Causal inference with competing events

Jessica Young, Mats Stensrud,
& L. Paloma Rojas-Saunero

SER 2023, Portland

Structure of the workshop

1. Review of fundamental definitions and terminology: *competing events*, *truncation events*, *causal effect*, *causal road map*
2. Historical notions of causal effect when competing/truncation events exist
3. Newer notions of causal effects when competing/truncation events exist: *separable effects*
4. Identification of some historical notions and a meaning of "censoring" that bridges the "causal road map" to the classical survival analysis literature
   - Leads to more transparent thought process around some familiar statistics
   - Brief code introduction
5. Identification and estimation of separable effects
   - Brief code introduction

The Fundamentals

First we need to be on same page about what we mean by

- “causal inference”
- “competing events”
- “ truncation events”

What is causal inference?

In this course, we mean broadly the process of

1. linking a particular causal effect we want to learn about
2. to a particular statistical method that we can apply to a data set
3. under explicit assumptions that we can reason through with subject matter (underlying causal model)
   - Causal diagrams (DAGs, SWIGs) useful tools for communicating subject matter assumptions

I like to refer to this process as "a causal road map".

What is a competing event?

A competing (risk) event is any event that makes it impossible for an event of interest (i.e. a study outcome of interest) to occur.

- Therefore, whether competing events exist (and what they are) depends not only on the data but also on what we are "interested in" – feature of the causal effect we choose to be at the "top of the road map"

Examples of competing events

- In population of men diagnosed with prostate cancer, outcome/event of interest is "death due to prostate cancer", some individuals die of stroke (fatal stroke is a competing event)
- In population of adults in midlife, outcome/event of interest is "diagnosis of dementia" in late life, some individuals die during follow-up prior to diagnosis (death due to any cause is a competing event).

"Semi-competing" risks

The second example where outcome of interest is "diagnosis is dementia" and deaths occur would traditionally be called a "semi-competing risk" setting.

- This makes no difference at all to the general arguments that follow.

It is only the existence of competing events for the outcome we care about that prompts these considerations. Once we lead with questions rather than statistics, it becomes clear that this separate designation is not necessary for causal inference.

- See: Stensrud MJ, Young JG, Martinussen T. Discussion on "Causal mediation of semicompeting risks" by Yen-Tsung Huang. *Biometrics*. 2021; 77(4): 1160-1164.

What is a truncation event?

A truncation event is any event that renders the study outcome of interest undefined.

- Similar to competing events, whether truncation events exist depends primarily on what we are interested in (as well as the data).

Examples of truncation events

- In population of people diagnosed with hypertension about to make a choice between two treatments (e.g. a thiazide diuretic versus ACE inhibitor), outcome of interest is weight change two years later, some individuals die prior to two year follow-up (death is a truncation event).
- In population of adults in midlife, outcome of interest is a cognitive score in later life, some individuals die (death is a truncation event).

Artificial distinction between competing and truncation event

In some ways the distinction is due to whether the investigator chooses to define a value of the study outcome after certain events occur

- We could choose to a assign a value of zero weight change or value of zero for a cognitive score in both examples for a dead individual.
- If we do that we have technically defined the outcome, however arguably the definition is not "meaningful" for these individuals.

Boils down to what we choose to put at the top of the road map, being transparent about that choice and why we made it.

The fundamental challenge of causal inference in the face of competing or truncation events

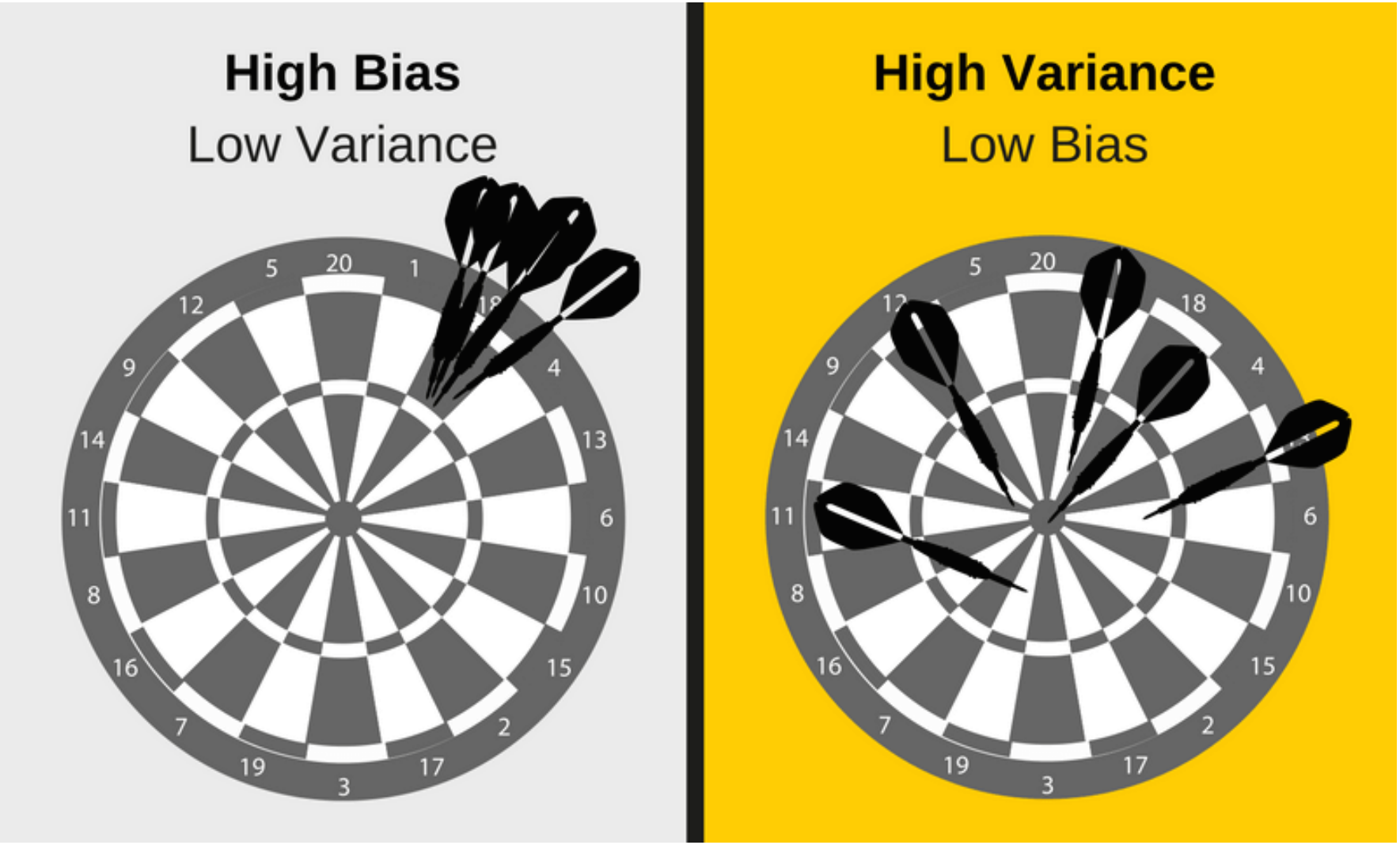
When these types of events occur, it can be challenging to define a causal effect of the study treatment on the study outcome of interest *that answers the question of interest*.

