Mediation Analysis for Time-to-Event Data

Deng, Yuhao
yuhaoden@umich.edu

University of Michigan

April 15, 2024

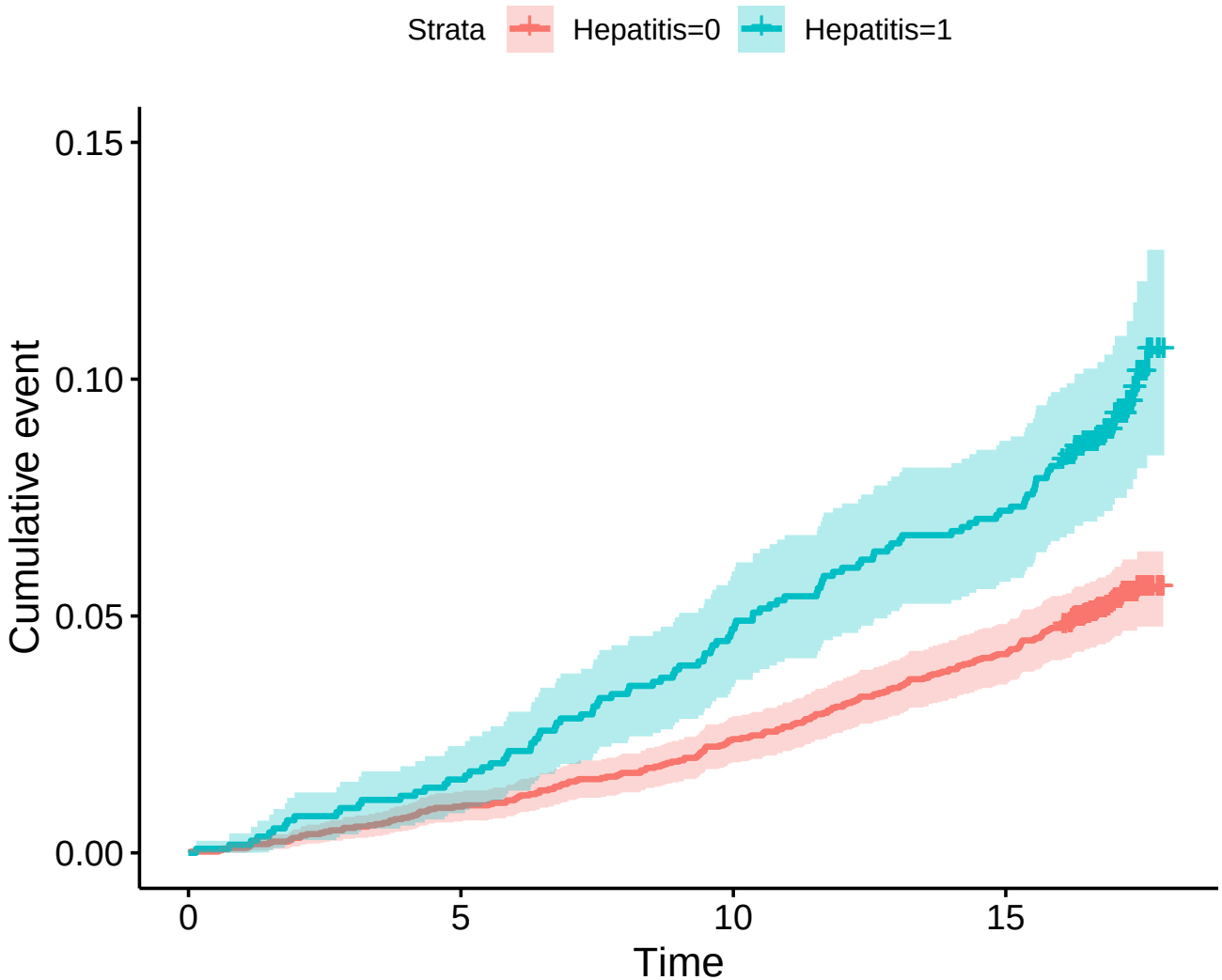
Outline

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?

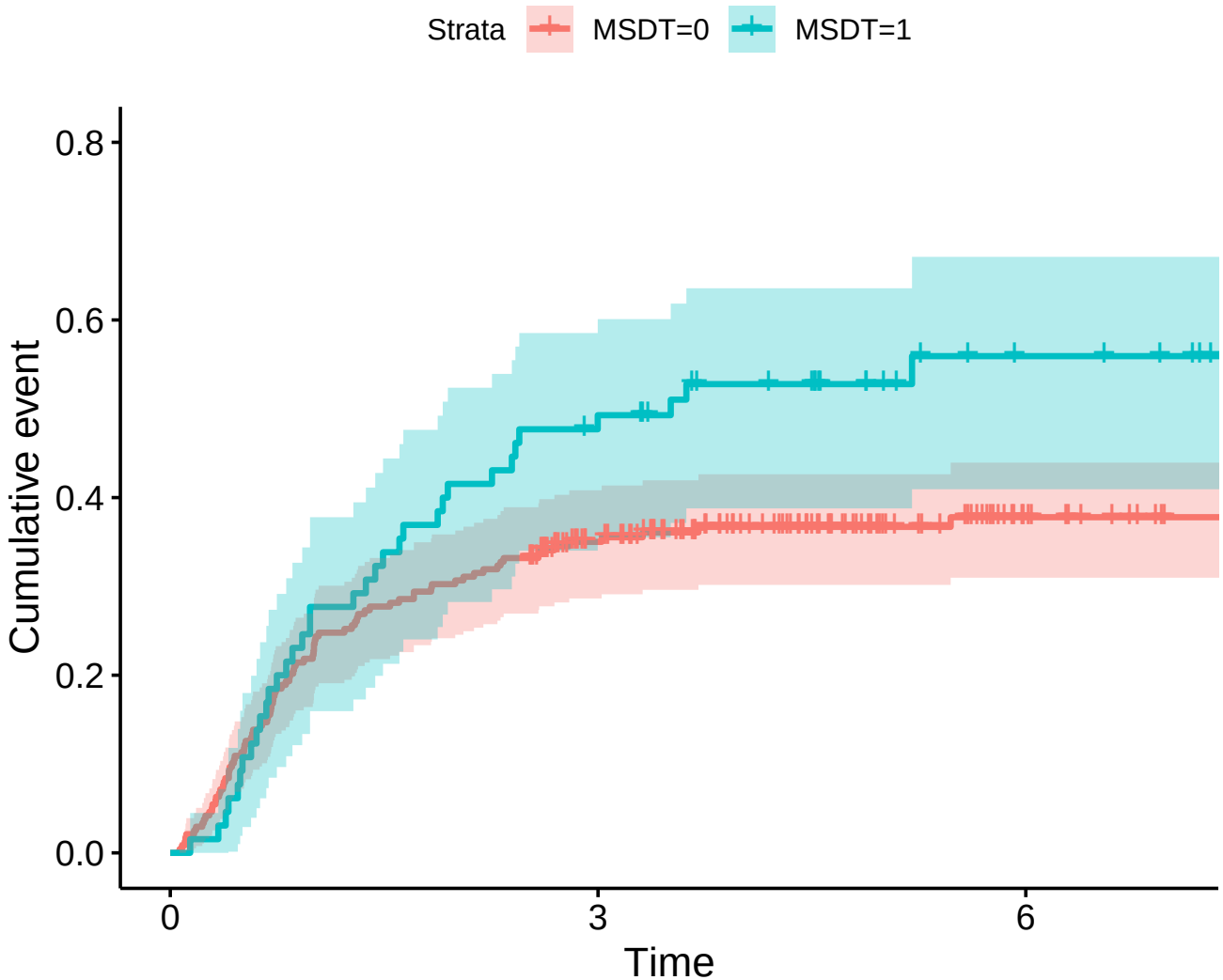
Motivating Data 1



Motivating Data 2

- Allogeneic stem cell transplantation is a commonly adopted approach to cure acute lymphoblastic leukemia.
- 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.
- Lower relapse rate means lower relapse-related mortality.
- What is the mechanism of transplant modality on overall mortality?

Motivating Data 2



Question

- The outcome is "time to event", subject to censoring.
- There is a terminal event and an intermediate event.
- 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 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" is an intermediate event and the "outcome" is a terminal event.
- The mediator process $N_1(t) \in \{0, 1\}$: whether the intermediate event has occurred at t.
- The outcome 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).
- The terminal event and non-terminal events have a competing nature, but their status is not equal.
- Some individuals develop terminal events after intermediate events, while others develop terminal events without intermediate events.
- For example, in the study of stem cell transplantation, mortality is a terminal event, and relapse is an intermediate event.

Formal Definition: Jumps

- We need to introduce some counting processes to formally define the semi-competing risks.
- Let $d\tilde{N}_1(t; z_1)$ be the jump of potential 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 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: Processes

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

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

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, consistent with the purpose of SUTVA — there is only one version of treatment.
- 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 $d\tilde{N}_2(t; z_2, n_1)$.

