Assessing treatment effect using registry data

Brice Ozenne

Dantrip 28-10-16

UNIVERSITY OF COPENHAGEN

Motivation

OBJECTIVE: decide whether a treatment is beneficial
→ for a give time horizon `1 year`

MATERIAL: registry data

- observational data (i.e. non-randomized)
- long term follow-up
- large number of patients

What do we mean by beneficial:

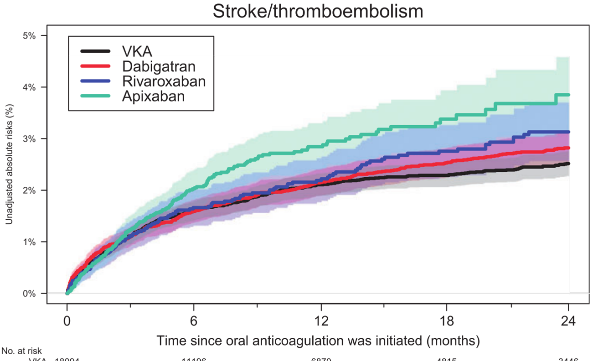
- does the treatment reduce the 1-year risk of developing the disease ?

Plan

1. Estimating a 1-year risk of a disease using registry data
   → model checking [Cox]

2. Estimating a treatment effect using registry data
   → model checking [new strategy]

   e.g. Staerk et al. 2016:

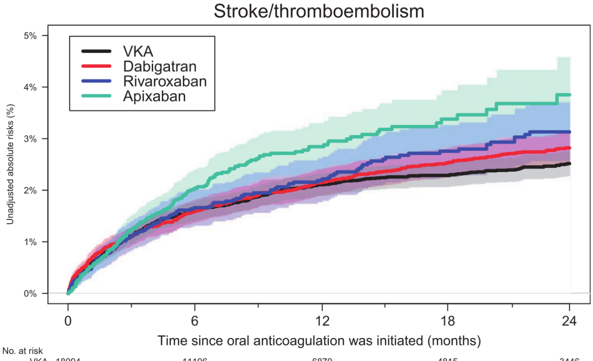


Plan

1. Estimating a 1-year risk of a disease using registry data
   → model checking [Cox]

2. Estimating a treatment effect using registry data
   → model checking [new strategy]

   e.g. Staerk et al. 2016:



Absolute risk

Definition

1-year absolute risk:

- chance that a person will be diagnosed with the event in 1 year
  $\rightarrow$ depends on the risk of the event λ_{event}
  $\rightarrow$ depends on the risk of death λ_{death}

$$r_{event}(t|X) = \underbrace{\int_0^t}_{\text{addition over time}} \underbrace{S_0(s|X)}_{\text{survival at to time s}} \underbrace{\lambda_{event}(s|X)}_{\substack{\text{immediate risk of} \\ \text{the event at time s}}} ds$$

X: covariates like age, gender ...

Considering registry data, are involved:

- the event of interest
- competing risks, e.g. death
  $\rightarrow$ will prevent the observation of the event

Definition

1-year absolute risk:

- chance that a person will be diagnosed with the event in 1 year
  → depends on the risk of the event λ_{event}
  → depends on the risk of death λ_{death}

$$r_{event}(t|X) = \underbrace{\int_0^t}_{\text{addition over time}} \underbrace{S_0(s|X)}_{\text{survival at to time s}} \underbrace{\lambda_{event}(s|X)}_{\text{immediate risk of the event at time s}} ds$$

X: covariates like age, gender …

Considering registry data, are involved:

- the event of interest
- competing risks, e.g. death
  → will prevent the observation of the event

Definition

1-year absolute risk:

- chance that a person will be diagnosed with the event in 1 year
  $\rightarrow$ depends on the risk of the event λ_{event}
  $\rightarrow$ depends on the risk of death λ_{death}

$$r_{event}(t|X) = \underbrace{\int_0^t}_{\text{addition over time}} \underbrace{S_0(s|X)}_{\text{survival at to time s}} \underbrace{\lambda_{event}(s|X)}_{\substack{\text{immediate risk of} \\ \text{the event at time s}}} ds$$

X: covariates like age, gender ...