- We can get stalled at the very top of the road map.

How we ultimately choose to define the target effect has important implications for the assumptions needed to unbiasedly estimate that effect with the study data and what the estimator/statistics look like.

- This is true even when the study data comes from a "perfectly executed" trial

The solution is not a statistical one – statistics is the last step of the road map



First we need to know what our target is to understand whether a statistical approach is "good" or "bad".

Classic pedagogic example

- Early trial (Byar and Green, 1980), stated interest in effect of estrogen therapy (versus placebo) on prostate cancer death in men diagnosed with prostate cancer
- Some men in both arms died of other causes
- Death from other causes are competing events for death from prostate cancer because an individual cannot die of prostate cancer once he has died of a cardiovascular event.

Implicit choices

There is a large literature with recommendations for applied researchers on how to analyze "competing risks data".

- Most of this has been presented with respect to choices in *statistical parameters* and methods rather than a choice of causal effect

What is the difference between a statistical parameter and a causal effect?

What is a statistical parameter?

We mean any function of the (joint) distribution of a set of variables that we could in principle measure/observe on a set of individuals. Can call these factual variables. Ex: suppose we conducted a "perfect" version of this estrogen trial

- A indicator of whether estrogen therapy (vs. placebo) at some baseline
- Y_k indicator of whether an individual dies of prostate cancer by any subsequent time (e.g. day) $k = 1, \ldots K = 365 \times 5$
- D_k indicator of whether an individual dies of something other than prostate cancer by $k = 1, \ldots, K = 365 \times 5$

Summary of the factual data structure

We can summarize the factual variables we measure in this study as $O = (A, \overline{Y}_K, \overline{D}_K)$

- overline means "history" e.g. $\overline{Y}_k = (Y_1, Y_2, \ldots, Y_k)$.
- underline means "future" e.g. $\underline{Y}_k = (Y_k, Y_2, \ldots, Y_K)$.

What is a statistical parameter?

Any function of the true joint distribution of these factual variables $O = (A, \overline{Y}_K, \overline{D}_K)$ is an example of a *statistical parameter*. E.g.

- $Pr[Y_{100} = 1|A = 1] - Pr[Y_{100} = 1|A = 0]$
- $P[D_{30} = 1|Y_{29} = 0]$

Can think of a statistical parameter as anything we can construct a statistic for.

Competing risks factual data structure

Key feature of this data structure which is what makes it an example of a "competing events data structure":

- If an individual is known to experience the competing event by interval $k > 0$ without history of the event of interest $(Y_{k-1} = 0, D_k = 1)$ then $\underline{Y}_k$ will by definition be deterministically zero
- This is what happens in the "observed" / "factual" world – when you die of a heart attack you cannot subsequently die of prostate cancer.

Link to familiar competing risks data structure

This *factual* data structure consistent with $O = (A, \tilde{T} = min(T, C), J)$ for T time to first failure from any cause, C "administrative censoring time", $J = 0$ if administratively censored, $J = 1$ if failed due to prostate cancer death and $J = 2$ if failed due to other death.

- Consistent in the sense that if I know O, I know the time-varying event history version $O = (A, \overline{Y}_K, \overline{D}_K)$
- Our less summarized way of defining data structure acknowledges time-varying nature of relevant event processes, crucial for transparency around interpretation and identification

Restrict discrete measurement process here but fundamental concepts don't really change if we wanted to conceptualize continuous time processes.

What is a causal effect?

Contrast (e.g. difference) of outcomes in the **same individuals** but under implementation of **different treatments (interventions)**.

- Quantifies *counterfactual* features of a population.