Consistency and Cross-Worlds

- The potential time to intermediate event $T_1(z_1)$ 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 $T_2(z_1, z_2)$ 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, $T_1 = T_1(Z)$ and $T_2 = T_2(Z, Z)$.

Consistency and Cross-Worlds

- For time-to-event outcomes, different from the classical mediation analysis, the intermediate event and terminal event are competing.
- Rigorously, $\tilde{N}_1(\cdot)$ (or T_1) is a potential outcome of z_1 and z_2. If we manipulate the treatment associated with the terminal event z_2, then the time to the intermediate event $T_1(z_1, z_2)$ is not $T_1(z_1)$ anymore.
- With a terminal event process $\tilde{n}_2(\cdot)$, the potential counting process for the intermediate event should be $\tilde{N}_1(t; z_1, \tilde{n}_2(\cdot))$, where $\tilde{n}_2(\cdot)$ can be $\tilde{N}_2(\cdot; z_2, \tilde{N}_1(\cdot; z_1))$ in the cross-world (z_1, z_2).
- However, we do not need to involve such complicated notations. As we will see later, only the single-world $\tilde{N}_1(t; z_1)$ $(T_1(z_1))$ is required for identification.

Estimand

- The *counterfactual cumulative incidence* of the terminal event

$$F(t; z_1, z_2) = P(T_2(z_1, z_2) \leq t) = P(\tilde{N}_2(t; z_2, \tilde{N}_1(\cdot; z_1)) = 1)$$

  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,

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

- The treatment Z is independent of potential values (generated from potential jumps).
- **Random censoring**:

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

- The censoring is independent of counting processes given Z.
- You may as well consider C as a potential outcome since C is a post-treatment value, in which case ignorability should be imposed for $C(z)$.

Assumptions for Identification

- **Positivity 1**:

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

  $z_1, z_2, n_1 \in \{0, 1\}, t \in [0, t^*]$.
- As long as the potential status is possible, we should have corresponding data.
- **Positivity 2**:

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

- The censoring time is large enough, so we have data to identify the functionals prior to t^*.

Observed Processes

- The censoring indicators for the intermediate event and terminal event $\delta_1 = I\{T_1 \leq C\}$ and $\delta_2 = I\{T_2 \leq C\}$, respectively.
- 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 $Y_*(t; z) = I\{T_1 \geq t, T_2 \geq t, C \geq t, Z = z\}$.
- 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 $Y_0(t; z) = I\{T_2 \geq t, T_1 \geq t, C \geq t, Z = z\}$.
- 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 $Y_1(t; z) = I\{T_2 \geq t, T_1 < t, C \geq t, Z = z\}$.
- $Y_*(t; z) = Y_0(t; z)$ because the intermediate event and the direct terminal event are a pair of competing events, sharing the same at-risk set.

Cause-Specific Hazards

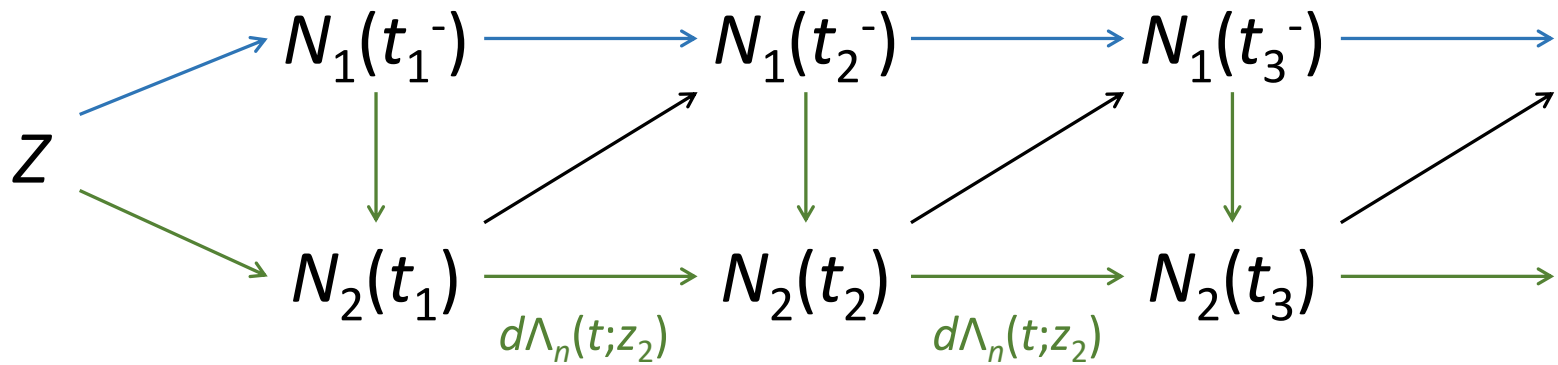
- Define the cause-specific hazard of the terminal event

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

- 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.

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**:

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

- 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 $\tilde{N}_1(\cdot; z_1)$.
- Whether $\tilde{N}_2(t)$ jumps does not affect when $\tilde{N}_1(\cdot; z_1)$ jumps in the future.
- 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,

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

Controlling the Prevalence

- What about the sequential ignorability for $\tilde{N}_1(t)$?
- Huang (2021) proposed the following version:

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

- The prevalence of intermediate events

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

  is irrelevant to z_2.
- An estimator for the prevalence:

$$\hat{w}_{n_1}(t; z_1) = \frac{I\{\bar{Y}_0(t; z_1) + \bar{Y}_1(t; z_1) > 0\}}{\bar{Y}_0(t; z_1) + \bar{Y}_1(t; z_1)} \bar{Y}_{n_1}(t; z_1).$$

Controlling the Prevalence

- Under all the assumptions above,

$$\frac{d}{dt}\Lambda(t; z_1, z_2) = \sum_{n_1 \in \{0,1\}} \frac{d}{dt}\Lambda_{n_1}(t; z_2) \cdot w_{n_1}(t; z_1).$$

- 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 *prevalence* of intermediate events.
- The natural indirect effect measures the treatment effect on the cumulative incidence of terminal event via changing the *prevalence* of intermediate events while controlling the cause-specific hazards of terminal events.

A Paradox

- In some scenarios, interpreting the natural direct effect as "controlling the prevalence of intermediate events" may not be meaningful.
- For example, if a novel therapy completely removes the terminal event, we may think that the direct effect should be large but the indirect effect is null.
- However, since the prevalence of intermediate events increases to 1 by removing terminal events, the indirect effect can also be large.

Sequential Ignorability, Part 2

- We replace the "controlling the prevalence" assumption with the following **sequential ignorability**:

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

- 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 $\tilde{N}_2(\cdot; z_2, \tilde{N}_1(\cdot; z_1))$.
- Whether $\tilde{N}_1(t)$ jumps does not affect when $\tilde{N}_2(\cdot; z_2, \tilde{N}_1(\cdot; z_1))$ jumps in the future.
- 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

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

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.

