Causal inference with semi-competing risks, and beyond

Deng, Yuhao
yuhaoden@umich.edu

Department of Biostatistics, University of Michigan

October 28, 2024

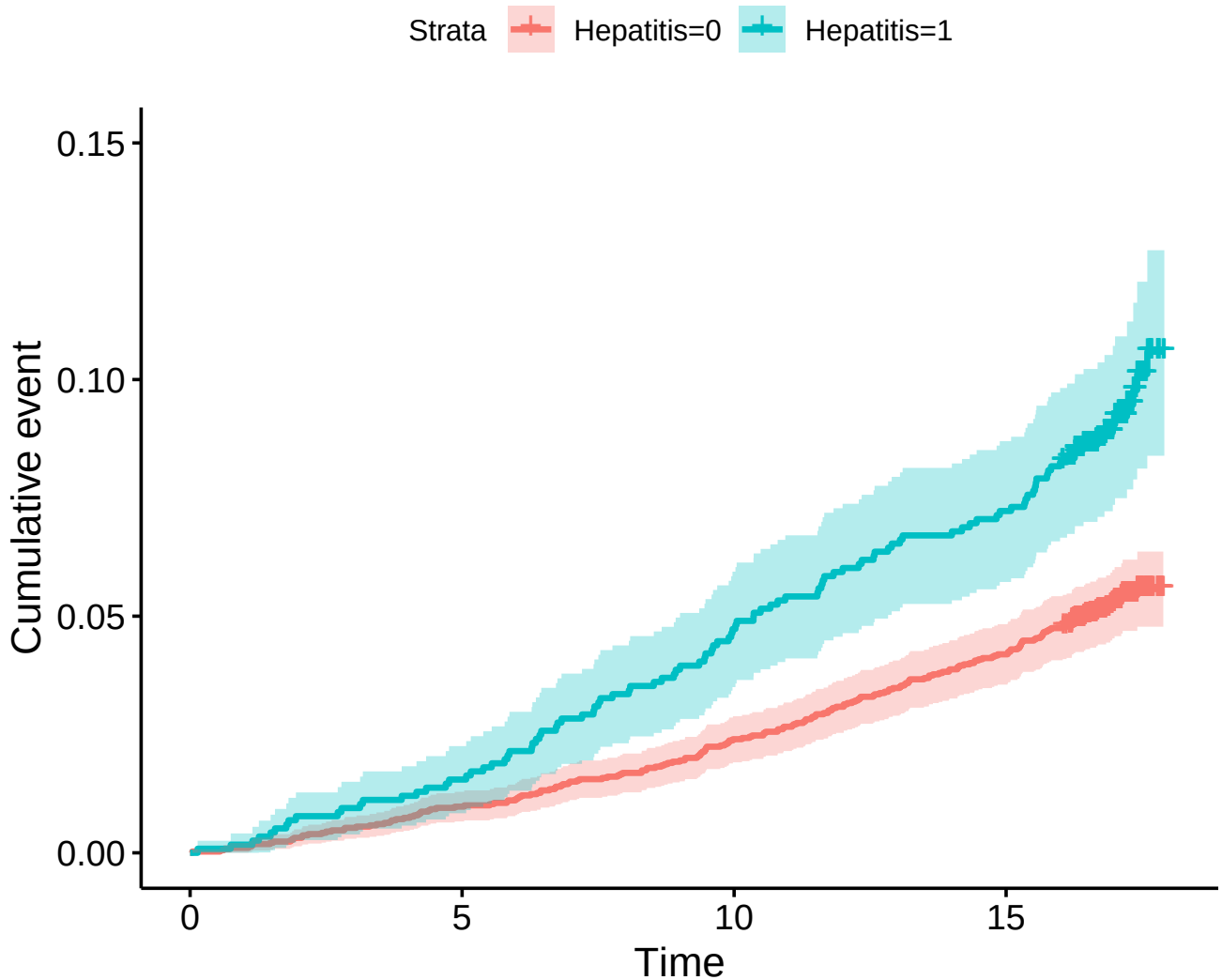
Outline

Outline

Motivating Data 1

- Hepatitis B causes serious public health burden worldwide.
- Exposed to hepatitis B, an individual has increasing risks of both liver cancer and mortality.
- What is the causal effect of hepatitis B on mortality mediated by liver cancer?

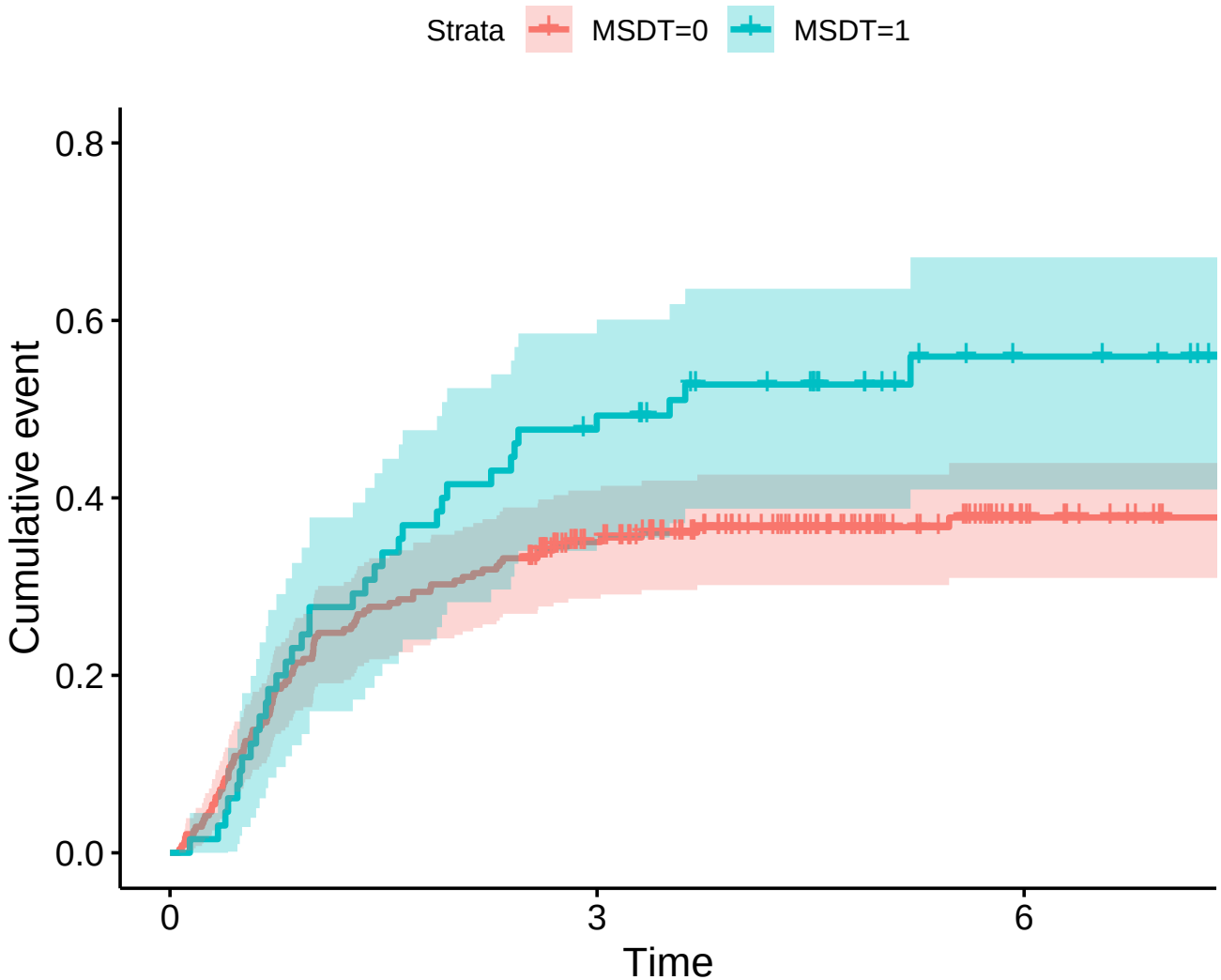
Motivating Data 1



Motivating Data 2

- Allogeneic stem cell transplantation is a commonly adopted approach to cure acute lymphoblastic leukemia (ALL).
- Human leukocyte antigen matched sibling donor transplantation (MSDT) has long been considered as the first choice of transplantation because of lower transplant-related mortality.
- In recent years, some benefits of haploidentical stem cell transplantation (haplo-SCT) have been found in that haplo-SCT results in lower relapse rate.
- What is the mechanism of transplant modality on overall mortality (transplant-related mortality, relapse-related mortality)?

