Causal Inference for Survival Outcomes:

An Introduction

Bianca L De Stavola for STRATOS TG 7

Great Ormond Street Institute of Child Health, University College London

LSHTM Centre for Statistical Methodology , 3rd Nov 2021

STRATOS INITIATIVE

Topic group 7: Causal Inference

Saskia le Cessie	Leiden University Medical Centre & Ghent University
Bianca De Stavola	University College London, UK
Vanessa Didelez	University of Bremen, Germany
Els Goetghebeur	Ghent University, Belgium
Erica Moodie	McGill University, Canada
Ingeborg Waernbaum	Uppsala University, Sweden

ofcaus.org

- To find a causal answer, start with a causal question.
  Then:
  1. specify exposure, outcome, population of interest, target causal effects (estimands)
  2. state assumptions for identification and estimation of effects from the data
  3. interpret results cautiously.

[Goetghebeur *et al.* 2020]

- To find a causal answer, start with a causal question.
  Then:
  1. specify exposure, outcome, population of interest, target causal effects (estimands)
  2. state assumptions for identification and estimation of effects from the data
  3. interpret results cautiously.

  [Goetghebeur *et al.* 2020]
- These are indeed the principles guiding RCTs and, for observational studies, are referred to as "target trial emulation" (TTE).

Estimands

- Accepted specification of estimands is via potential outcomes
- Y_a: outcome that would arise had we set A to take the value a by a well-defined (hypothetical) intervention.

Estimands

- Accepted specification of estimands is via potential outcomes
- Y_a: outcome that would arise had we set A to take the value a by a well-defined (hypothetical) intervention.
- For example (for a binary exposure A):
  - ACE=$E(Y_1) - E(Y_0)$
  - ATT=$E(Y_1|A = 1) - E(Y_0|A = 1)$

Estimands

- Accepted specification of estimands is via potential outcomes
- Y_a: outcome that would arise had we set A to take the value a by a well-defined (hypothetical) intervention.
- For example (for a binary exposure A):
  - ACE=$E(Y_1) - E(Y_0)$
  - ATT=$E(Y_1|A = 1) - E(Y_0|A = 1)$
- Identification requires linking data to these hypothetical quantities, *e.g.* invoking assumptions of no interference, consistency, and positivity.

Estimands

- Accepted specification of estimands is via potential outcomes
- Y_a: outcome that would arise had we set A to take the value a by a well-defined (hypothetical) intervention.
- For example (for a binary exposure A):
  - ACE=$E(Y_1) - E(Y_0)$
  - ATT=$E(Y_1|A = 1) - E(Y_0|A = 1)$
- Identification requires linking data to these hypothetical quantities, *e.g.* invoking assumptions of no interference, consistency, and positivity.
- Choice of estimation methods: each requiring additional assumptions (*e.g.* no unmeasured confounding, correct (semi-)parametric models).

Estimands for time-to-event outcomes

- Well known challenges [Andersen *et al.* 2020]:
  - Time origin and scale: from birth/entry/surgery?
  - Censoring: information on whether event is observed.

Estimands for time-to-event outcomes

- Well known challenges [Andersen *et al.* 2020]:
  - Time origin and scale: from birth/entry/surgery?
  - Censoring: information on whether event is observed.

Which aspects of time should we focus on when comparing alternative worlds?

Outline

1. Basic Principles
2. Causality and Survival Analysis
   - Estimands
   - Dealing with censoring
3. Identification and Estimation
   - Dealing with censoring
4. The Simulation Learner
   - Results
5. Summary
   - References

Survival Estimands (Version I)

Let T_a be the potential survival time if we set A to take the value a by a well-defined (hypothetical) intervention.

Survival Estimands (Version I)

Let T_a be the potential survival time if we set A to take the value a by a well-defined (hypothetical) intervention.

(A) Risk scale:
Differences in survival probabilities at relevant times

$$\text{ACE}(t) \;=\; P(T_1 > t) - P(T_0 > t), \quad t \;\text{in}[0, \tau]$$

This is the difference in (marginal) survival functions of POs,

Survival Estimands (Version I)

Let T_a be the potential survival time if we set A to take the value a by a well-defined (hypothetical) intervention.

(A) Risk scale:
Differences in survival probabilities at relevant times

$$\text{ACE}(t) \;=\; P(T_1 > t) - P(T_0 > t), \quad t \;\text{in}[0, \tau]$$

This is the difference in (marginal) survival functions of POs,

- Interpretation:
risk difference for no event by time t had random patient been treated versus not.

Survival Estimands (Version I)

(B) Hazard scale:

Contrast of hazards, for example $\frac{\lambda_1(t)}{\lambda_0(t)}$, where

$$\lambda_a(t) = lim_{h \to 0} \frac{1}{h} P(t \leq T_a < t + h | T_a \geq t)$$

Survival Estimands (Version I)

(B) Hazard scale:

Contrast of hazards, for example $\frac{\lambda_1(t)}{\lambda_0(t)}$, where

$$\lambda_a(t) = lim_{h \to 0} \frac{1}{h} P(t \leq T_a < t + h | T_a \geq t)$$

- Complication:

  Interpretational difficulties because of the built-in selection due to the conditioning on different subgroups ($T_0 \geq t$ and $T_1 \geq t$)).