Considering registry data, are involved:

- the event of interest
- competing risks, e.g. death
  $\rightarrow$ will prevent the observation of the event

Definition

1-year absolute risk:

- chance that a person will be diagnosed with the event in 1 year
  $\rightarrow$ depends on the risk of the event λ_{event}
  $\rightarrow$ depends on the risk of death λ_{death}

$$r_{event}(t|X) = \underbrace{\int_0^t}_{\text{addition over time}} \underbrace{S_0(s|X)}_{\text{survival at to time s}} \underbrace{\lambda_{event}(s|X)}_{\substack{\text{immediate risk of} \\ \text{the event at time s}}} ds$$

X: covariates like age, gender ...

Considering registry data, are involved:

- the event of interest
- competing risks, e.g. death
  $\rightarrow$ will prevent the observation of the event

Cause specific Cox model

One Cox regression for each competing risk:

$$\lambda_{event}(t|X) = \lambda_{0,event}(t)\exp(X\beta_{event})$$
$$\lambda_{death}(t|X) = \lambda_{0,death}(t)\exp(X\beta_{death})$$

We can then estimate the overall survival.

Cause specific Cox model

One Cox regression for each competing risk:

$$\lambda_{event}(t|X) = \lambda_{0,event}(t)\exp(X\beta_{event})$$
$$\lambda_{death}(t|X) = \lambda_{0,death}(t)\exp(X\beta_{death})$$

We can then estimate the overall survival.

- no event
- not dead

Cause specific Cox model

One Cox regression for each competing risk:

$$\lambda_{event}(t|X) = \lambda_{0,event}(t)\exp(X\beta_{event})$$
$$\lambda_{death}(t|X) = \lambda_{0,death}(t)\exp(X\beta_{death})$$

We can then estimate the overall survival.

$$S_0(t|X) = \exp\left(-\int_0^t \lambda_{death}(s|X) + \lambda_{event}(s|X)ds\right)$$

In R

```
> library(riskRegression)
> data(Melanoma)
> fit1 <- CSC(formula=Hist(time,status)~sex+invasion+age,
+              data=Melanoma)
fit1$models$`Cause 1`
Call:
survival::coxph(...)
                     coef exp(coef) se(coef)    z       p
sexMale           0.66338   1.94135  0.26632 2.49 0.01274
invasionlevel.1 1.03717   2.82122  0.32824 3.16 0.00158
invasionlevel.2 1.40323   4.06830  0.38074 3.69 0.00023
age               0.00982   1.00987  0.00834 1.18 0.23884
```

In R

```
> head(Melanoma[1:2,c("sex","invasion","age")])
      sex invasion age
    1 Male  level.1  76
    2 Male  level.0  56
> predictRisk(fit1, newdata = Melanoma[1:2,],
                  cause = 1, time = 365.25)
         365.25
[1,] 0.06441670
[2,] 0.01992289
```

Summary

We can easily compute the absolute risk

- using one Cox model for the event of interest
- using another Cox model for the competing events

But now we have to check the assumptions for each Cox model !

Cox model assumptions

Assumptions:

1. proportional hazard (PH) assumption
2. (linear) functional form
3. (absence of) interaction

[Not covered] non-informative censoring, influential observations

Checking Cox model assumptions

Model checking is more complex compared to a linear regression

- several types of residuals
- many different diagnostic tools
    - validity of the null hypothesis

      `e.g. PH vs. non PH`
    - against a specific alternative hypothesis

      `e.g. quadratic vs. linear effect age`

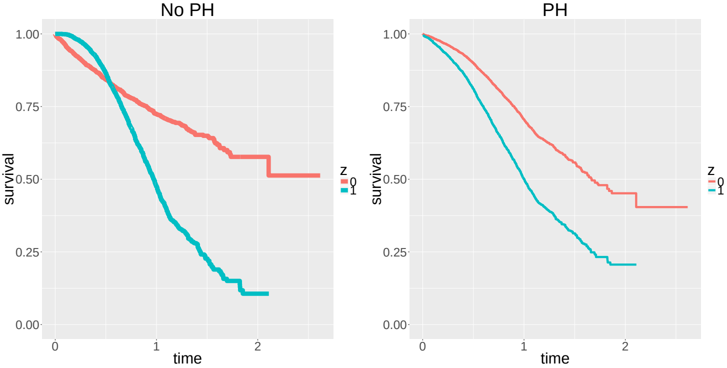
(1) Checking Proportional hazard assumption

