Marginal hazard ratios and covariate adjustment: a causal inference perspective

Rhian Daniel, Cardiff University

PSI Conference 2025 · Wembley, London

Outline

Outline

Potential outcomes and estimands

- In causal inference, estimands are often expressed using potential outcomes: e.g. Y^1 is the outcome that would be seen under active treatment, Y^0 is the outcome that would be seen under placebo. [Hernán and Robins (2025) *What If?*]

- Estimands are contrasts of (some aspects of) the distributions of these, e.g.
  - (marginal) mean (risk) difference or ratio:

$$\mathbb{E}(Y^1 - Y^0) \quad \text{or} \quad \frac{\mathbb{E}(Y^1)}{\mathbb{E}(Y^0)}$$

  - (marginal) odds ratio:

$$\frac{\mathbb{P}(Y^1 = 1)/\mathbb{P}(Y^1 = 0)}{\mathbb{P}(Y^0 = 1)/\mathbb{P}(Y^0 = 0)}$$

  - (marginal) hazard ratio:

$$\frac{\lim_{\Delta t \to 0} \mathbb{P}(t < Y^1 \leq t + \Delta t | Y^1 > t)/\Delta t}{\lim_{\Delta t \to 0} \mathbb{P}(t < Y^0 \leq t + \Delta t | Y^0 > t)/\Delta t}$$

Conditional estimands

- Each of these could be made conditional on covariates, e.g. $X =$ smoking status
  - conditional (on smoking) mean (risk) difference or ratio:

$$\mathbb{E}(Y^1 - Y^0|X) \quad \text{or} \quad \frac{\mathbb{E}(Y^1|X)}{\mathbb{E}(Y^0|X)}$$

  - conditional (on smoking) odds ratio:

$$\frac{\mathbb{P}(Y^1 = 1|X)/\mathbb{P}(Y^1 = 0|X)}{\mathbb{P}(Y^0 = 1|X)/\mathbb{P}(Y^0 = 0|X)}$$

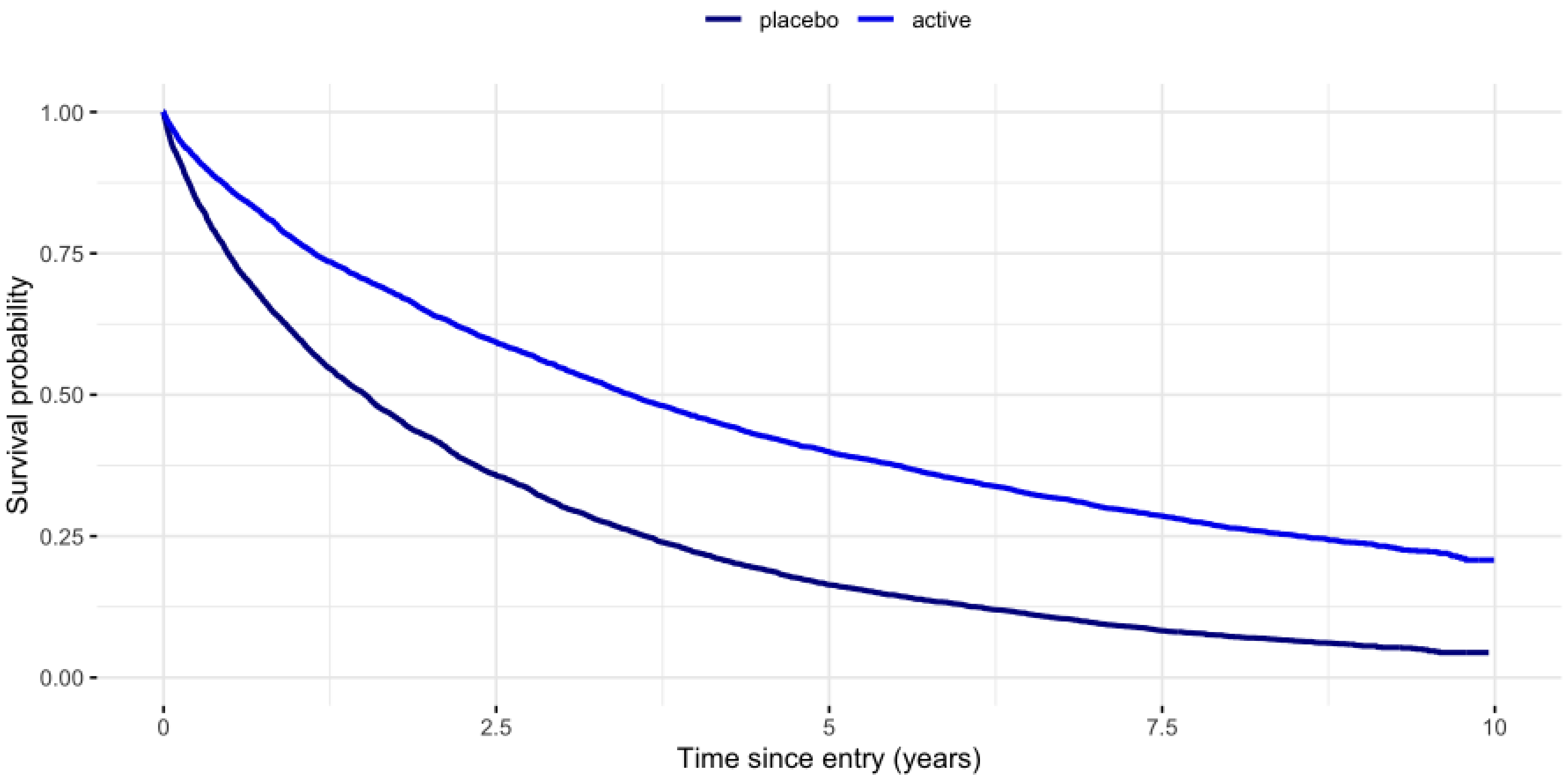
  - conditional (on smoking) hazard ratio:

$$\frac{\lim_{\Delta t \to 0} \mathbb{P}(t < Y^1 \leq t + \Delta t | Y^1 > t, X)/\Delta t}{\lim_{\Delta t \to 0} \mathbb{P}(t < Y^0 \leq t + \Delta t | Y^0 > t, X)/\Delta t}$$

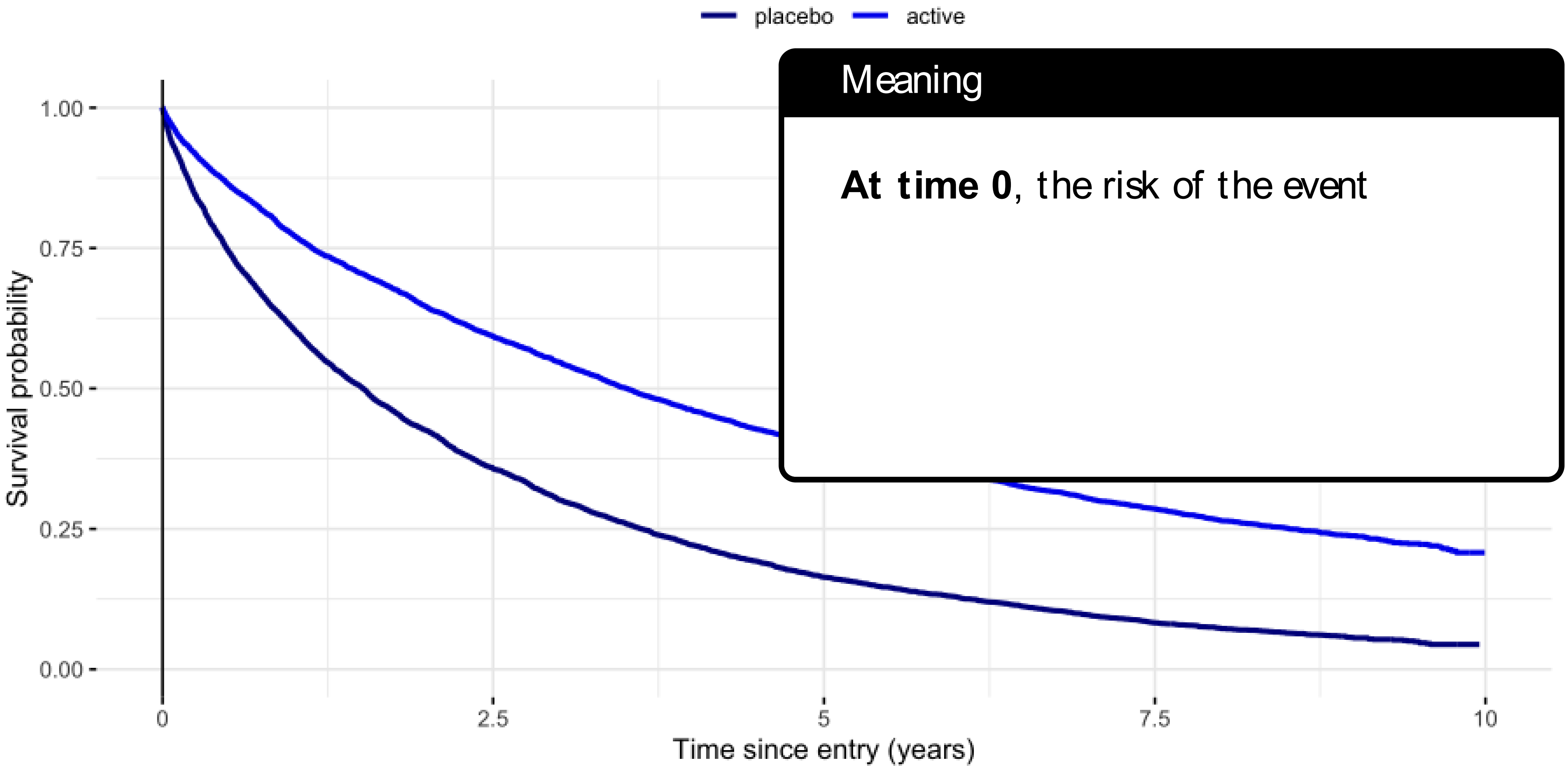
- More on conditional estimands later.

