Emulating a target trial from observational data in the presence of immortal-time bias

Clémence Leyrat, Camille Maringe, Sara Benitez Majano, Matthew Smith, Aimilia Exarchakou, Bernard Rachet, Aurélien Belot

Cancer Survival Group, London School of Hygiene and Tropical Medicine, UK
Department of Medical Statistics, London School of Hygiene and Tropical Medicine, UK

Research supported by **Cancer Research UK** (grant number C7923/A18525)

Clinical context

Non-small-cell lung carcinoma (NSCLC):

- Most common type of lung cancer (≈85%)
- Leading cause of cancer death in the UK

Clinical context

Non-small-cell lung carcinoma (NSCLC):

- Most common type of lung cancer (≈85%)
- Leading cause of cancer death in the UK

Surgery: recommended treatment for early stage NSCLC

Clinical context

Non-small-cell lung carcinoma (NSCLC):

- Most common type of lung cancer ($\approx$85%)
- Leading cause of cancer death in the UK

Surgery: recommended treatment for early stage NSCLC

Mean age of NSCLC patients at diagnosis $\approx$73 years BUT:

- the chance of **receiving surgery decreases with age**[1]
- older patients often **excluded from clinical trials**

Clinical context

Non-small-cell lung carcinoma (NSCLC):

- Most common type of lung cancer (≈85%)
- Leading cause of cancer death in the UK

Surgery: recommended treatment for early stage NSCLC

Mean age of NSCLC patients at diagnosis ≈73 years BUT:

- the chance of **receiving surgery decreases with age**[1]
- older patients often **excluded from clinical trials**

$\Longrightarrow$ **Lack of available evidence** of the benefits of surgery on survival among older NSCLC patients

Using observational data

Observational studies: valuable sources of information for **causal inference**

Using observational data

Observational studies: valuable sources of information for **causal inference**

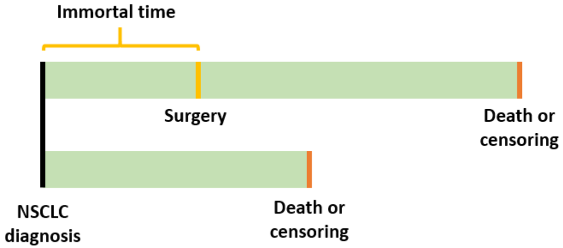
Challenge 1: confounding

Using observational data

Observational studies: valuable sources of information for **causal inference**

Challenge 1: confounding

Challenge 2: immortal-time bias



Trial emulation

Framework to **emulate trials from observational data**[2]

Trial emulation

Framework to **emulate trials from observational data**[2]

Involves the definition of a **target trial** (ideal trial) we would like to conduct

Trial emulation

Framework to **emulate trials from observational data**[2]

Involves the definition of a **target trial** (ideal trial) we would like to conduct

Then, a causal analysis of observational data can be viewed as an **attempt to emulate this target trial**

Trial emulation

Framework to **emulate trials from observational data**[2]

Involves the definition of a **target trial** (ideal trial) we would like to conduct

Then, a causal analysis of observational data can be viewed as an **attempt to emulate this target trial**

OXFORD JOURNALS American Journal of Epidemiology

Am J Epidemiol. 2016 Apr 15; 183(8): 758–764.
Published online 2016 Mar 18. doi: 10.1093/aje/kwv254

PMCID: PMC4832051
PMID: 26994063

Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available

Miguel A. Hernán* and James M. Robins

Emulation & immortal-time

In observational studies, immortal-time bias occurs when T_0 **and treatment initiation do not coincide**

Emulation & immortal-time

In observational studies, immortal-time bias occurs when T_0 **and treatment initiation do not coincide**

In our study: median time between NSCLC diagnosis (T_0) and surgery$=$ 29 days [0;49]

Emulation & immortal-time

In observational studies, immortal-time bias occurs when T_0 **and treatment initiation do not coincide**

In our study: median time between NSCLC diagnosis (T_0) and surgery$=$ 29 days [0;49]

