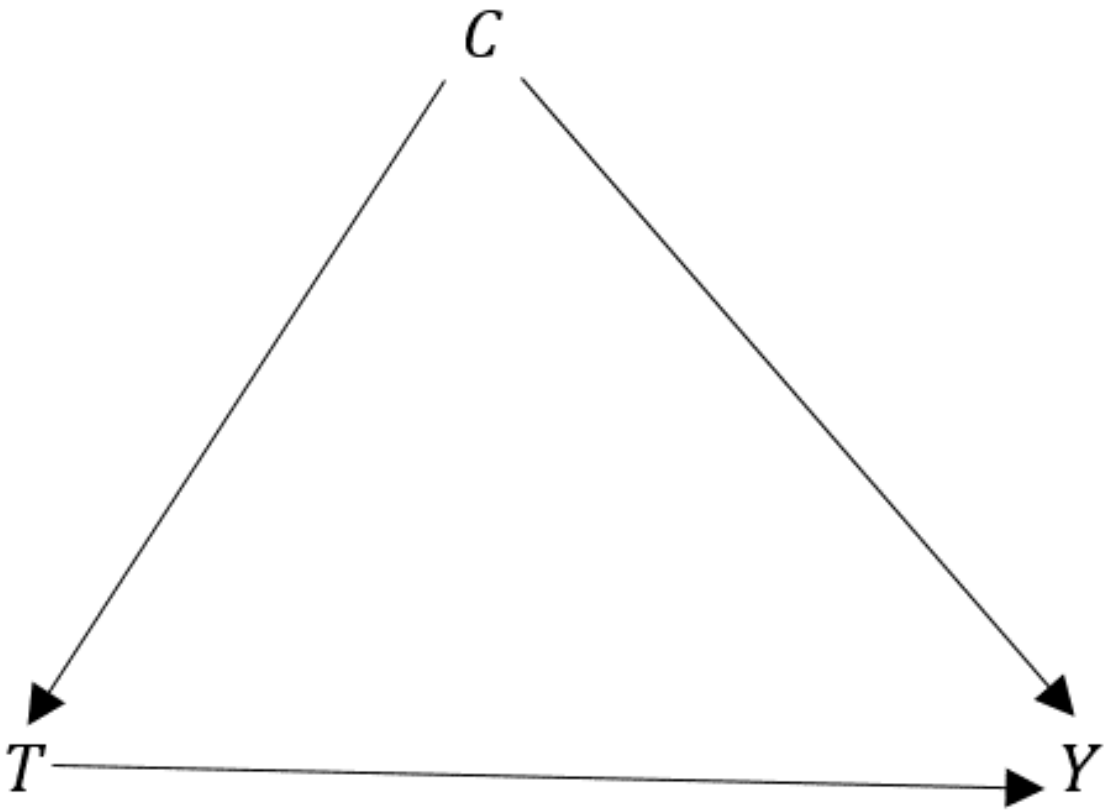


T causes M and M causes Y

The relationship between
T and Y is mediated by D

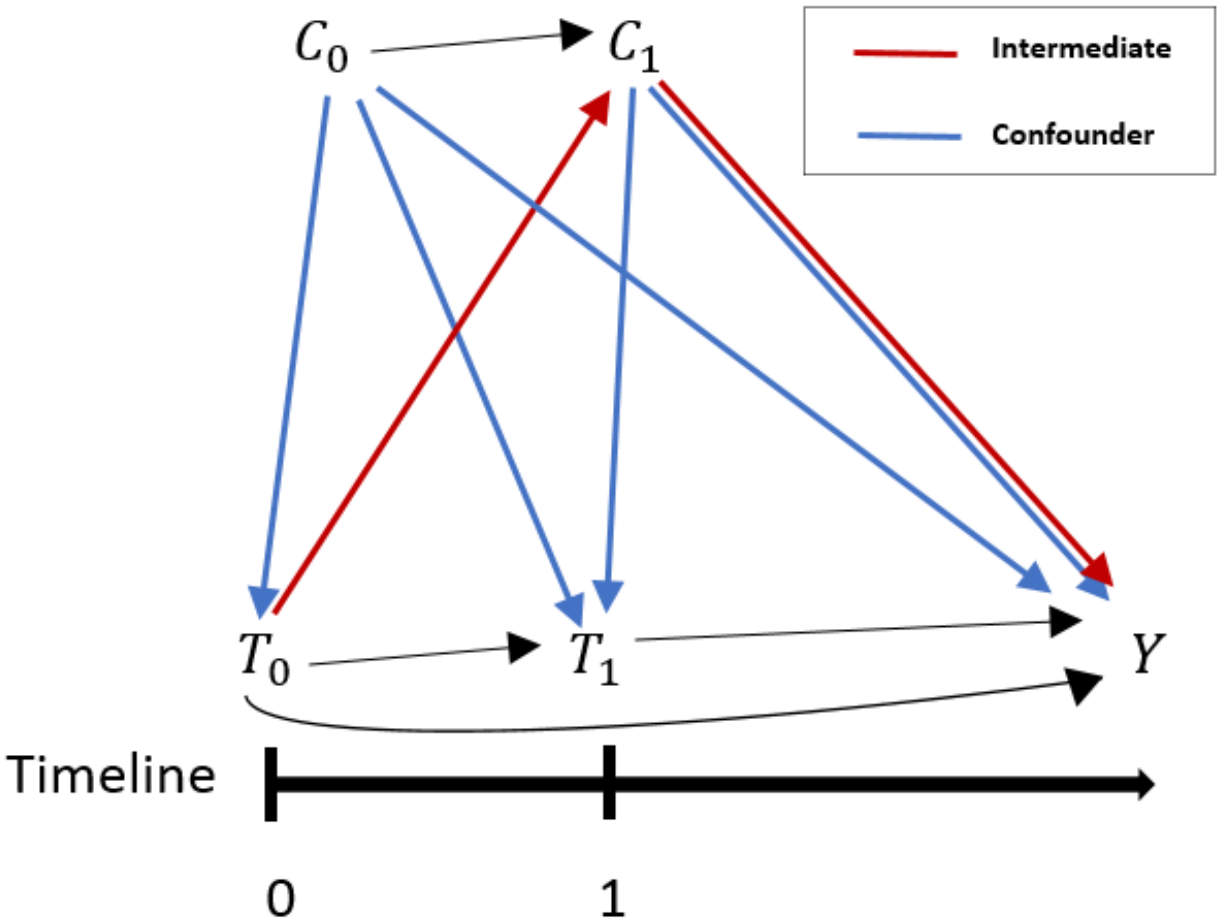
Time-varying confounders

Confounder for a point exposure



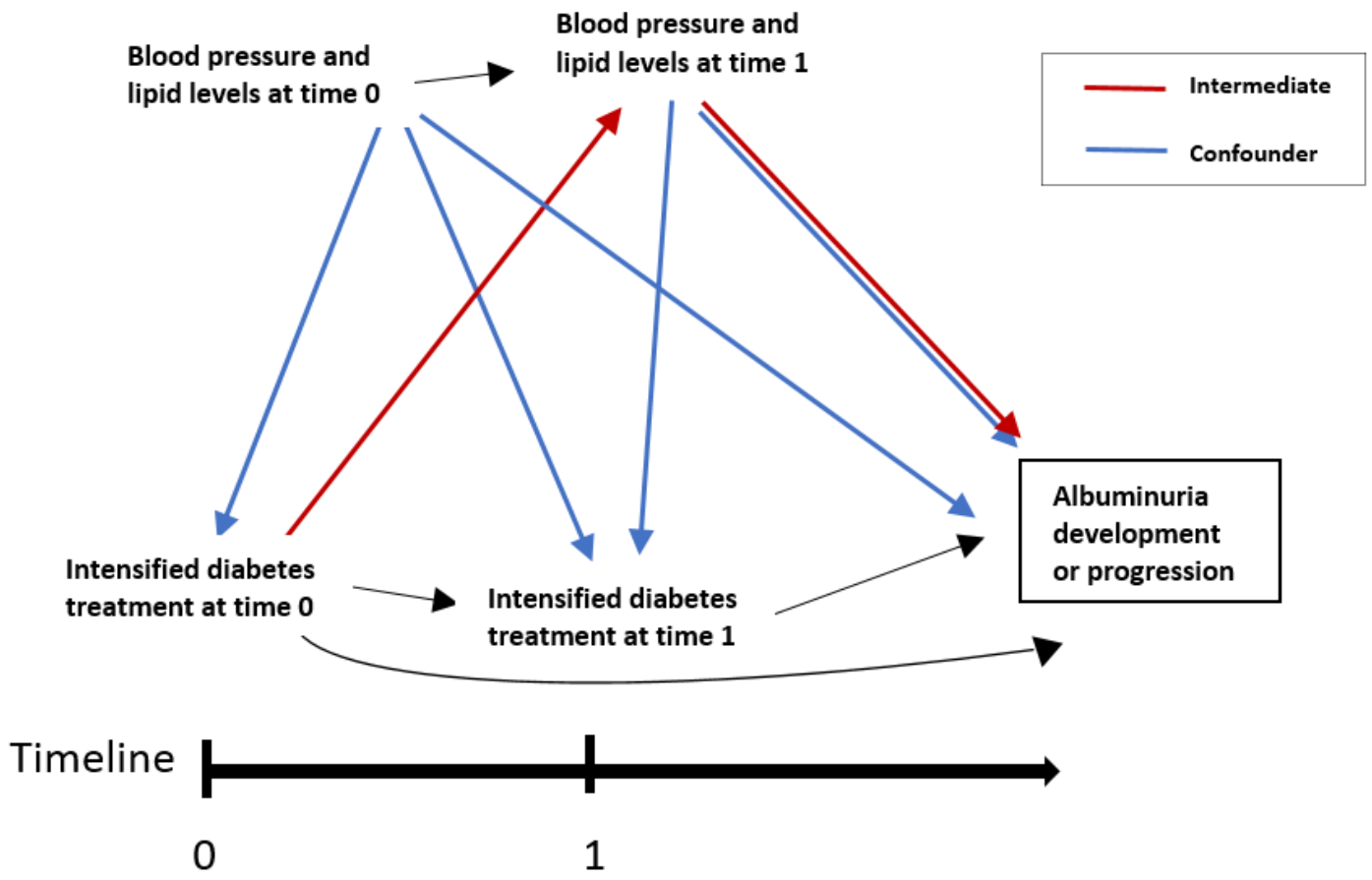
Time-varying confounders

Confounder for a time-varying exposure



Time-varying confounders

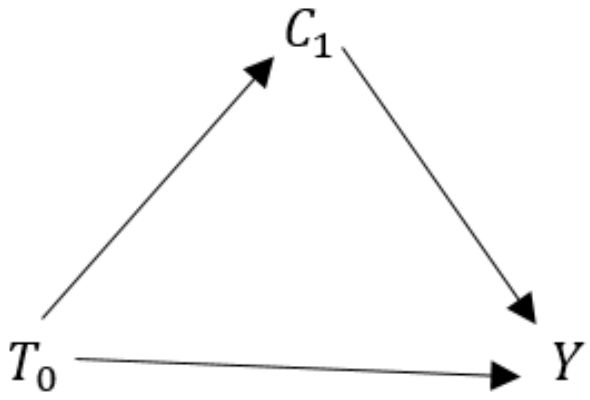
Clinical example



Time-varying confounders

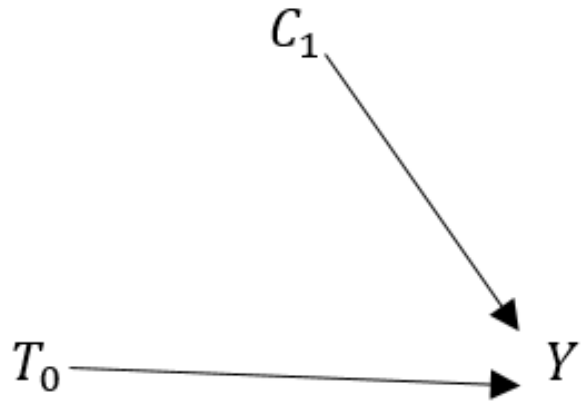
Adjustment for a time-varying confounder using different methods – considering or not the treatment-confounder feedback

MSMs and SNFTMs



Adjustment considering all the arrows in the DAG

Time-varying covariates and propensity score based methods



Adjustment not considering the arrow between T_0 and C_1

Time-varying confounders

Failure of traditional methods and past history

- Traditional methods fail because the conditioning on the level C_1 of the confounder creates an association between the previous treatment level T_0 and Y that does not have a causal interpretation.
- The **past history** of the treatment is fundamental to define the ATE but traditional methods cannot properly adjust for it in presence of tretment-confounder feedback.
- To avoid such bias in cross-sectional studies, one can adjust for prior treatment history or restrict the analysis to individuals with a specific treatment history, which is the idea behind the **"new-user design"**.