Survival Estimands (Version I)

(C) Restricted Mean Survival Times (RMST):

$$\int_0^{\tau} P(T_1 > t)dt \quad - \quad \int_0^{\tau} P(T_0 > t)dt$$

These mean times are specific to τ

Survival Estimands (Version I)

(C) Restricted Mean Survival Times (RMST):

$$\int_0^{\tau} P(T_1 > t)dt \quad - \quad \int_0^{\tau} P(T_0 > t)dt$$

These mean times are specific to τ

(D) Mean Survival Time:

$$\int_0^{\infty} P(T_1 > t)dt \quad - \quad \int_0^{\infty} P(T_0 > t)dt$$

Preferred in econometrics, *e.g.* in `stteffects` in Stata

Survival Estimands (Version I)

(C) Restricted Mean Survival Times (RMST):

$$\int_0^{\tau} P(T_1 > t)dt \quad - \quad \int_0^{\tau} P(T_0 > t)dt$$

These mean times are specific to τ

(D) Mean Survival Time:

$$\int_0^{\infty} P(T_1 > t)dt \quad - \quad \int_0^{\infty} P(T_0 > t)dt$$

Preferred in econometrics, *e.g.* in `stteffects` in Stata

(E) Other scales, *e.g.* speed from Accelerated Failure Time Models.

Comments

- The choice between these estimands should be guided by their clinical relevance.
- In most settings these are contrasts on risk scale.
- Note however that hazard models are useful to derive such contrasts.

Comments

- The choice between these estimands should be guided by their clinical relevance.
- In most settings these are contrasts on risk scale.
- Note however that hazard models are useful to derive such contrasts.
- Whichever one is chosen, definitions above have no consideration of the impact of censoring.

Censoring

- Do we want to quantify causal effects in the absence in censoring?

Censoring

- Do we want to quantify causal effects in the absence in censoring?

  It depends on the source of censoring: for some it does not always make clinical sense to remove them,

  - Administrative reasons
  - Loss to follow-up
  - Treatment switching
  - Competing event

Censoring

- Do we want to quantify causal effects in the absence in censoring?

  It depends on the source of censoring: for some it does not always make clinical sense to remove them,

  - Administrative reasons ✓
  - Loss to follow-up ✓
  - Treatment switching
  - Competing event

Censoring

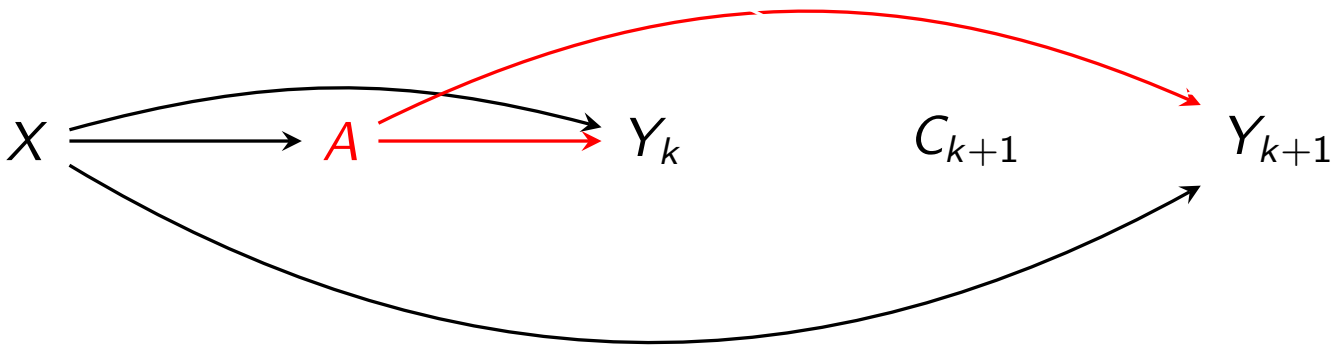
- Do we want to quantify causal effects in the absence in censoring?

  It depends on the source of censoring: for some it does not always make clinical sense to remove them,

  - Administrative reasons
  - Loss to follow-up
  - Treatment switching **??**
  - Competing event **??**

The problem with censoring

Consider a subset of the follow-up with administrative censoring:



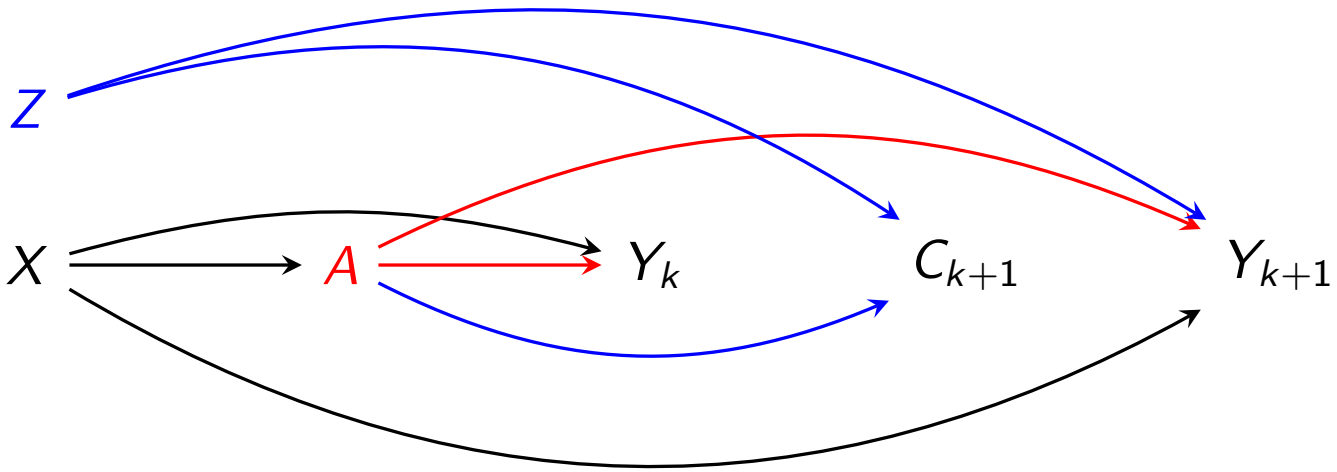
[Adapted from Young *et al.* , 2020]