Motivating Data 2



Question

- The question is about mediation analysis.
- The mediator and outcome are "time to event", subject to censoring.
- We want to know
  (1) the treatment effect delivered *directly* to the terminal event and
  (2) the treatment effect delivered *indirectly* through (*mediated* by) the intermediate event to the termianl event.
- How to define, identify, estimate, infer and intepret the causal effect?

Outline

Mediation Analysis at Varying Time Points

- For time-to-event data, the mediator and outcome are processes.
- The mediator (intermediate event) process $N_1(t) \in \{0, 1\}$: whether the intermediate event has occurred at t.
- The outcome (terminal event) process $N_2(t) \in \{0, 1\}$: whether the terminal event has occurred at t.

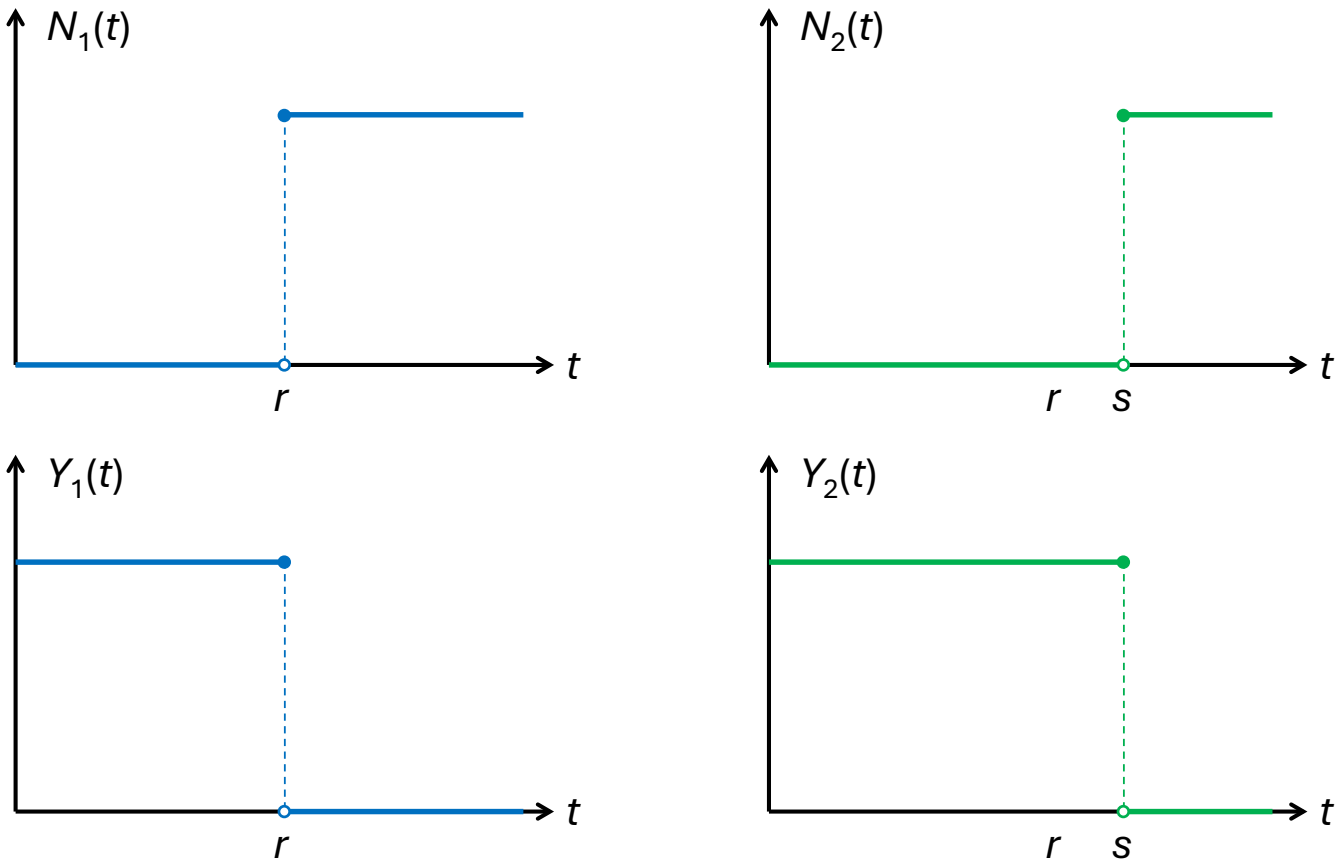
$N_1(t^-)$	$N_2(t)$	Description at time t
0	0	No intermediate event, no terminal event
0	1	No intermediate event, a terminal event
1	0	An intermediate event, no terminal event
1	1	An intermediate event, a terminal event

Semi-Competing Risks

- *Semi-competing risks* refer to the phenomenon that the terminal event can truncate the non-terminal event but not vice versa (Fine et al, 2001; Huang, 2021).
- The terminal event and non-terminal events have a competing nature, but their roles are not the same.
- Related to the truncation-by-death problem.