Outline

An example: constant HR of 0.5



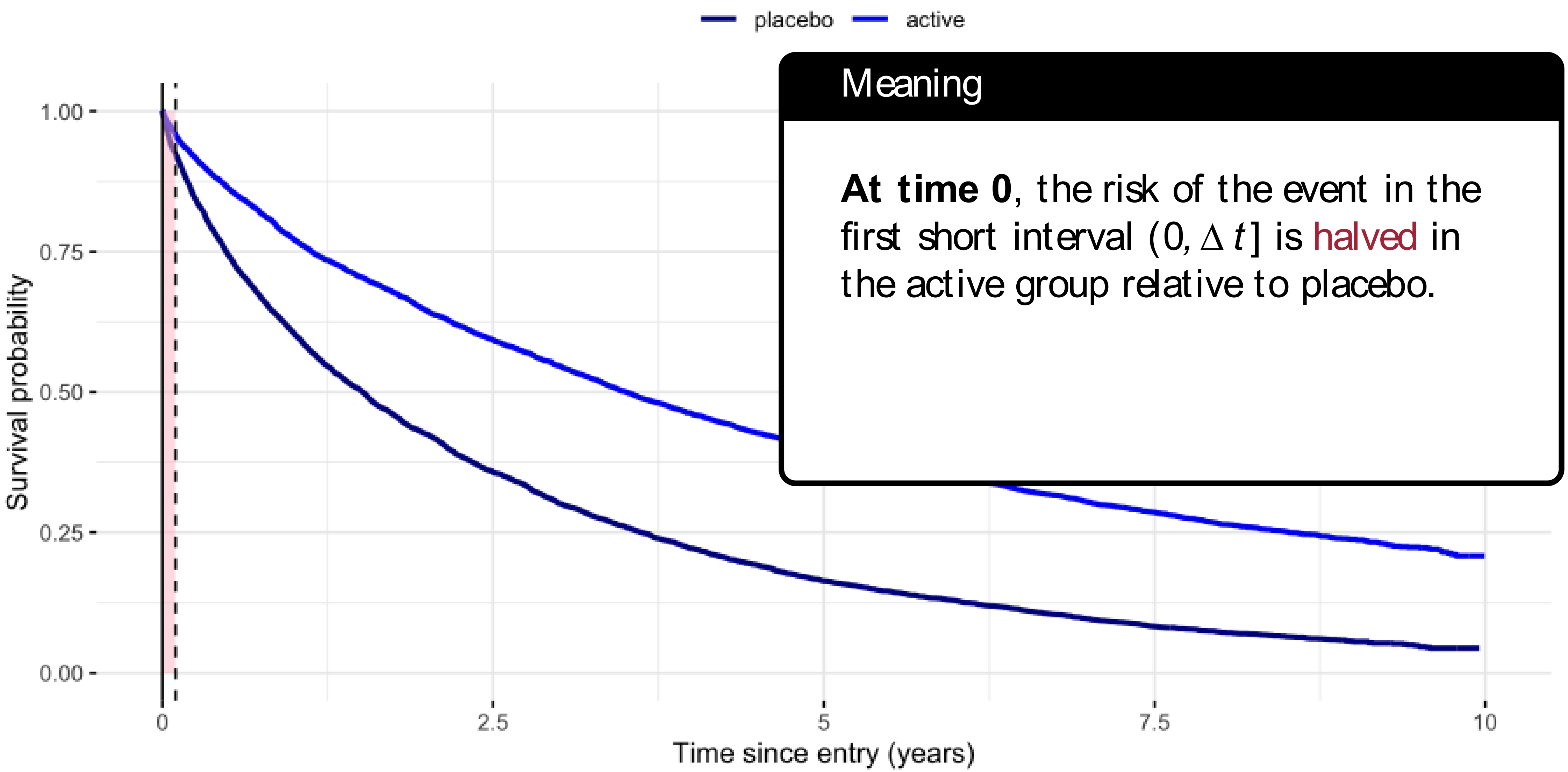
An example: constant HR of 0.5



Meaning

At time 0, the risk of the event

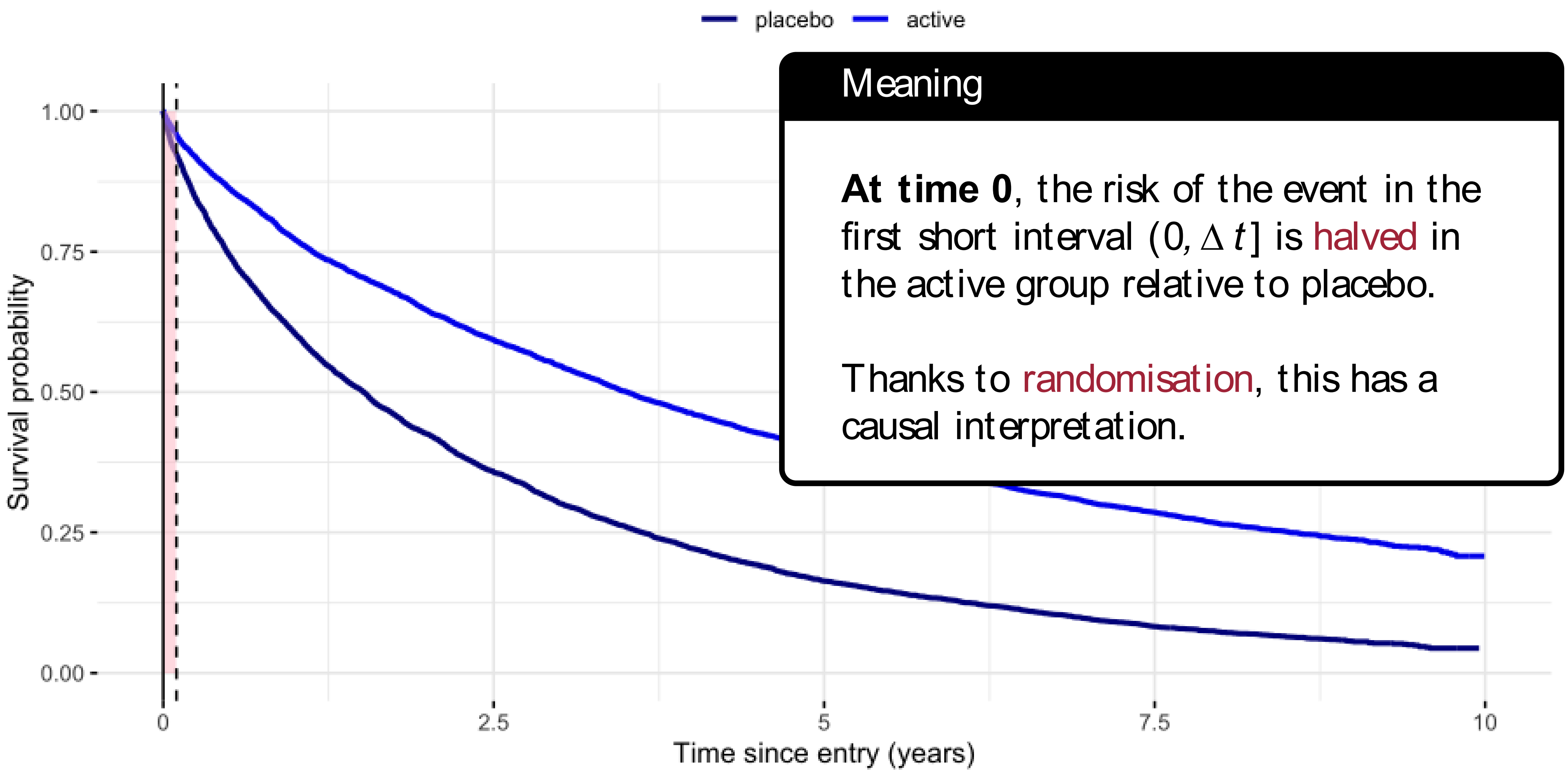
An example: constant HR of 0.5



Meaning

At time 0, the risk of the event in the first short interval $(0, \Delta t]$ is halved in the active group relative to placebo.

An example: constant HR of 0.5

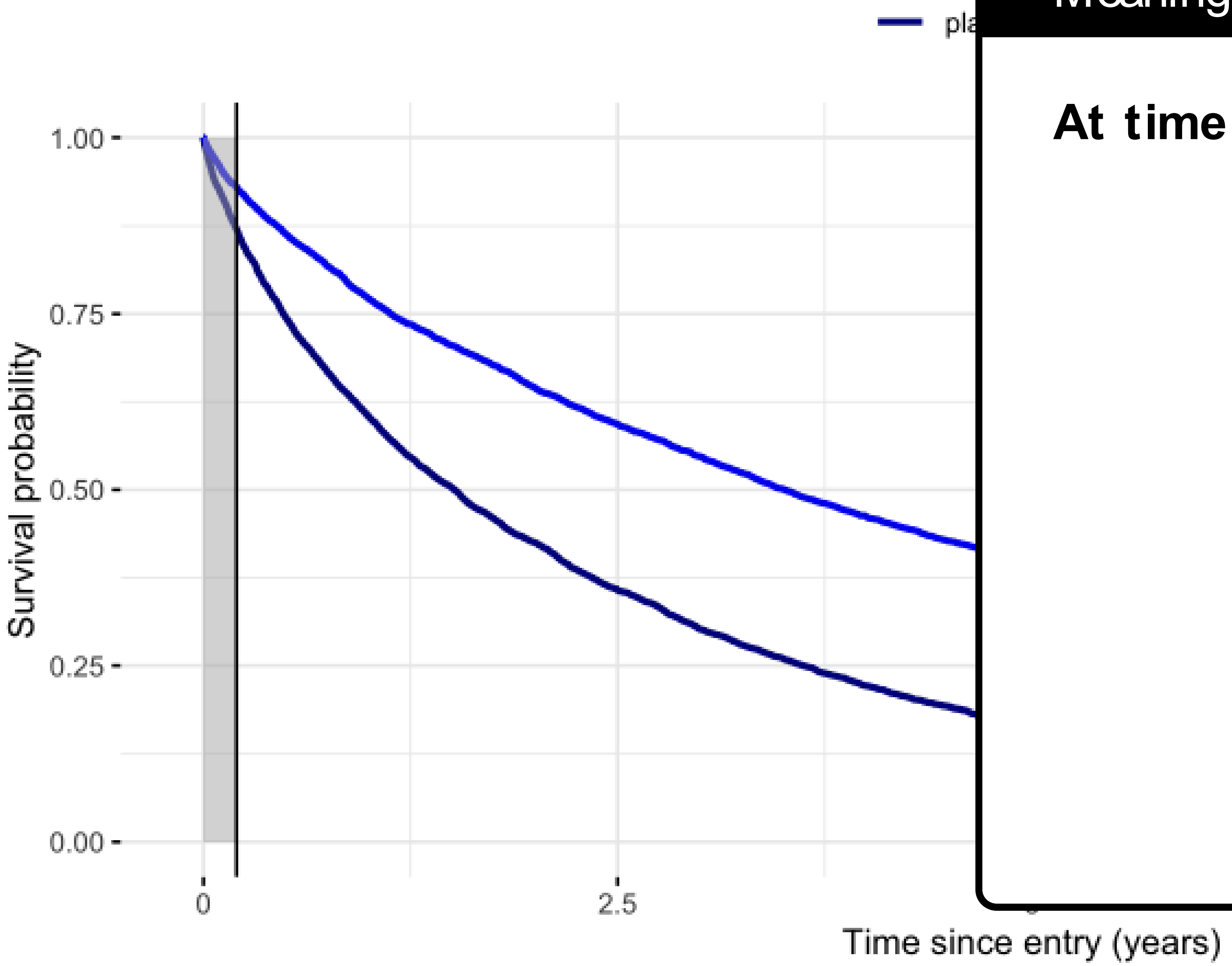


Meaning

At time 0, the risk of the event in the first short interval $(0, \Delta t]$ is halved in the active group relative to placebo.

Thanks to randomisation, this has a causal interpretation.