A: baseline treatment
Y_k: outcome at time t_k
X: baseline confounders

C_k: censoring indicator at time t_k

The problem with censoring

Consider a subset of the follow-up with informative censoring:



[Adapted from Young *et al.* , 2020]

A: baseline treatment
Y_k: outcome at time t_k
X: baseline confounders
C_k: censoring indicator at time t_k Z: predictors of censoring and outcome

Survival Estimands (Version II)

- If happy to phrase the question "in the absence of censoring",
- Risk scale:

$$\text{ACE}(t) \;=\; P(T_{A=1,C=0} > t) - P(T_{A=0,C=0} > t), \quad t \;\text{in}[0,\tau]$$

where $T_{A=a,C=0}$ is the potential survival time if we set A to take the value a and $C = 0$ over the entire follow-up.

Survival Estimands (Version II)

- If happy to phrase the question "in the absence of censoring",
- Risk scale:

$$\text{ACE}(t) \;=\; P(T_{A=1,C=0} > t) - P(T_{A=0,C=0} > t), \quad t \;\text{in}[0,\tau]$$

  where $T_{A=a,C=0}$ is the potential survival time if we set A to take the value a and $C = 0$ over the entire follow-up.
- and similarly for the other estimands.

Identification

For identification of causal effects for a time-to-event outcome, invoke:

- No interference, consistency and positivity of the exposure
- No unmeasured confounding (NUC), *i.e.* : sufficient covariate information regarding treatment assignment confounding

Identification

For identification of causal effects for a time-to-event outcome, invoke:

- No interference, consistency and positivity of the exposure
- No unmeasured confounding (NUC), *i.e.* : sufficient covariate information regarding treatment assignment confounding
- In the presence of censoring we also require: sufficient covariate information regarding (possibly time-varying) 'common causes' of censoring and event.

Estimation of ACE(t)

With administrative censoring

$$\text{ACE}(t) \;=\; P(T_1 > t) – P(T_0 > t), \quad t \;\text{in}[0, \tau]$$

1. Model-based marginal counterfactual survival curves:

Estimation of ACE(t)

With administrative censoring

$$\text{ACE}(t) \;=\; P(T_1 > t) – P(T_0 > t), \quad t \;\text{in}[0, \tau]$$

1. Model-based marginal counterfactual survival curves:
   - (Sufficiently) flexible survival models (*e.g.* proportional or additive hazard model)

Estimation of ACE(t)

With administrative censoring

$$\text{ACE}(t) \;=\; P(T_1 > t) – P(T_0 > t), \quad t \;\text{in}[0, \tau]$$

1. Model-based marginal counterfactual survival curves:
   - (Sufficiently) flexible survival models (*e.g.* proportional or additive hazard model)
   - Derive individual-level predicted potential survival curves

Estimation of ACE(t)

With administrative censoring

$$\text{ACE}(t) \;=\; P(T_1 > t) – P(T_0 > t), \quad t \;\text{in}[0, \tau]$$

1. Model-based marginal counterfactual survival curves:
   - (Sufficiently) flexible survival models (*e.g.* proportional or additive hazard model)
   - Derive individual-level predicted potential survival curves
   - Standardisation to the distribution of the observed confounders *("empirical standardization")*

Estimation of ACE(t)

With administrative censoring

$$\text{ACE}(t) \;=\; P(T_1 > t) – P(T_0 > t), \quad t \;\text{in}[0, \tau]$$

1. Model-based marginal counterfactual survival curves:
   - (Sufficiently) flexible survival models (*e.g.* proportional or additive hazard model)
   - Derive individual-level predicted potential survival curves
   - Standardisation to the distribution of the observed confounders *("empirical standardization")*
   - Compute difference at selected values of t.

Estimation of ACE(t)

With administrative censoring

$$\text{ACE}(t) \;=\; P(T_1 > t) – P(T_0 > t), \quad t \;\text{in}[0, \tau]$$

2. IPTW of Marginal Structural Models (MSM):

Estimation of ACE(t)

With administrative censoring

$$\text{ACE}(t) \;=\; P(T_1 > t) – P(T_0 > t), \quad t \;\text{in}[0, \tau]$$

2. IPTW of Marginal Structural Models (MSM):
   - Fit propensity score (PS) model for treatment A

Estimation of ACE(t)

With administrative censoring

$$\text{ACE}(t) \;=\; P(T_1 > t) – P(T_0 > t), \quad t \;\text{in}[0, \tau]$$

2. IPTW of Marginal Structural Models (MSM):
   - Fit propensity score (PS) model for treatment A
   - Derive stabilized weights w from the predicted PS (IPTW)

Estimation of ACE(t)

With administrative censoring

$$\text{ACE}(t) \;=\; P(T_1 > t) – P(T_0 > t), \quad t \;\text{in}[0, \tau]$$

2. IPTW of Marginal Structural Models (MSM):
   - Fit propensity score (PS) model for treatment A
   - Derive stabilized weights w from the predicted PS (IPTW)
   - Fit survival MSM for treatment by reweighing the data using w

Estimation of ACE(t)

With administrative censoring

$$\text{ACE}(t) \;=\; P(T_1 > t) – P(T_0 > t), \quad t \;\text{in}[0, \tau]$$

2. IPTW of Marginal Structural Models (MSM):
   - Fit propensity score (PS) model for treatment A
   - Derive stabilized weights w from the predicted PS (IPTW)
   - Fit survival MSM for treatment by reweighing the data using w
   - Use the model to predict potential survival curves

Estimation of ACE(t)

With administrative censoring

$$\text{ACE}(t) \;=\; P(T_1 > t) – P(T_0 > t), \quad t \;\text{in}[0, \tau]$$

2. IPTW of Marginal Structural Models (MSM):
   - Fit propensity score (PS) model for treatment A
   - Derive stabilized weights w from the predicted PS (IPTW)
   - Fit survival MSM for treatment by reweighing the data using w
   - Use the model to predict potential survival curves
   - Compute difference at selected values of t.