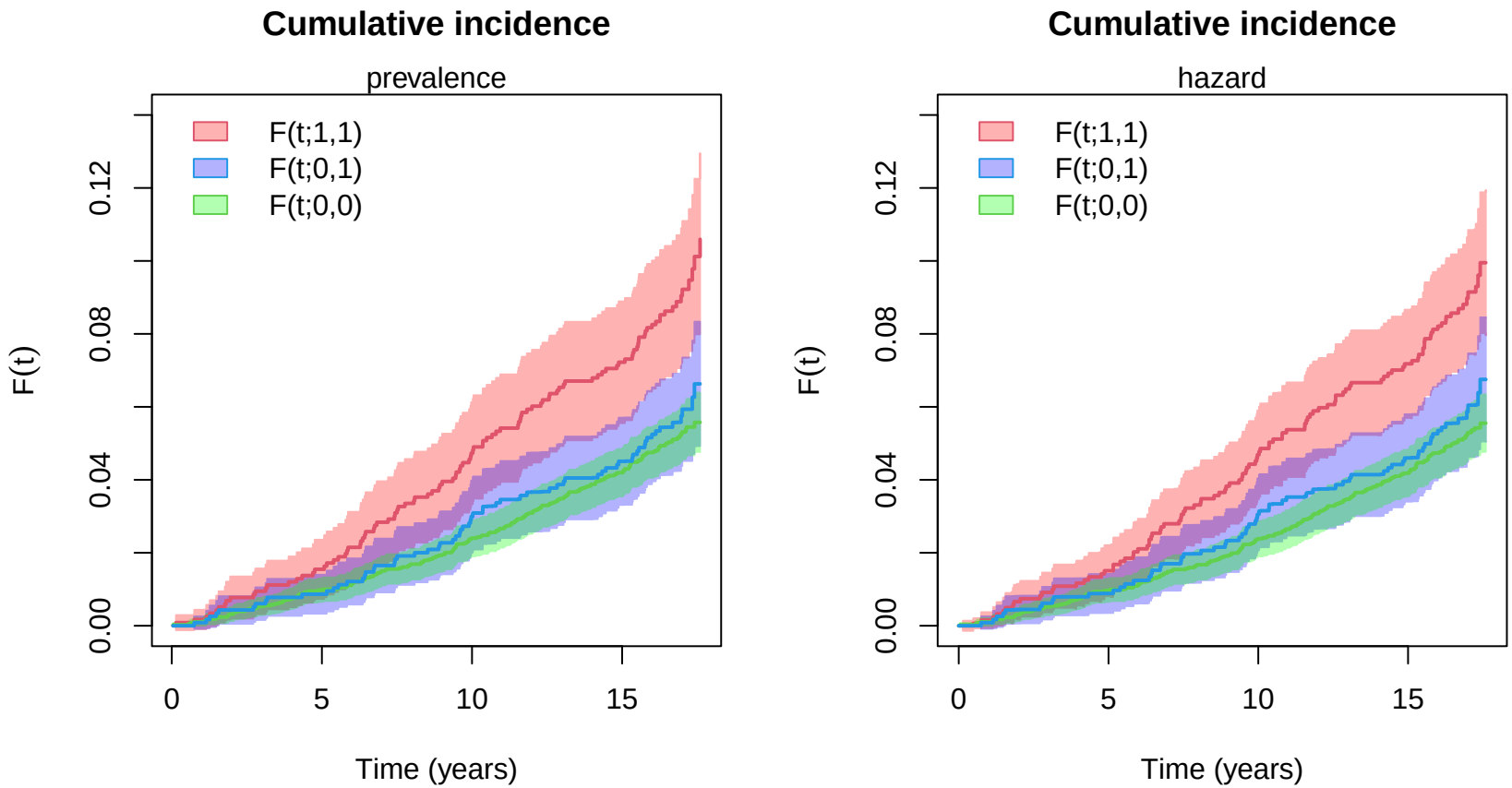
Outline

Hepatitis B Data

- Hepatitis B causes serious public health burden worldwide.
- We are interested in the mechanism of hepatitis B on overall mortality in order to prevent the negative consequences of hepatitis B.
- Treatment: hepatitis B (1 for positive and 0 for negative).
- Intermediate event: cancer.
- Terminal event: mortality.
- Sample size 4954.

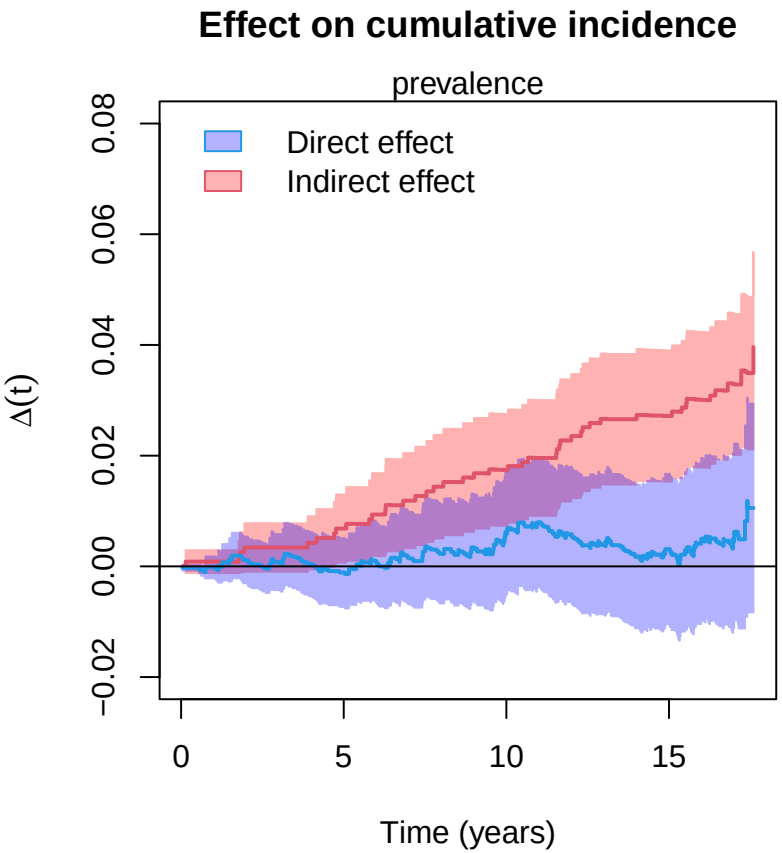
Hepatitis B Data

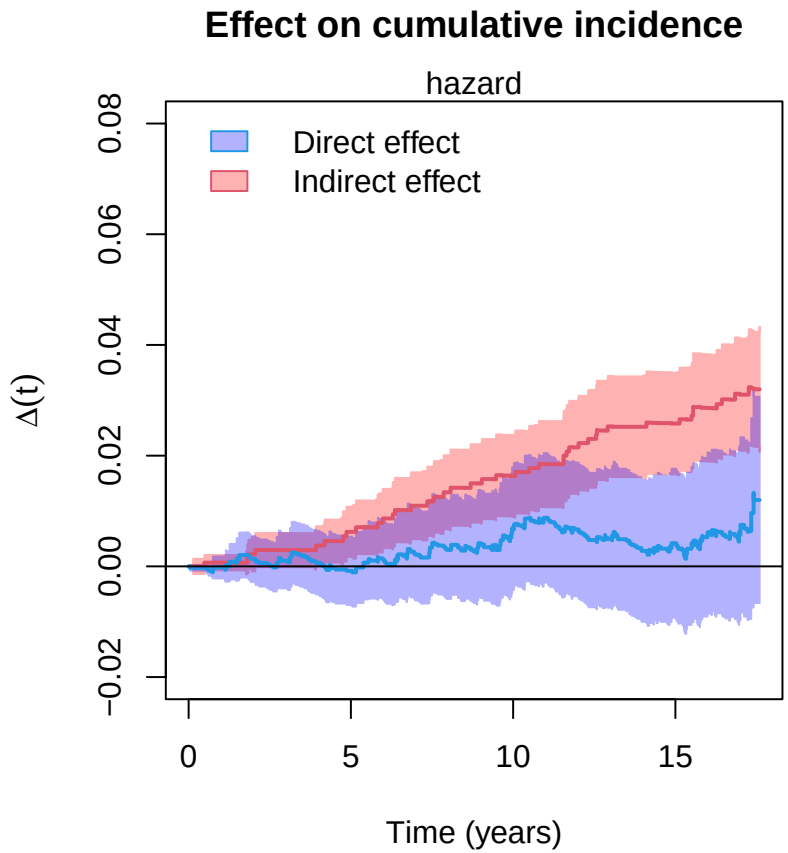
- Hepatitis B increases mortality.
- $F(t; 0, 1)$: the incidence of mortality under the prevalence/hazard of cancer without hepatitis B, hazard of mortality with hepatitis B.



Hepatitis B Data

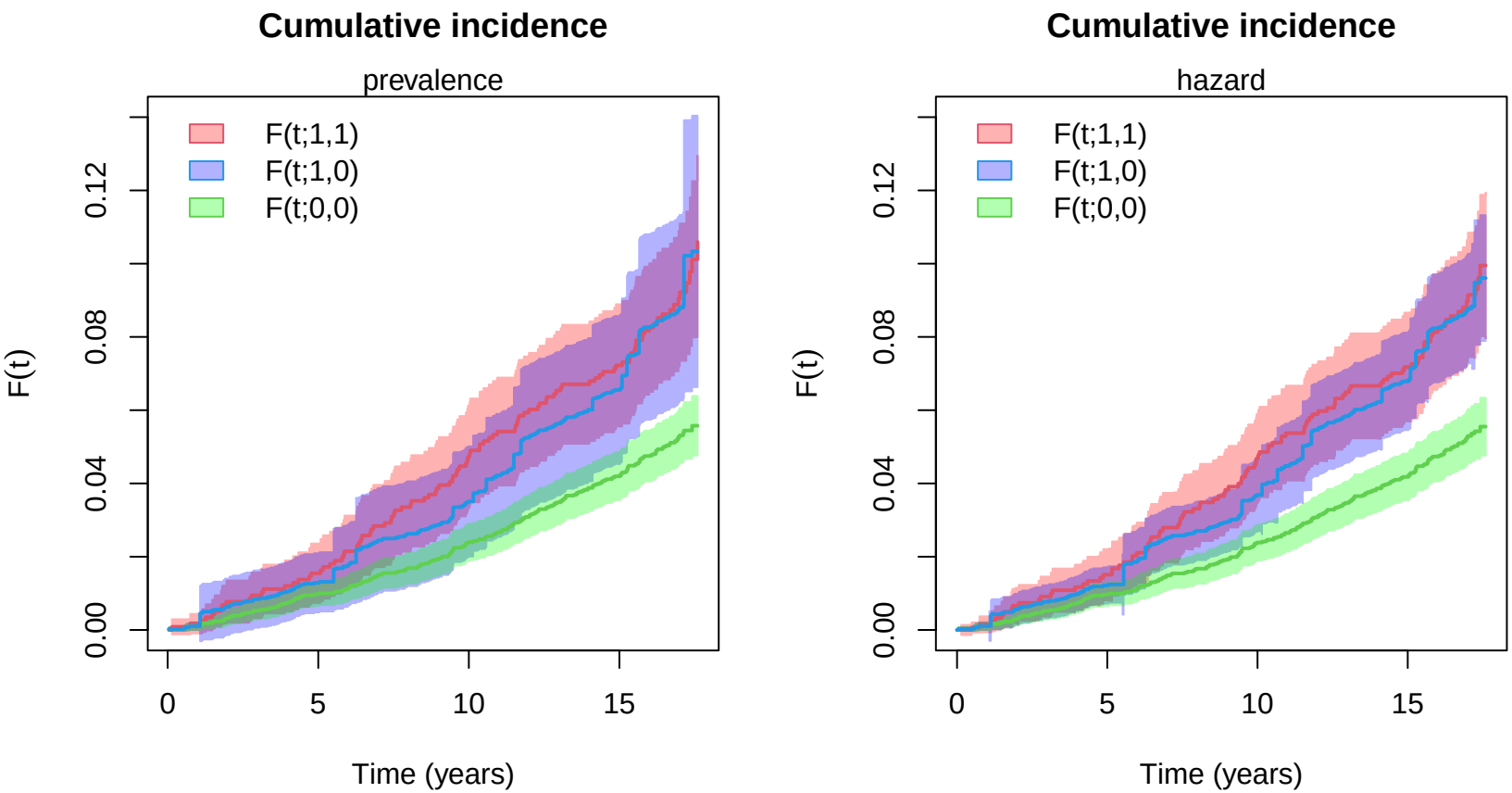
- Hepatitis B increases mortality through increasing the risk of cancer.
- The direct effect of hepatitis B on mortality is not significant.