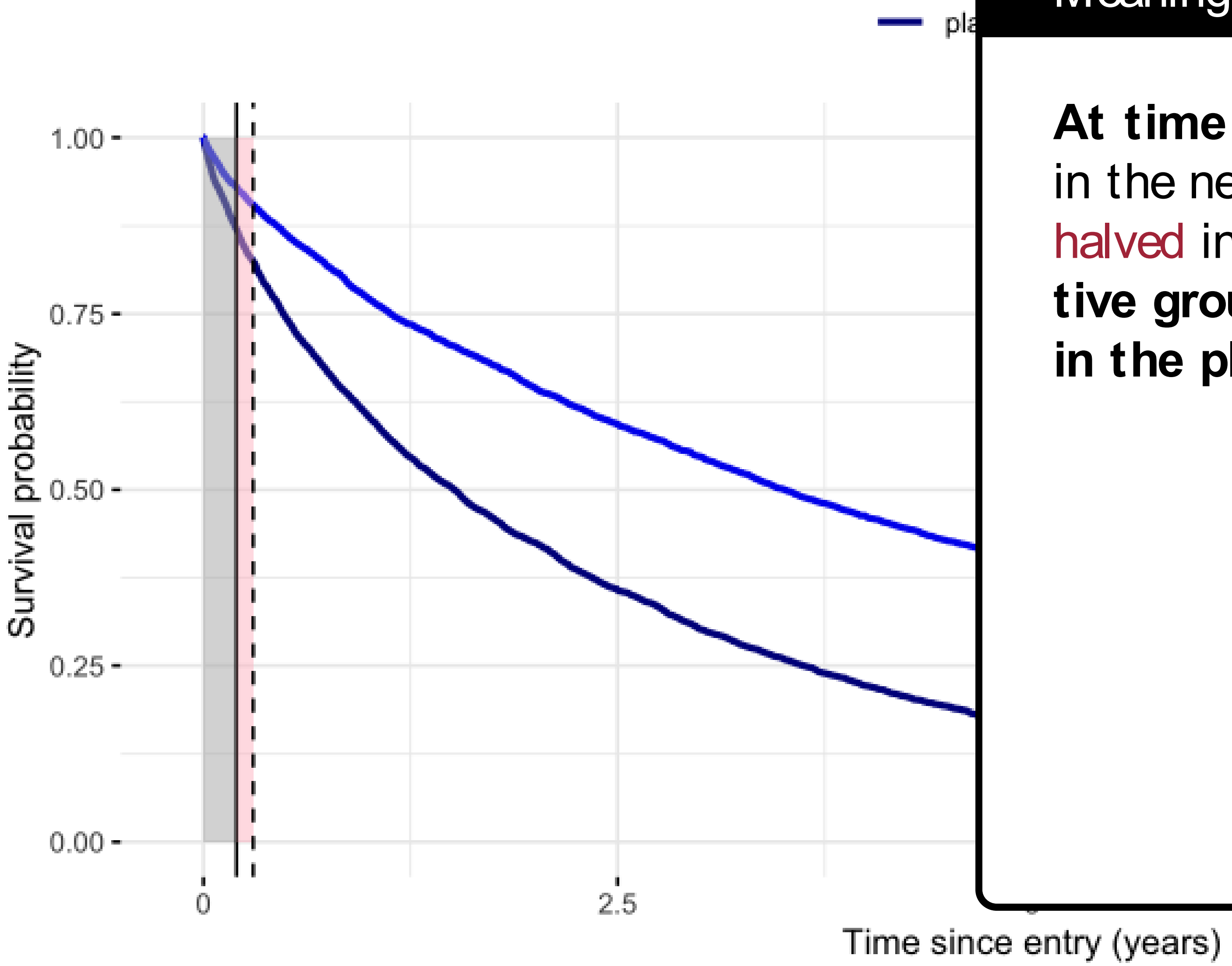
An example: constant HR of 0.5



Meaning?

At time $t > 0$, the risk of the event

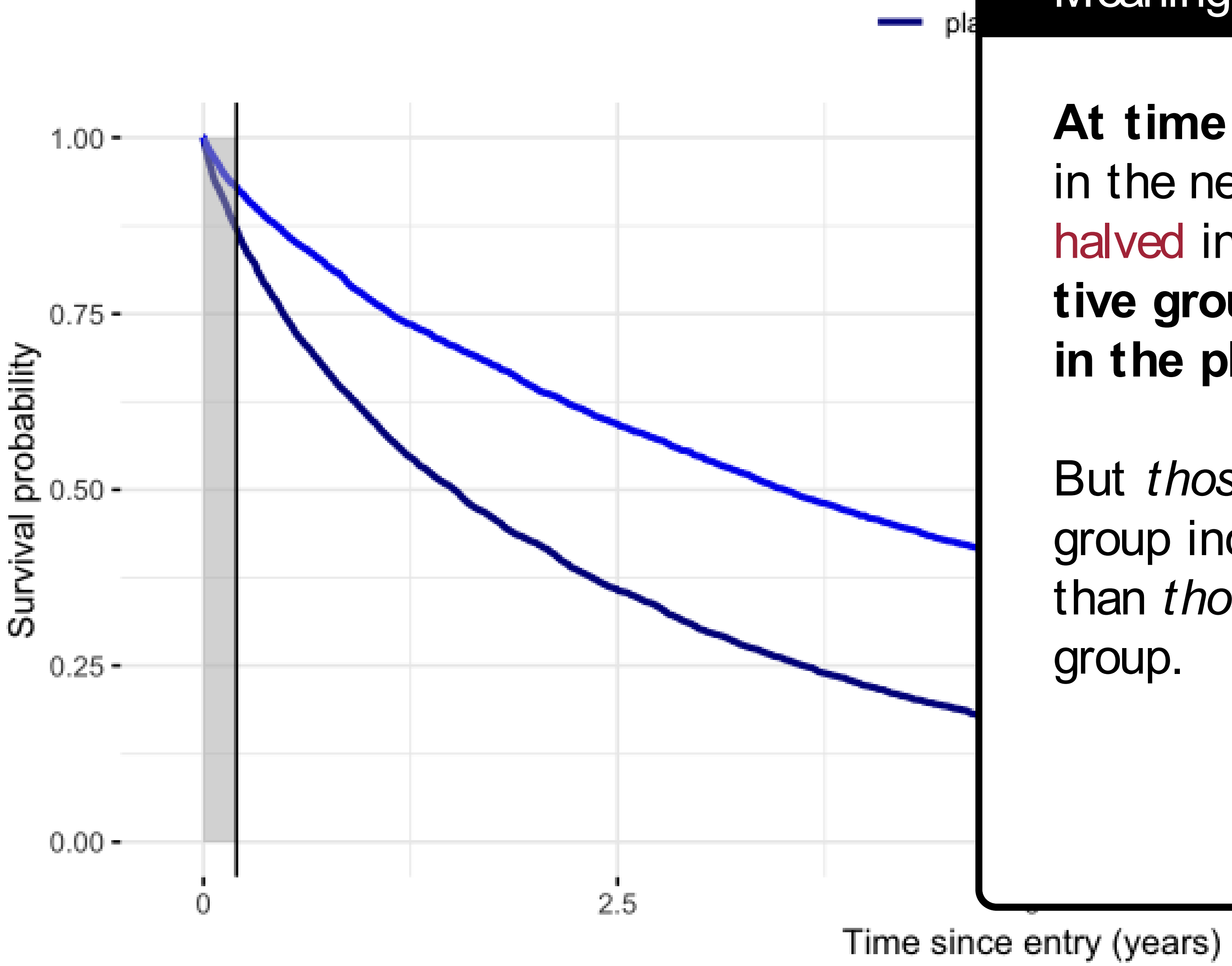
An example: constant HR of 0.5



Meaning?

At time $t > 0$, the risk of the event in the next short interval $(t, t + \Delta t]$ is halved in **those remaining in the active group** relative to **those remaining in the placebo group**.

An example: constant HR of 0.5

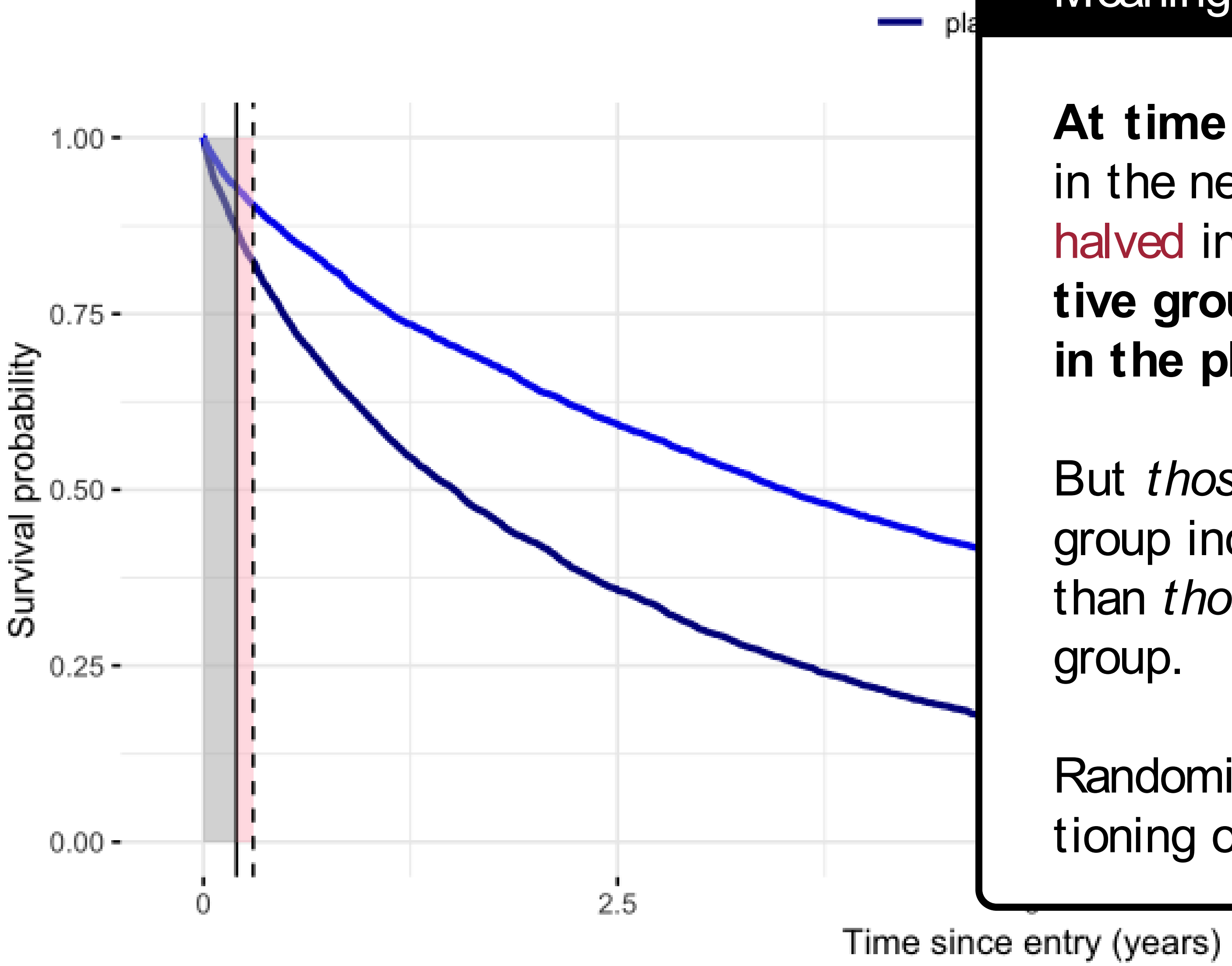


Meaning?

At time $t > 0$, the risk of the event in the next short interval $(t, t + \Delta t]$ is halved in **those remaining in the active group** relative to **those remaining in the placebo group**.

But *those remaining* in the active group include more frailer participants than *those remaining* in the placebo group.