Treatment switching

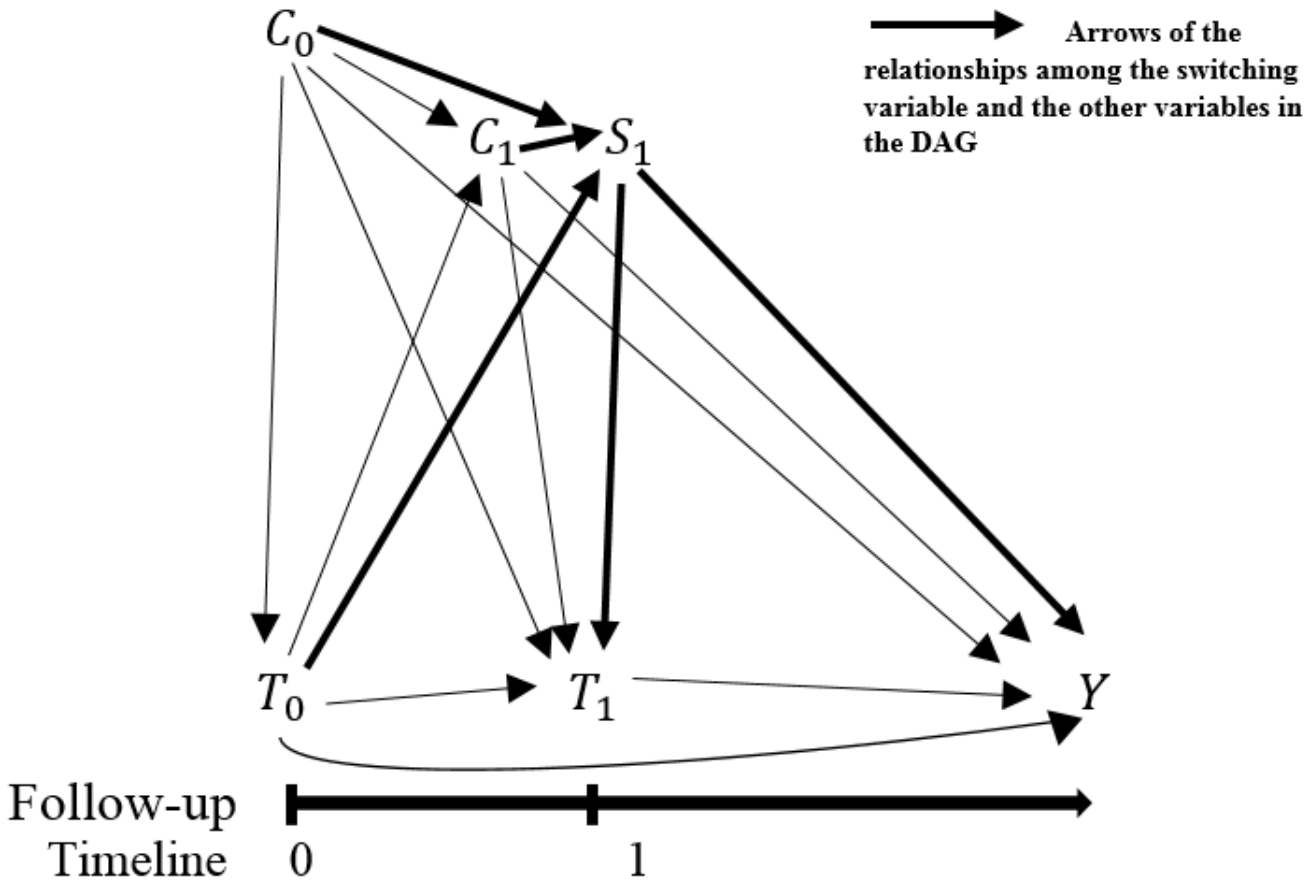
The reasons of the switching of a medical treatment can be several, ranging from adverse reactions to the drug, to the personal behaviour in the compliance to the treatment.

The mechanism behind the switching process is often not at random, and this should be taken into account in the analysis.

In longitudinal studies the switching indicator becomes a time-varying variable $S_1, S_2, \ldots, S_{K+1}$.