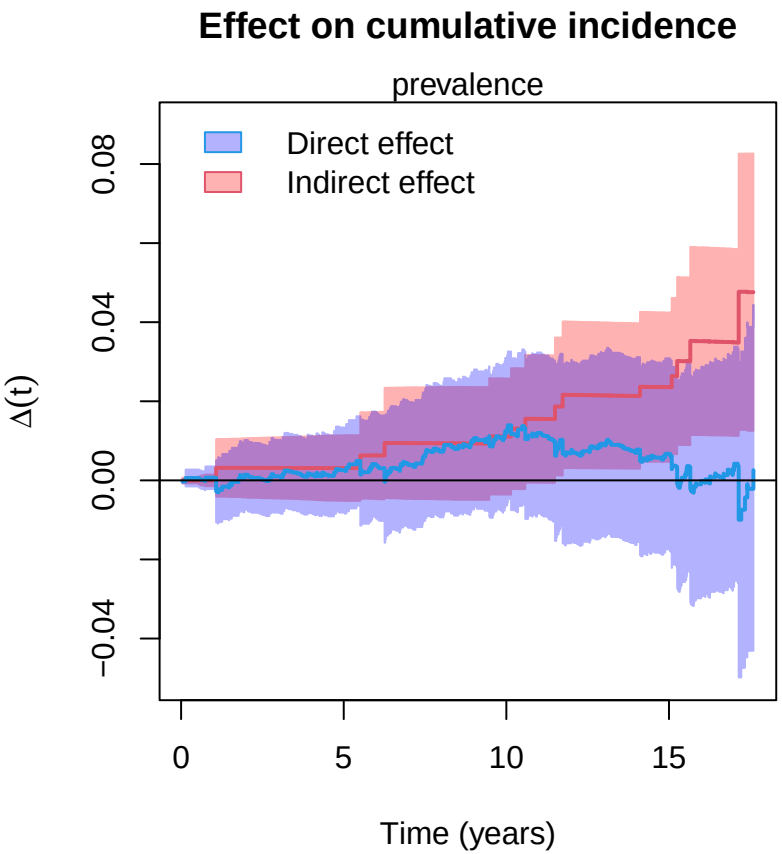
Hepatitis B Data: Alternative Decomposition

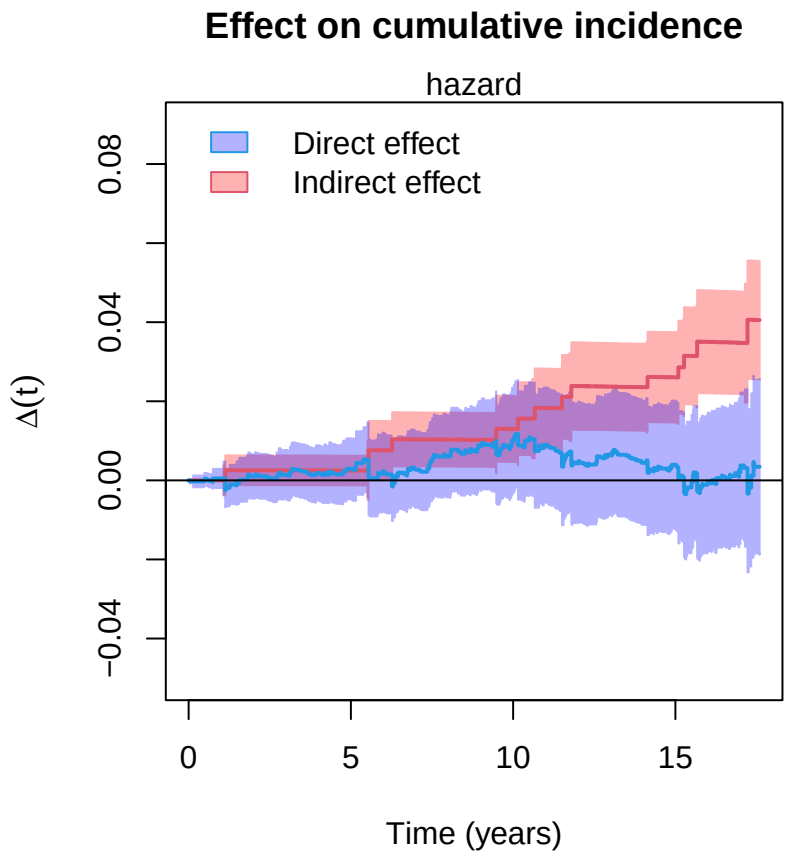
- If we reverse the treatment and control (1 for negative and 0 for positive)... which can serve as a sensitivity analysis on the *interaction effect*.



Hepatitis B Data: Alternative Decomposition

- Hepatitis B increases mortality through increasing the risk of cancer.
- The substantial conclusion still holds.