Estimation of ACE(t)

With informative censoring

$$\text{ACE}(t) \;=\; P(T_{A=1,C=0} > t) - P(T_{A=0,C=0} > t), \quad t \;\text{in}[0,\tau]$$

Estimation requires dealing with the drivers of censoring:

Estimation of ACE(t)

With informative censoring

$$\text{ACE}(t) \;=\; P(T_{A=1,C=0} > t) - P(T_{A=0,C=0} > t), \quad t \;\text{in}[0,\tau]$$

Estimation requires dealing with the drivers of censoring:

Approach 1. Include predictors of missingness among the conditioning factors

Estimation of ACE(t)

With informative censoring

$$\text{ACE}(t) \;=\; P(T_{A=1,C=0} > t) - P(T_{A=0,C=0} > t), \quad t \;\text{in}[0,\tau]$$

Estimation requires dealing with the drivers of censoring:

Approach 1. Include predictors of missingness among the conditioning factors

Approach 2. Estimate inverse probability of censoring weights (IPCW) and combine them with IPTW.

The Simulation Learner

Inspiration: the Rotterdam study [Royston & Lambert, 2009; Sjolander, 2016]

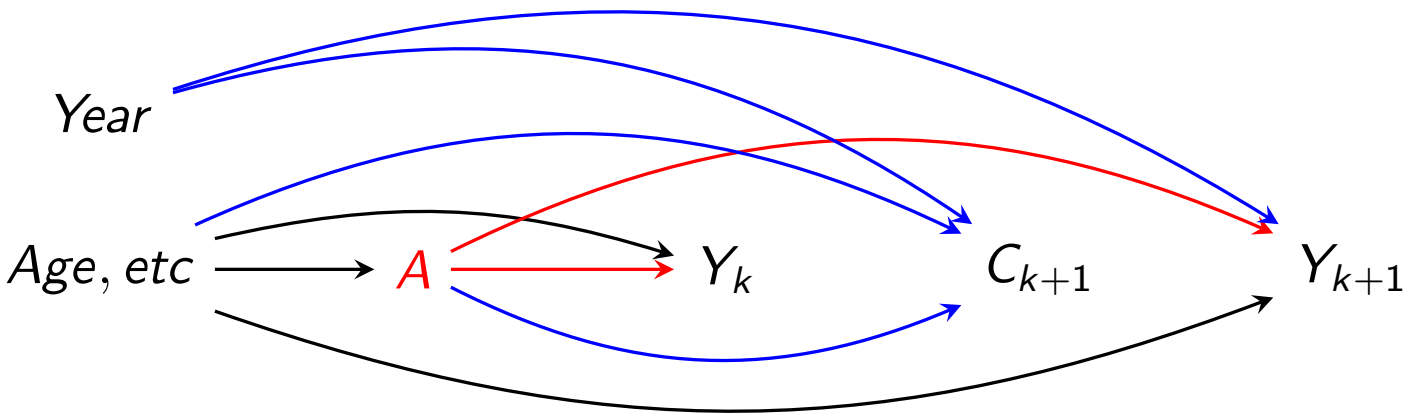
- About 3000 women who had undergone surgery for breast cancer and, for some, hormonal therapy was offered in 1978 to 1993
- Outcome of interest: overall mortality
- Strong negative confounding of the association between therapy and mortality
- Informative loss to follow-up driven by age and year of surgery
- Lack of positivity for younger women and those treated before 1985