Cox model:

$$\lambda(t|X) = \lambda_0(t)e^{\beta X}$$

Here we assume $\beta \perp\!\!\!\perp t$

- Visual checking with Kaplan Meier



(1) Checking Proportional hazard assumption

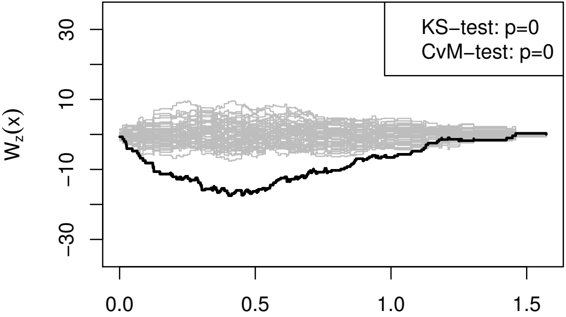
Cox model:

$$\lambda(t|X) = \lambda_0(t)e^{\beta X}$$

Here we assume $\beta \perp\!\!\!\perp t$

- Statistical test: $(\mathcal{H}_0)$ the PH assumption holds, i.e. the cumulative score process follows a brownian bridge

```
> plot(gof::cumres(coxph))
```

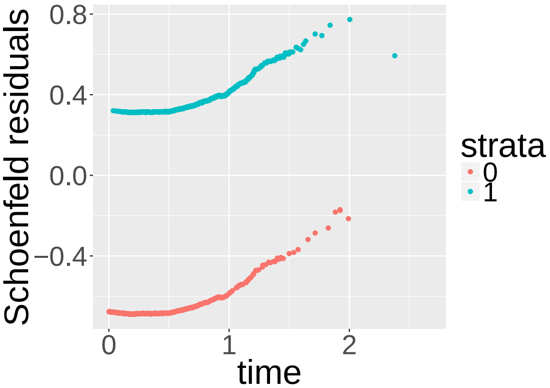


(1) Remedies for non proportional hazard

Strategy 1: find the problematic variable and the type of time dependency

- Display of the Schoenfeld residuals (Grambsch et al. 1994)

$$\mathbb{E}\left[r_{ij}\right] \approx \beta_j(t_i) - \hat{\beta}_j$$

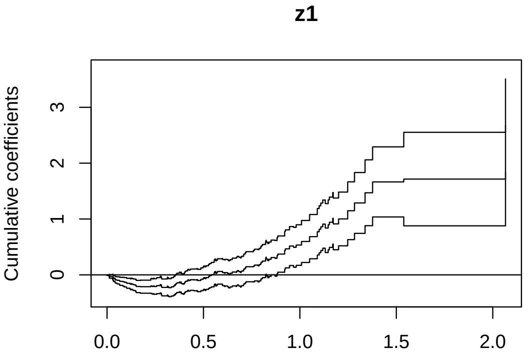


(1) Remedies for non proportional hazard

Strategy 1: find the problematic variable and the type of time dependency

- Use a Cox model with time varying effects

$$\lambda(t|X) = \lambda_0(t)e^{\beta(t)X}$$



```
> plot(timereg::timecox(Surv(time,status) ∼ z,
```

(1) Remedies for non proportional hazard

Strategy 1: find the problematic variable and the type of time dependency

Strategy 2: stratification
Cox model:

$$\lambda(t|X, treatment) = \lambda_0(t)e^{\beta X+\gamma treatment}$$

Stratified Cox model:

$$\lambda(t|X, treatment) = \lambda_{0,treatment}(t)e^{\beta X}$$

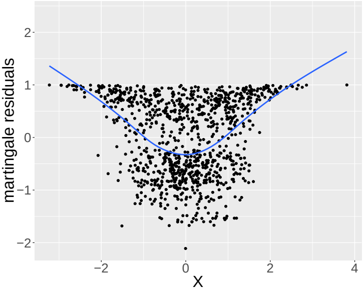
(2) Checking the functional form

$$\lambda(t|X, T) = \lambda_0 e^{\beta X}$$

Here we assume the log of the risk increase linearly with X, e.g. with age.