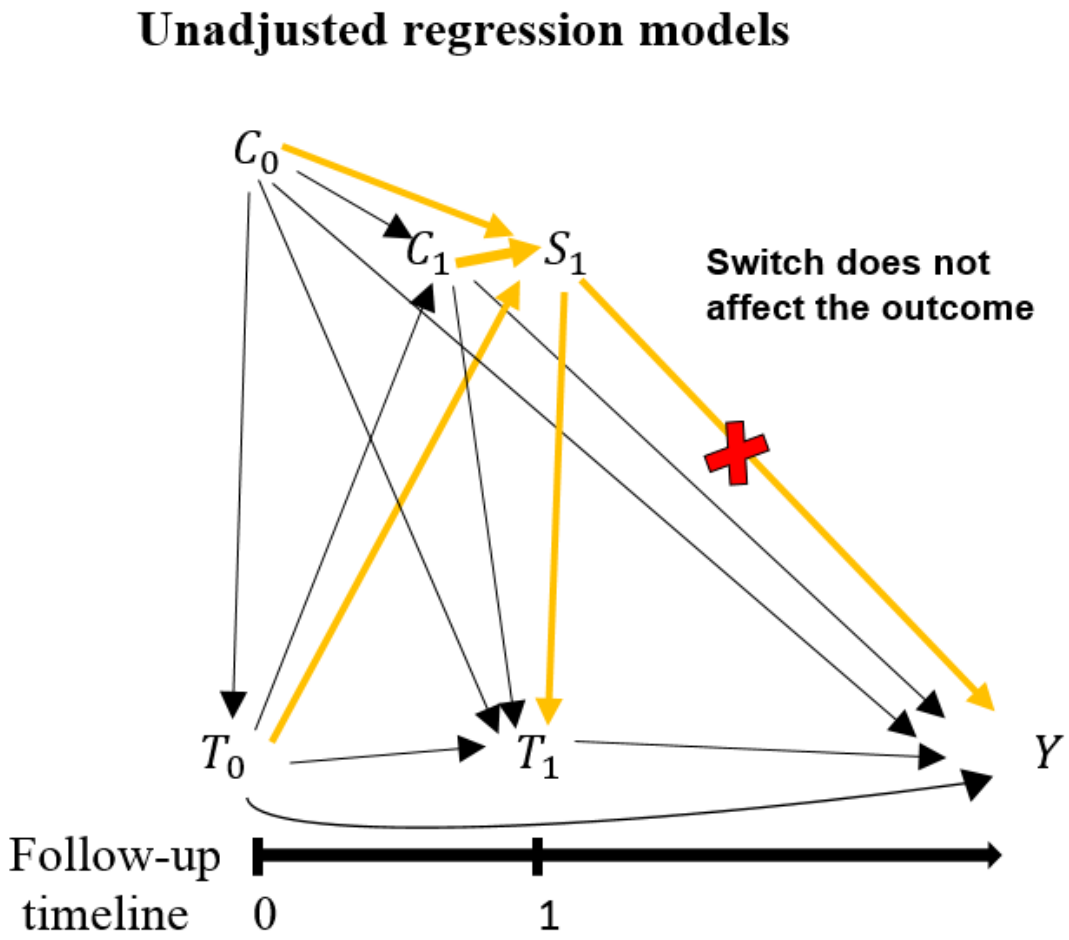
Treatment switching

Treatment switching



Treatment switching

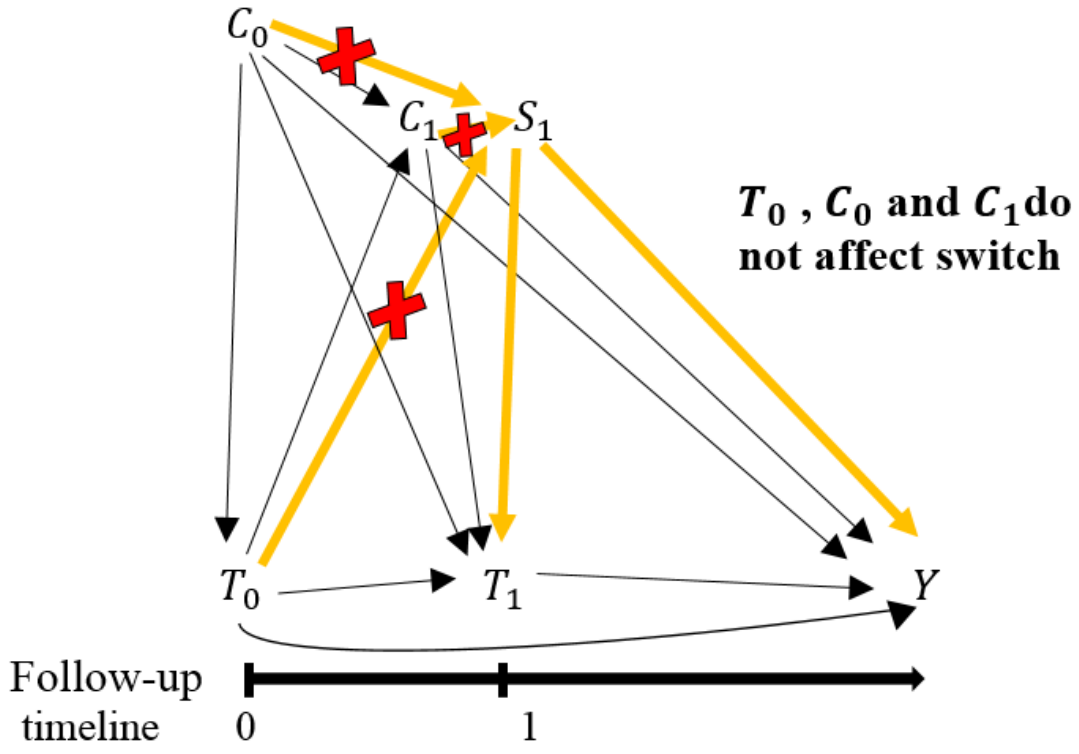
Graphical example of treatment switching