An example: constant HR of 0.5



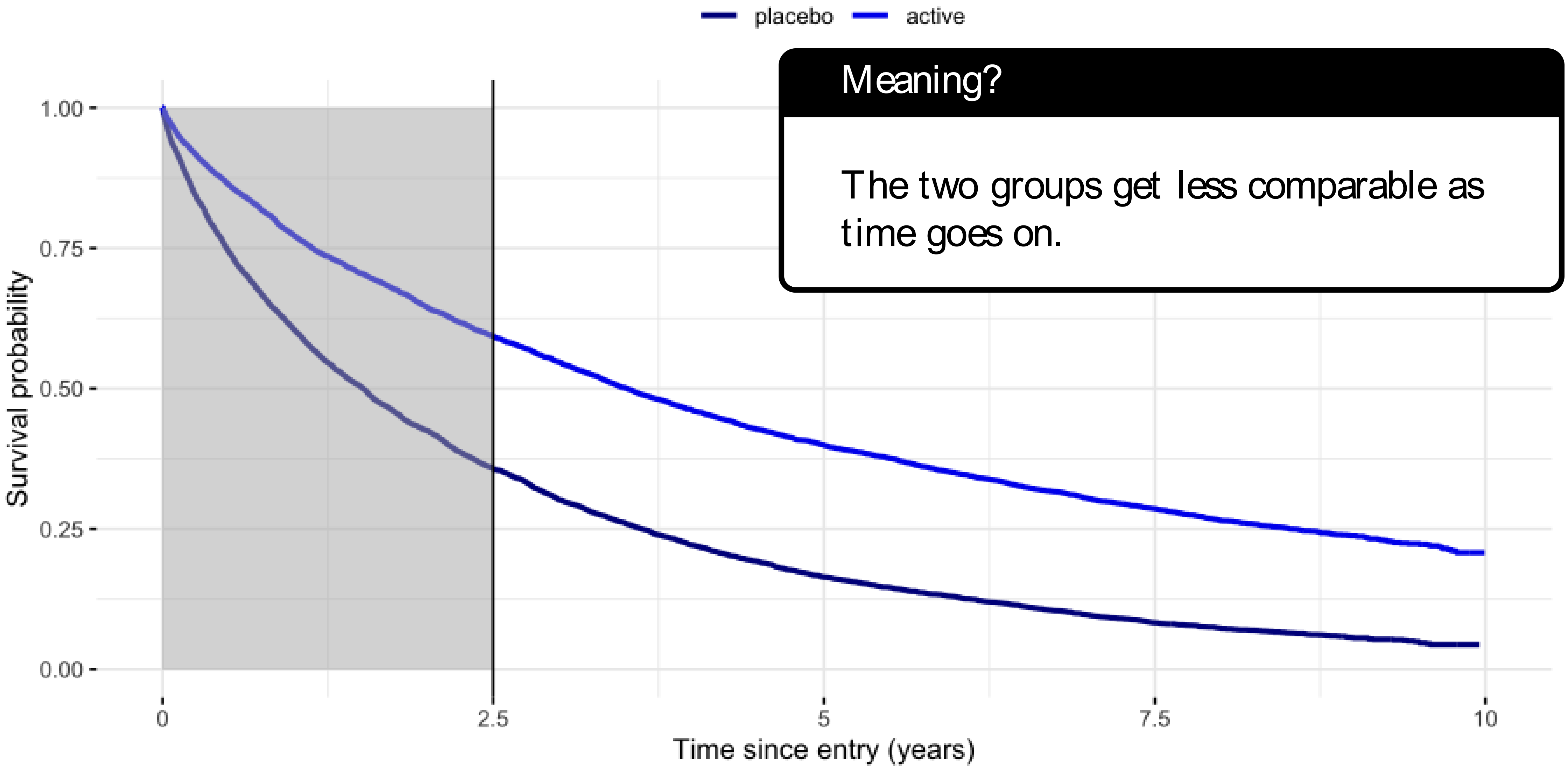
Meaning?

At time $t > 0$, the risk of the event in the next short interval $(t, t + \Delta t]$ is halved in **those remaining in the active group** relative to **those remaining in the placebo group**.

But *those remaining* in the active group include more frailer participants than *those remaining* in the placebo group.

Randomisation is '*broken*' by conditioning on survival to time t.

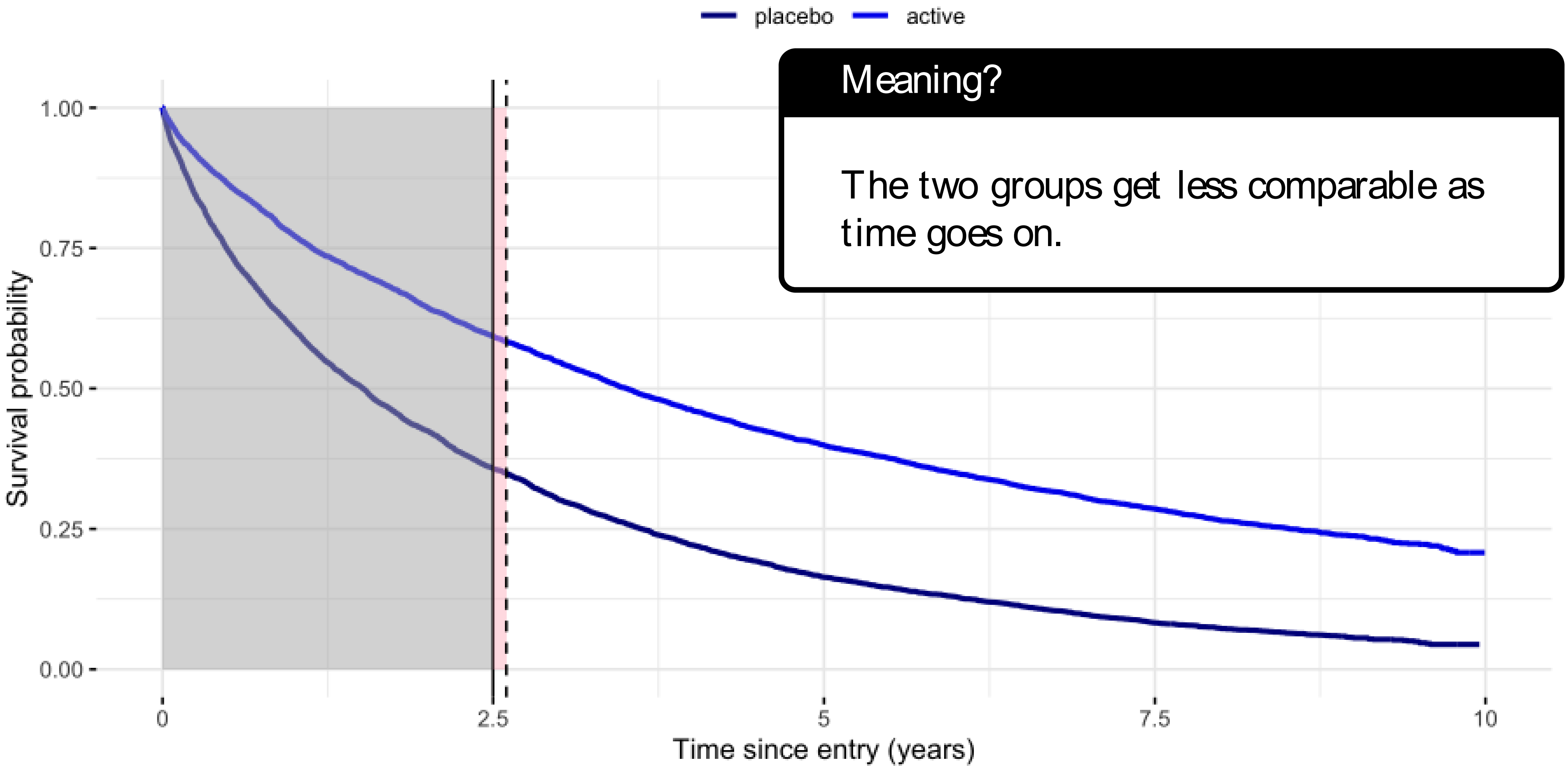
An example: constant HR of 0.5



Meaning?

The two groups get less comparable as time goes on.

An example: constant HR of 0.5



Meaning?

The two groups get less comparable as time goes on.

The hazard of hazard ratios

- Tempting to interpret a hazard ratio as a causal effect at each t: "given survival to time t, active treatment halves the instantaneous risk of the event at t relative to placebo".
  - But this is wrong: 'survival to time t' refers to different sub-groups in the two arms.
  - ...the so-called hazard of hazard ratios. [Hernán, 2010]
- All is not lost if the HR is truly constant, because then the HR can be shown to equal

$$\frac{\log\{\mathbb{P}(Y_1 > t)\}}{\log\{\mathbb{P}(Y_0 > t)\}}$$

which does not vary with t. [Martinussen *et al.*, 2018]

  - "Active treatment halves the log of the probability of surviving event-free to time t, relative to placebo (for all t)".
  - But this ratio of log survival probabilities requires a shift from the usual interpretation.
  - And the equivalence does not hold for non-constant HRs.
- The interpretation of time-varying HRs is thus **very hazardous**. [Bartlett *et al.*, 2020]
  - More on this later.

Outline

Imagined RCT of new wonderdrug

- For the remainder of the talk, I will focus on a hypothetical placebo-controlled RCT of a wonderdrug purported to reduce the incidence of cardiovascular disease, type II diabetes, and other metabolic conditions.
- The trial recruited $n = 10,000$ healthy adults and followed them up for up to 10 years.
- The outcome is "time to onset of any cardiovascular or metabolic disease".
- The trial collected one baseline covariate, namely **smoking status** (current/ex/never).

Outline

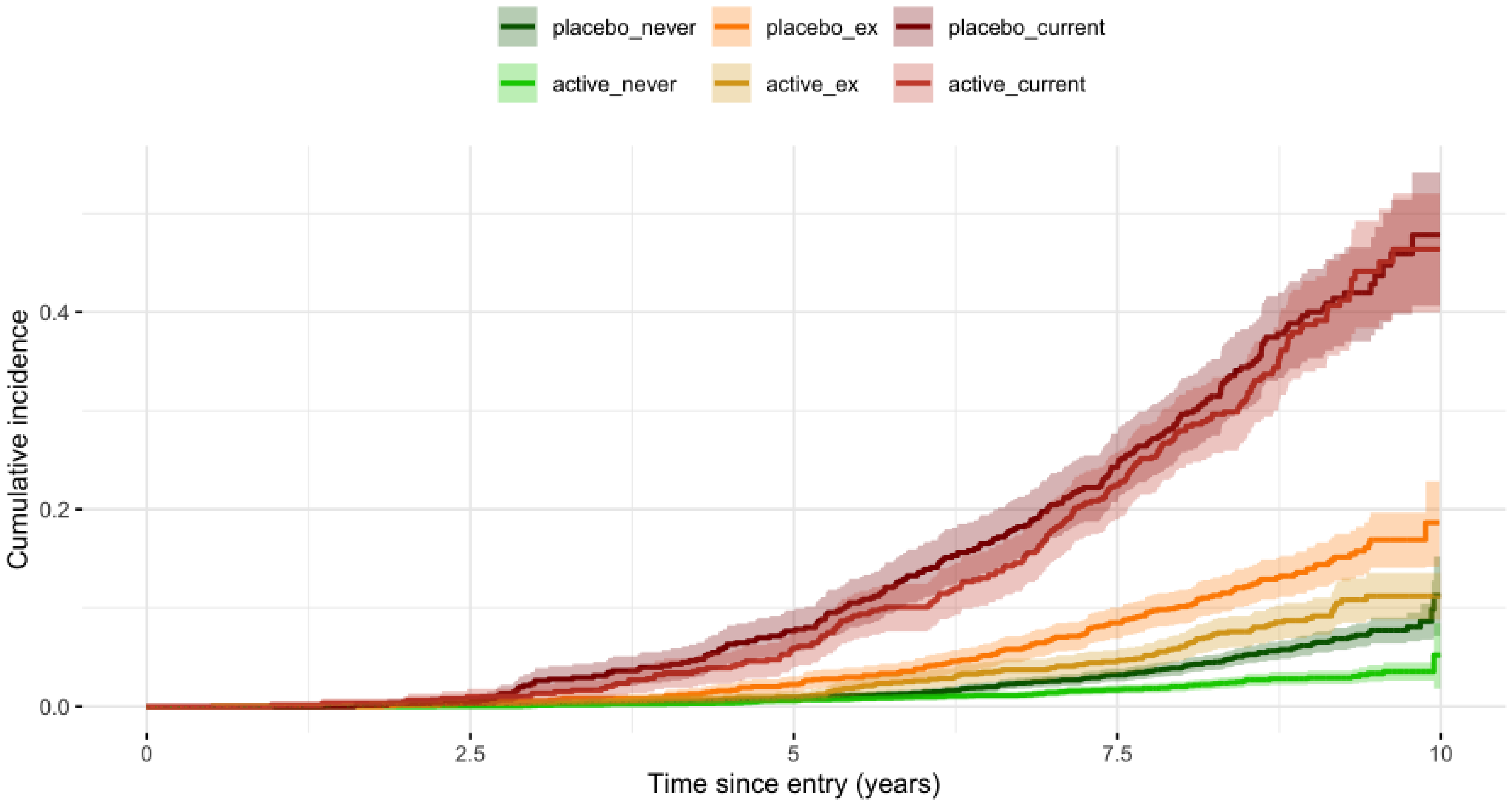
Conditional and marginal estimands

- A focus on conditional estimands **can** mean separately estimating a different treatment contrast for each level of the baseline covariate.
  - What is the hazard ratio in never-smokers? What is it in ex-smokers, and in current smokers?
- **Or** it can mean simply including smoking status as a covariate in the analysis (e.g. Cox PH model), but estimating a single conditional treatment constrast, assumed constant over smoking status.