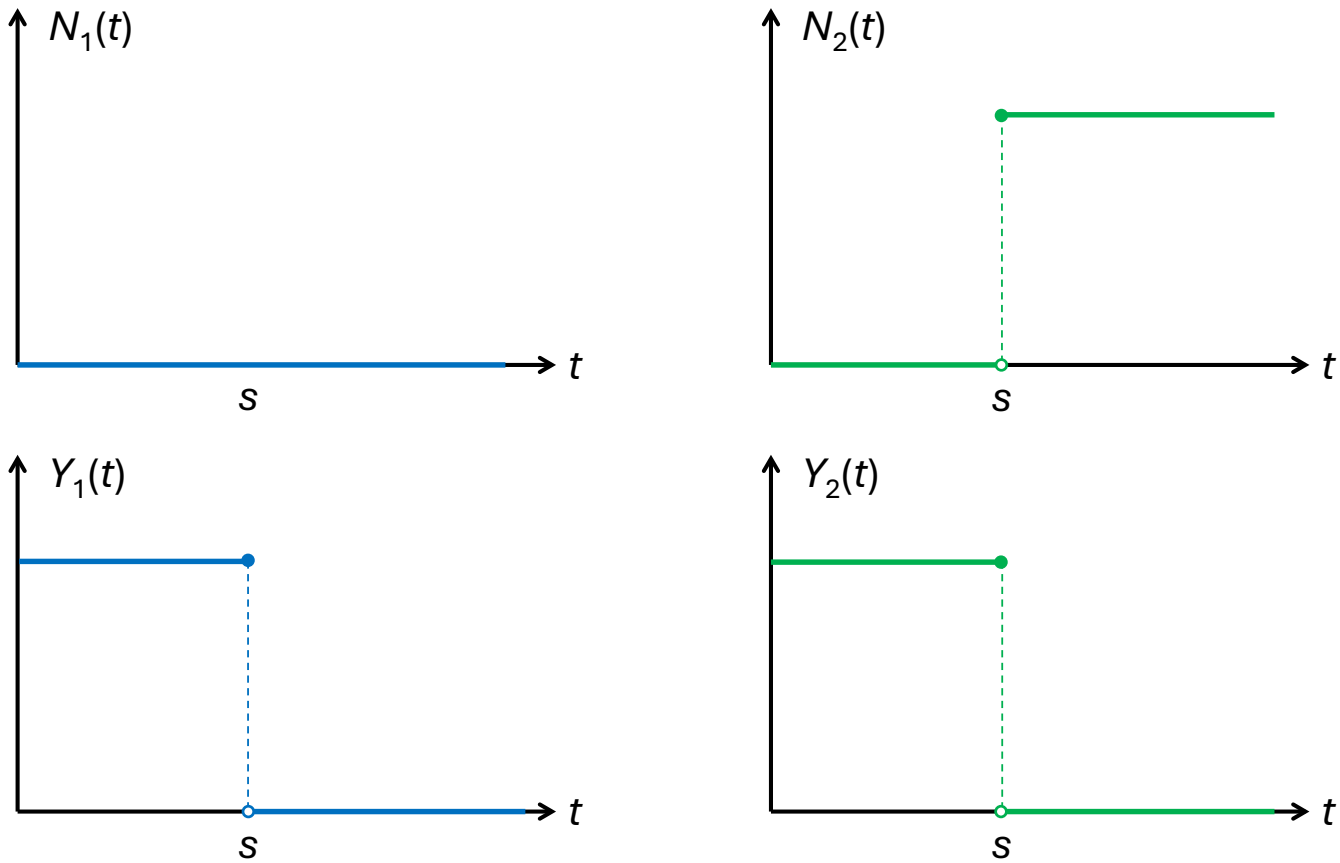
Event Processes and At-Risk Processes

- Intermediate event occurs first.



Event Processes and At-Risk Processes

- Intermediate event does not occur.



Formal Definition: Jumps

- Let $d\tilde{N}_1(t; z_1)$ be the jump of potential event counting process for the intermediate event during $[t, t + dt)$ when the treatment is set at z_1.
- Let $d\tilde{N}_2(t; z_2, n_1)$ be the jump of potential event counting process for the terminal event during $[t, t + dt)$ when the treatment is set at z_2 and the counting process for the intermediate event at t^- is set at n_1.

Formal Definition: Well Definedness

- **Markovness**: the hazard of jumps only relies on the current status, not the history,

$$\begin{aligned} & P(d\tilde{N}_2(t; z_2, n_1) = 1 \mid \tilde{N}_2(t^-; z_2, \tilde{n}_1(\cdot)) = 0, \tilde{n}_1(t^-) = n_1, \tilde{n}_1(s^-) = n_1^*) \\ = & P(d\tilde{N}_2(t; z_2, n_1) = 1 \mid \tilde{N}_2(t^-; z_2, \tilde{n}_1(\cdot)) = 0, \tilde{n}_1(t^-) = n_1), \ s < t. \end{aligned}$$

- Markovness ensures that the notation $d\tilde{N}_2(t; z_2, n_1)$ is well defined.
- We can manipulate the treatment z_2, or we can manipulate the status of the intermediate event n_1, in order to generate a potential jump $d\tilde{N}_2(t; z_2, n_1)$.

Formal Definition: Processes

- We integrate the jumps to obtain potential counting processes.
- The potential event counting process for the intermediate event

$$\tilde{N}_1(t; z_1) = \int_0^t d\tilde{N}_1(s; z_1).$$

- Given the full intermediate event process $\tilde{n}_1(\cdot)$, the potential event counting process for the terminal event

$$\tilde{N}_2(t; z_2, \tilde{n}_1(\cdot)) = \int_0^t d\tilde{N}_2(s; z_2, \tilde{n}_1(s^-)).$$

Consistency and Cross-Worlds

- The potential time to intermediate event $\tilde{T}_1(z_1)$ is the time that $\tilde{N}_1(t; z_1)$ jumps.
- Substituting $\tilde{n}_1(\cdot)$ with $\tilde{N}_1(\cdot; z_1)$, the potential time to terminal event $\tilde{T}_2(z_1, z_2)$ is the time that $\tilde{N}_2(t; z_2, \tilde{N}_1(\cdot; z_1))$ jumps.
- **Causal consistency**:

$$\tilde{N}_1(t) = \tilde{N}_1(t; Z),\ \tilde{N}_2(t) = \tilde{N}_2(t; Z, \tilde{N}_1(\cdot; Z)),\ t \in [0, t^*].$$

- Under causal consistency, $\tilde{T}_1 = \tilde{T}_1(Z)$ and $\tilde{T}_2 = \tilde{T}_2(Z, Z)$.

Estimand

- The *counterfactual cumulative incidence* of the terminal event

$$F(t; z_1, z_2) = P(\tilde{T}_2(z_1, z_2) \leq t) = P(\tilde{N}_2(t; z_2, \tilde{N}_1(\cdot; z_1)) = 1)$$

  is of primary interest.
- To identify $F(t; z_1, z_2)$, it is equivalent to identify the hazard

$$d\Lambda(t; z_1, z_2) = d \log\{1 - F(t; z_1, z_2)\}.$$

- Total treatment effect: $F(t; 1, 1) - F(t; 0, 0)$.
- Natural direct effect: $F(t; 0, 1) - F(t; 0, 0)$.
- Natural indirect effect: $F(t; 1, 1) - F(t; 0, 1)$.

Outline

Assumptions for Identification

- **Ignorability**: the treatment is randomized,

$$\{d\tilde{N}_1(t; z_1), d\tilde{N}_2(t; z_2, n_1) : 0 \leq t \leq t^*\} \perp\!\!\!\perp Z.$$

- **Random censoring**:

$$\{\tilde{N}_1(t; z), \tilde{N}_2(t; z, \tilde{N}_1(\cdot; z)) : 0 \leq t \leq t^*\} \perp\!\!\!\perp C \mid Z.$$

- You may consider C as a potential outcome since C is a post-treatment value, in which case ignorability should be assumed for $C(z)$.

Assumptions for Identification

- **Positivity 1**:

$$P(\tilde{N}_1(t; z_1) = n_1, \tilde{N}_2(t; z_2, \tilde{N}_1(\cdot; z_1)) = 0) > 0$$
$$\Rightarrow P(Z = z \mid \tilde{N}_1(t; z) = n_1, \tilde{N}_2(t; z, \tilde{N}_1(\cdot; z)) = 0) > 0, \ z \in \{0, 1\},$$

  $z_1, z_2, n_1 \in \{0, 1\}, t \in [0, t^*]$.

- **Positivity 2**:

$$P(C > t^* \mid Z) > 0.$$

Observed Processes

- The event times $T_1 = \tilde{T}_1 \wedge C$ and $T_2 = \tilde{T}_2 \wedge C$.
- The censoring indicators $\delta_1 = I\{\tilde{T}_1 \leq C\}$ and $\delta_2 = I\{\tilde{T}_2 \leq C\}$.
- The observed counting process for the intermediate event $N_*(t; z) = I\{T_1 \leq t, T_2 \geq T_1, \delta_1 = 1, Z = z\}$.
- The observed at-risk process for the intermediate event $Y_*(t; z) = I\{T_1 \geq t, T_2 \geq t, Z = z\}$.
- Hereafter we use the subscript "$*$" to represent the intermediate event, "1" to represent the terminal event with prior intermediate event (direct terminal event), and "0" to represent the terminal event without prior intermediate event (indirect terminal event).