Treatment switching

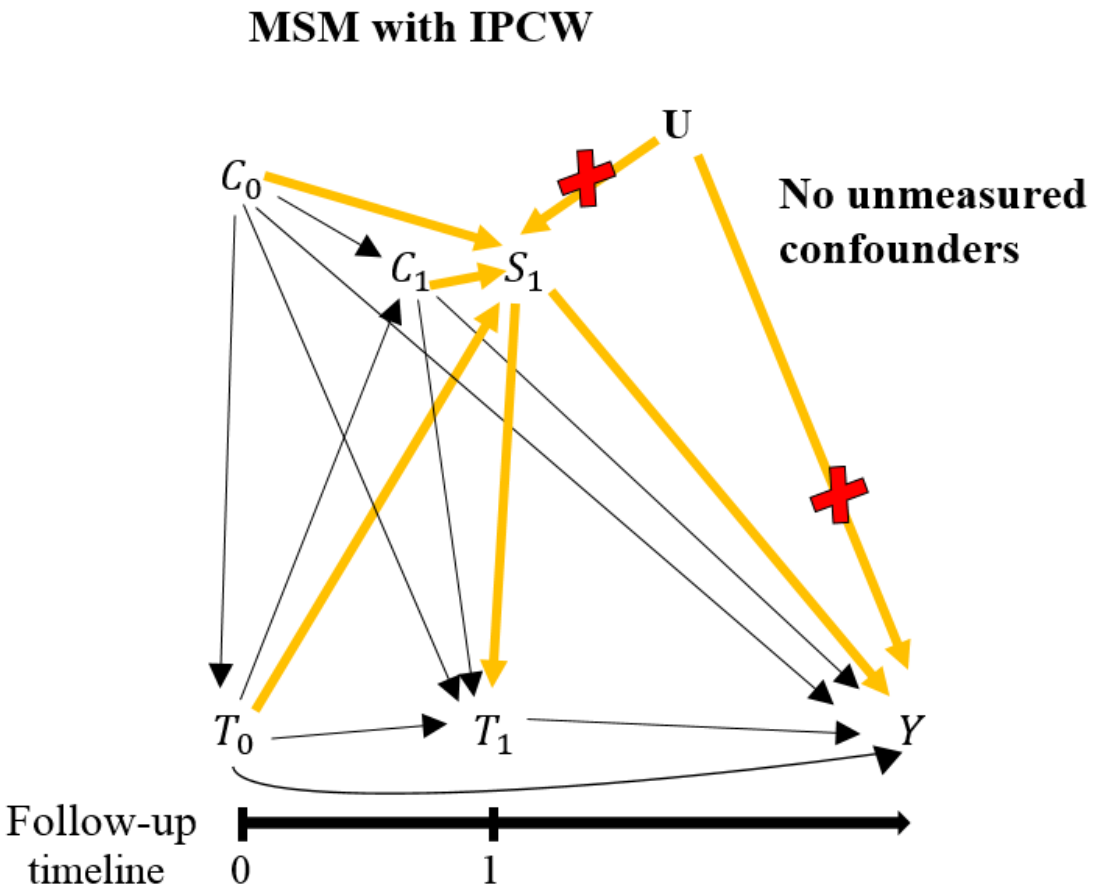
Graphical example of treatment switching

Excluding, censoring switchers and time-varying covariate for switch or treatment



Treatment switching

Graphical example of treatment switching



Marginal structural models

- In longitudinal studies MSMs are both observational and randomized experiments methods proposed by Robins to adjust for time-varying confounders in presence of treatment-confounder feedback.
- These models work under an extension for longitudinal data of the potential outcomes framework proposed by Rubin.
- The IPW method creates a pseudo-population in which the assignment of the treatment is like randomized.
- The inverse probability of censoring weighted estimator (IPCW) creates a pseudo-population in which the switching of the treatment is like randomized.