The Simulation Learner

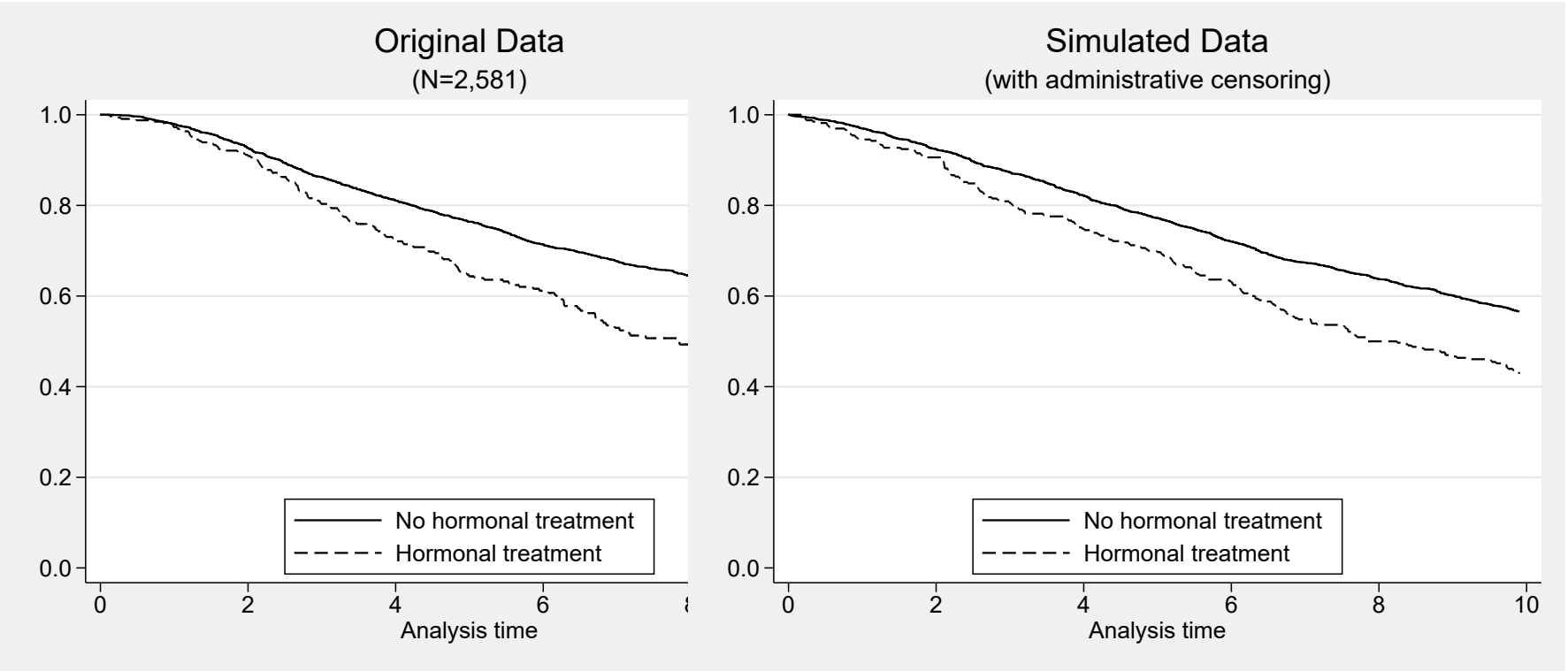
- Excluded women younger than 40 and with surgery before 1982
- Retained all the original confounders data

The Simulation Learner

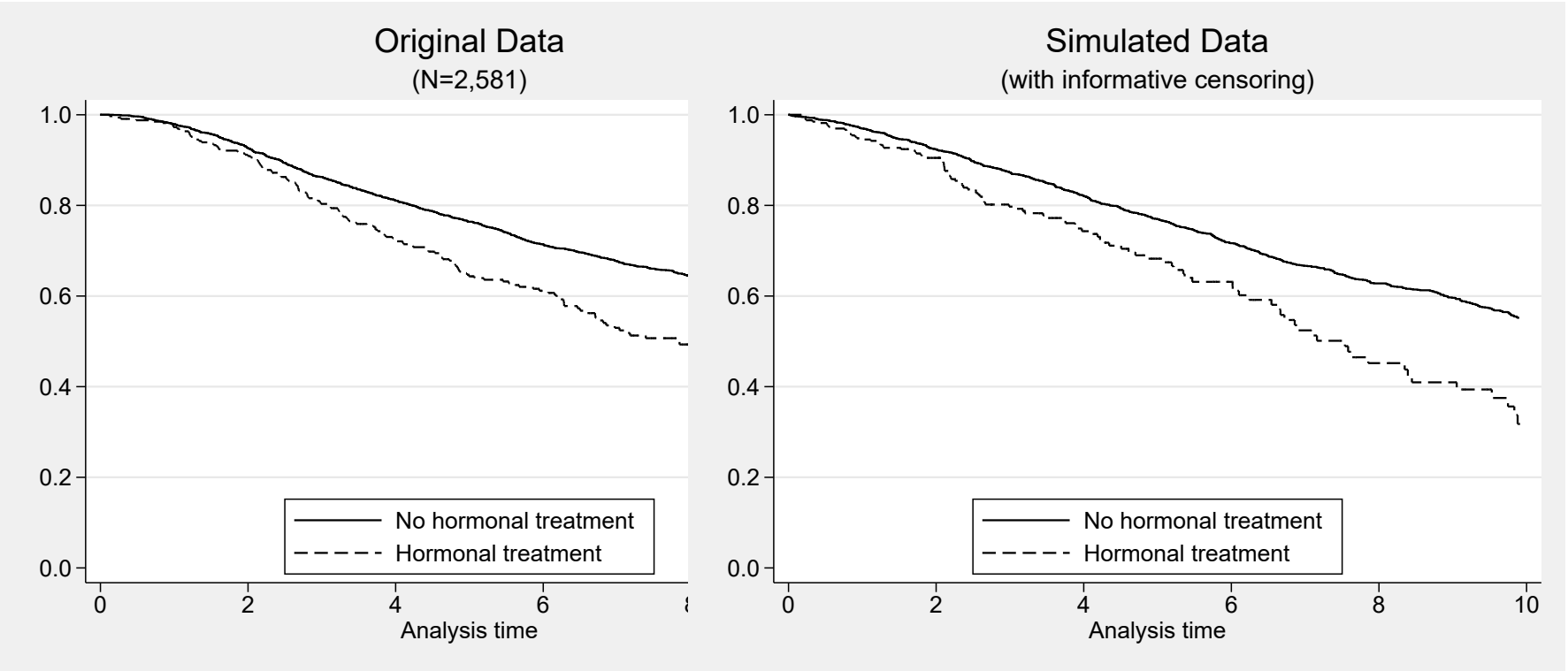
- Excluded women younger than 40 and with surgery before 1982
- Retained all the original confounders data
- Generated:
  - two versions of observed data: with / without informative censoring
  - potential survival times: $T_{A=0,C=0}$ and $T_{A=1,C=0}$