Diagnostic tools:

- Display martingale residuals
- Comparison with model including a quadratic term or spline



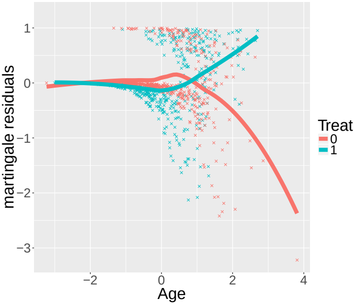
(3) Checking possible interactions

$$\lambda(t|X, T) = \lambda_0 e^{\beta X + \gamma \mathit{treatment}}$$

Here we assume that the risk increase independently with X and with *treatment*

Diagnostic tools:

- Display martingale residuals
- Comparison with a model with interactions



Limits

In practice model validation is tedious:

- large number of tests
  - at least 2 per variables + PH
    (i.e. linearity and interaction with treatment)
  - competing risks: two Cox models to check
- unclear alternative hypothesis
  - residual plot can be hard to interpret
- large *n* small *p*
  - overpowered tests (Grøn et al. 2016)
    → may detect unimportant deviations to hypothesis

Average treatment effect

Observational vs. randomized study

Randomized experiment

- eliminates confounding
- balances all risk factors: known **AND** unknown

→ causal interpretation

Observational studies

- can **ONLY** account for known and measured risk factors

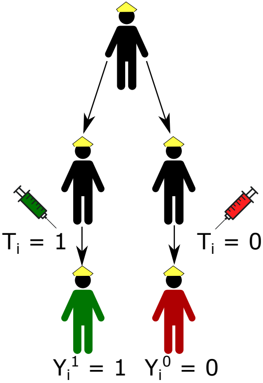
→ establish associations

Causal inference theory:

- causal interpretation (under hypothesis) in observational studies

Counterfactual outcomes

$\mathcal{H}$ypothetical world



Counterfactual outcomes

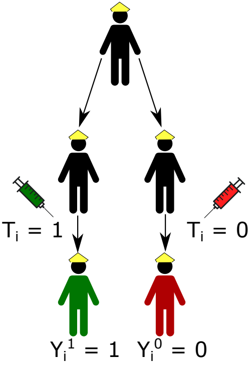
$\mathcal{H}$ypothetical world

We can measure for individual i at time t:
${Y_i}^{T=1}(t)$, outcome using intervention 1
${Y_i}^{T=0}(t)$, outcome using intervention 0

We can estimate

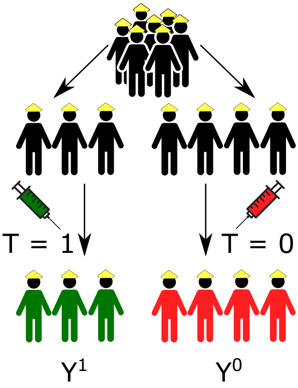
$${Y_i}^{T=1}(t) - {Y_i}^{T=0}(t)$$

the individual causal effect at t



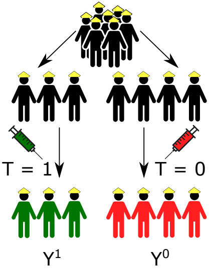
Counterfactual outcomes

$\mathcal{R}$eal world



Counterfactual outcomes

$\mathcal{R}$eal world



We only can measure :
$Y_i^{T=1}(t)$ OR $Y_i^{T=0}(t)$

We can infer the average causal effect:

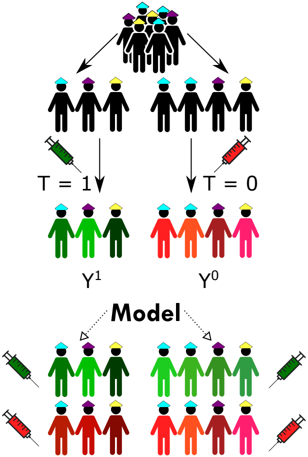
$$\begin{aligned} ACE(t) &= \mathbb{E}\left[Y^{T=1}(t) - Y^{T=0}(t)\right] \\ &= \mathbb{E}\left[Y^{T=1}(t)\right] - \mathbb{E}\left[Y^{T=0}(t)\right] \end{aligned}$$

e.g. (no confounder)

$$\widehat{ACE}(t) = \sum_{i=1}^{n_1} Y_i^{T=1}(t) - \sum_{j=1}^{n_2} Y_j^{T=0}(t)$$