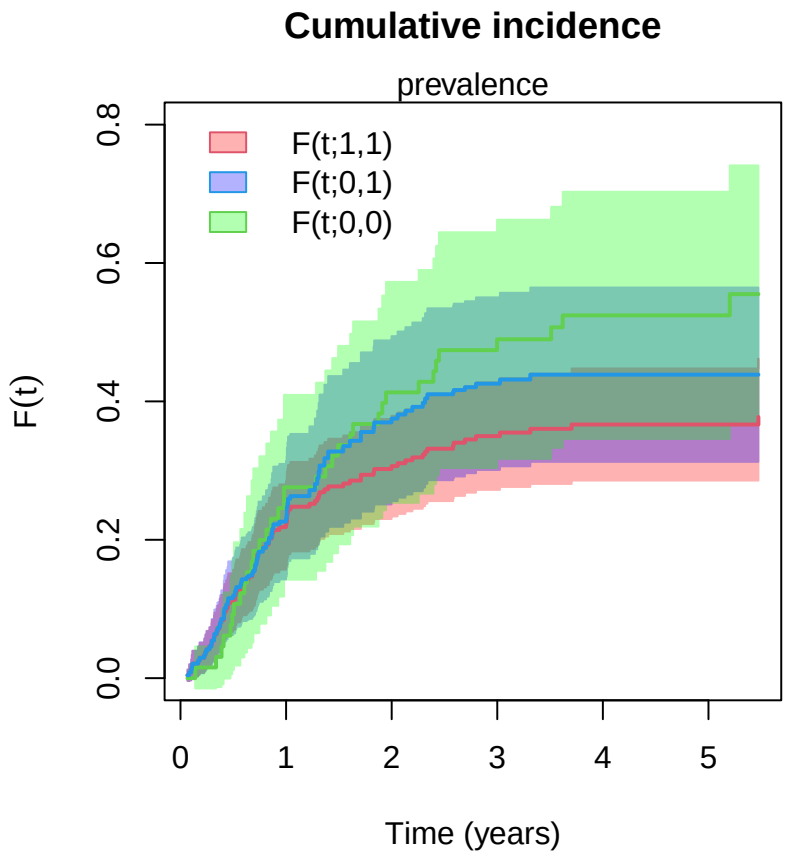


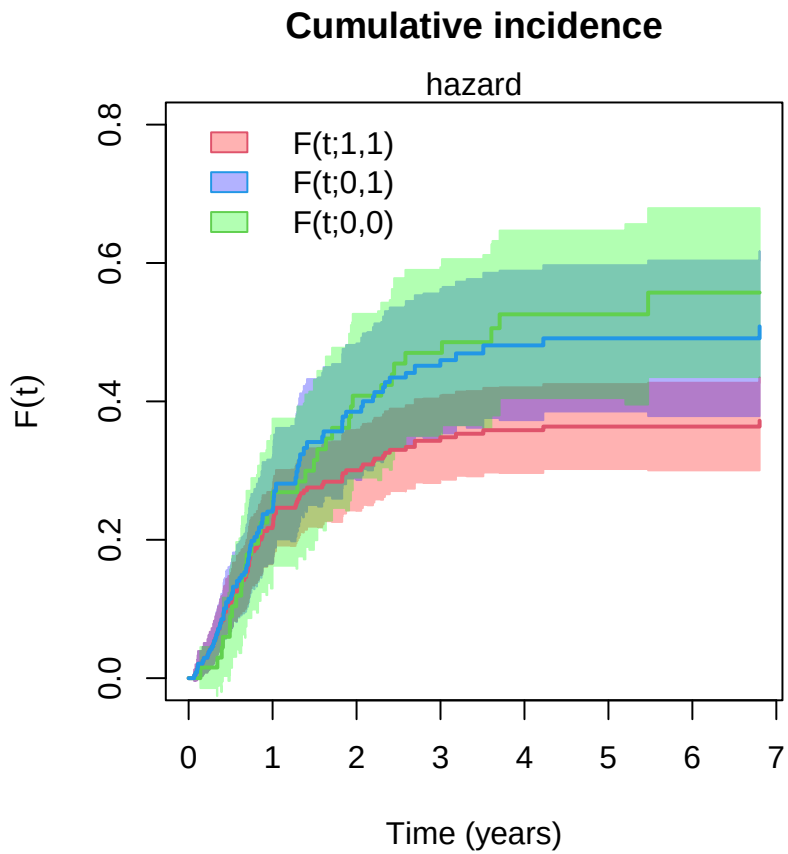
Leukemia Data

- 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.
- Treatment: transplantation type (1 for haplo-SCT and 0 for MSDT).
- Intermediate event: relapse.
- Terminal event: mortality.
- Sample size 303.

Leukemia Data

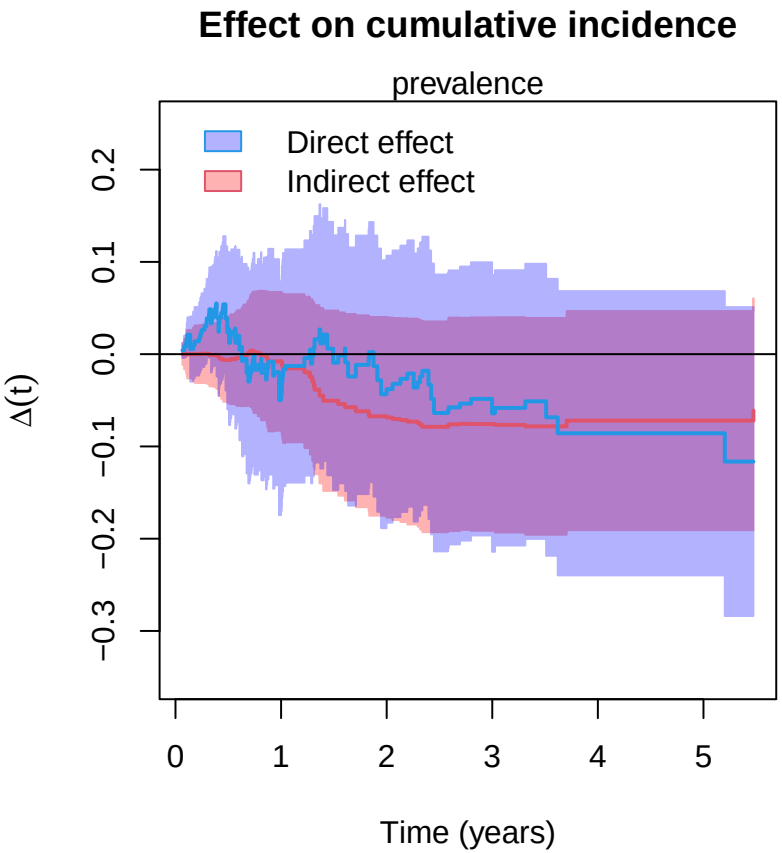
- Haplo-SCT reduces mortality.
- $F(t; 0, 1)$: the incidence of mortality under the prevalence/hazard of relapse undergoing MSDT, hazard of mortality undergoing haplo-SCT.

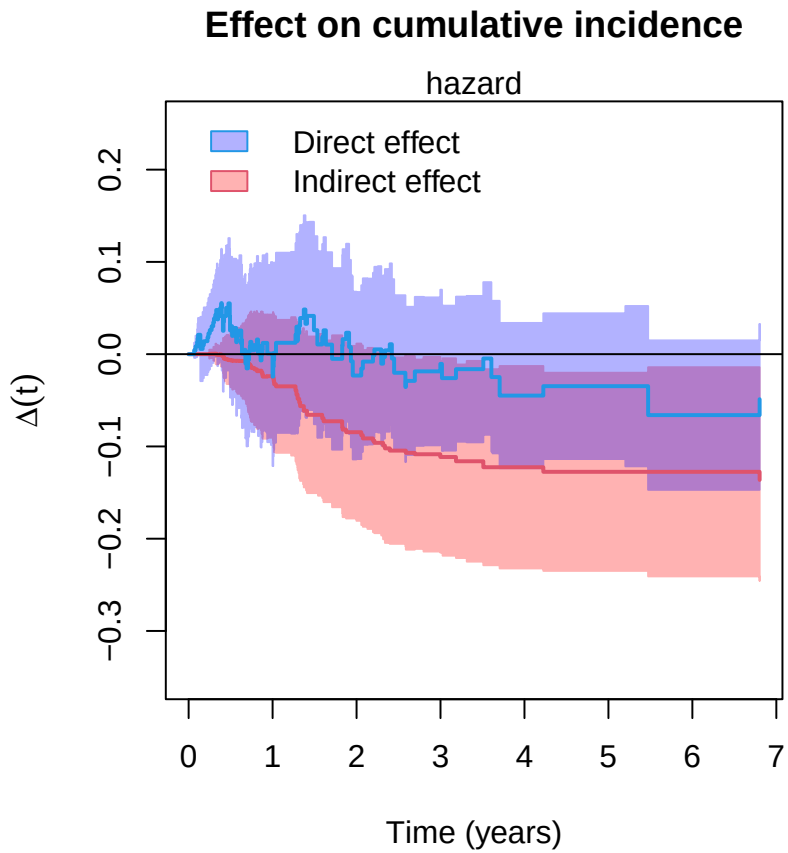




Leukemia Data

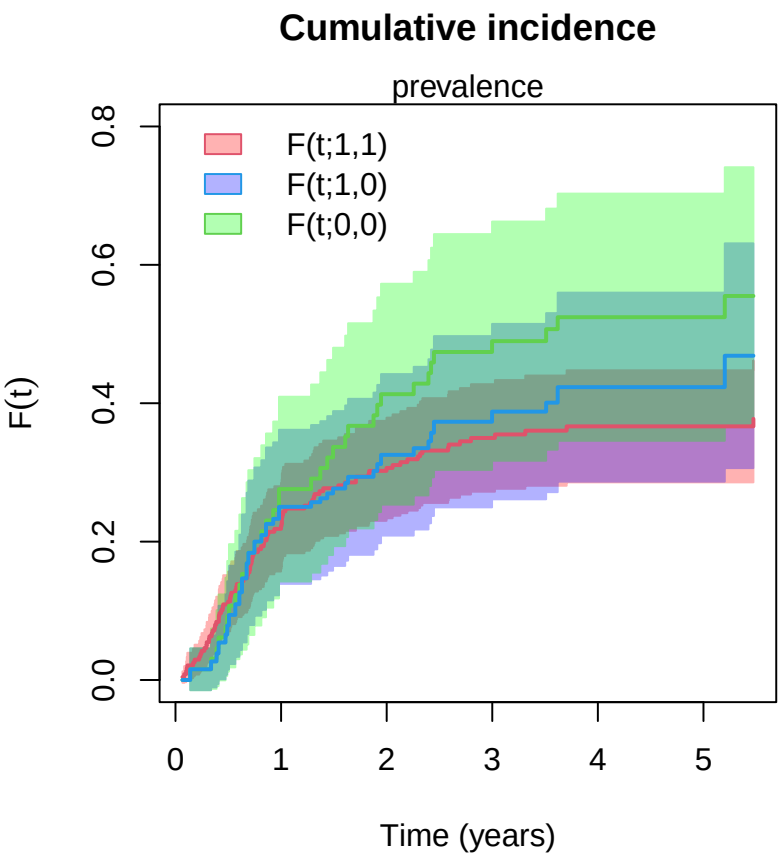
- Haplo-SCT reduces mortality through reducing the risk of relapse.
- Two different assumptions give slightly different conclusions!