Observed Processes

- The observed counting process for the direct terminal event $N_0(t;z) = I\{T_2 \leq t, T_1 \geq T_2, \delta_2 = 1, Z = z\}$.
- The observed at-risk process for the direct terminal event $Y_0(t;z) = I\{T_2 \geq t, T_1 \geq t, Z = z\}$.
- The observed counting process for the indirect terminal event $N_1(t;z) = I\{T_2 \leq t, T_1 < t, \delta_2 = 1, Z = z\}$.
- The observed at-risk process for the indirect terminal event $Y_1(t;z) = I\{T_2 \geq t, T_1 < t, Z = z\}$.
- $Y_*(t;z) = Y_0(t;z)$ because the intermediate event and the direct terminal event are a pair of competing events, sharing the same at-risk set.

Cause-Specific Hazards

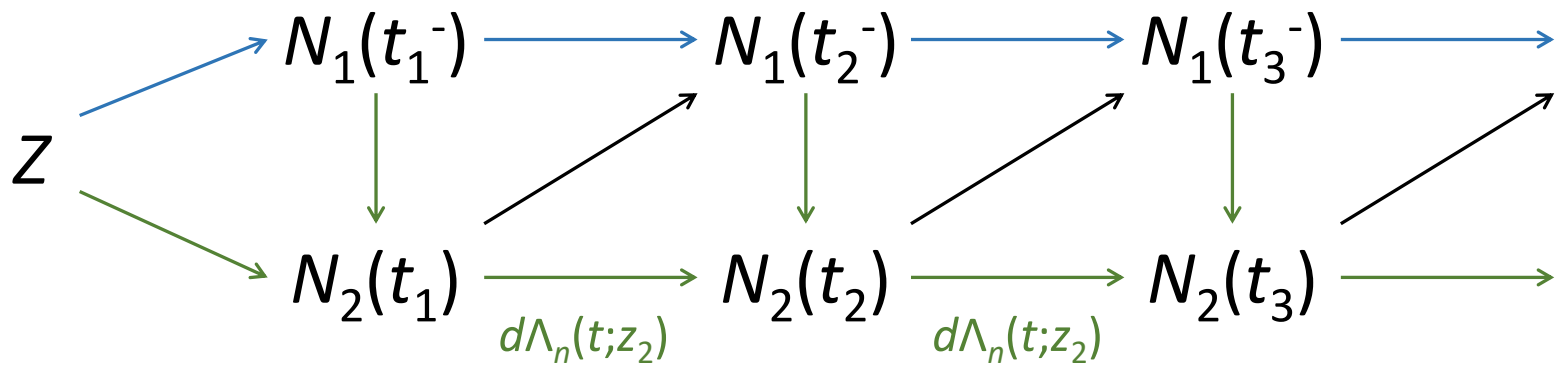
- Define the cause-specific hazard of the terminal event

$$
\begin{aligned}
&d\Lambda_{n_1}(t; z_1, z_2) \\
&= P(d\tilde{N}_2(t; z_2, n_1) = 1 \mid \tilde{N}_1(t^-; z_1) = n_1, \tilde{N}_2(t^-; z_2, \tilde{N}_1(z_1)) = 0).
\end{aligned}
$$

- It involves cross-world values. How to identify the cause-specific hazard?
- Recall the assumptions we have made: Markovness, consistency, ignorability, random censoring, and positivity.
- Compared to mediation analysis, what assumption is lost? — Sequential ignorability!

Competing Events in the Cross-World

- Note that $\tilde{N}_1(\cdot)$ and $\tilde{N}_2(\cdot)$ are competing. In the cross-world (z_1, z_2), $\tilde{N}_2(\cdot)$ relies on z_1 and $\tilde{N}_1(\cdot)$ relies on z_2.
- We need two sequential ignorability assumptions, one for $\tilde{N}_2(\cdot)$ and the other for $\tilde{N}_1(\cdot)$.



Sequential Ignorability, Part 1

- **Sequential ignorability (1)**:

$$P(d\tilde{N}_2(t; z_2, n_1) = 1 \mid Z = z_2, \tilde{N}_1(t^-; z_1) = n_1, \tilde{N}_2(t^-; z_2, \tilde{N}_1(\cdot; z_1)) = 0)$$
$$= P(d\tilde{N}_2(t; z_2, n_1) = 1 \mid Z = z_2, \tilde{N}_1(t^-) = n_1, \tilde{N}_2(t^-) = 0).$$

- At time t, given the "baseline" Z, the intermediate event status $\tilde{N}_1(t^-) = n_1$, and the fact that the terminal event has not occurred $\tilde{N}_2(t^-) = 0$, the "treatment" $d\tilde{N}_2(t; Z, n_1)$ is independent of the potential process $\tilde{N}_1(\cdot; z_1)$.
- This sequential ignorability excludes the cross-world reliance of $d\tilde{N}_2(t; z_2, n_1)$ on z_1.

Sequential Ignorability, Part 1

- The cause-specific hazard (of the terminal event) is identifiable

$$
\begin{aligned}
& d\Lambda_{n_1}(t; z_1, z_2) \\
&= P(d\tilde{N}_2(t; z_2, n_1) = 1 \mid \tilde{N}_1(t^-; z_1) = n_1, \tilde{N}_2(t^-; z_2, \tilde{N}_1(\cdot; z_1)) = 0) \\
&= P(d\tilde{N}_2(t; z_2, n_1) = 1 \mid Z = z_2, \tilde{N}_1(t^-) = n_1, \tilde{N}_2(t^-) = 0)
\end{aligned}
$$

and can be estimated by Nelson–Aalen estimators,

$$
d\hat{\Lambda}_{n_1}(t; z_2) = \frac{I\{\bar{Y}_{n_1}(t; z_2) > 0\}}{\bar{Y}_{n_1}(t; z_2)} d\bar{N}_{n_1}(t; z_2).
$$

Sequential Ignorability, Part 2

- **sequential ignorability (2)**:

$$\begin{aligned} & P(d\tilde{N}_1(t; z_1) = 1 \mid Z = z_1, \tilde{N}_1(t^-) = 0, \tilde{N}_2(t^-; z_2, \tilde{N}_1(\cdot; z_1)) = 0) \\ = \; & P(d\tilde{N}_1(t; z_1) = 1 \mid Z = z_1, \tilde{N}_1(t^-) = 0, \tilde{N}_2(t^-) = 0). \end{aligned}$$

- At time t, given the "baseline" Z, the intermediate event status $\tilde{N}_1(t^-) = 0$, and the fact that the terminal event has not occurred $\tilde{N}_2(t^-) = 0$, the "treatment" $d\tilde{N}_1(t; Z)$ is independent of the potential process $\tilde{N}_2(\cdot; z_2, \tilde{N}_1(\cdot; z_1))$.
- This sequential ignorability excludes the cross-world reliance of $d\tilde{N}_1(t; z_1)$ on z_2.