G formula

$\mathcal{R}$eal world: confounders



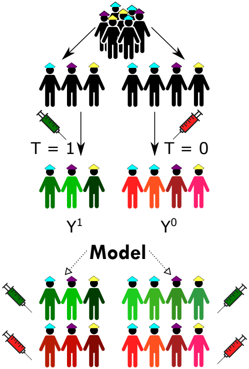
G formula

$\mathcal{R}$eal world: confounders

Statistical model: $\mathbb{E}[Y|X,T]$

$$ACE(t) = \sum_{i=1}^{n} \mathbb{E}[Y_i(t)|X_i, T=1] \\ -\mathbb{E}[Y_i(t)|X_i, T=0]$$

Here $Y(t)|X,T$ is the absolute risk



Workflow (Christiansen et al. 2015)

1. Define the population of interest

   `Patients with first-time ischemic stroke (n=19223)`

   `Exclusion criteria: atrial fibrillation ...`
2. Define the intervention (T)
3. Define the event of interest (Y)
4. Identify the possible competing events (D)
5. Identify the possible confounders/pronostic variable (X)
6. Define a statistical model for relating Y, T, and X

Workflow (Christiansen et al. 2015)

1. Define the population of interest
2. Define the intervention (T)

```
e.g.: antiplatelet regimens for secondary stroke prevention
      T=0:  ASA
      T=1:  Clopidogrel
      T=2:  ASA+Clopidogrel
```

3. Define the event of interest (Y)
4. Identify the possible competing events (D)
5. Identify the possible confounders/pronostic variable (X)
6. Define a statistical model for relating Y, T, and X

Workflow (Christiansen et al. 2015)

1. Define the population of interest
2. Define the intervention (T)
3. Define the event of interest (Y)

   `e.g.: fatal or non fatal ischemic stroke`
4. Identify the possible competing events (D)
5. Identify the possible confounders/pronostic variable (X)
6. Define a statistical model for relating Y, T, and X

Workflow (Christiansen et al. 2015)

1. Define the population of interest
2. Define the intervention (T)
3. Define the event of interest (Y)
4. Identify the possible competing events (D)

   `e.g.: death not related to a stroke event`
5. Identify the possible confounders/pronostic variable (X)
6. Define a statistical model for relating Y, T, and X

Workflow (Christiansen et al. 2015)

1. Define the population of interest
2. Define the intervention (T)
3. Define the event of interest (Y)
4. Identify the possible competing events (D)
5. Identify the possible confounders/pronostic variable (X)
   `e.g. age, hypertension, ...`
6. Define a statistical model for relating Y, T, and X

Workflow (Christiansen et al. 2015)

1. Define the population of interest
2. Define the intervention (T)
3. Define the event of interest (Y)
4. Identify the possible competing events (D)
5. Identify the possible confounders/pronostic variable (X)
6. Define a statistical model for relating Y, T, and X

A two-cause specific Cox model:

$$\begin{aligned}\lambda^Y(t|X,T) &= \lambda_0^Y e^{\beta^Y X+\gamma^Y T}\\ \lambda^D(t|X,T) &= \lambda_0^D e^{\beta^D X+\gamma^D T}\end{aligned}$$

Computation of the G-formula - in R

Package: riskRegression https://github.com/tagteam/riskRegression
Function: ate

Arguments

- `object`: outcome model which describes how event risk depends on treatment and covariates
- `data`
- `treatment`: name of the treatment variable in `data`
- `times`: time points at which to evaluate risks
- `cause`: the cause of interest
- B: the number of bootstrap replications used to compute the confidence interval.