Marginal structural models

- Moreover MSMs allow to estimate the joint causal effect of several time-varying treatment or exposures on an outcome of interest.
- MSMs can properly adjust for both baseline and time-varying confounders via IPW method.
- The method can be implemented in **two steps**:

  (1) one model for each exposure (or treatment) variable to get the estimated probabilities of receiving the exposure variable.

  (2) the outcome model in a sample weighted by inverse probability of receiving the exposure variable.

- The idea is adopted from the sample theory (Horvitz−Thompson estimator 1952)

Marginal structural models

Assumptions

- **Sequential positivity**

$$\eta < P(T_k = 1 \mid C_{k-1}, B) < 1 - \eta$$

where T_k is the treatment at time k, C_{k-1} the time-varying confounder at time $k-1$ and B is a set of baseline confounders. It must hold for each time point of the follow-up.

- **Global sequential exchangeability**

$$Y(g) \perp\!\!\!\perp T_k | \overline{T}_{k-1} = g(\overline{T}_{k-2}, \overline{C}_{k-1}), C_k$$

for all dynamic strategies g and $k = 0, 1 \ldots K$

- Correct model specification for the weights distribution.

Marginal structural models

IPW estimator

MSMs adjust for time-varying confounders by weighting observations at time t with the stabilized weights (Cole and Hernán 2008; Howe et al. 2012)

$$SW_{it}{}^{T} = \prod_{k=0}^{t} \frac{P(T_k \mid \overline{T}_{k-1}, B)}{P(T_k \mid \overline{T}_{k-1}, \overline{C}_{k-1}, B)}$$

where k is the time indicator, B the set of baseline confounders, $\overline{T}_{k-1}$ and $\overline{C}_{k-1}$ respectively the history of treatment and confounders up to time $k-1$.

Marginal structural models

IPW estimator

In the formula of stabilized weights the factor in the numerator

$$P(T_k \mid \overline{T}_{k-1}, B)$$

is the probability at time k of the exposure T conditioned on the history of the exposures T up to time $k-1$ and on the vector of baseline confounders B. The factor in the denominator

$$P(T_k \mid \overline{T}_{k-1}, \overline{C}_{k-1}, B)$$

is the probability at time k of the exposure T conditioned on the history of the exposure T and time-varying confounders C up to time $k-1$ and the vector of baseline confounders B.

Marginal structural models

IPCW estimator

MSMs can adjust for treatment switching (censoring) via the weight at time t

$$SW_{it}{}^{S} = \prod_{k=0}^{t+1} \frac{P(S_k = 0 \mid \overline{S}_{k-1} = \overline{0}, \overline{T}_k, B)}{P(S_k = 0 \mid \overline{T}_k, \overline{C}_{k-1}, B)}$$

where k is the time indicator, S is the switching indicator, B is the set of baseline confounders, $\overline{T}_k$ and $\overline{C}_{k-1}$ are respectively the histories of time-varying treatment and confounders until time k and $k-1$.

Marginal structural models

IPCW estimator

- In a MSM with IPCW estimator the weights for the non-switchers (uncensored subjects) are constructed to create a pseudo population in which all individuals have the counterfactual outcomes as if they have never switched treatment.
- **Switchers are censored** in the final analysis.
- **Bigger weights** are given to non-switchers with similar history of the censored switchers to represent them when estimating the ATE.

Marginal structural models

Total weights

The total weight at time t for subject i is given by the product of the stabilized weigths for treatment and switching

$$SW_{it} = {SW_{it}}^{T} * {SW_{it}}^{S}$$

A simplified version of SW_{it} is given by the unstabilized weigths W_{it} which have 1 at the numerator. Both produce unbiased estimates of the ATE but stabilized weights are recommended because of the reduction in variability.

Marginal structural models

Estimation of the ATE with a MSM

Let $\bar{t}$ and $\bar{t}_c$ be two possible realizations of $\overline{T}$ and $\overline{T}_c$, **treatment strategies**. The conditional joint ATE is the expected difference between the outcome under $\bar{t}$ and the outcome under $\bar{t}_c$ conditional or non on the set of baseline confounders B

$$E[Y_{\bar{t}} - Y_{\bar{t}_c} \mid B] = f(\bar{t}; \beta, B) - f(\bar{t}_c; \beta, B),$$

where β is a vector of the coefficients for a function $f(\cdot)$. To estimate $f(\bar{t}; \beta, B)$ can be used a weighted logistic regression.