ELSEVIER

Journal of Clinical Epidemiology

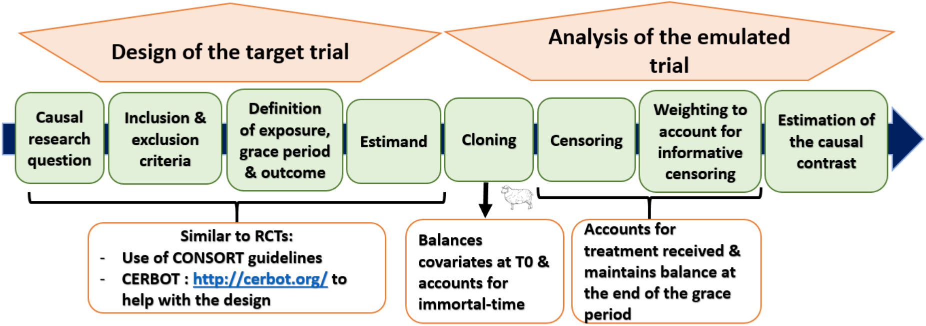
Volume 79, November 2016, Pages 70-75

Original Article

Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses

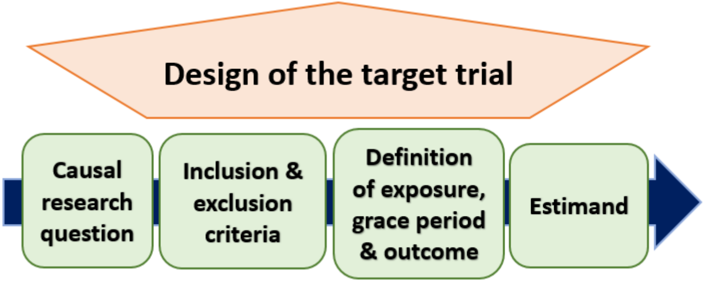
Miguel A. Hernán [a, b, c], Brian C. Sauer [d], Sonia Hernández-Díaz [a], Robert Platt [e, f, g], Ian Shrier [g]

Emulation & immortal-time

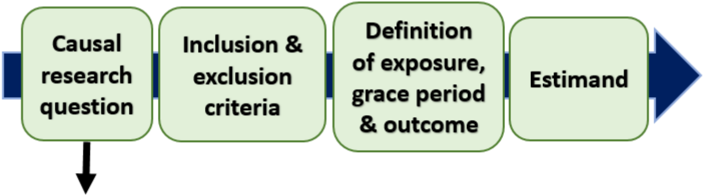


Data

Data: England Cancer registry linked to EHR data

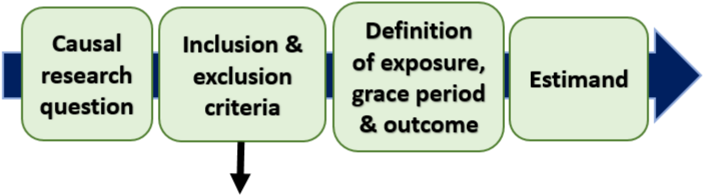


Question



What is the causal effect on 1 year survival of surgery within 6 months following NSCLC diagnosis among patients >70?

Inclusion criteria



- ✓ **NSCLC patients diagnosed in 2012**
- ✓ **70-89 years-old at diagnosis**
- ✓ **High performance status (<=2)**
- ✓ **Charlson's comorbidity index <=3**
- ✓ **TNM stage I-II**

Criteria for surgery among younger patients

Exposure and outcome

Causal research question → Inclusion & exclusion criteria → Definition of exposure, grace period & outcome → Estimand

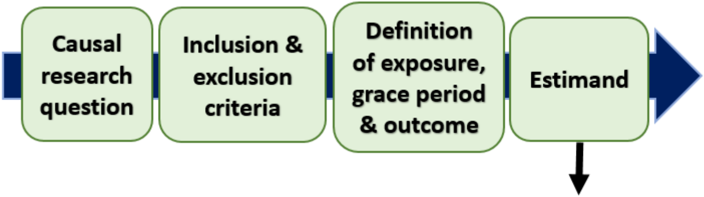
Study entry: NSCL diagnosis (n=2309)
Exposure: surgery following NSCLC diagnosis (n=1241)
Grace period: 6 months
Trial arms:

- Surgery within 6 months
- No surgery or surgery after 6 months

Outcome: time to death:

- 507 events in total
- 156 patients died within 6 months without surgery

Estimand

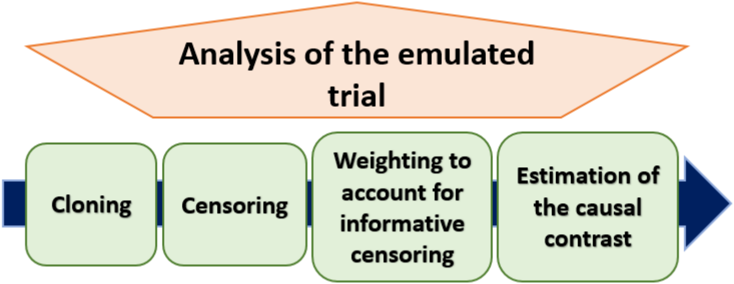


Per protocol effect: "intention" to treat unknown

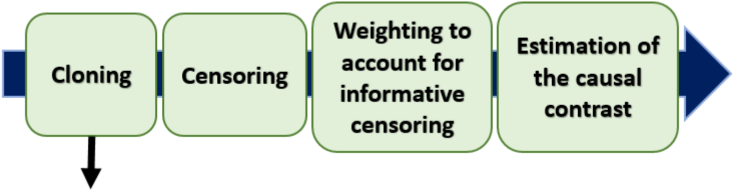
Causal contrasts:

- Difference in 1 year survival probabilities
- 1 year difference in restricted mean survival time

Analysis


Analysis of the emulated trial
Cloning
Censoring
Weighting to account for informative censoring
Estimation of the causal contrast

Estimand

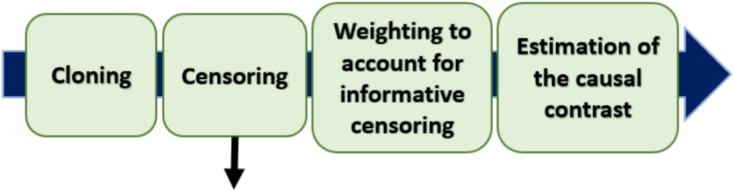


At T0, every patient is cloned

One clone assigned to each arm

Trial arms identical at baseline

Estimand

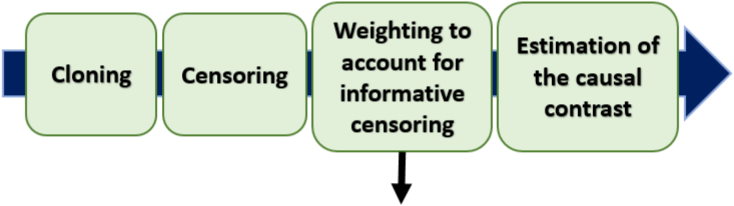


Observations are then censored when the treatment received is no longer compatible with the trial arm definition:

- **Surgery in the control group**
- **No surgery at 6 months in the surgery group**

Induces imbalance between arms over time

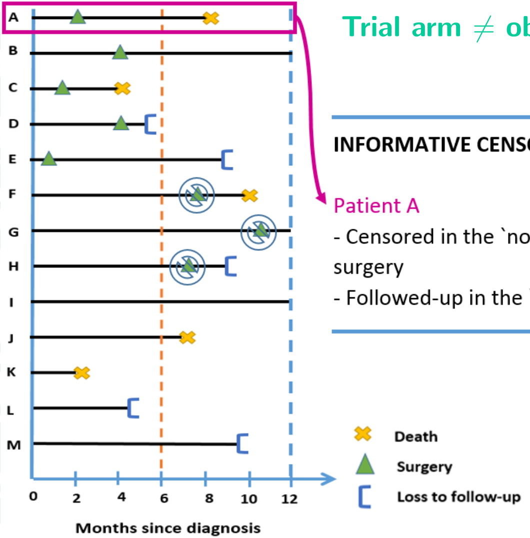
Estimand



Similar to propensity score weighting

Clones are reweighted based on their inverse probability of remaining uncensored at time t