Sequential Ignorability, Part 2

- The hazard of the intermediate event is identifiable

$$
\begin{aligned}
& d\Lambda_*(t; z_1, z_2) \\
&= P(d\tilde{N}_1(t; z_1) = 1 \mid \tilde{N}_1(t^-; z_1) = 0, \tilde{N}_2(t^-; z_2, \tilde{N}_1(\cdot; z_1)) = 0) \\
&= P(d\tilde{N}_1(t; z_1) = 1 \mid Z = z_1, \tilde{N}_1(t^-) = 0, \tilde{N}_2(t^-) = 0)
\end{aligned}
$$

  and can be estimated by Nelson–Aalen estimators,

$$
d\hat{\Lambda}_*(t; z_1) = \frac{I\{\bar{Y}_*(t; z_1) > 0\}}{\bar{Y}_*(t; z_1)} dN_*(t; z_1).
$$

Subdistribution

- We partition the terminal event into a direct event (which does not have a history of intermediate event) and an indirect event (which has a history of intermediate event).
- The cumulative incidence of the terminal event consists of two subdistributions,

$$F(t; z_1, z_2) = F_0(t; z_1, z_2) + F_1(t; z_1, z_2),$$

where

$$\begin{aligned} & F_0(t; z_1, z_2) = P(\tilde{N}_2(t; z_2, \tilde{N}_1(\cdot; z_1)) = 1, \tilde{N}_1(t; z_1) = 0) \\ = & \int_0^t \exp\{-\Lambda_*(s^-; z_1) - \Lambda_0(s^-; z_2)\} d\Lambda_0(s; z_2), \\ & F_1(t; z_1, z_2) = P(\tilde{N}_2(t; z_2, \tilde{N}_1(\cdot; z_1)) = 1, \tilde{N}_1(t; z_1) = 1) \\ = & \int_0^t \exp\{-\Lambda_*(s^-; z_1) - \Lambda_0(s^-; z_2)\}[1 - \exp\{-\Lambda_1(t; z_2) + \Lambda_1(s; z_2)\}] d\Lambda_*(s; z_1). \end{aligned}$$

Interpretation: Controlling the Hazard

- The natural direct effect measures the treatment effect on the cumulative incidence of terminal event via changing the cause-specific hazards of terminal events while controlling the *hazard* of intermediate events.
- The natural indirect effect measures the treatment effect on the cumulative incidence of terminal event via changing the *hazard* of intermediate events while controlling the cause-specific hazards of terminal events.
- Nevertheless, since we have no information of the cross-world, whether the assumption is appropriate is worthy of extensive discussion.

Outline

Limitation of Natural Effects

- Sequential ignorability is untestable in principal.
- The simple notation of potential counting processes stands on Markovness, which is not required for identification.
- The notation in sequential ignorability (Part 2) is confusing:

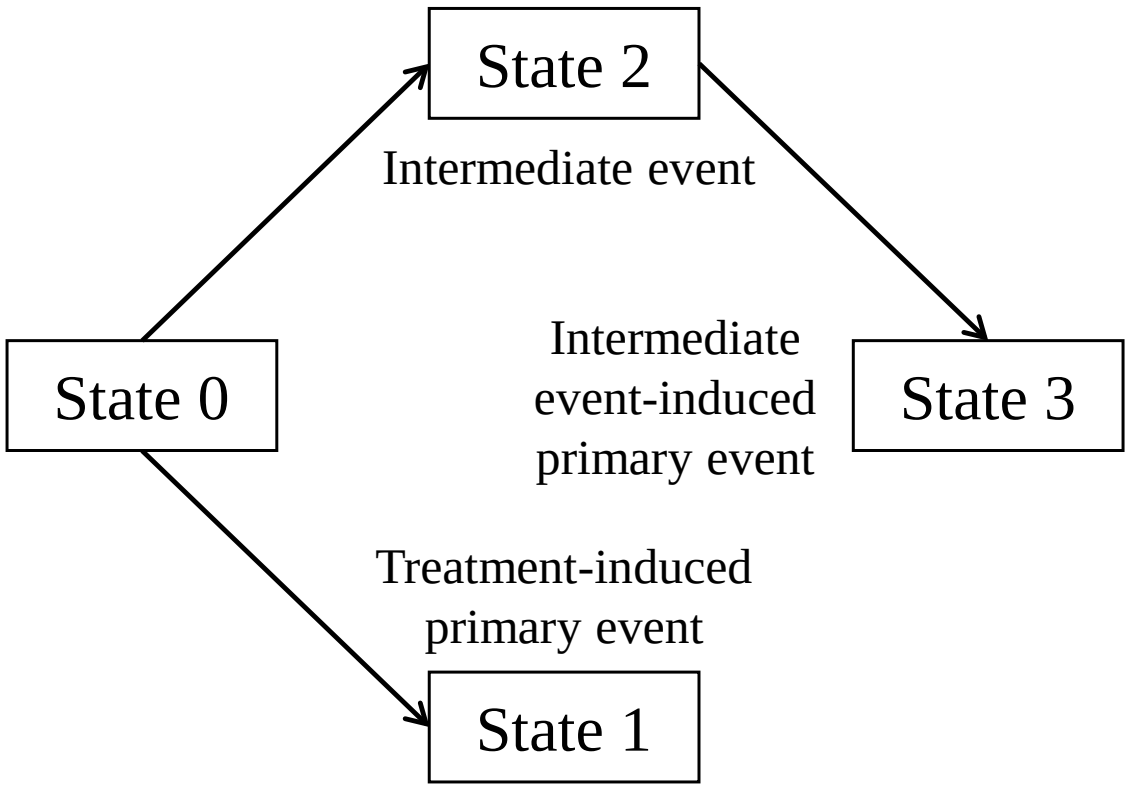
$$P(d\tilde{N}_1(t; z_1) = 1 \mid \tilde{N}_1(t^-; z_1) = 0, \tilde{N}_2(t^-; z_2, \tilde{N}_1(\cdot; z_1)) = 0).$$

- The natural effects are hard to interpret.

Separable Effects

- We adopt the separable effects framework (Stensrud et al, 2021, 2022).
- A: Binary treatment.
- $\tilde{T}^a$: Potential time to primary event.
- $\tilde{R}^a$: Potential time to intermediate event.
- C^a: Potential censoring time.
- X: Baseline covariates.
- The treatment effect should be defined by contrasting well-defined distributions of $\tilde{T}^1$ and $\tilde{T}^0$, appropriately adjusting $\tilde{R}^1$ and $\tilde{R}^0$.

Terminal Events Developed from Different Sources



Assumptions

- Assumption 1 (Ignorability):

$$(\tilde{T}^a, \tilde{R}^a, C^a) \perp\!\!\!\perp A \mid X.$$

- Assumption 2 (Random censoring):

$$I(t \leq C^a < t + dt) \perp\!\!\!\perp (\tilde{T}^a, \tilde{R}^a) \mid \mathcal{F}(t).$$

- Assumption 3 (Positivity):

$$c < P(A = a \mid X) < 1 - c,$$

$$P(\tilde{T}^a > t^*, C^a > t^* \mid A = a, \mathcal{F}(t)) > 0.$$

Potential Hazards (Transition Rates)

- Potential hazards (transition rates) of direct outcome event (State 1), intermediate event (State 2) and indirect outcome event through intermediate event occurring at time r (State 3) at time t:

$$d\Lambda_1^a(t \mid \mathcal{F}(t)) := P(t \leq \tilde{T}^a < t + dt, \tilde{T}^a < \tilde{R}^a \mid \tilde{T}^a \geq t, \tilde{R}^a \geq t, \mathcal{F}(t)),$$
$$d\Lambda_2^a(t \mid \mathcal{F}(t)) := P(t \leq \tilde{R}^a < t + dt, \tilde{R}^a \leq \tilde{T}^a \mid \tilde{T}^a \geq t, \tilde{R}^a \geq t, \mathcal{F}(t)),$$
$$d\Lambda_3^a(t \mid \mathcal{F}(t)) := P(t \leq \tilde{T}^a < t + dt, \tilde{R}^a \leq t \mid \tilde{T}^a \geq t, \tilde{R}^a \leq t, \mathcal{F}(t)).$$

Decomposing Treatment Components

- Suppose the treatment can be decomposed into three components (A_1, A_2, A_3), where A_j only has an effect on the hazard of State j.
- Then, potential hazards $\{d\Lambda_j^a(\cdot) : j = 1, 2, 3\}$ can be written as functions of treatment components $a = (a_1, a_2, a_3)$.
- In a realized trial, observed treatment components $A_1 = A_2 = A_3$, equal to the actual treatment A.
- In a hypothetical world, treatment components can take different values.