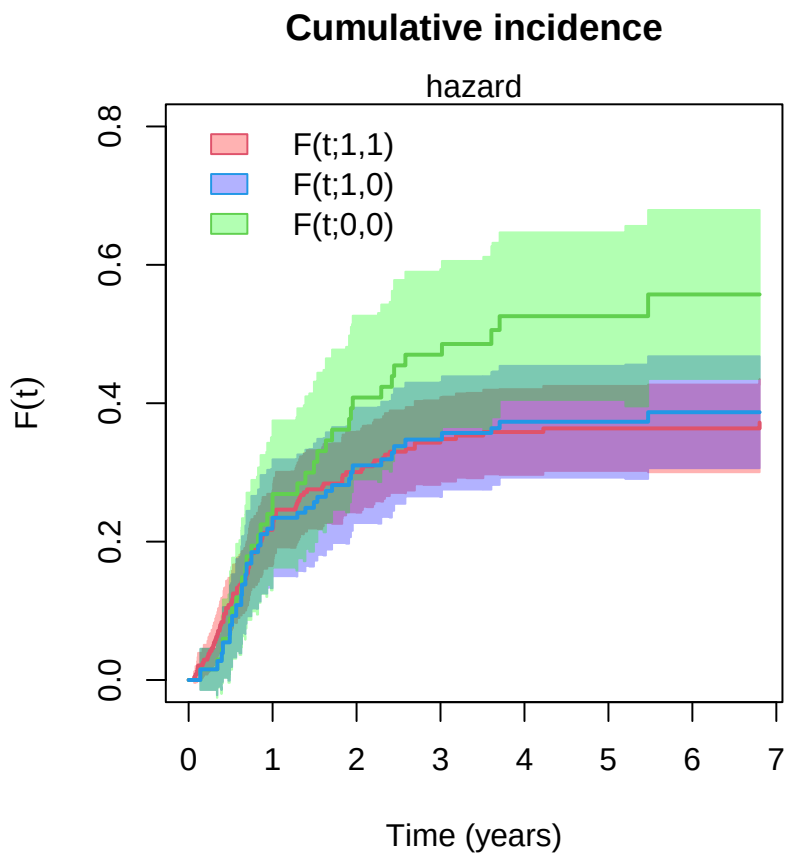




Leukemia Data: Alternative Decomposition

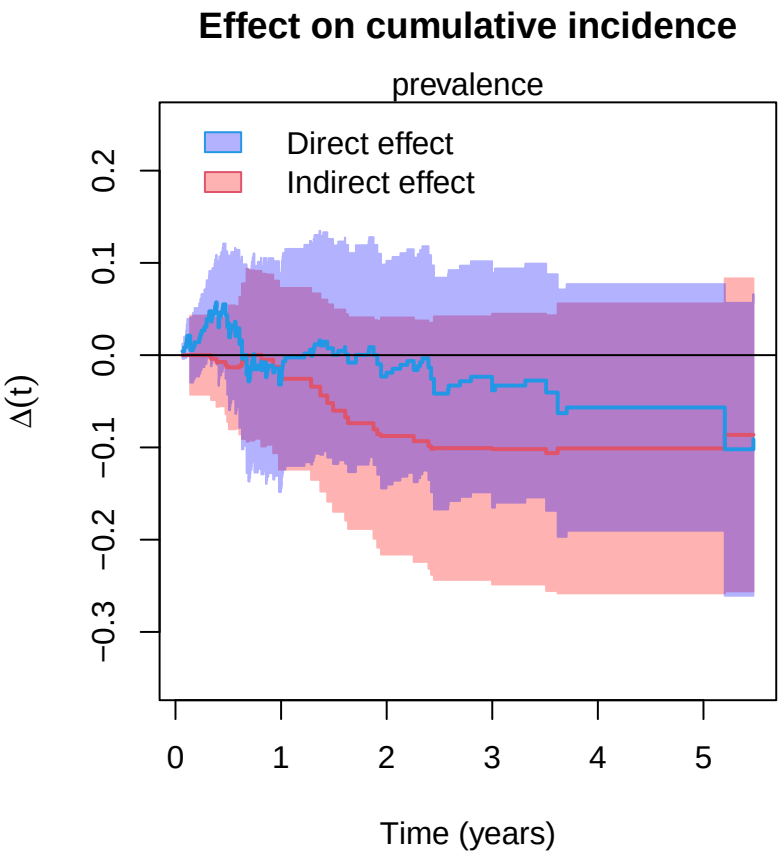
- If we reverse the treatment and control (1 for MSDT and 0 for haplo-SCT)... which can serve as a sensitivity analysis on the *interaction effect*.

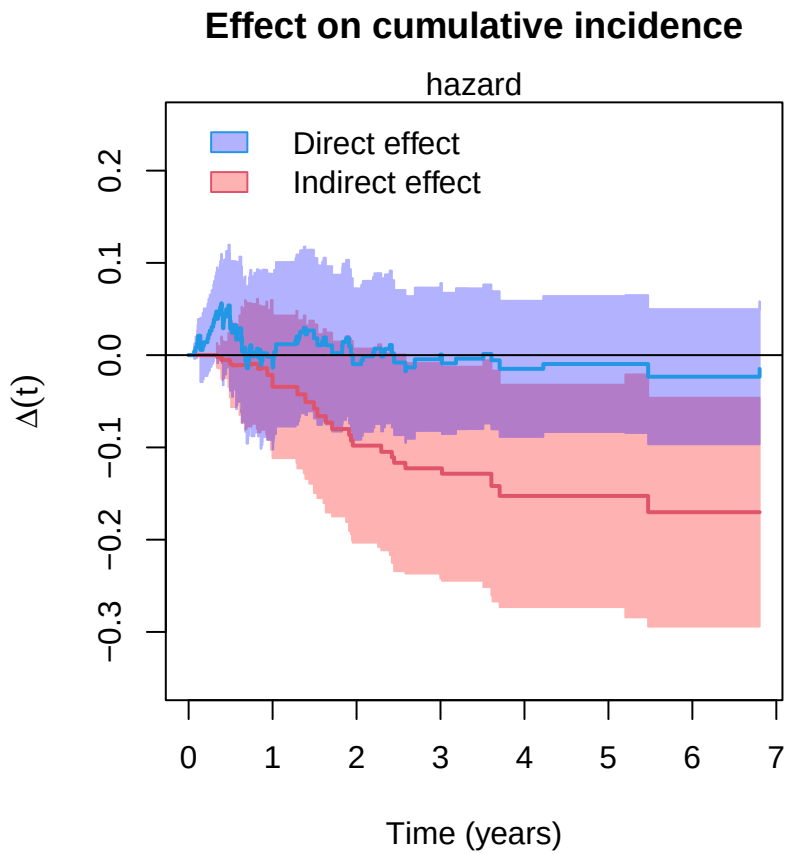




Leukemia Data: Alternative Decomposition

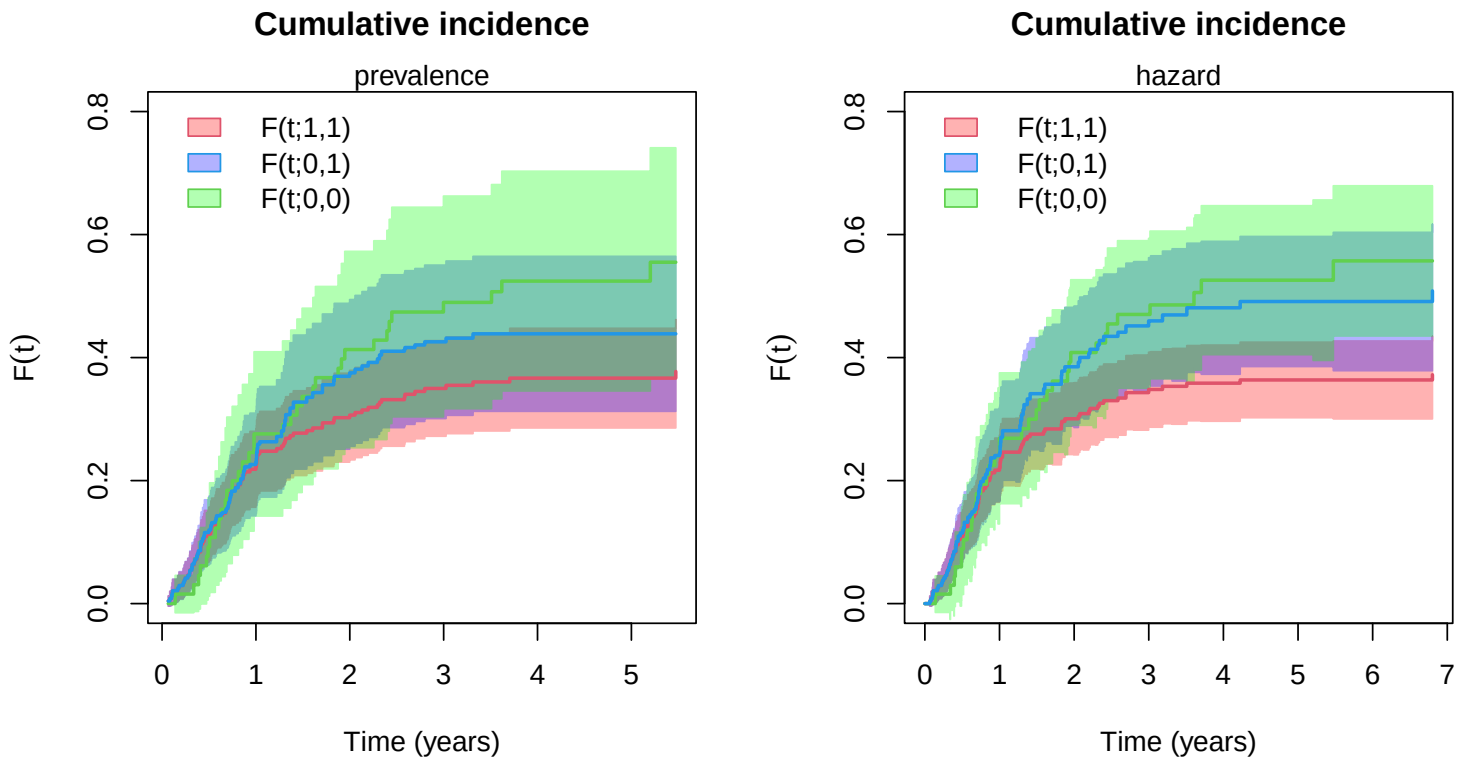
- Haplo-SCT reduces mortality through reducing the risk of relapse.
- The substantial conclusion still holds.