These are not statistical parameters in the sense that we cannot even in principle measure or observe outcomes under both of these scenarios.

Counterfactual notation

Typically we use superscripts to denote counterfactual variables to distinguish them from measureable (factual) variables.

- Could write causal effect had we made everyone initiate active treatment versus placebo on the risk of dying of prostate cancer over the next five years as:

$$Pr[Y^{a=1}_{k=365\times5} = 1] - Pr[Y^{a=0}_{k=365\times5} = 1]$$

- Unlike a statistical parameter, this is not something that we can directly estimate with statistics.
- However, under assumptions we might be able to equate it to (*identify it by*) a statistical parameter, a function of only factuals.

Popular targets

The most widely discussed and taught statistical methods for competing risk analysis have been motivated in the survival analysis literature from the following targets:

- the *cause-specific cumulative incidence* (or *crude risk*)
- the *marginal cumulative incidence* (*net risk*)
- various definitions of hazard (e.g. cause-specific, subdistribution)

Turns out even contrasts in *counterfactual versions* of any of these parameters may not constitute what an investigator is really (implicitly) interested in when causal inference is the goal.

Causal inference literature

Arguably, the causal inference literature has not helped fill this gap:

- even within the formal literature on counterfactual causal inference, notions of causal effects historically posed also may not constitute what the investigator is really interested in.

How do I know what investigators are interested in?

This is going to be investigator-specific and study-specific.

- The argument that other notions of causal effect are needed to "top the road map" is evidenced by frequent misinterpretation of the targets mentioned above in terms of something else.

Counterfactual contrast in crude risks

Contrast in (discrete time) *cause-specific cumulative incidence functions* for the event of interest under different treatments a at some k is

$$\Pr[Y_k^{a=1} = 1] \text{ vs. } \Pr[Y_k^{a=0} = 1]$$

- This is a causal effect because it compares distributions of counterfactual outcomes under different treatment interventions but in the same individuals
- Under randomization of A, identified simply by $\Pr[Y_k = 1|A = 1]$ vs. $\Pr[Y_k = 1|A = 0]$.

Why consider anything else?

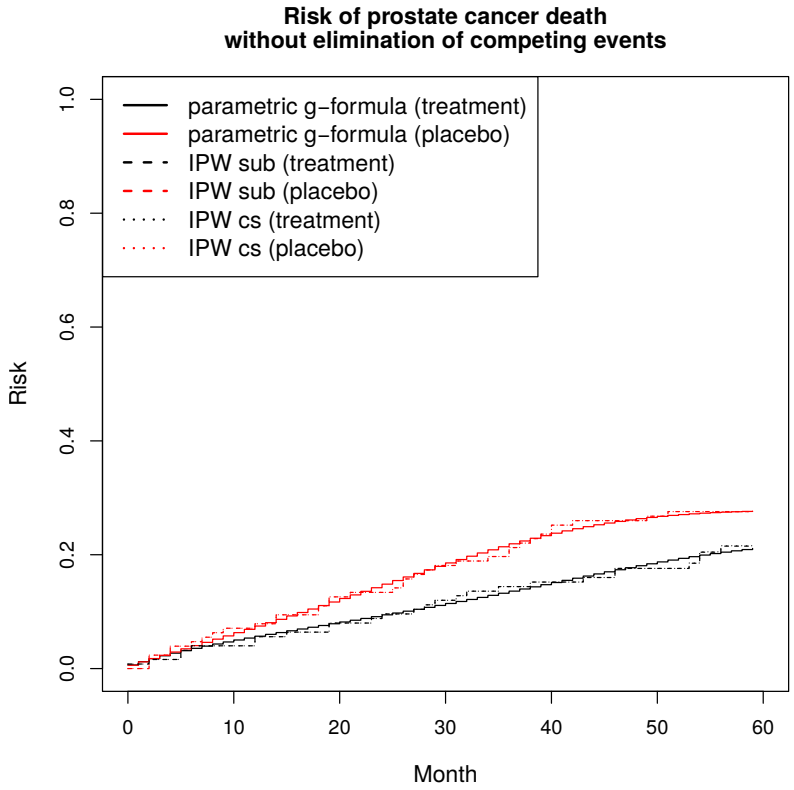
Counterfactual contrast in cause-specific cumulative incidence

1. quantifies a causal effect of treatment on the event of interest to investigators
2. guaranteed identified and super easy to estimate – at least in a "perfect" trial

So are we done? Why not always top the road map with this?