Dismissible Components Condition

- Assumption 4 (Dismissible components):

$$d\Lambda_j^{(a_1,a_2,a_3)}(t \mid \mathcal{F}(t)) = d\Lambda_j^{a_j}(t \mid \mathcal{F}(t)).$$

- When $a_1 = a_2 = a_3$, this assumption is naturally satisfied because no hypothetical worlds are involved.
- Covariates isolation: $\mathcal{F}(t)$ only includes prior states history rather than covariates,

$$\begin{aligned}
d\Lambda_1^a(t \mid \mathcal{F}(t)) &= d\Lambda_1^a(t \mid \emptyset) := d\Lambda_1^a(t), \\
d\Lambda_2^a(t \mid \mathcal{F}(t)) &= d\Lambda_2^a(t \mid \emptyset) := d\Lambda_2^a(t), \\
d\Lambda_3^a(t \mid \mathcal{F}(t)) &= d\Lambda_3^a(t \mid R^a = r) := d\Lambda_3^a(t; r).
\end{aligned}$$

Counterfactual Cumulative Incidences

- Denote counterfactual cumulative incidences for three states by $F_1^a(t) = P(\tilde{T}^a \leq t, \tilde{T}^a < \tilde{R}^a)$, $F_2^a(t) = P(\tilde{R}^a \leq t, \tilde{R}^a \leq \tilde{T}^a)$ and $F_3^a(t) = P(\tilde{T}^a \leq t, \tilde{R}^a \leq \tilde{T}^a)$, respectively.
- The counterfactual cumulative incidence of primary event becomes $F^a(t) = P(\tilde{T}^a \leq t) = F_1^a(t) + F_3^a(t)$, with

$$F_1^{a=(a_1,a_2,a_3)}(t) = \int_0^t \exp\{-\Lambda_1^{a_1}(s) - \Lambda_2^{a_2}(s)\} d\Lambda_1^{a_1}(s),$$

$$F_2^{a=(a_1,a_2,a_3)}(t) = \int_0^t \exp\{-\Lambda_1^{a_1}(s) - \Lambda_2^{a_2}(s)\} d\Lambda_2^{a_2}(s),$$

$$F_3^{a=(a_1,a_2,a_3)}(t) = \int_0^t \exp\{-\Lambda_1^{a_1}(s) - \Lambda_2^{a_2}(s)\}[1 - \exp\{-\Lambda_3^{a_3}(t;s)\}] d\Lambda_2^{a_2}(s).$$

Definition of Separable Pathway Effects

- The total treatment effect is decomposed as

$$
\begin{aligned}
& F^{a=(1,1,1)}(t) - F^{a=(0,0,0)}(t) \\
&= \{F^{a=(1,0,0)}(t) - F^{a=(0,0,0)}(t)\} + \{F^{a=(1,1,1)}(t) - F^{a=(1,0,0)}(t)\} \\
&:= \mathrm{SPE}_{0\to 1}(t;0,0) + \mathrm{SPE}_{0\to 3}(t;1) \\
&= \{F^{a=(1,0,0)}(t) - F^{a=(0,0,0)}(t)\} \\
&\quad + \{F^{a=(1,1,0)}(t) - F^{a=(1,0,0)}(t)\} \\
&\quad + \{F^{a=(1,1,1)}(t) - F^{a=(1,1,0)}(t)\} \\
&:= \mathrm{SPE}_{0\to 1}(t;0,0) + \mathrm{SPE}_{0\to 2}(t;1,0) + \mathrm{SPE}_{2\to 3}(t;1,1).
\end{aligned}
$$

Weighted Counting Processes

- Let $w_i(a_j) = I\{A_i = a_j\}/P(A_i = a_j \mid X_i)$.
- Define weighted counting processes, at-risk processes and residuals with respect to $d\Lambda_1^{a_1}(t)$ and $d\Lambda_2^{a_2}(t)$ as follows:

$$N_1(t; a_1) = \sum_{i=1}^{n} w_i(a_1) I\{T_i \le t, R_i > t, \delta_i^T = 1\},$$

$$N_2(t; a_2) = \sum_{i=1}^{n} w_i(a_2) I\{R_i \le t, T_i \ge t, \delta_i^R = 1\},$$

$$Y_j(t; a_j) = \sum_{i=1}^{n} w_i(a_j) I\{T_i \ge t, R_i \ge t\},$$

$$Y_j^*(t; a_j) = \sum_{i=1}^{n} w_i(a_j)^2 I\{T_i \ge t, R_i \ge t\},$$

$$M_j(t; a_j) = \int_0^t \left\{ dN_j(s; a_j) - Y_j(s; a_j) d\Lambda_j^{a_j}(s) \right\}, \ j = 1, 2.$$

Weighted Counting Processes

- To yield well-defined (nonparametric) estimators for $\Lambda_3^{a_3}(t;s)$, processes $N_3(t;r,a_3)$ and $Y_3(t;r,a_3)$ should be refined so that $Y_3(t;r,a_3)$ is nonzero and

$$M_3(t;r,a_3) = \int_r^t \{dN_3(s;r,a_3) - Y_3(s;r,a_3)d\Lambda_3^{a_3}(s;r)\}$$

  is a martingale with respect to some filter.
- For example, we can assume Markovness or semi-Markovness for the transition from State 2 to State 3.

Markovness

- Under Markov assumption $d\Lambda_3^{a_3}(t;r) = d\Lambda_{3,\text{ma.}}^{a_3}(t)$, let

$$N_3(t;r,a_3) = N_{3,\text{ma.}}(t;a_3) = \sum_{i=1}^{n} w_i(a_3) I\{T_i \le t, \delta_i^T \delta_i^R = 1\},$$

$$Y_3(t;r,a_3) = Y_{3,\text{ma.}}(t;a_3) = \sum_{i=1}^{n} w_i(a_3) I\{T_i \ge t, R_i \le t, \delta_i^R = 1\},$$

$$Y_3^*(t;r,a_3) = Y_{3,\text{ma.}}^*(t;a_3) = \sum_{i=1}^{n} w_i(a_3)^2 I\{T_i \ge t, R_i \le t, \delta_i^R = 1\},$$

$$M_3(t;r,a_3) = M_{3,\text{ma.}}(t;a_3) = \int_0^t \left\{ dN_{3,\text{ma.}}(s;a_3) - Y_{3,\text{ma.}}(s;a_3) d\Lambda_{3,\text{ma.}}^{a_3}(s) \right\}.$$