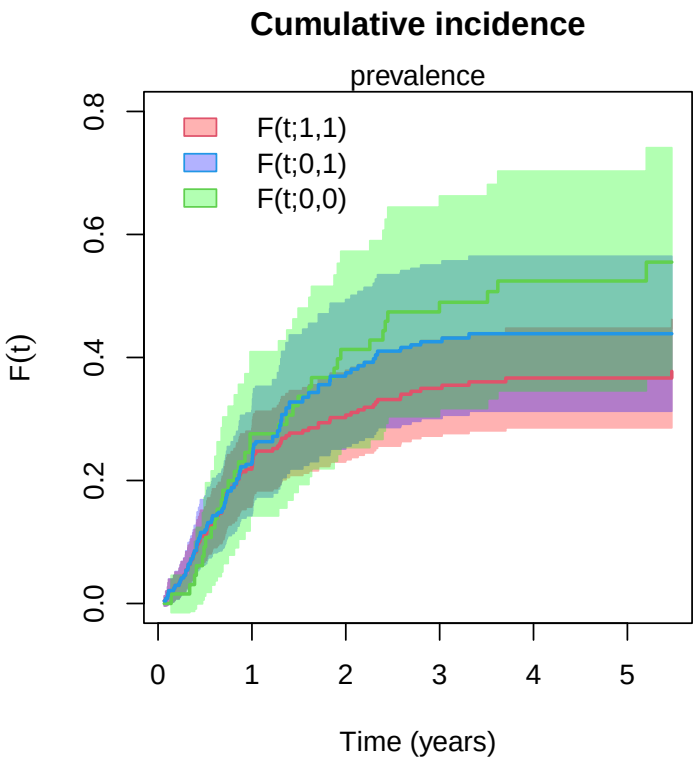
Why Different Results

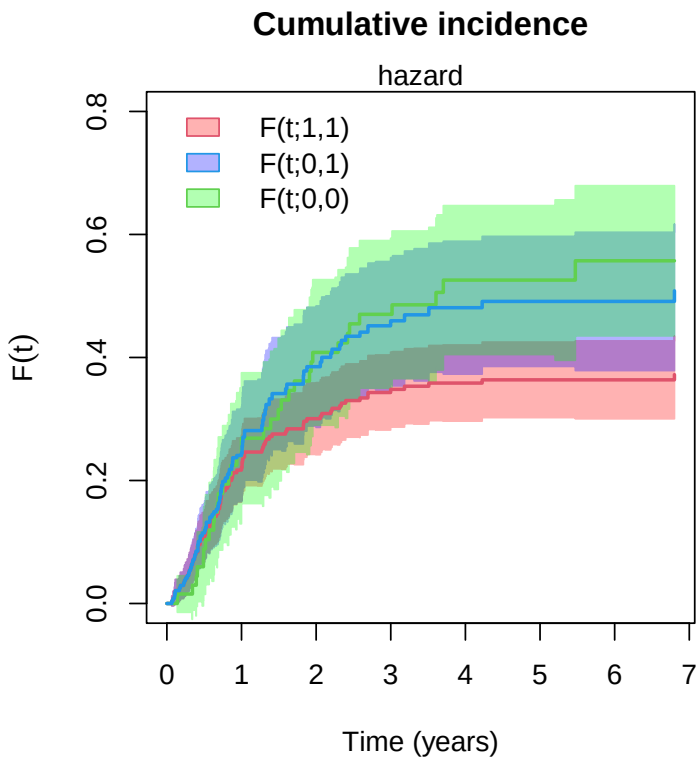
- When envisioning $F(t; 0, 1)$, Decomposition 1 tries to leave the hazard of death as natural while holding the prevalence of relapse unchanged among alive patients.
- Unfortunately, this task is impossible.



Why Different Results

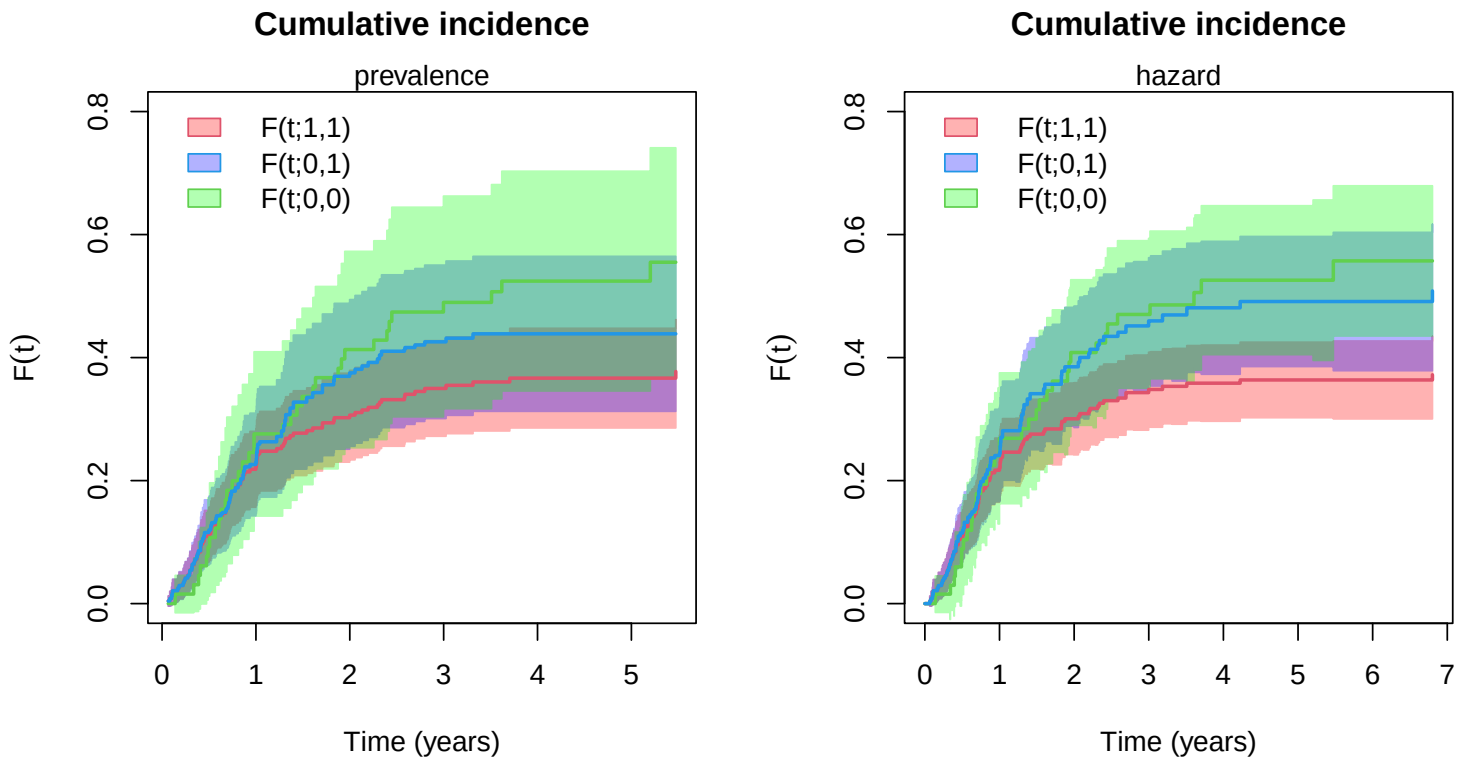
- When switching the treatment from 0 (MSDT) to 1 (haplo-SCT), more individuals would experience transplant-related mortality, so the prevalence of relapse tends to get higher.
- Therefore, Decomposition 1 controlled the prevalence of relapse at a level lower than natural.