The Data

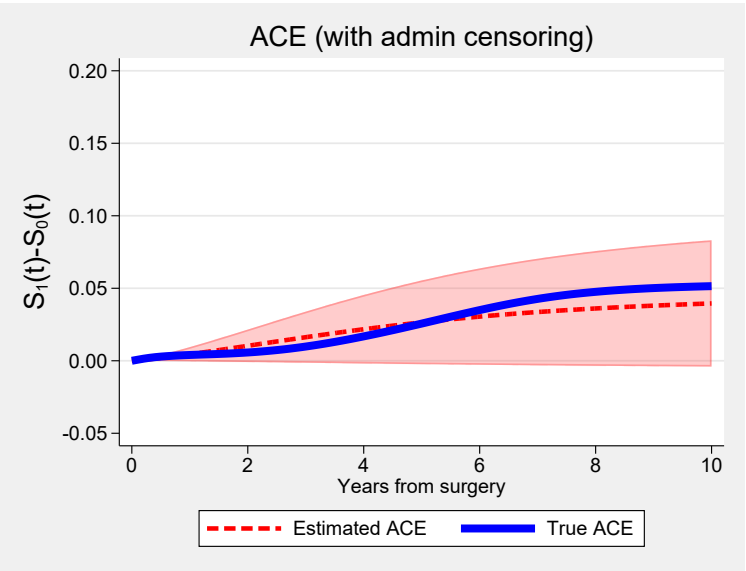

Original Data
(N=2,581)
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
Analysis time
No hormonal treatment
Hormonal treatment
Simulated Data
(with administrative censoring)
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
Analysis time
No hormonal treatment
Hormonal treatment

The Data


Original Data
(N=2,581)
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
Analysis time
No hormonal treatment
Hormonal treatment
Simulated Data
(with informative censoring)
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
Analysis time
No hormonal treatment
Hormonal treatment

Results: administrative censoring

1- Model-based marginal counterfactual survival curves

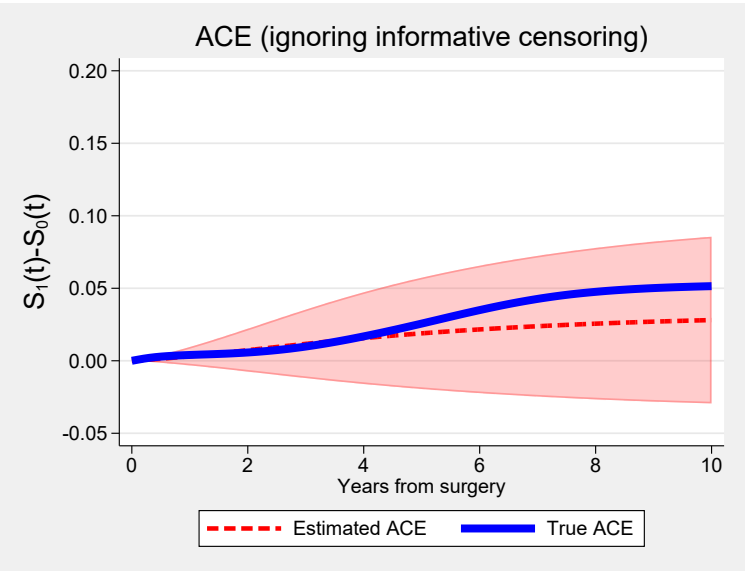


[Estimated in Stata using P. Lambert's `standsurv`]

Time	True ACE(t)	Estimated
5 yr	0.025	0.026
		(-0.002, 0.055)
10 yr	0.052	0.040
		(-0.004, 0.083)

Results: informative censoring

1- Model-based marginal counterfactual survival curves

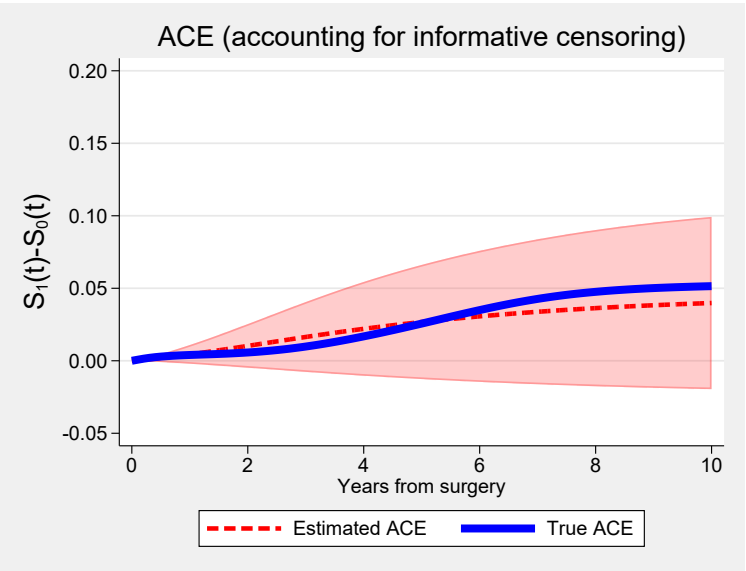


[Estimated in Stata using P. Lambert's `standsurv`]

Time	True ACE(t)	Estimated
5 yr	0.025	0.019
		(-0.019, 0.057)
10 yr	0.052	0.028
		(-0.029, 0.086)