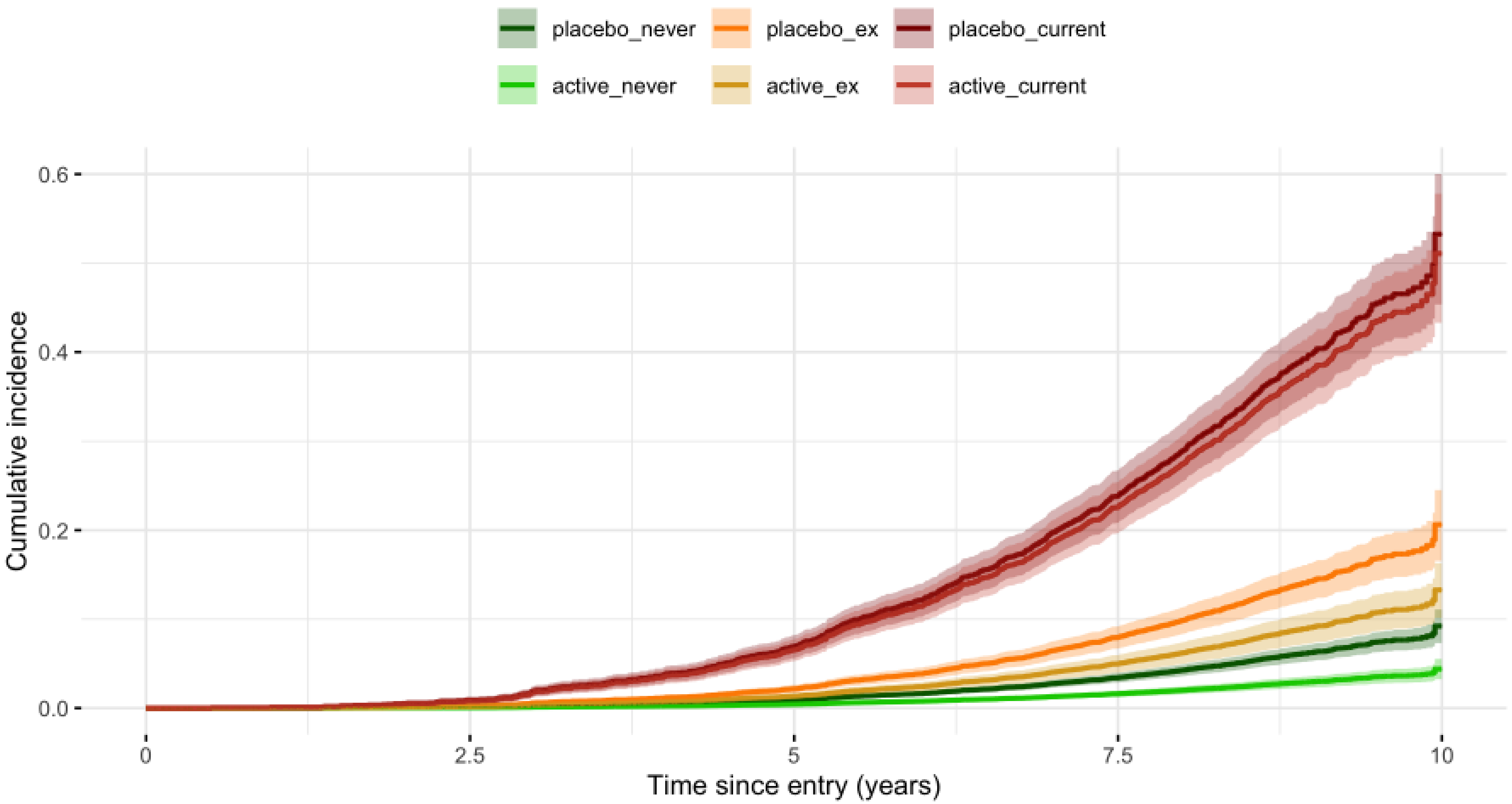
! **These two meanings are clearly distinct !**

- A conditional estimand (correctly) assumed homogeneous across levels of a covariate may or may not be equal to the corresponding marginal estimand.
  - (Non)collapsibility.
  - Hazard ratios are non-collapsible.
  - Sjölander *et al.*, 2015; Daniel *et al.*, 2021

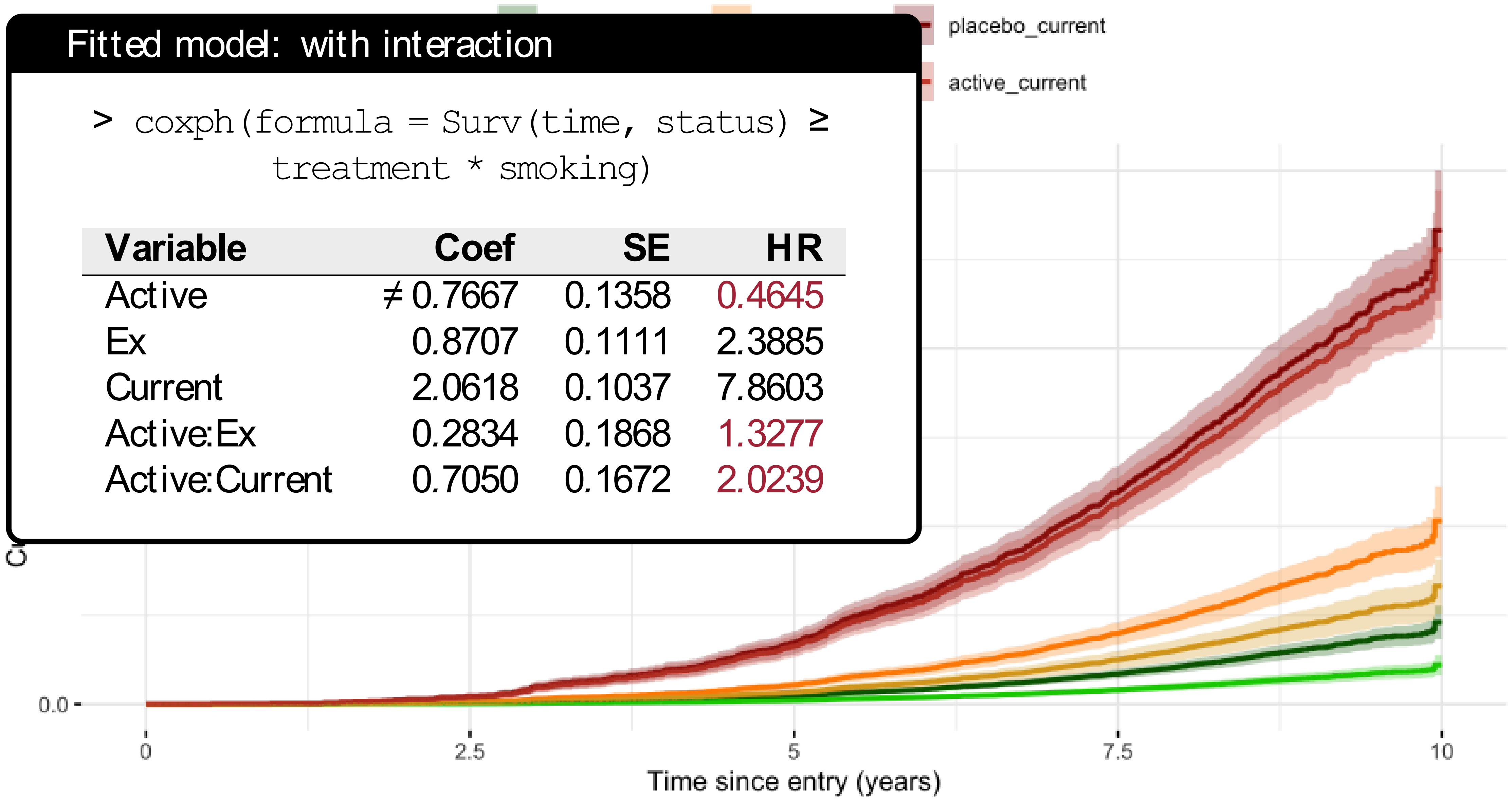
Example: heterogeneous conditional HRs: Kaplan–Meier