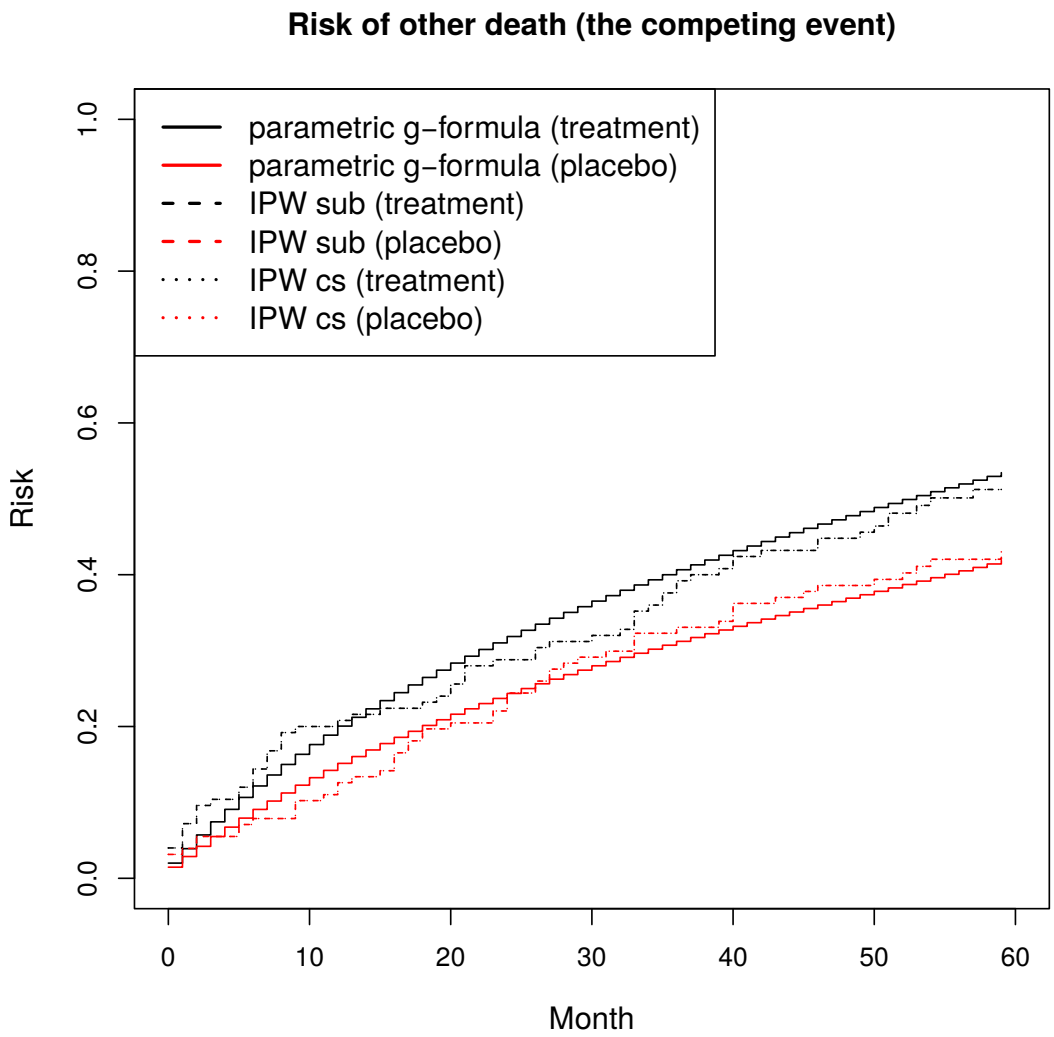
- The issue is that this effect is a case of a **total effect** in that it *captures all paths* by which A affects Y_k

Total effect estimates on prostate cancer death from the original estrogen trial



Young et al., Statistics in Medicine, 2020.

Total effect estimates on other death from the original estrogen trial



Young et al., Statistics in Medicine, 2020.

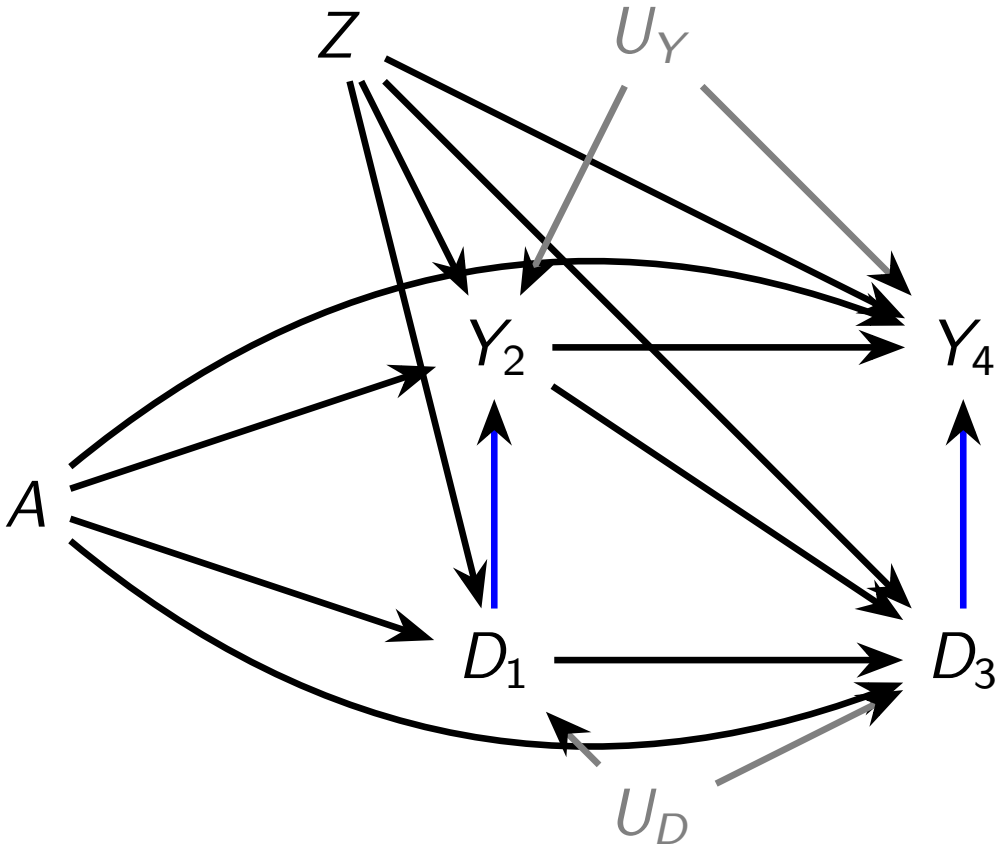
Total effects do not answer questions about mechanism

A total protective effect of estrogen therapy on prostate cancer death doesn't distinguish between the following scenarios:

- Is this only due to an action of estrogen that prohibits proliferation of prostate cancer cells?
- OR is this only due to an action of estrogen that kills in other ways? Estrogen is known to increase risk of cardiovascular events. You are forever protected from dying of prostate cancer if you die of something else first.
- OR both?