Duplication and censoring



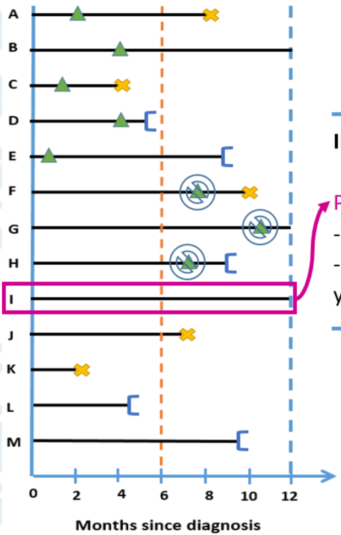
Trial arm $\neq$ observed treatment

INFORMATIVE CENSORING

Patient A

- Censored in the `no surgery' arm at time of surgery
- Followed-up in the `Surgery' arm until death

Duplication and censoring



Trial arm ≠ observed treatment

INFORMATIVE CENSORING

Patient I

- Censored in the `Surgery' arm at 6 months
- Followed-up in the `No Surgery' arm for 1 year

Censoring weights

Model 1: weight model

Weights **at each time of event** including: age, sex, deprivation, stage, performance status, comorbidities, emergency presentation

Censoring weights

Model 1: weight model

Weights **at each time of event** including: age, sex, deprivation, stage, performance status, comorbidities, emergency presentation

To estimate these weights we compared:

- **Cox proportional hazards** model
- **Flexible Cox** model using GAM: Cox model with smooth functions (penalized splines) for continuous covariates[3]
- **Survival forest**: ensemble method to estimate non parametrically the survivor function[4]

Outcome model

Model 2: outcome model

Comparison of **survival functions**

Outcome model

Model 2: outcome model

Comparison of **survival functions**

On the original dataset:

- Naive approach: unweighted **Kaplan-Meier**
- **G-computation** using a flexible hazard regression model (B-splines)[4]

Outcome model

Model 2: outcome model

Comparison of **survival functions**

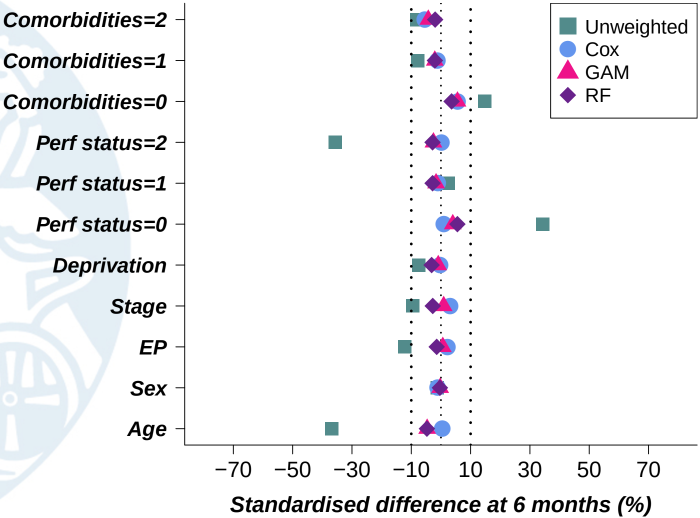
On the original dataset:

- Naive approach: unweighted **Kaplan-Meier**
- **G-computation** using a flexible hazard regression model (B-splines)[4]

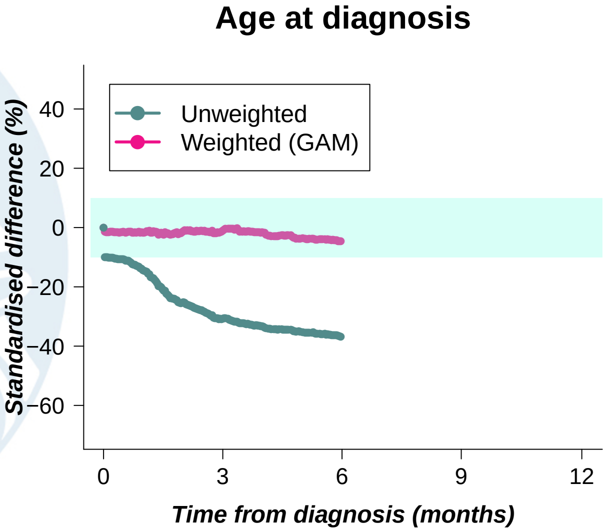
On the duplicated dataset:

- Unweighted **Kaplan-Meier**
- **Weighted Kaplan-Meier** (with the 3 sets of weights)

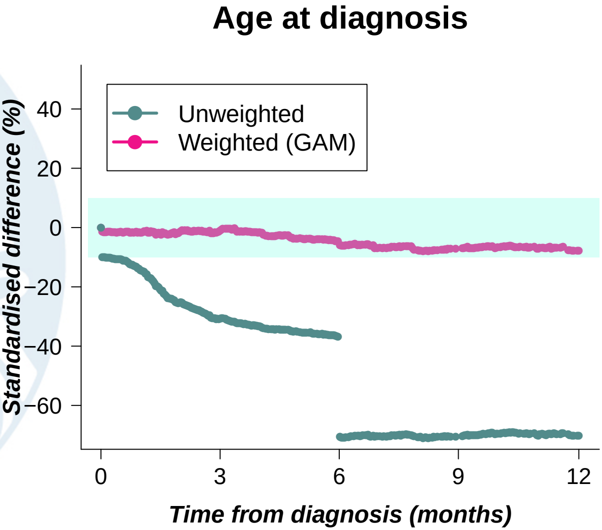
Balance at 6 months



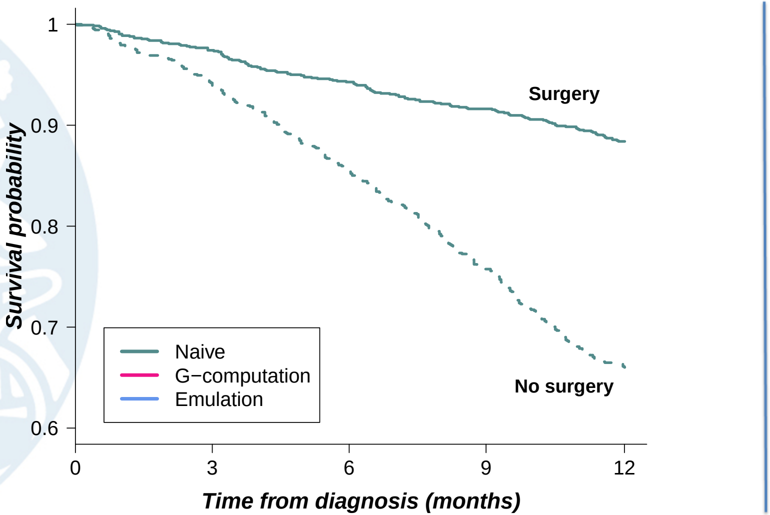
Balance over time: an example



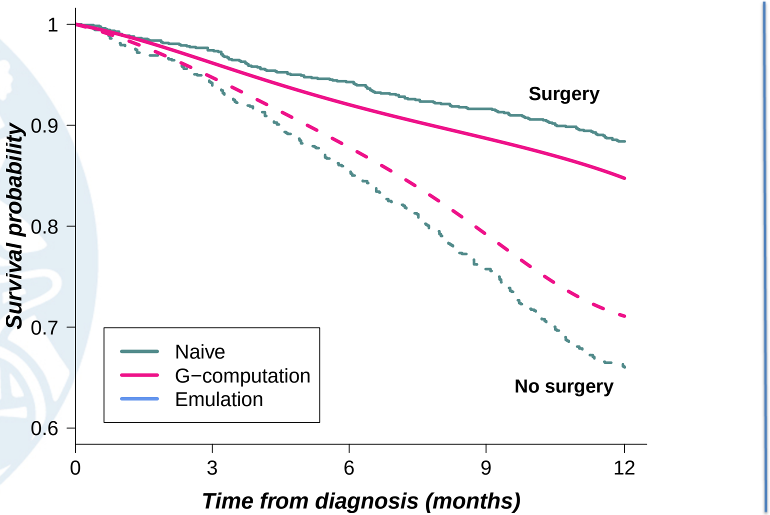
Balance over time: an example



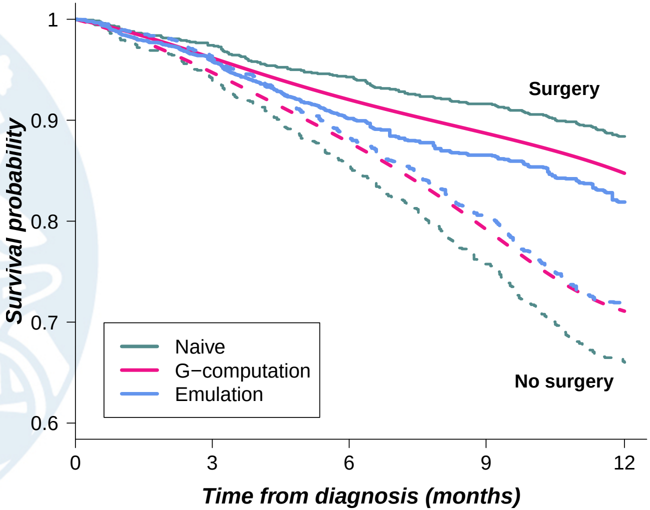
1-year survival



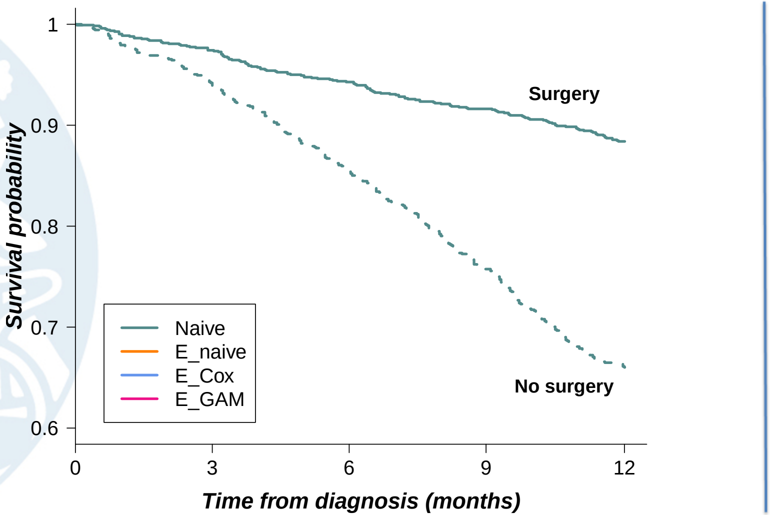
1-year survival



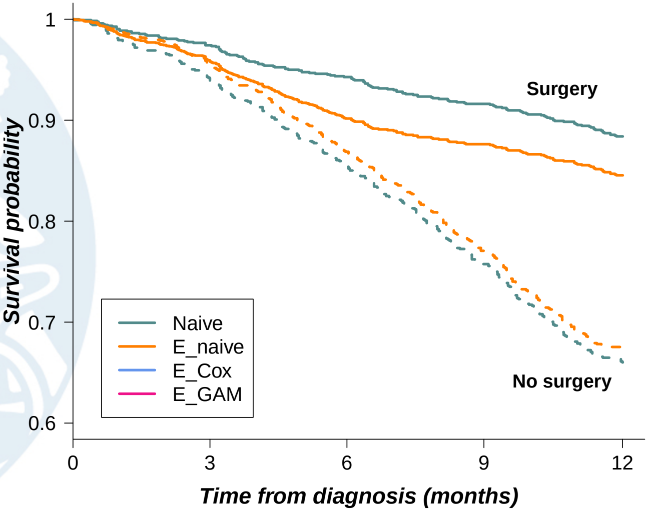
1-year survival



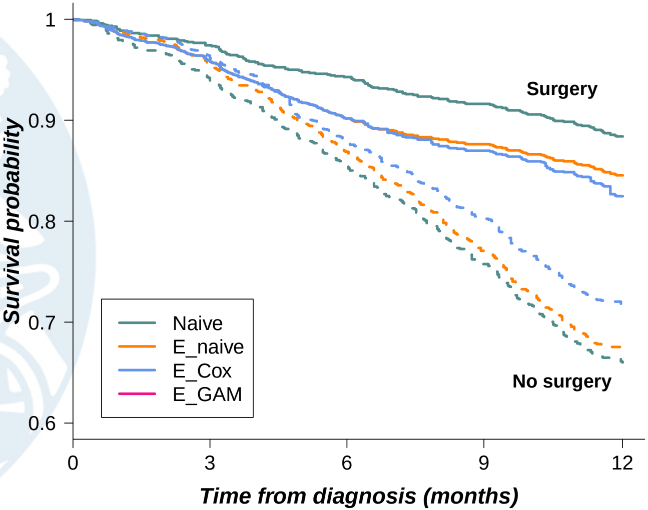
Impact of the weight model



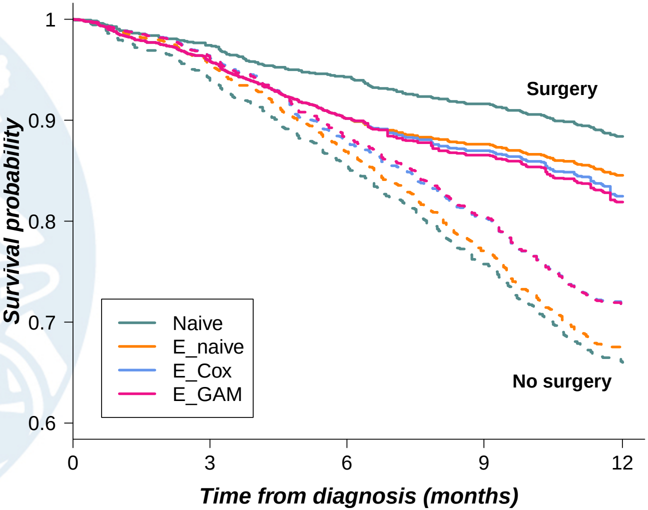
Impact of the weight model



Impact of the weight model



Impact of the weight model



HRs and Differences in Survival

Method	n	Difference in 1 year Survival (%)
Naïve (unadjusted)	2309	22.4 [18.1;26.9]
G-computation*	2309	13.7 [10.2;18.0]
Emulation**	4618	
Unweighted		17.4 [14.6; 20.1]
IPCW-Cox weights		10.9 [7.9; 15.3]
IPCW-GAM weights		10.4 -
IPCW-SF weights		10.7 ?