G-formula (software)

No competing risks:

```
> head(dtSurv)
         time strokeEvent Treatment      Age
1:   4.901849       FALSE        T0 59.78796
2:   4.555159        TRUE        T0 60.66406
3:   6.681136       FALSE        T1 58.76296
> mCox <- coxph(Surv(time,strokeEvent)~ Treatment + Age,
+               data = dtSurv)

> ate(mCox, data = dtSurv, treatment = "Treatment",
+     times = 12, B = 1000)
```

G-formula (software)

Competing risks:

```
> head(dtCR)
   time eventtype eventtypeNum Treatment      Age
1:  2.9    stroke            1        T0 58.96060
2:  9.3 censoring            0        T0 59.37469
3:  2.0     death            2        T0 59.36296
```

```
> mCSC <- CSC(
+           list(Hist(time,eventtypeNum)~ Treatment + Age,
+           Hist(time,eventtypeNum)~ Age),
+             data = dtCR
+ )

> ate(mCSC,data = dtCR, treatment = "Treatment",
+     times = 12, cause = 1, B = 1000)
```

G-formula (software output)

Absolute risk of stroke relapse

```
   Treatment meanRisk  meanRiskBoot lower upper n.boot
1:        T0    0.111         0.111 0.101 0.123 1000
2:        T1    0.080         0.080 0.071 0.090 1000
3:        T2    0.078         0.078 0.073 0.082 1000
```

Difference in absolute risk of stroke between treatments:

```
  Treatment.A  Treatment.B  time diff
1:          T1           T0    12 0.032
2:          T2           T0    12 0.034
3:          T2           T1    12 0.002
   diffMeanBoot diff.lower  diff.upper n.boot
1:        0.032      0.017       0.046   1000
2:        0.033      0.022       0.046   1000
3:        0.002      0.002       0.013   1000
```

Assumptions

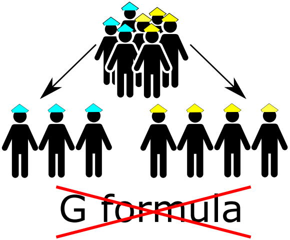
- no unmeasured confounders
- positivity
- well-defined intervention
- **correctly specified model**
  - ▷ proportional hazard assumption
  - ▷ (linear) functional form
  - ▷ (absence of) interaction

[Not covered] non-informative censoring, influential observations

Assumptions

- no unmeasured confounders
- positivity
- well-defined intervention
- **correctly specified model**
  - ▷ proportional hazard assumption
  - ▷ (linear) functional form
  - ▷ (absence of) interaction

[Not covered] non-informative censoring, influential observations



Assumptions

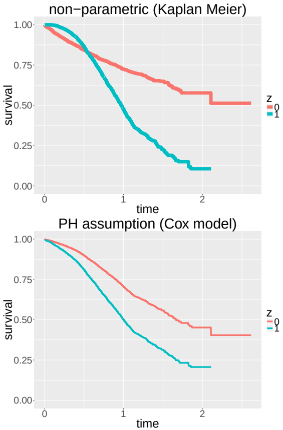
- no unmeasured confounders
- positivity
- well-defined intervention
- **correctly specified model**
    - proportional hazard assumption
    - (linear) functional form
    - (absence of) interaction

[Not covered] non-informative censoring, influential observations

Assumptions

- no unmeasured confounders
- positivity
- well-defined intervention
- **correctly specified model**
  - ▷ proportional hazard assumption
  - ▷ (linear) functional form
  - ▷ (absence of) interaction

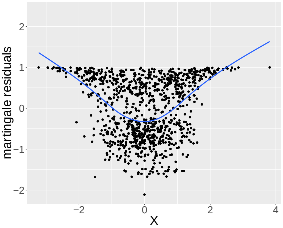
[Not covered] non-informative censoring, influential observations



$$\widehat{ATE} : -0.513[-0.571; -0.441] \text{ vs } -0.244[-0.281; -0.206]$$

Assumptions

- no unmeasured confounders
- positivity
- well-defined intervention
- **correctly specified model**
  - ▷ proportional hazard assumption
  - ▷ (linear) functional form
  - ▷ (absence of) interaction

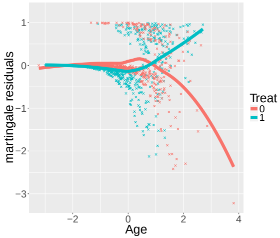


[Not covered] non-informative censoring, influential observations

$\widehat{ATE}$: $-0.164[-0.221; -0.107]$ vs $-0.200[-0.264; -0.132]$

Assumptions

- no unmeasured confounders
- positivity
- well-defined intervention
- **correctly specified model**
  - ▷ proportional hazard assumption
  - ▷ (linear) functional form
  - ▷ (absence of) interaction