Why Different Results

- Since relapse is strongly associated with relapse-related mortality, underestimation of the prevalence of relapse is reflected by an underestimation of relapse-related mortality.
- Thus, the total incidence of mortality $F(t; 0, 1)$ is underestimated.



Outline

Why Interventionist Approach

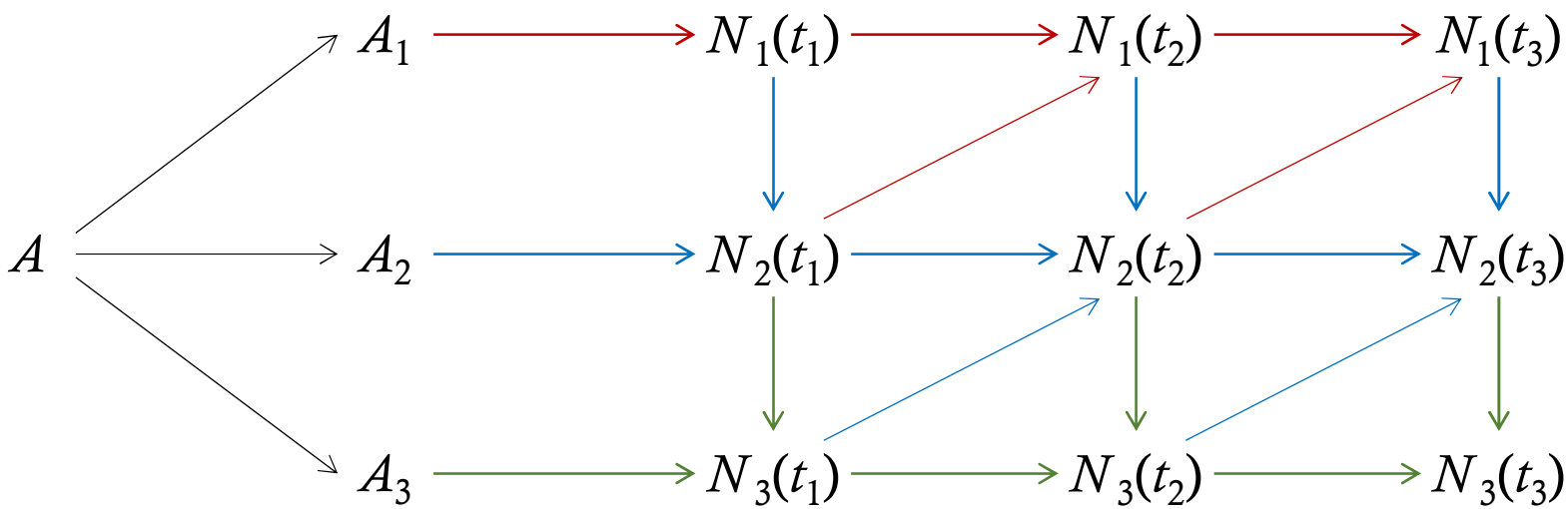
- Recall the separable effects framework for mediation analysis.
- There are three advantages for the separable effects framework over the standard mediation analysis using sequential ignorbaility.
- First, the notation is simpler, and the interpretation is easier, although the implications may remain the same.
- Second, the dismissible treatment components assumption, compared with sequential ignorability, is testable in principle.
- Third, when the sequential ignorability is violated, the separable effects estimand is still meaningful from the interventionist view as long as we can envision dismissible treatment components.

Separable Effects for Time-to-Event Outcomes

- We decompose the treatment into $A = (A_1, A_2, A_3)$.
- A_1 influences the hazard of the direct terminal event.
- A_2 influences the hazard of the intermediate event.
- A_3 influences the hazard of the indirect terminal after the intermediate event.
- In the realized trial, $A = A_1 = A_2 = A_3$. But in future experiments, A_1, A_2 and A_3 can be unequal.

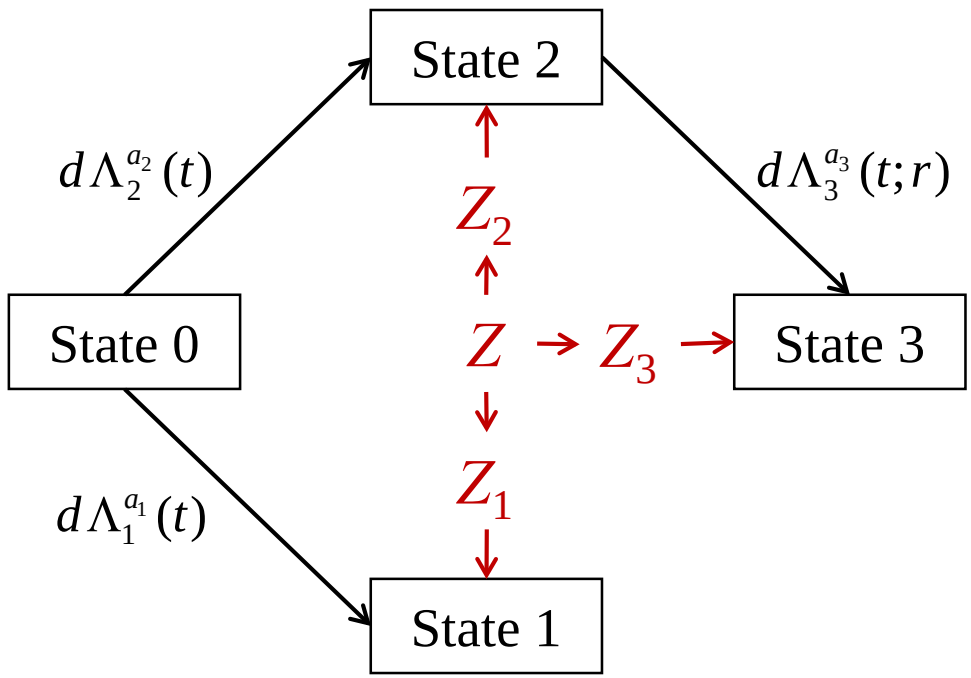
Separable Effects for Time-to-Event Outcomes

- N_1: counting process of the direct outcome event.
- N_2: counting process of the intermediate event.
- N_3: counting process of the indirect outcome event.



A Multi-State Model

- State 0: original status
- State 1: direct outcome (terminal) event
- State 2: intermediate event
- State 3: indirect outcome (terminal) event following intermediate event



New Estimand and Assumption

- Potential time to the terminal event $T^{a=(a_1,a_2,a_3)}$.
- Potential time to the intermediate event $R^{a=(a_1,a_2,a_3)}$.
- The quantity of interest is the counterfactual cumulative incidence of the terminal event

$$F^{a=(a_1,a_2,a_3)}(t) = P(T^{a=(a_1,a_2,a_3)} \leq t).$$

- Denote the transition hazard to State j by $d\Lambda_j^{a=(a_1,a_2,a_3)}(t)$.
- **Dismissible treatment components** (under Markovness):

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

- The transitions are "stable". It means there is no confounding between events, similar with sequential ignorability, but notationally so concise!

Separable Pathway Effects

- The total effect

$$TE(t) = F^{a=(1,1,1)}(t) - F^{a=(0,0,0)}(t)$$

can be decomposed into three separable pathway effects,

$$\begin{aligned} TE(t) &= F^{a=(1,0,0)}(t) - F^{a=(0,0,0)}(t) \\ &\quad + F^{a=(1,1,0)}(t) - F^{a=(1,0,0)}(t) \\ &\quad + F^{a=(1,1,1)}(t) - F^{a=(1,1,0)}(t) \\ &= SPE_{0\to1}(t) + SPE_{0\to2}(t) + SPE_{2\to3}(t). \end{aligned}$$

- $SPE_{0\to1}(t)$ serves as the direct effect, $SPE_{0\to2}(t)$ serves as the indirect effect, and $SPE_{2\to3}(t)$ serves as the interaction effect.

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 separable effects framework provides an easy-to-interpret approach to mediation analysis.
- The pointwise variance of estimators can be derived by the stochastic process theory, or simply by bootstrap.

Extensions

- Interaction effects.
  This problem seems to have been overcome by separable effects.
- Alternatives to mediation analysis, for example, by principal stratification.
  Principal stratification provides a tool to study the pure effect in a principal stratum.
  Untestable assumptions like principal ignorability cannot be avoided.

Extensions

- Baseline confounders.
  We need to condition on confounders in assumptions.
  Then we can derive the efficient influence functions, and propose efficient estimators (Martinussen and Stensrud, 2023).
- Time-varying confounders.
  Identification can be achieved by g-formula.
  Efficient estimation could be difficult.
- Multiple intermediate events.
  Using the separable effects framework, the idea is straightforward.
  However, if the model is too complicated, estimation is accompanied by large variation.

References

- Deng Y, Wang Y, Zhou X-H. Direct and Indirect Effects in the presence of Semi-competing risks. *Biometrics*. 2024.
- Deng Y, Wang Y, Zhan X, Zhou X-H. Separable Pathway Effects of Semi-competing Risks via Multi-State Models. *ArXiv*. 2023.

If you are interested in this topic, feel free to discuss with me

yuhaoden@umich.edu