We can allow for this in the analysis: Cox model with interactions



We can allow for this in the analysis: Cox model with interactions

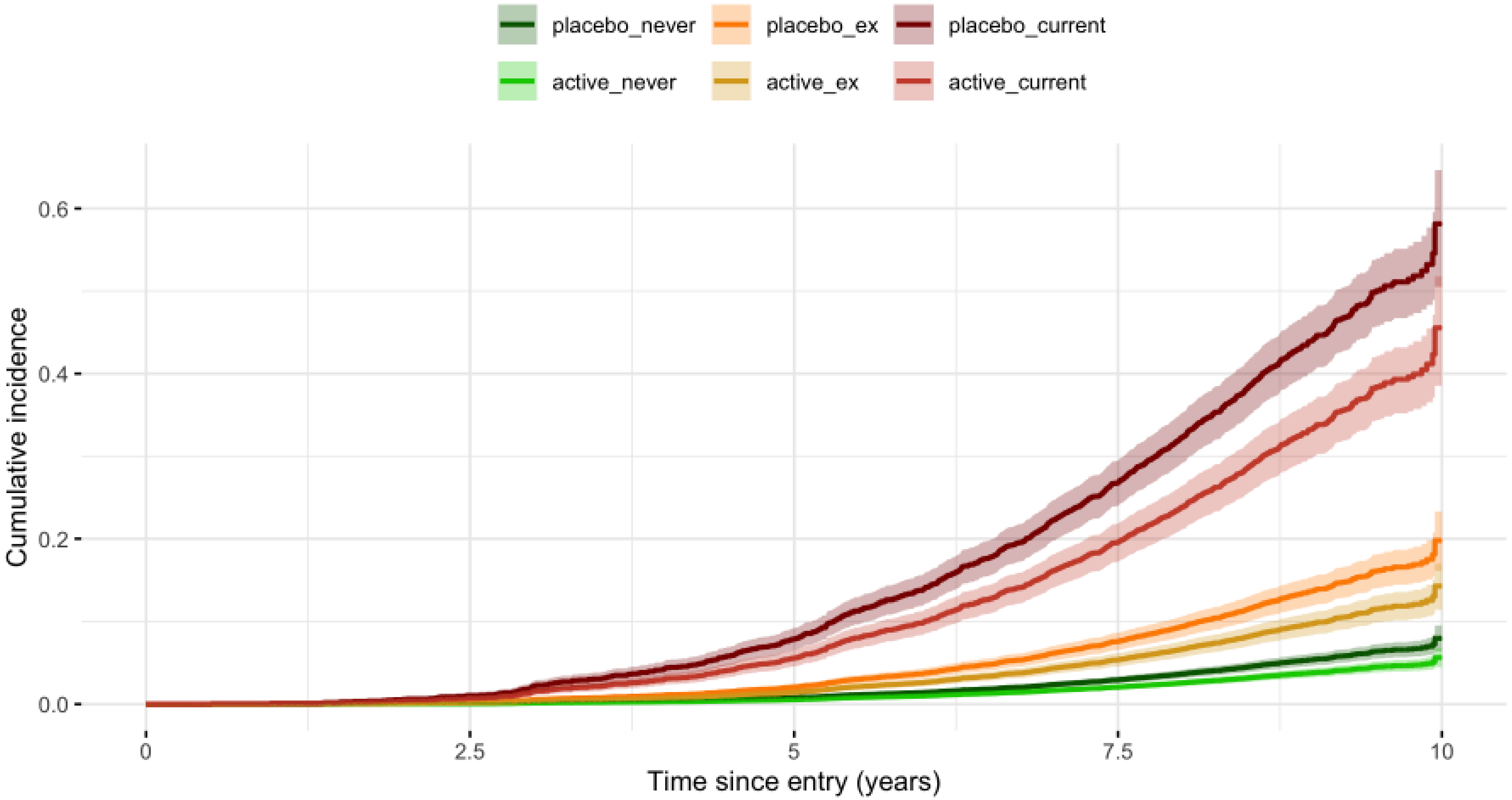


Fitted model: with interaction

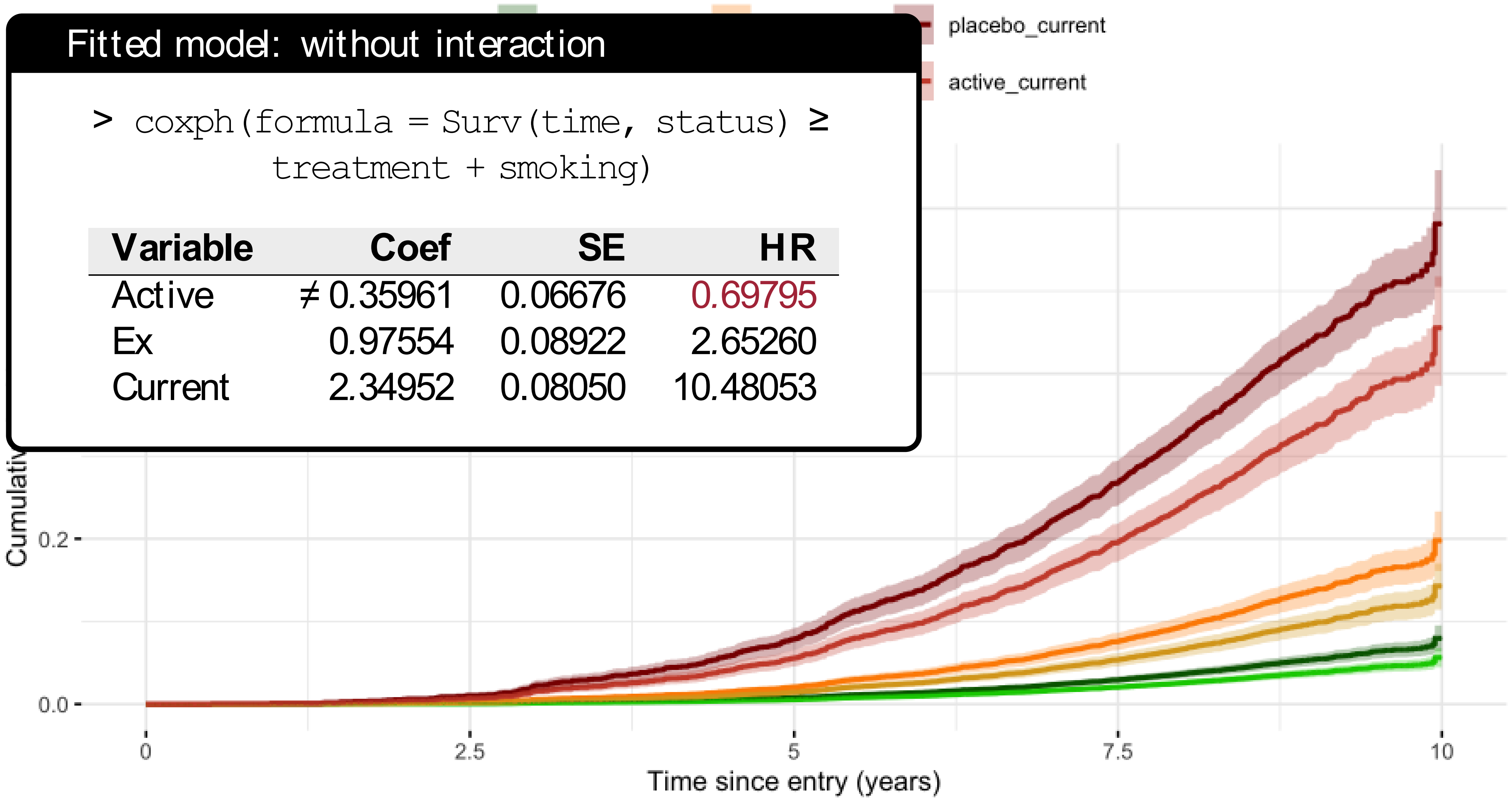
```
> coxph(formula = Surv(time, status) ≥
          treatment * smoking)
```

Variable	Coef	SE	HR
Active	≠ 0.7667	0.1358	0.4645
Ex	0.8707	0.1111	2.3885
Current	2.0618	0.1037	7.8603
Active:Ex	0.2834	0.1868	1.3277
Active:Current	0.7050	0.1672	2.0239