[Not covered] non-informative censoring, influential observations

$\widehat{ATE}$: $-0.110[-0.169; -0.049]$ vs $-0.267[-0.302; -0.228]$

Assumptions

- no unmeasured confounders
- positivity
- well-defined intervention
- **correctly specified model**
  - proportional hazard assumption
  - (linear) functional form
  - (absence of) interaction

[Not covered] non-informative censoring, influential observations

Proposed approach

- Compare alternative modelling strategies to the Cox model
  → check result sensitivity to model assumptions

Alternative models:

✓ increased flexibility:
→ less biased

✗ increased complexity:
→ harder to interpret
→ increased variance of the estimates

Extensions/alternatives to Cox model

- relax PH assumption:

-Cox strata	stratified Cox model
-Others	Cox model with time varying effects
	logistic risk regression

- include non-linear relationships/interactions

-Cox spline	regression spline in Cox model
-RF	random survival Forest

Simulation study

Investigate the bias/variance trade-off

Three scenari:

(I) violation of proportional hazard assumption

(II) mispecification of the functional form of a risk factor

(III) missing interaction with the treatment variable

Results - proportional hazard

Scenario I	Non proportional effect of treatment		
$OR_T(t \leq 5)$	0.8	1.7	7.4
$OR_T(t \geq 5)$	0.4	0.4	0.4
ATE	0.181	-0.002	-0.068
Cox	-0.049 (0.05)	-0.082 (0.083)	-0.062 (0.062)
Cox strata	0.001 (0.011)	0 (0.01)	-0.01 (0.013)
Random Forest	0.001 (0.01)	0 (0.009)	0.005 (0.007)

Table: last three rows:
bias (root mean square error) of the ATE estimated by the models
OR_T odd ratio for the treatment effect

Results - functional form

Scenario II	Non linear effect of covariate		
OR_T	0.4	0.4	0.4
OR_{Age}	0.8	2.7	7.4
ATE	0.338	0.338	0.338
Cox	0.001 (0.01)	0 (0.01)	0 (0.01)
Cox strata	0.001 (0.011)	-0.001 (0.011)	-0.001 (0.011)
Random Forest	-0.003 (0.011)	-0.004 (0.012)	-0.004 (0.012)

Table: last three rows:
bias (root mean square error) of the ATE estimated by the models
OR_T odd ratio for the treatment effect
OR_{Age} odd ratio for the non linear effect of the risk factor
(increased risk after 50 years)

Results - interaction

Scenario III	Interaction between treatment and covariate		
OR_T	0.4	0.4	0.4
$OR_{T*gender}$	1.6	2.7	7.4
ATE	0.075	-0.011	-0.097
Cox	-0.01 (0.014)	-0.04 (0.041)	-0.146 (0.146)
Cox strata	0.001 (0.011)	0 (0.011)	0.001 (0.01)
Random Forest	0.001 (0.011)	0 (0.011)	0 (0.01)

Table: last three rows:
bias (root mean square error) of the ATE estimated by the models
OR_T odd ratio for the treatment effect
$OR_{T*gender}$ odd ratio for interaction between gender and treatment

Discussion

Alternative models:

- ✓ No noticeable increase of the variance of the estimates
- ✓ For categorical variables, a fully stratified Cox model is robust against:
  - non PH
  - interaction between variables
- ✓ For dealing with continuous variables, use random Forests
  - ✗ extra-parameters to be tuned (e.g. number of trees)

✗ Increased computation time

Application

Objective:

- to compare 3 antiplatelet regimens using the danish registry
- n = 19223 patients
- time horizon: 1 year

Outcome:

- date of first stroke event (n=1610, 8.4%)

Competing event:

- death (n=677, 3.5%)

Possible confounders:

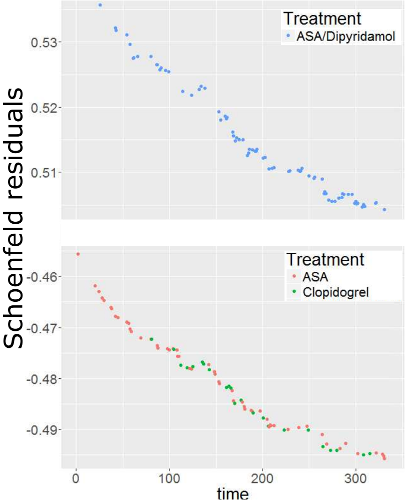
- many (p=10) including age, gender, . . .