IF the investigators wanted to inform this, they have not with the total effect. The total effect does not elucidate anything about treatment mechanism.

Causal DAG representing an assumption on data generating process in our trial



Blue arrows there when D a competing event for Y. *If* cannot rule out arrows from A to D nodes (not intersected by Y nodes), then the total effect will capture paths from A to Y_k through the blue arrows.

When we care about mechanism

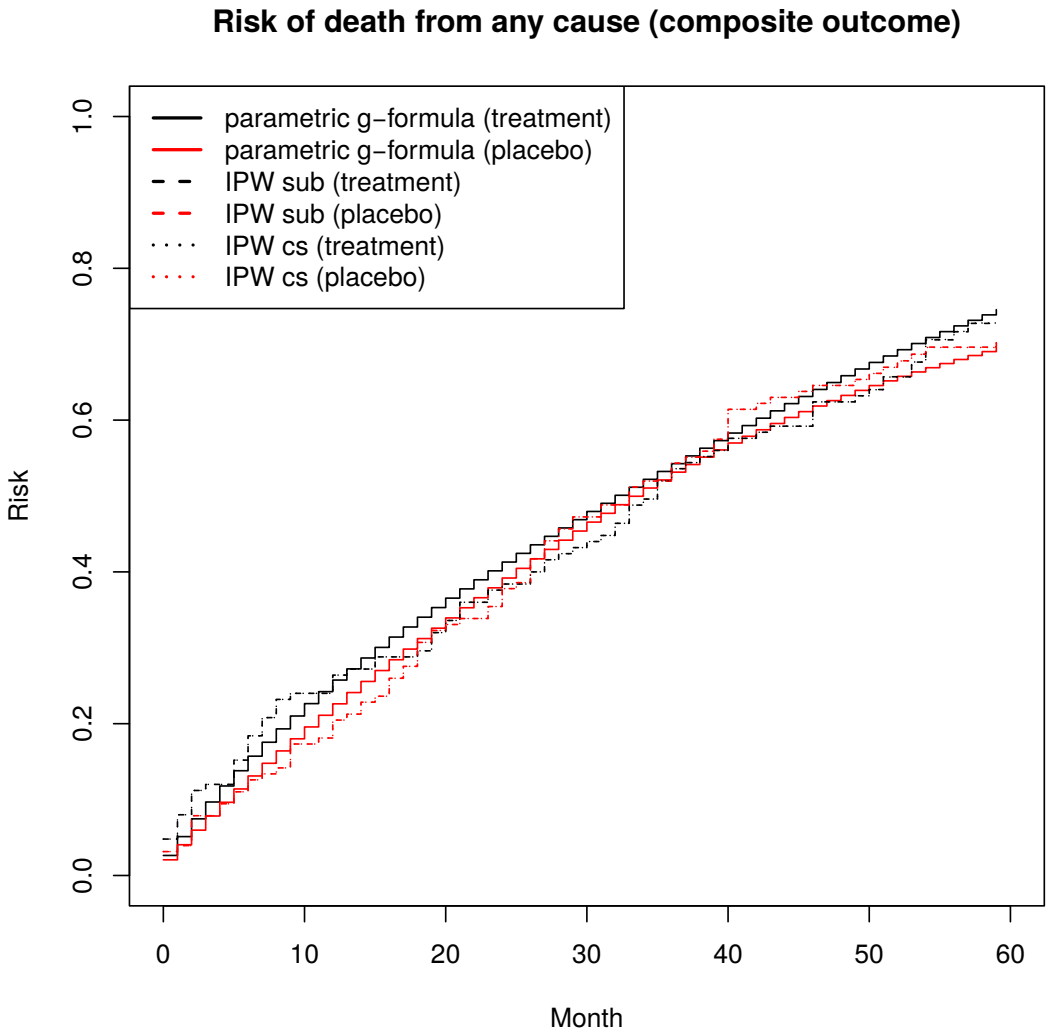
In many cases, when the investigators pose interest specifically in treatment effects on an event subject to competing events, IMPLICITLY, the underlying causal question is about some type of "direct" or "path-specific" effect that avoids capturing these uninteresting paths through the blue arrows.

Again...total effects don't answer questions about mechanism

Importantly, estimating other types of total effects doesn't solve this because, again, total effects do not elucidate mechanism. This includes

- Total effect on competing event – certainly interesting that we estimate a harmful effect of treatment on other death in the trial but it *still doesn't explain the mechanism by which we see a protective total effect on prostate cancer death!*
- Total effect on composite outcome is even less informative! It just combines the total effect on competing event and event of interest.

Total effect estimates on composite outcome (all-cause mortality) from the actual trial



Young et al., Statistics in Medicine, 2020.

Truncation events

- When D is a truncation event, the blue arrows reflect the determinism that if $D = 1$ then Y is not defined
- Here the total effect is not even defined so we have no choice but to consider something else

Am I saying don't go after a total effect?