Or not: Cox model without interactions



Or not: Cox model without interactions

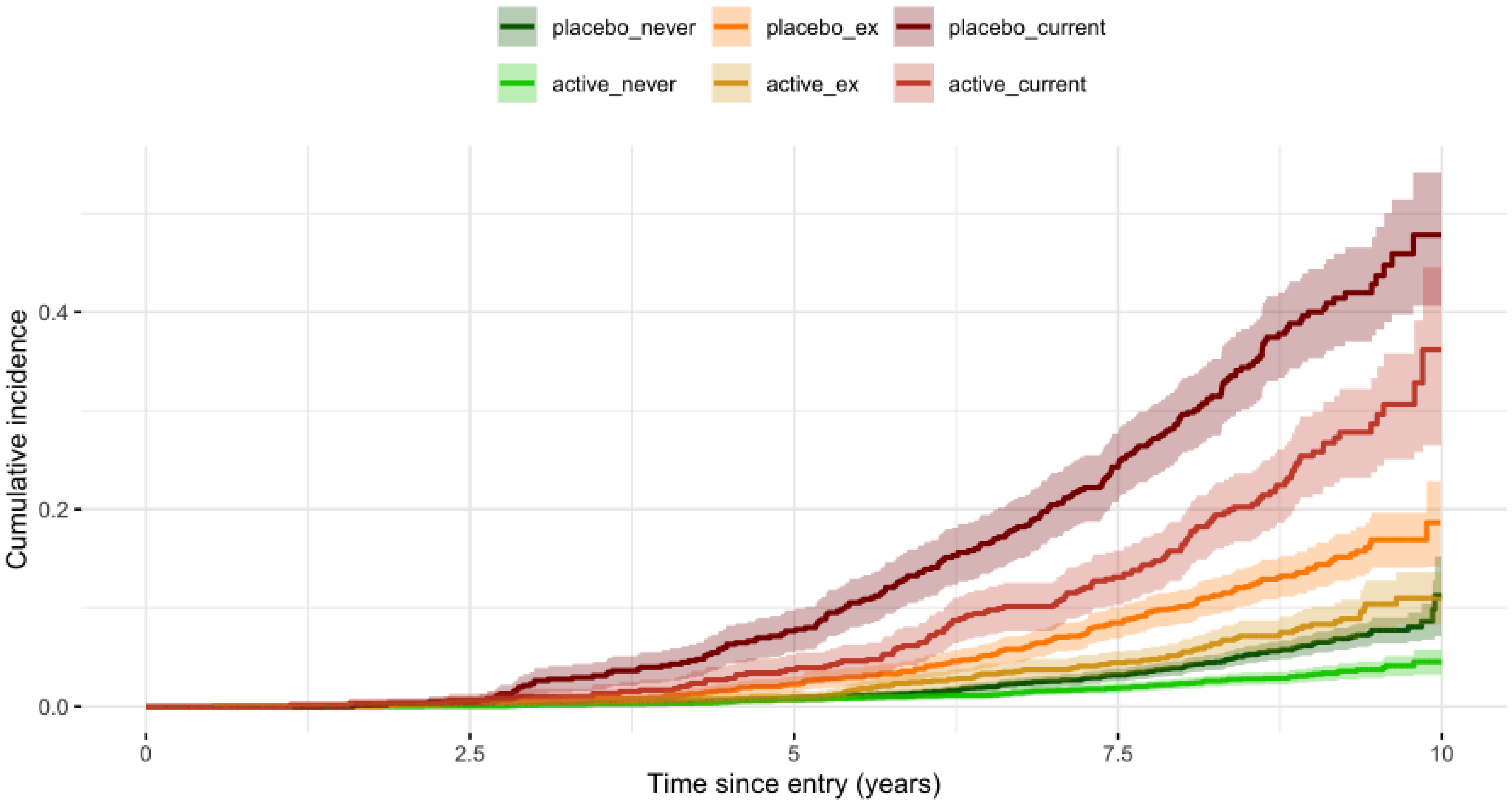


Fitted model: without interaction

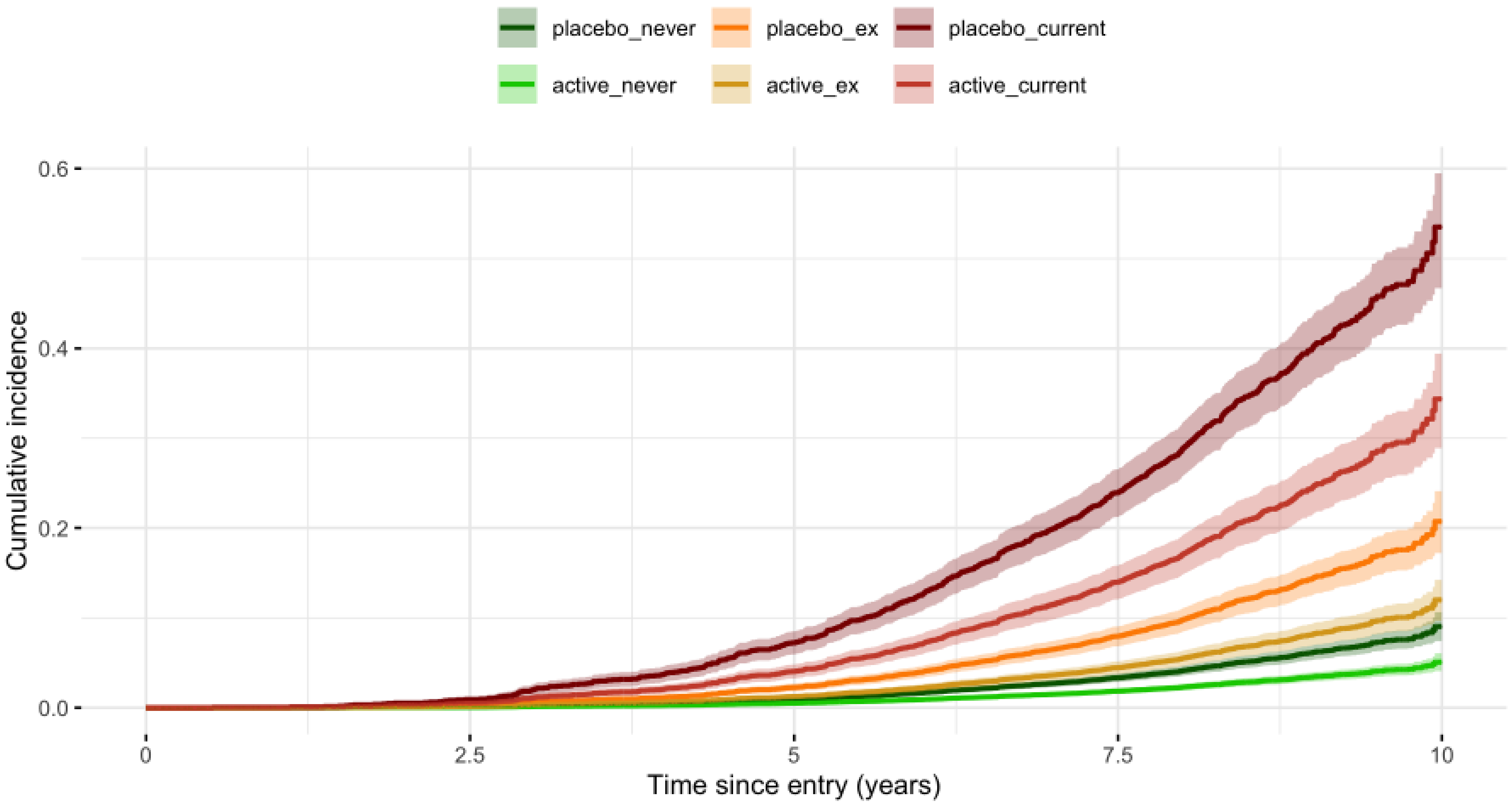
```
> coxph(formula = Surv(time, status) ≥
        treatment + smoking)
```

Variable	Coef	SE	HR
Active	≠0.35961	0.06676	0.69795
Ex	0.97554	0.08922	2.65260
Current	2.34952	0.08050	10.48053

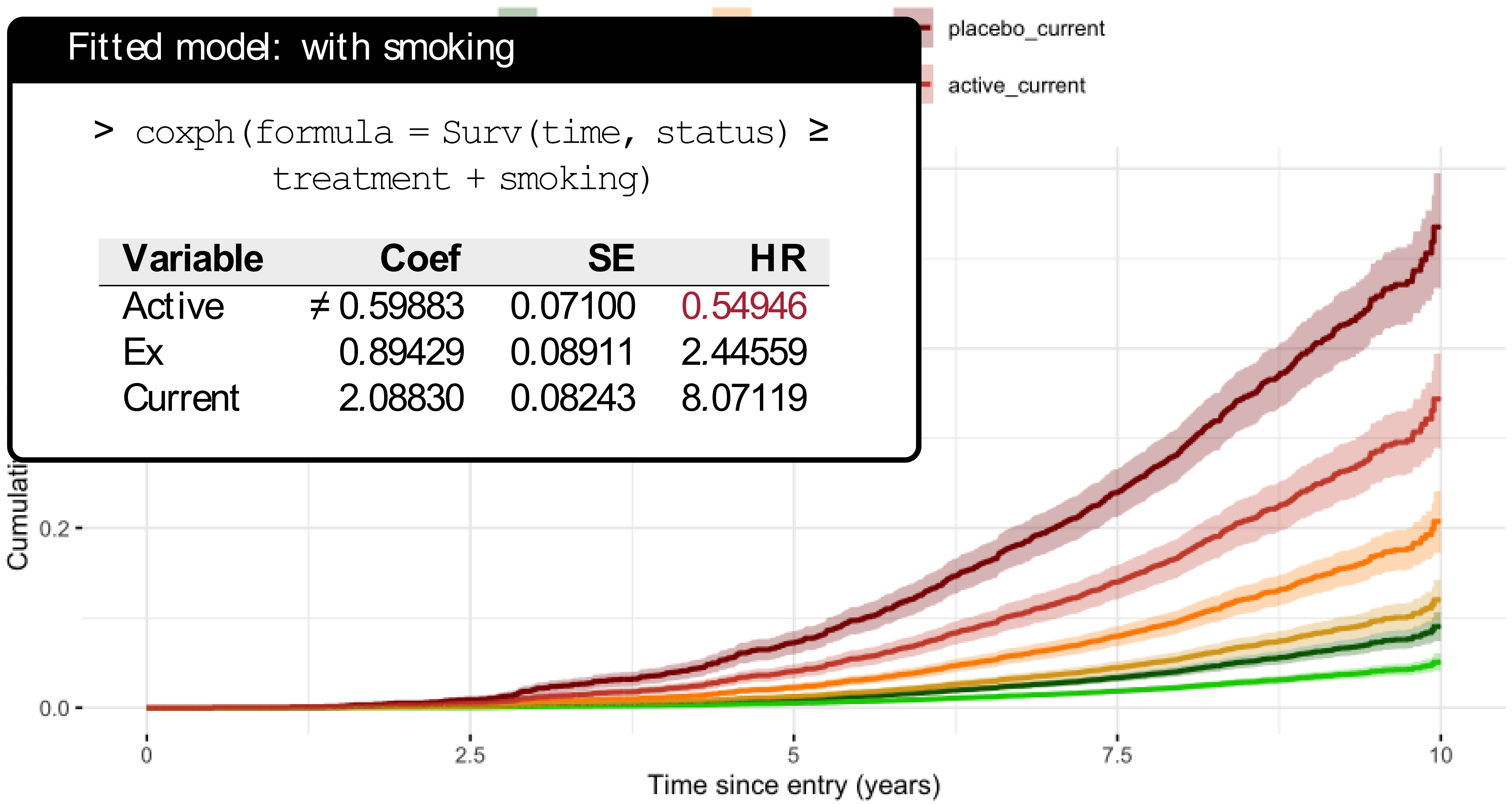
Now an example with homogeneous cond. HRs: Kaplan–Meier