Marginal structural models

Example of survival analysis and administrative censoring

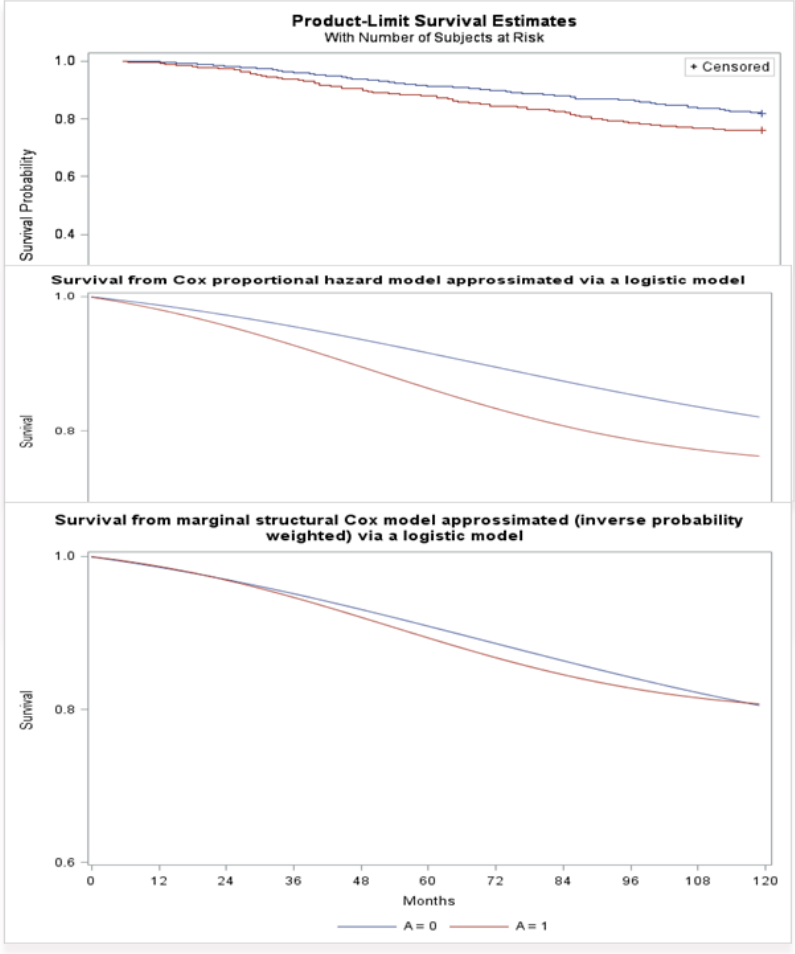
- From Chapter 12 of the Causal inference book (Jamie Robins and Miguel Hernán).
- NHEFS: National Health Epidemiologic Follow-up Study (NHEFS) of National Center for Health Statistics.
- The sample for the survey is selected to represent the U.S. population of all ages.
- Survival analysis among 1629 cigarette smokers, with start of follow-up in January 1983 and administrative end in December 1992.

Marginal structural models

Example of survival analysis and administrative censoring

Table of death by quit smoking

Death between1983 and 1992	Quit smoking between baseline and 1982		
	Non quitters	Quitters	Total
alive	985	326	1311
dead	216	102	318
Total	1201	428	1629



Marginal structural models

- MSM is an useful tool for time-varying data structures.
- A MSM for the odds ratio is just a weighted logistic regression model.
- MSM may provides consistent estimates of the joint causal effect of several exposures under study.
- Under the assumptions of sequential positivity, global sequential exchangeability, and correct specification of the weights distribution, each coefficient in a MSM can be interpreted as the causal effect of the corresponding exposure on the outcome.

Software

- In the website of the Harvard School of Public Health (https://www.hsph.harvard.edu/causal/software/) an example of the code to implement a marginal structural Cox model is provided in Stata and SAS.

- The R package ipw is available in the CRAN website (https://cran.r-project.org/web/packages/ipw/) for the estimation of the weights of a MSM.

Summary

- Different methods have been proposed for time-varying data but the most advanced are not commonly used.
- No statistical method is ”best” in all situations. Involved assumptions should be evaluated considering the particular context.
- Guidelines and more statistical packages for the implementation of the most advanced methods as MSM and SNM are needed.

References

D. B. Rubin (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66, 688.

M.A. Hernán, J.M. Robins (2017). Causal Inference. Boca Raton: Chapman & Hall/CRC, forthcoming.

S. R. Cole, M.A. Hernán (2008). Constructing inverse probability weights for marginal structural models. *American journal of epidemiology*, 168, 656-664.

J. Pearl (1995). Causal diagrams for empirical research. *Biometrika*, 82:669-688.

P.R. Rosenbaum , D.B. Rubin (1973). The central role of the propensity score in observational studies for causal effect. *Biometrika*, 70:41-55.

Thank you for the attention