Provided you define it...No!

- A total effect of A on Y is the most identifiable type of effect we can consider.
- In fact, this is all that a perfect executed trial guarantees we can identify.

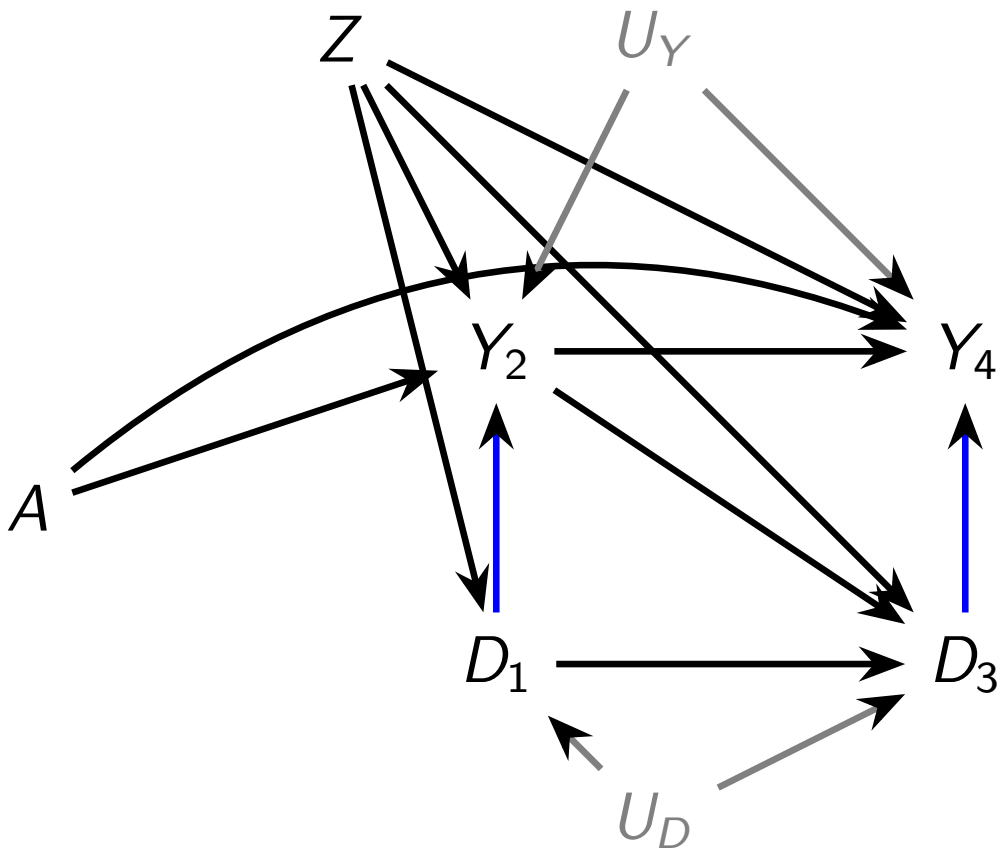
When the total effect is not enough?

When you care about treatment mechanism.

- In this case, if you estimate a total effect, you have to at least *think* about a mechanism story which is beyond what the magnitude or direction of the effect estimate.

Example: maybe by your background subject matter knowledge you can reasonably assume no "direct" arrows from A to D (not intersected by Y nodes).

Example: total effect is a "direct effect"



Assumes arrows from A into D nodes absent. E.g. A is PrEP for HIV Y is HIV incidence, D is death, reasonable to assume that treatment affects death only via its protection against HIV incidence.

When the total effect is not enough?

In the PrEP example, saying the total effect is "not enough" just means we need this additional story to interpret our estimates of total effects in terms of estimates of a "direct effect"

- this interpretation doesn't come only from the data, it's coupled with our story (our working knowledge of how the treatment works)

When we can't assume arrows from A to D absent, then the total effect is not a direct effect and we would need to target something else if we wish to estimate that.

What are our options for something other than the total effect in these settings: maybe because it isn't defined or we want an effect that isolates a particular treatment mechanism that the total effect doesn't isolate.

- We'll start with a popular measure that is often mistaken for a *direct treatment effect on the event of interest* but actually is not a causal effect at all.

Cause-specific hazard

Many authors have advocated reporting a cause-specific hazard ratio when competing events are present and interest is in "etiology"

- This is problematic because hazard ratios are not causal effects
- Even when they contrast counterfactual hazards
- Additional assumptions are always needed to claim that a hazard comparison equal some causal effect.
  - These assumptions are not going to be reasonable in routine settings, they are typically going to be made for convenience.
- So starting with a hazard contrast (whether ratio or difference) does not constitute articulating a causal effect.