Checking Proportional hazard assumption

Tests:[1]

- not significant
  except for one treatment modality

We can stratify on treatment!

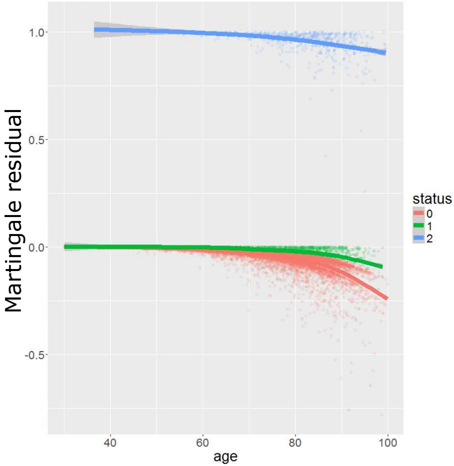


[1] based on the score process

Checking functional form

Variable age

- additional risk after 75 years
- approx. linear



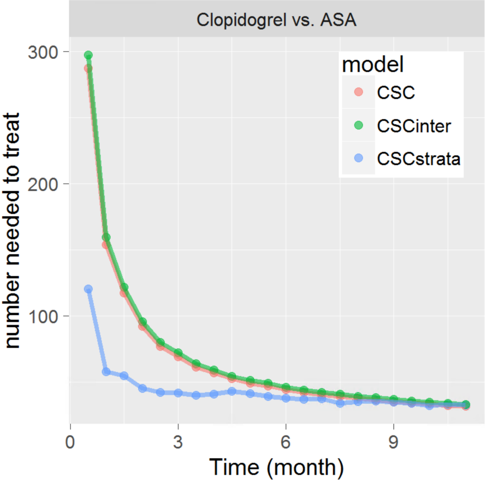
Models

Variables:

- Continuous: age
- Categorical: treatment, gender, year

CSC	Two cause specific Cox model
CSC inter	CSC + interactions between treatment and gender, age, year + cubic spline on age
CSC strata	stratified CSC on treatment, gender, year

naive Cox model vs. alternatives



- violation of PH assumption impacts the estimate of NTT at early times

Summary

Cox models can be used to assess :

- disease incidence (absolute risk)
- average treatment effect

Relies on several assumptions, e.g.:

- proportional hazard
- linear effect, no interaction between variables

Model checking:

- **Hope**: no unmeasured confounders
- usual diagnostic tools are of limited interest for large p or n
- *proposal*: assess the impact of Cox model assumptions using alternative models

Bibliography I

Christiansen, C. B. et al. (2015). ‘Comparison of antiplatelet regimens in secondary stroke prevention: a nationwide cohort study’. In: *BMC Neurology* 15.1, p. 225. ISSN: 1471-2377. URL: http://www.biomedcentral.com/1471-2377/15/225.

Grambsch, P. M. et al. (1994). ‘Proportional hazards tests and diagnostics based on weighted residuals’. In: *Biometrika* 81.3, pp. 515–526. ISSN: 00063444.

Grøn, R. et al. (2016). ‘Misspecified poisson regression models for large-scale registry data: inference for 'large n and small p'.’ In: *Statistics in medicine* 35.7, pp. 1117–29. ISSN: 1097-0258. URL: http://www.ncbi.nlm.nih.gov/pubmed/26423319.

Staerk, L. et al. (2016). ‘Ischaemic and haemorrhagic stroke associated with non-vitamin K antagonist oral anticoagulants and warfarin use in patients with atrial fibrillation: a nationwide cohort study’. In: *European Heart Journal*. URL: http://eurheartj.oxfordjournals.org/content/early/2016/10/12/eurheartj.ehw496.abstract.