Results: informative censoring

1- Model-based marginal counterfactual survival curves

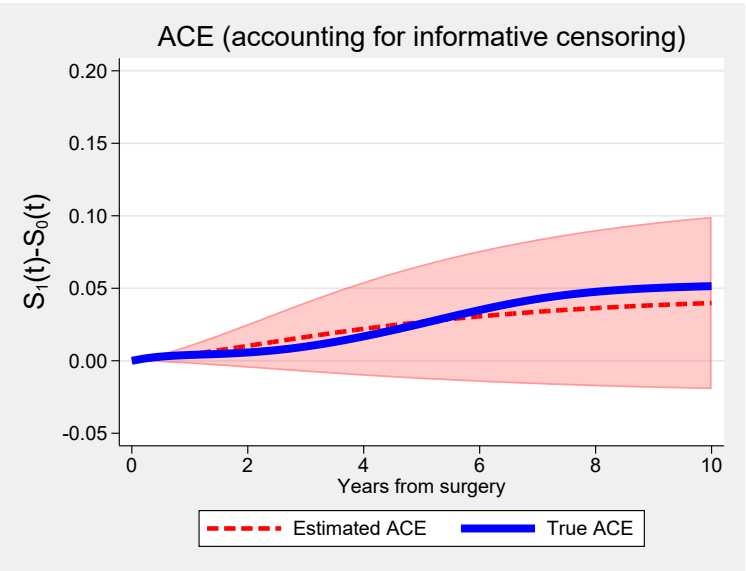


[Estimated in Stata using P. Lambert's `standsurv`]

Time	True ACE(t)	Estimated
5 yr	0.025	0.027
		(-0.013, 0.066)
10 yr	0.052	0.040
		(-0.020, 0.099)

Results: informative censoring

1- Model-based marginal counterfactual survival curves

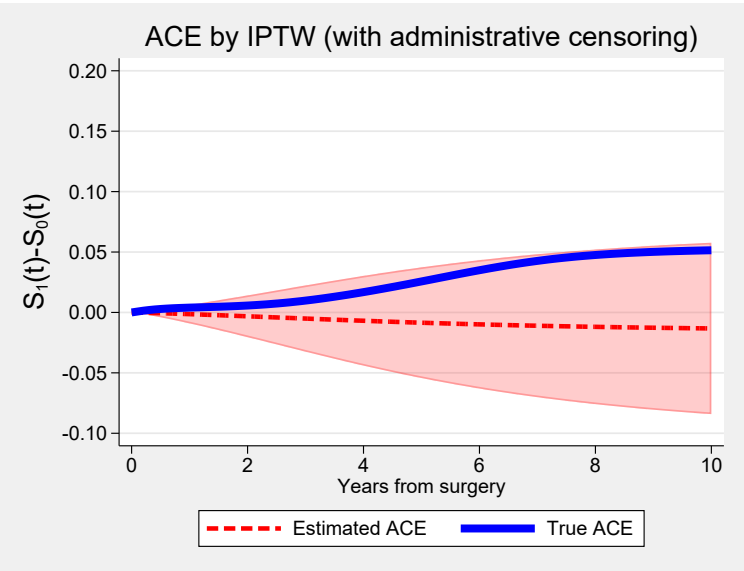


[Estimated in Stata using P. Lambert's `standsurv`]

Time	True RMST	Estimated
5 yr	0.049	0.066
		(-0.031, 0.164)
10 yr	0.264	0.240
		(-0.114, 0.592)