Counterfactual Cause-Specific Hazard

Contrast of (discrete-time) *cause-specific hazards* under different treatments a is

$$\Pr[Y_k^{a=1} = 1 | D_{k-1}^{a=1} = Y_{k-1}^{a=1} = 0] \text{ vs. } \Pr[Y_k^{a=0} = 1 | D_{k-1}^{a=0} = Y_{k-1}^{a=0} = 0]$$

- Does not compare outcomes under different treatments in the same people
- Subdistribution hazard generally has the same problem.

$$\Pr[Y_k^{a=1} = 1 | Y_{k-1}^{a=1} = 0] \text{ vs. } \Pr[Y_k^{a=0} = 1 | Y_{k-1}^{a=0} = 0]$$

And it further doesn't remove the problems of the total effect: not defined when there are truncation events, captures paths through the blue arrows when arrows from A to D nodes

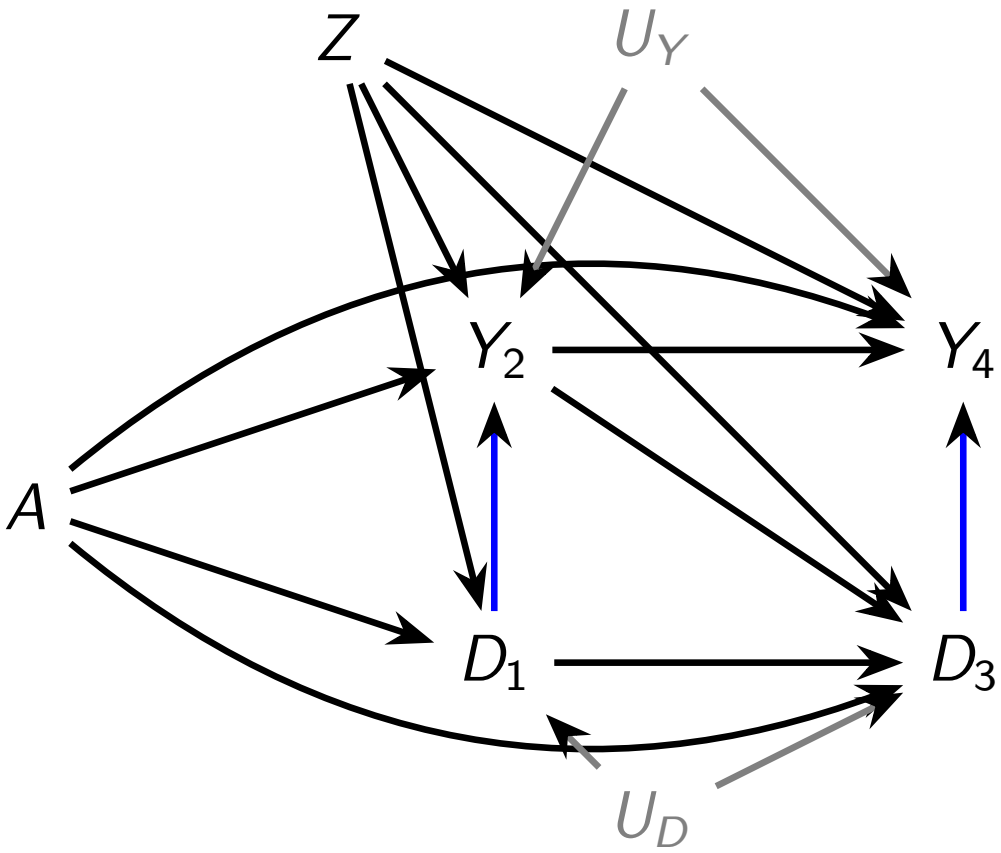
A case where cs-hazard comparison is a causal effect

One special case where the a contrast of *cause-specific hazards* under different treatments a is a causal effect is under the null (no arrows out of A on the causal DAG!)

$$\Pr[Y_k^{a=1} = 1 | D_{k-1} = Y_{k-1} = 0] \text{ vs. } \Pr[Y_k^{a=0} = 1 | D_{k-1} = Y_{k-1} = 0]$$

Conditioning sets are now the same because we can remove the counterfactual superscripts on each side. Of course, we don't typically do studies under the *a priori* assumption of no treatment effect, this is the point of the study.

Conditioning on all common causes



In principle, a counterfactual cause-specific hazard contrast might be equated to some causal effect if we could measure and condition on ALL common causes of all event processes: on this graph that would include Z, U_Y and U_D.

More than we need for direct effects

We have yet to actually define a direct effect when there are arrows from A to D not intersected by Y on the causal DAG. This is coming up in a moment!

- It turns to identify any of the notions of direct effect that we will consider, we actually don't need to measure or account for common causes of the each separate event processes over time (U_Y and U_D on our DAG)
- All we need to worry about are shared causes of the Y and D processes (Z on our DAG).

Hazards do not belong at the top of a causal road map, but they can play a role at the bottom!

- We can make special assumptions that estimating contrasts in hazards constitute unbiased estimates of causal effects.
- We will also see that in trials with "censoring" (or in observational studies), we can justify statistics for causal effects that involve estmating hazards as interim steps in the algorithm (what we code). This includes statistics for total effects!

Take home: if we conduct an analysis that articulates nothing more than interest in some hazard contrast, we have not clarified what causal question we have!

What are alternatives to the total effect?

If not a hazard ratio (or some hazard contrast), then what are our options for grounding analysis in a causal effect other than a total effect when

- there are arrows from A to D and we care about mechanism OR
- the total effect is not defined (there are truncation events in our study).

What other effects can we consider in these cases?