*Normal-based bootstrap confidence interval.

The **difference in RMST** is another useful measure in this context

Discussion

Illustration of the use of **trial emulation** to establish the causal effect of early surgery among older NSCLC patients

Discussion

Illustration of the use of **trial emulation** to establish the causal effect of early surgery among older NSCLC patients

Balance between arms over time using **graphical methods**

Discussion

Illustration of the use of **trial emulation** to establish the causal effect of early surgery among older NSCLC patients

Balance between arms over time using **graphical methods**

Better balance obtained using a **flexible** weight model

Discussion

Illustration of the use of **trial emulation** to establish the causal effect of early surgery among older NSCLC patients

Balance between arms over time using **graphical methods**

Better balance obtained using a **flexible** weight model

Further work needed to:

- Develop more flexible **weighted analysis models**
- Determine how to **estimate the variance** of different measures of interest, accounting for uncertainty in weight estimation
- Investigate the performance of **survival forests** in this context

References

[1] Belot A *et al*. **Association between age, deprivation and specific comorbid conditions and the receipt of major surgery in patients with non-small cell lung cancer in England: A population-based study**. Epub ahead of print: 2018. doi:10.1136/ thoraxjnl-2017-211395

[2] Hernan M and Robins J. **Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available**. Am J Epidemiol. 2016 Apr 15; 183(8): 758-764.

[3] Wood S. **Generalized Additive Models: An Introduction with R**, Second Edition. Chapman & Hall/CRC Texts in Statistical Science.

[4] Charvat H *et al*. **A multilevel excess hazard model to estimate net survival on hierarchical data allowing for non-linear and non-proportional effects of covariates**. Stat Med 2016. doi: 10.1002/sim.6881