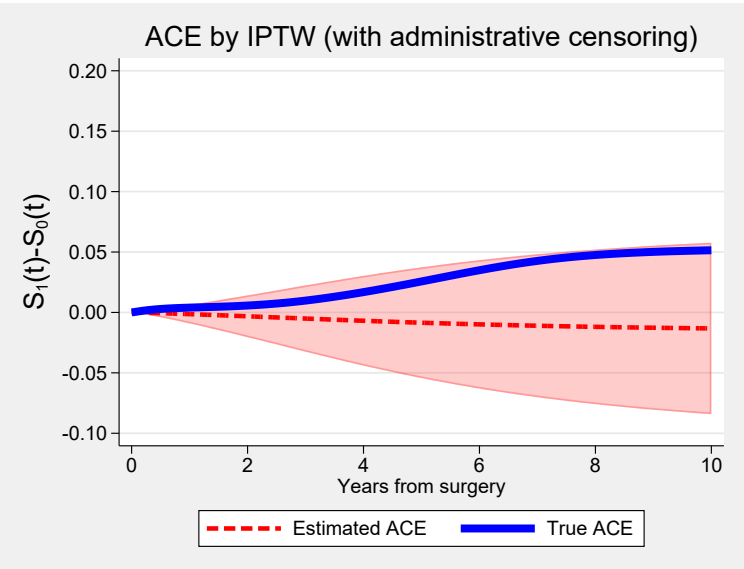
Results: administrative censoring

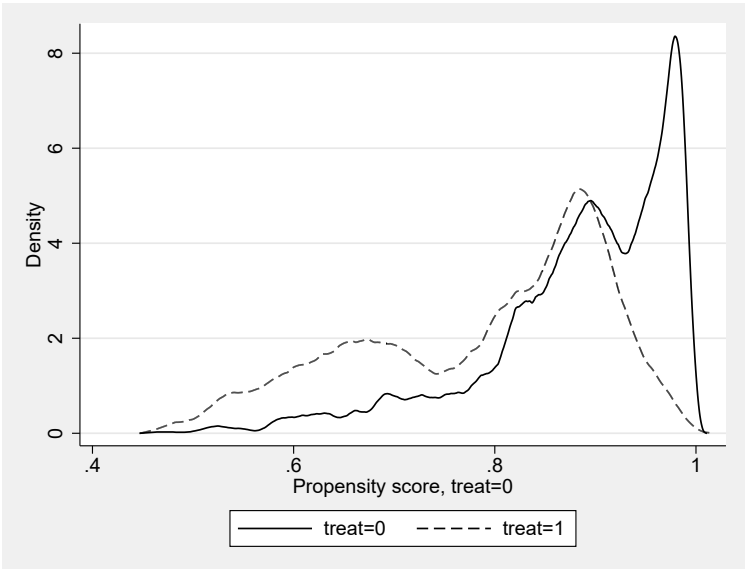
2- IPTW of MSMs



Results: administrative censoring

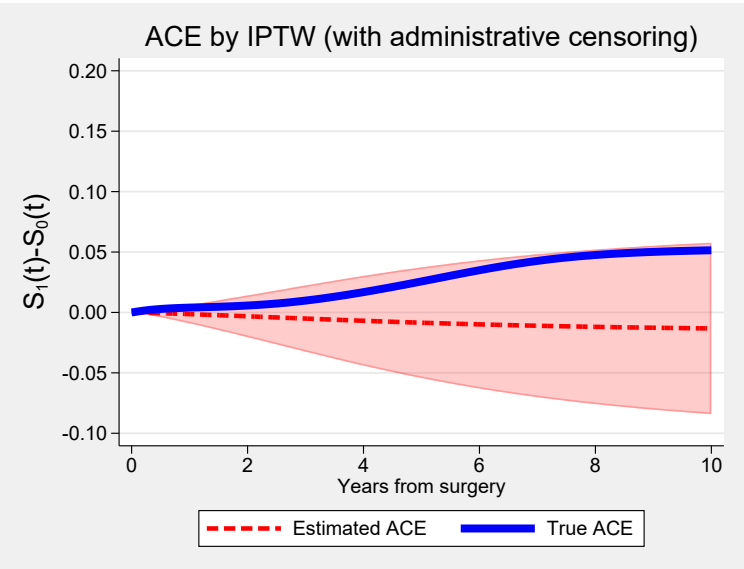
2- IPTW of MSMs

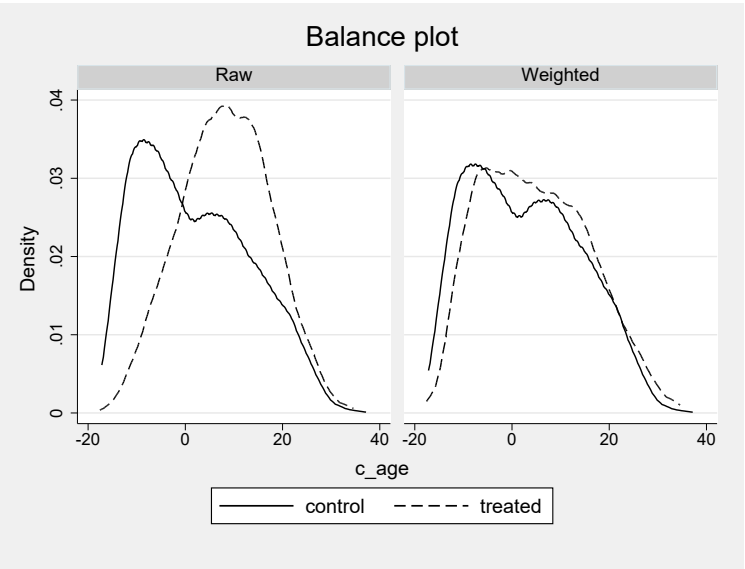




Results: administrative censoring

2- IPTW of MSMs





Summary

- Counterfactual-based causal inference has shifted the focus from model-based parameters to estimands defined irrespectively of any model.

Summary

- Counterfactual-based causal inference has shifted the focus from model-based parameters to estimands defined irrespectively of any model.
- This should free us from necessarily wanting to express causal effects on the hazard scale and choose meaningful, clinically relevant quantities.

Summary

- Counterfactual-based causal inference has shifted the focus from model-based parameters to estimands defined irrespectively of any model.
- This should free us from necessarily wanting to express causal effects on the hazard scale and choose meaningful, clinically relevant quantities.
- Dealing with censoring calls upon a careful choice of potential outcomes and of appropriate estimation approach.

Summary

- Counterfactual-based causal inference has shifted the focus from model-based parameters to estimands defined irrespectively of any model.
- This should free us from necessarily wanting to express causal effects on the hazard scale and choose meaningful, clinically relevant quantities.
- Dealing with censoring calls upon a careful choice of potential outcomes and of appropriate estimation approach.
- This was just a basic introduction:
  the real fun will start with the next speakers!

Summary

- Counterfactual-based causal inference has shifted the focus from model-based parameters to estimands defined irrespectively of any model.
- This should free us from necessarily wanting to express causal effects on the hazard scale and choose meaningful, clinically relevant quantities.
- Dealing with censoring calls upon a careful choice of potential outcomes and of appropriate estimation approach.
- This was just a basic introduction:
  the real fun will start with the next speakers!

THANK YOU!

References

- Andersen PK *et al.* (2021) Analysis of time-to-event for observational studies: Guidance to the use of intensity models, *Statistics in Medicine*, 40, 185–211
- Goetghebeur *et al.* (2020) Formulating causal questions and principled statistical answers, *Statistics in Medicine*, 39, 4922–4948
- Paul Lambert's blog: `https://pclambert.net/software/standsurv/`
- Young *et al.* (2020) A causal framework for classical statistical estimands in failure-time settings with competing events. *Statistics in Medicine* 39: 1199–1236