Semi-Markovness

- Under semi-Markov assumption $d\Lambda_3^{a_3}(t;r) = d\Lambda_{3,\text{sm.}}^{a_3}(t-r)$, let

$$N_3(t;r,a_3) = N_{3,\text{sm.}}(u;a_3) = \sum_{i=1}^{n} w_i(a_3) I\{T_i - R_i \le u, \delta_i^T \delta_i^R = 1\},$$

$$Y_3(t;r,a_3) = Y_{3,\text{sm.}}(u;a_3) = \sum_{i=1}^{n} w_i(a_3) I\{T_i - R_i \ge u, \delta_i^R = 1\},$$

$$Y_3^*(t;r,a_3) = Y_{3,\text{sm.}}^*(u;a_3) = \sum_{i=1}^{n} w_i(a_3)^2 I\{T_i - R_i \ge u, \delta_i^R = 1\},$$

$$M_3(t;r,a_3) = M_{3,\text{sm.}}(u;a_3) = \int_0^u \left\{ dN_{3,\text{sm.}}(s;a_3) - Y_{3,\text{sm.}}(s;a_3) d\Lambda_{3,\text{sm.}}^{a_3}(s) \right\}.$$

Estimators for Cumulative Hazards

- The generalized Nelson–Aalen estimators for cumulative hazards are

$$\hat{\Lambda}_1^{a_1}(t) = \int_0^t \frac{dN_1(s; a_1)}{Y_1(s; a_1)}, \ \hat{\Lambda}_2^{a_2}(t) = \int_0^t \frac{dN_2(s; a_2)}{Y_2(s; a_2)},$$
$$\hat{\Lambda}_3^{a_3}(t; r) = \int_r^t \frac{dN_3(s; r, a_3)}{Y_3(s; r, a_3)}.$$

- To use the martingale theory to establish asymptotic properties for the incidence estimator, we need to assume the propensity score is known.
- Then

$$\sqrt{n}\{\hat{F}^a(t) - F^a(t)\} \xrightarrow{d} G_1^a(t) + G_2^a(t) + G_3^a(t).$$

Outline

Realxed Assumptions

- We may relax the assumptions and improve efficiency using semiparametric estimation.
- Censoring is random conditional on covariates.
- The treatment components are dismissible conditional on covariates:

$$
\begin{aligned}
d\Lambda_1^a(t \mid \mathcal{F}(t)) &= d\Lambda_1^a(t; x), \\
d\Lambda_2^a(t \mid \mathcal{F}(t)) &= d\Lambda_2^a(t; x), \\
d\Lambda_3^a(t \mid \mathcal{F}(t)) &= d\Lambda_3^a(t; r, x).
\end{aligned}
$$

- We assume that there is a subset $\mathcal{H}(t; r) \subset \{(\tilde{T}, \tilde{R})\}$ such that $d\Lambda_3^{a_3}(t; r, x)$ is identical for $(t, r) \in \mathcal{H}(t, r)$ given $X = x$.
- Markvoness implies $\mathcal{H}(t; r) = \{(\tilde{T}, \tilde{R}) : \tilde{T} \geq t, \tilde{R} \leq t\}$.
- Semi-Markovness implies $\mathcal{H}(t; r) = \{(\tilde{T}, \tilde{R}) : \tilde{T} - \tilde{R} \geq t - r\}$.

Efficient Influence Function

- The efficient influence function (EIF) of $F^{(a_1,a_2,a_3)}(t)$ is

$$
\begin{aligned}
\varphi^a(t) = \int_0^t \exp\{-\Lambda_1^{a_1}(s;X)\} \Bigg\{ & \frac{I(A=a_1)}{P(A=a_1 \mid X)} \frac{dM_1(s;A,X)}{P(T \wedge R \geq s \mid A,X)} \\
& - \sum_{j\in\{1,2\}} \frac{I(A=a_j)}{P(A=a_j \mid X)} \int_0^s \frac{dM_j(u;A,X)}{P(T \wedge R \geq u \mid A,X)} d\Lambda_1^{a_1}(s;X) \Bigg\} \\
& + \int_0^t \int_0^s \exp\{-\Lambda_1^{a_1}(r;X) - \Lambda_2^{a_2}(r;X) - \Lambda_3^{a_3}(s;r,X)\} \\
& \Bigg\{ \frac{I(A=a_2)}{P(A=a_2 \mid X)} \frac{dM_2(r;A,X)}{P(T \wedge R \geq r)} d\Lambda_3^{a_3}(s;r,X) \\
& - \sum_{j\in\{1,2\}} \frac{I(A=a_j)}{P(A=a_j \mid X)} \int_0^r \frac{dM_j(u;A,X)}{P(T \wedge R \geq u \mid A,X)} d\Lambda_2^{a_2}(r;X) d\Lambda_3^{a_3}(s;r,X) \\
& + \frac{I(A=a_3)}{P(A=a_3 \mid X)} \frac{dM_3(s;R,A,X)}{P(\mathcal{H}(s,R) \mid A,X)} d\Lambda_2^{a_2}(s;X) \\
& - \frac{I(A=a_3)}{P(A=a_3 \mid X)} \int_{\mathcal{H}(s,R)} \frac{dM_3(u;R,A,X)}{P(\mathcal{H}(u,R) \mid A,X)} d\Lambda_2^{a_2}(r;X) d\Lambda_3^{a_3}(s;r,X) \Bigg\} \\
& + F^{(a_1,a_2,a_3)}(t;X) - F^{(a_1,a_2,a_3)}(t).
\end{aligned}
$$

Estimator Based on EIF

- The EIF-based estimator $\tilde{F}^a(t)$ is obtained by solving $P_n\{\hat{\varphi}^a(t)\} = 0$.
- Asymptotic normality under some regularity conditions:

$$\sqrt{n}\{\tilde{F}^a(t) - F^a(t)\} \xrightarrow{d} N\left(0, E[\varphi^a(t)]^2\right).$$

- The EIF-based estimator has multiple robustness:
  (1) one of the three hazards is misspecified;
  (2) the propensity score and censoring hazard are misspecified.