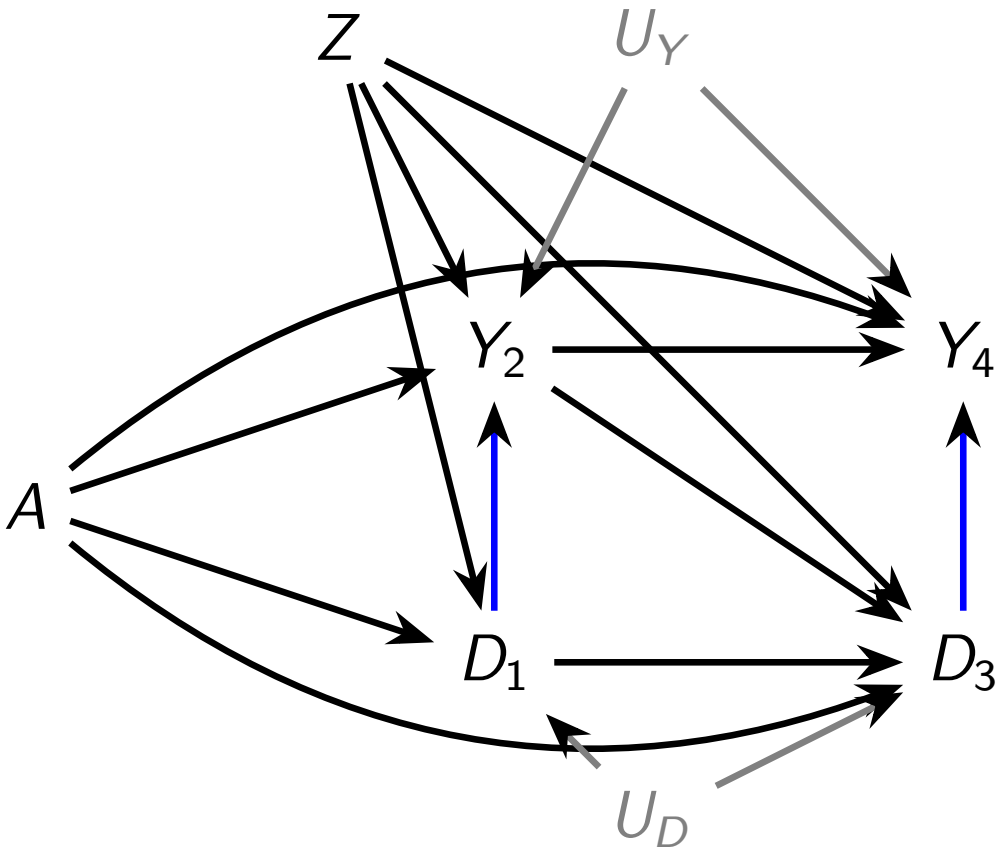
Existing proposals

We're now going to review some alternative notions of causal effect that do not capture treatment actions on the outcome of interest Y via treatment effects on competing or truncation events D when the total effect does capture this

- with some also, unlike a total effect, still defined in the presence of truncation events.

Crucial to remember: ALL of these proposals require stronger assumptions for identification than a total effect, with all requiring adjustment for covariates like Z (common causes of Y and D)

If you want something other than total effect, you need to measure Z



Particularly important to understand this when planning randomized trials where competing or truncation events are likely to occur.

Counterfactual contrast in net risks

Contrast in (discrete time) *marginal cumulative incidence functions* under different treatments a at some k is

$$\Pr[Y_k^{a=1,\overline{d}=\overline{0}} = 1] \text{ vs. } \Pr[Y_k^{a=0,\overline{d}=\overline{0}} = 1]$$

- This is a causal effect (compares outcomes in same units under different treatments)
- It refers to a counterfactual scenario where somehow competing events are eliminated over the follow-up. Therefore, by definition, this is a direct effect in that it cannot capture the treatment effect on death
- Example of a **controlled direct effect** – effect when we fix mediator to a constant (Robins & Greenland, 1992)

Controlled direct effect

- Unlike total effect, not guaranteed identified even in our perfect trial – must measure rich set of adjustment covariates to hope to identify this.
- Appealing feature of this notion is that the "road map" leads to familiar/popular estimation procedures for right-censored data – e.g. weighted Kaplan Meier
- Also defined when D is a truncation event

But even if we felt confident we estimated it well, is this really what motivated the investigator to do this study? How does it inform any action/decision/policy?

Survivor Average Causal Effect (SACE)

SACE (Robins, 1986; Frangakis & Rubin, 2002)– Total effect but in subset of original population who would never experience the competing event under either treatment level

$$\Pr[Y_k^{a=1} = 1 | \overline{D}_{k-1}^{a=1} = \overline{D}_{k-1}^{a=0} = 0] \text{ vs. } \Pr[Y_k^{a=0} = 1 | \overline{D}_{k-1}^{a=1} = \overline{D}_{k-1}^{a=0} = 0]$$

This does not refer to a setting involving impossible interventions on competing events like death. Also defined when D is a truncation event.

- Like the controlled direct effect it relies on assumptions not guaranteed even in a perfect trial.

But even if we could argue those assumptions hold, refers to a subpopulation who we cannot observe and may not exist. Is this really what motivated the investigator to do this study? How does it inform any action/decision/policy?

Pure (natural) effects

The pure (natural) direct effect (Robins & Greenland; Pearl) is another definition of direct effect, when death is the relevant "mediator" this is

- $$\Pr[Y_k^{a=1,\overline{D}_{k-1}^{a=0}} = 1] \text{ vs. } \Pr[Y_k^{a=0,\overline{D}_{k-1}^{a=0}} = 1]$$

Indirect analogue together sums to total effect. Refers to an effect under intervention where the other death process is "set" for each individual to the value it would take under placebo. Is this really the question? What could it inform now or in future?

What is the real question?

- If we have a causal question but the story motivating us does not justify a total effect or any of these notions of direct effect then what other options are there?
- Next Mats will introduce another way to define a causal effect when competing/truncation events exist that formalizes stories about what is of interest that are often told by investigators in these types of settings.