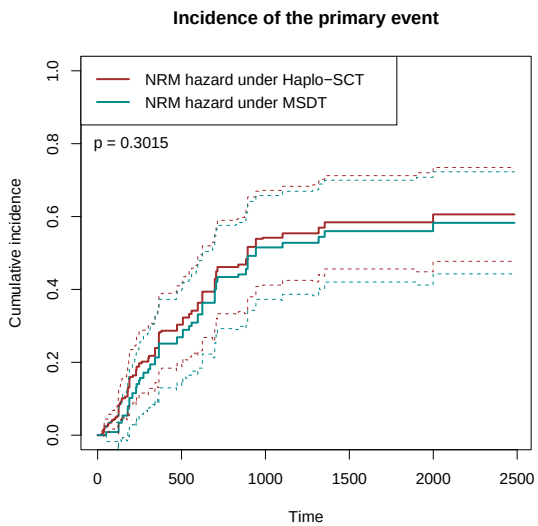
Outline

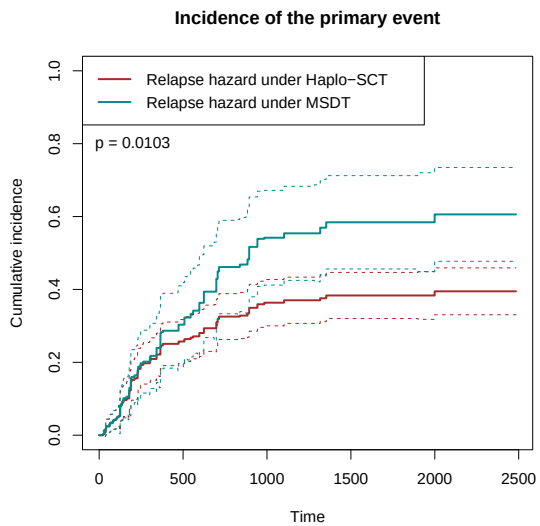
Leukemia Data

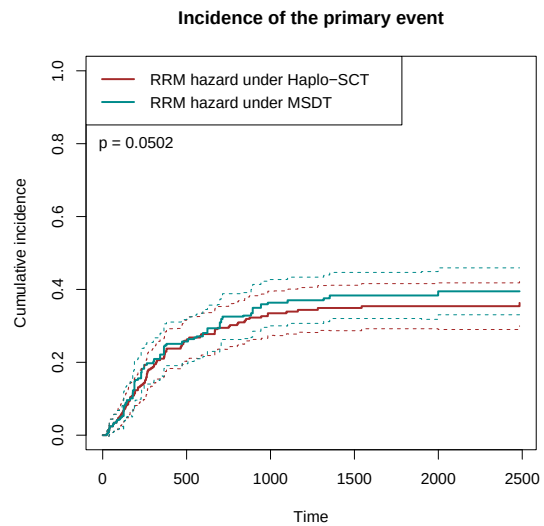
- Allogeneic stem cell transplantation is a commonly adopted approach to cure acute lymphoblastic leukemia.
- Human leukocyte antigen matched sibling donor transplantation (MSDT) has long been considered as the first choice of transplantation because of lower transplant-related mortality.
- In recent years, some benefits of haploidentical stem cell transplantation (haplo-SCT) have been found in that haplo-SCT results in lower relapse rate.
- We are interested in the mechanism of transplantation types on mortality.
- Sample size $n = 303$.

Estimating Separable Effects

- Estimation based on the covariates isolation (Markovian).
- *P*-values given by the weighted logrank test.
- Haplo-SCT reduces the overall mortality through reducing the risk (hazard) of relapse.

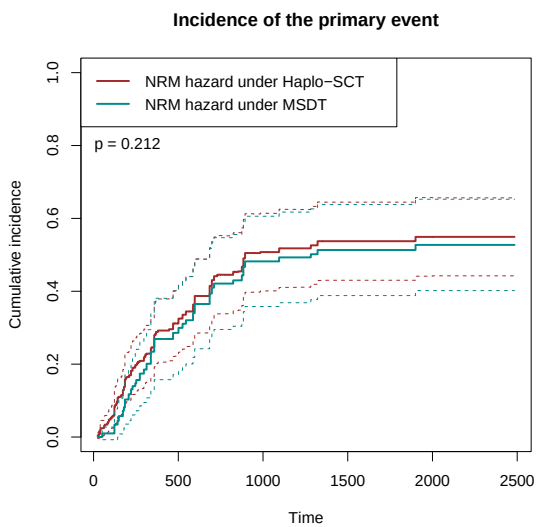


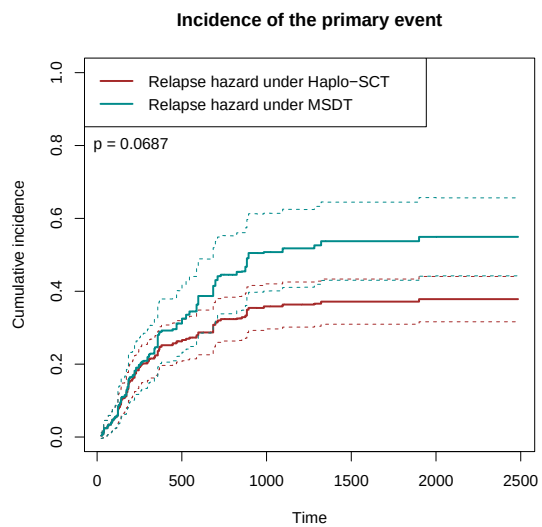


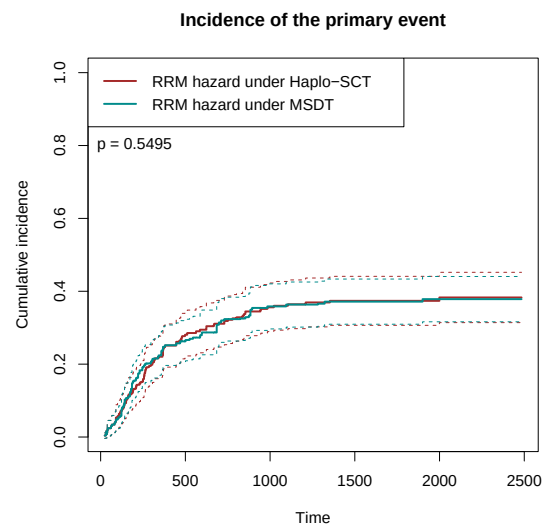


Estimating Separable Effects

- Estimation based on the efficient influence function (Markovian).
- P-values given by the test statistic $\int_0^{t^*} \{\hat{F}^a(s) - \hat{F}^{a'}(s)\} d\{\hat{F}^a(s) + \hat{F}^{a'}(s)\}$.
- Haplo-SCT reduces the overall mortality through reducing the risk (hazard) of relapse.







Testing Separable Effects

Test	Interpretation	P-value
Total	The total effect on mortality	0.5341
$\text{SPE}_{0\to1}$	The separable pathway effect via NRM	0.2120
$\text{SPE}_{0\to2}$	The separable pathway effect via relapse	0.0687
$\text{SPE}_{2\to3}$	The separable pathway effect via RRM	0.5495
$\text{SPE}_{0\to3}$	The separable pathway effect through the path relapse–RRM (via relapse and RRM)	0.0676
$\text{SPE}_{0\to1,2\to3}$	The separable pathway effect on mortality (via NRM and RRM)	0.4281

Outline

Summary

- We discussed the mediation analysis for time-to-event data, also known as semi-competing risks.
- A significant challenge for time-to-event data is that the outcome variables are processes. The potential outcomes are defined on counting processes.
- The pointwise variance of estimators can be derived by the stochastic processes theory.
- An interventional approach to mediation analysis, the separable effects framework, provides an easy-to-interpret solution to semi-competing risks.

Extensions

- Interaction effects.
  This problem seems to have been overcome by separable effects.
- Time-varying confounders.
  Identification can be achieved by g-formula with dismissible components of confounders.
- Unobserved confounders.
  We may assume a frailty (random effect).
- Interventional effects.
  A key difference with conventional mediation analysis: the risk of the intermediate event is undefined after the terminal event.
- Multiple intermediate events (multi-state model).
  The separable effects framework can simplify problems.

References

- Natural effects
  Deng Y, Wang Y, Zhou X-H. Direct and indirect effects in the presence of semi-competing risks. *Biometrics*. 2024.
- Separable effects
  Deng Y, Wang Y, Zhan X, Zhou X-H. Separable pathway effects of semi-competing risks using multi-State models. *ArXiv*. 2023.
- The influence of estimated propensity score in hazard estimation
  Deng Y, Wang R. The generalized Nelson–Aalen estimator by inverse treatment probability weighting with a possibly unknown propensity score. *ArXiv*. 2024.
- Seperable effects with time-varying confounders
  Deng Y, Wei H, Xiao X, Zhang Y, Huang Y. Sequential ignorability and dismissible treatment components to identify mediation effects. *Mathematics*. 2024.

Acknowledgements

- Credits to co-authors:
  Xiao-Hua Zhou (Peking University),
  Yi Wang (Shanghai University of Business and Economics),
  Rui Wang (University of Washington),
  Haoyu Wei (University of California, San Diego).
- We thank Dr. Yingjun Chang and Dr. Leqing Cao for collecting and cleaning the data.

- Contact: yuhaoden@umich.edu