A model including smoking (no interaction) gives one treatment coef



A model including smoking (no interaction) gives one treatment coef

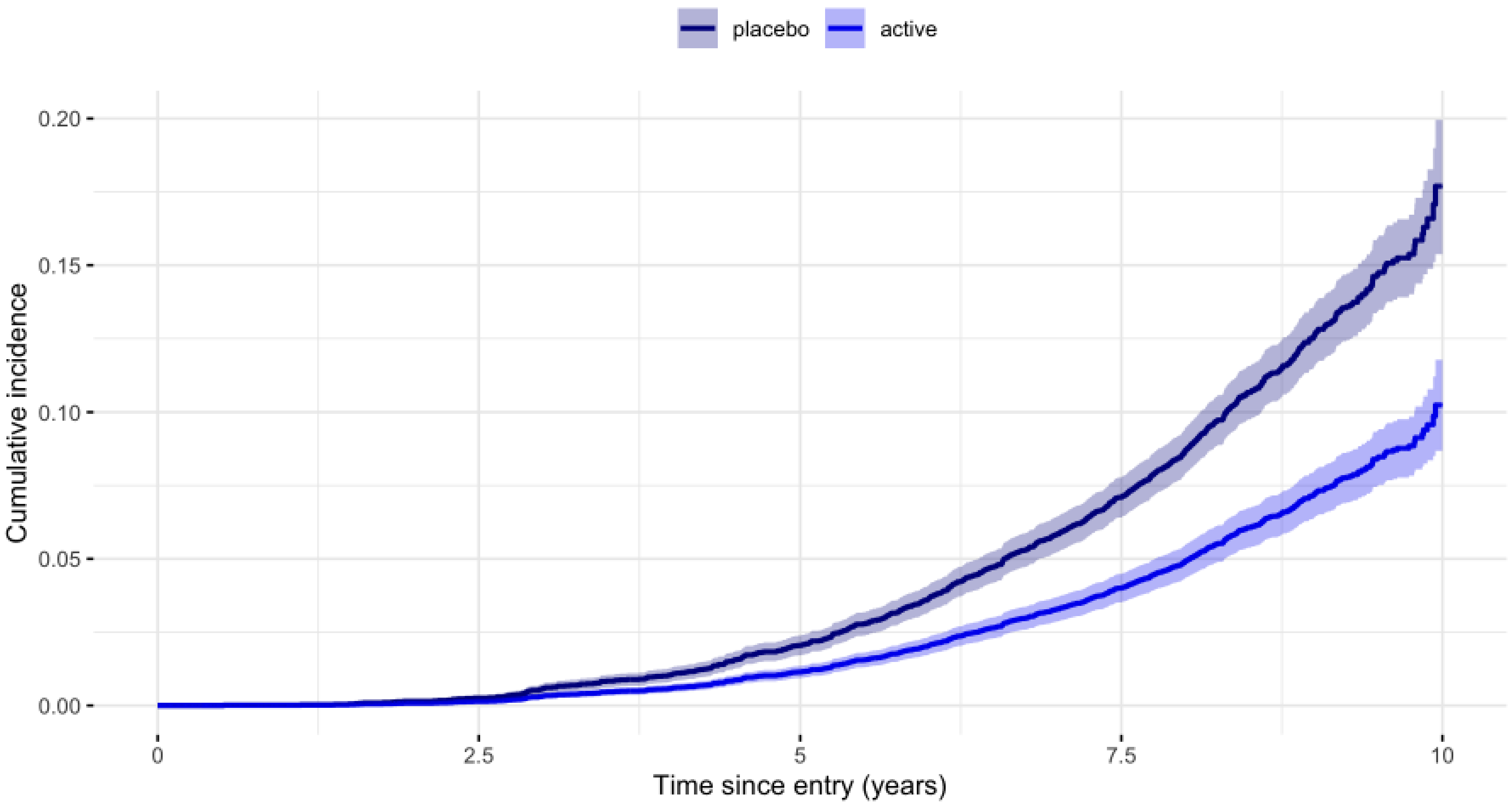


Fitted model: with smoking

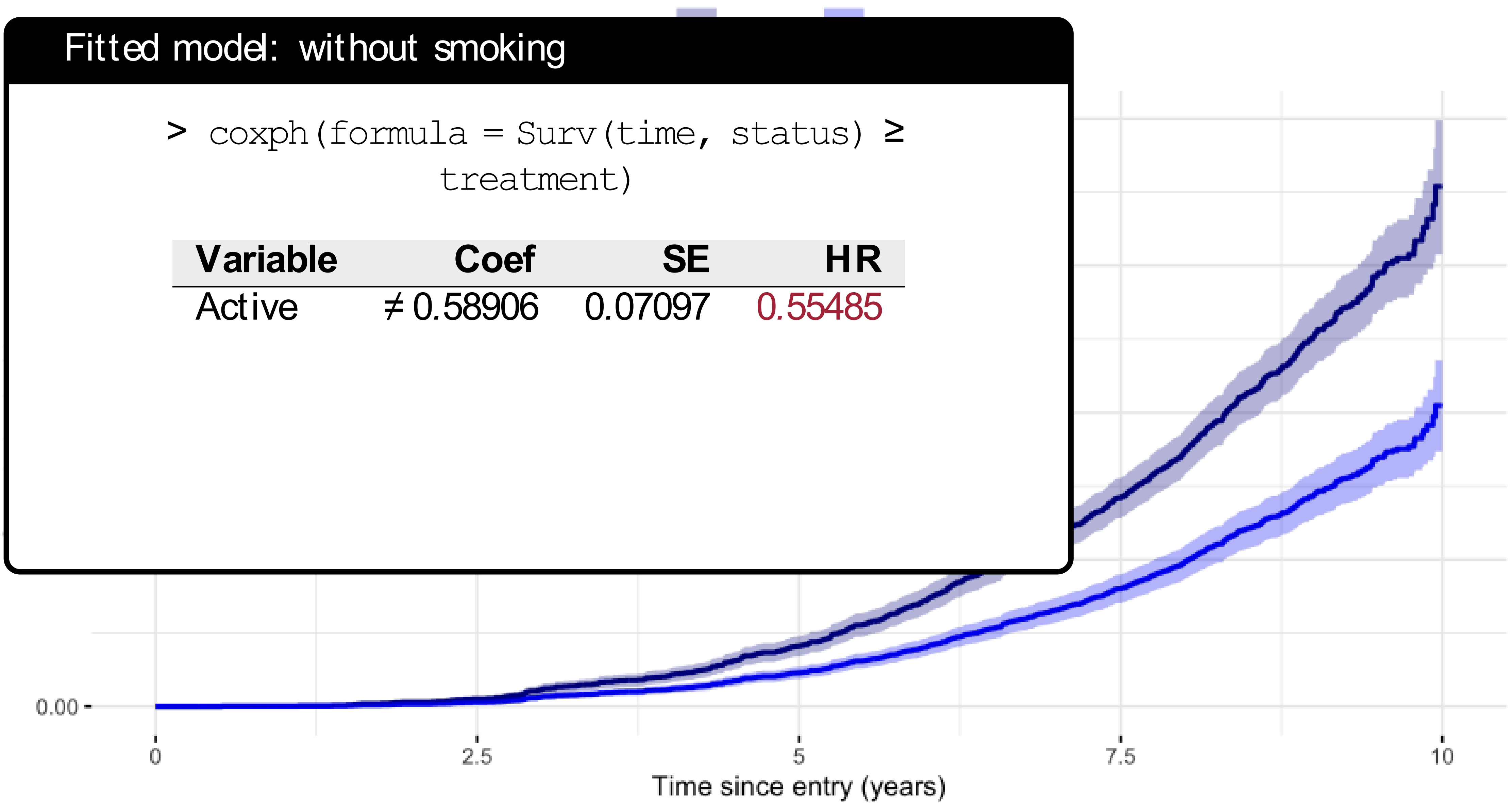
```
> coxph(formula = Surv(time, status) ≥
        treatment + smoking)
```

Variable	Coef	SE	HR
Active	≠0.59883	0.07100	0.54946
Ex	0.89429	0.08911	2.44559
Current	2.08830	0.08243	8.07119

But so does a model without smoking



But so does a model without smoking

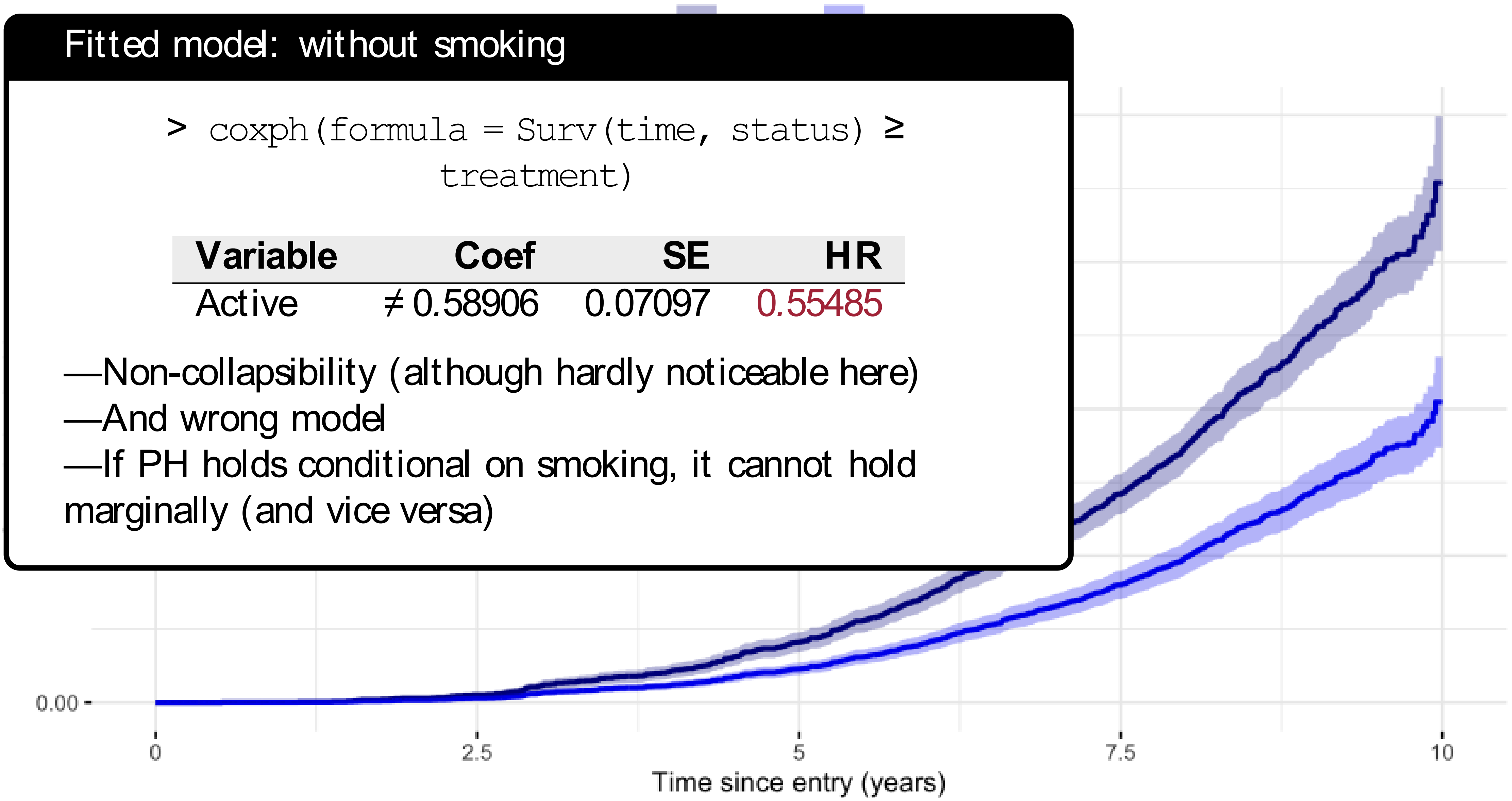


Fitted model: without smoking

```
> coxph(formula = Surv(time, status) ≥
treatment)
```

Variable	Coef	SE	HR
Active	≠0.58906	0.07097	0.55485

But so does a model without smoking



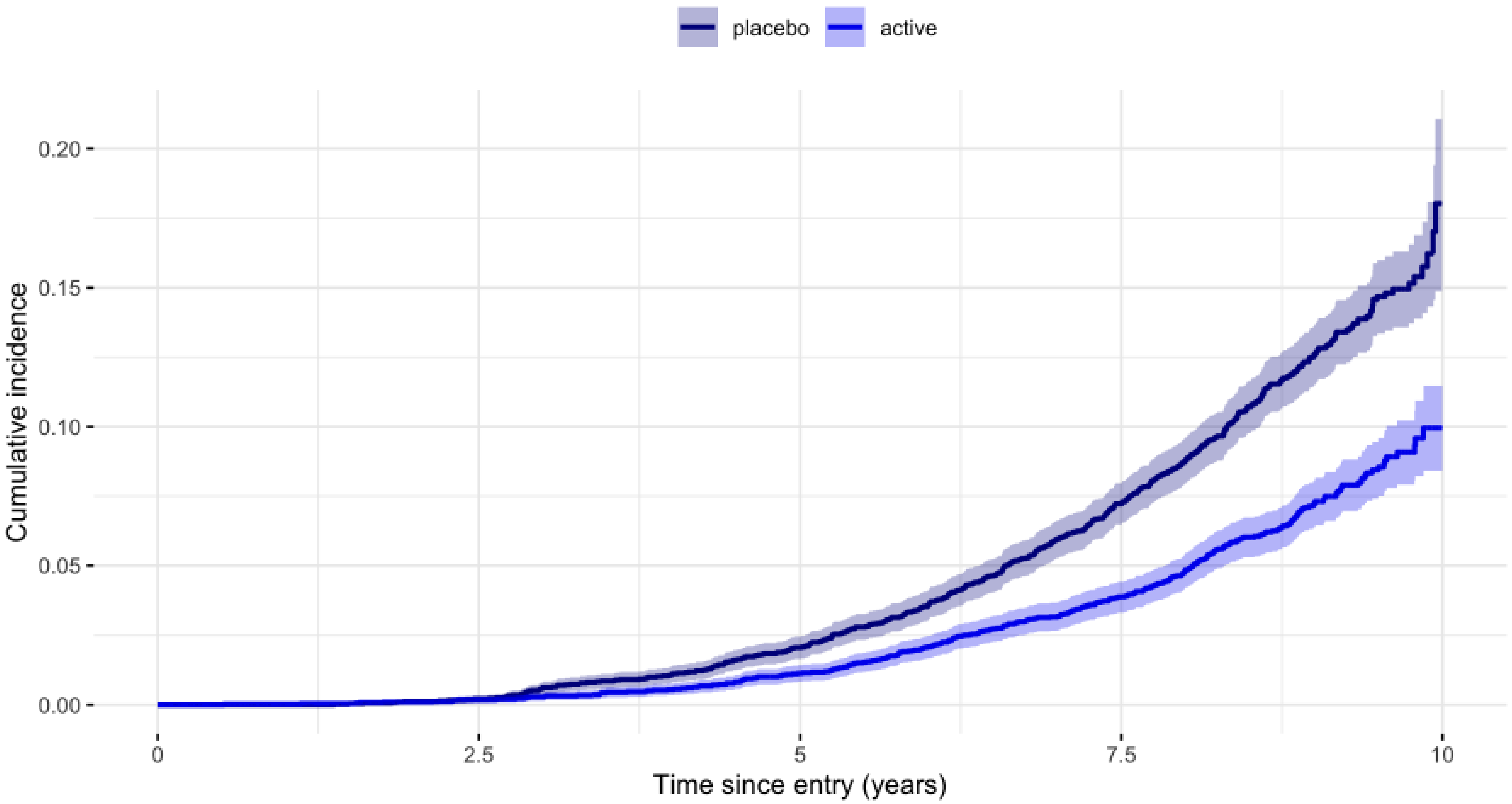
Fitted model: without smoking

```
> coxph(formula = Surv(time, status) ≥
                 treatment)
```

Variable	**Coef**	**SE**	**HR**
Active	≠0.58906	0.07097	0.55485

—Non-collapsibility (although hardly noticeable here)
—And wrong model
—If PH holds conditional on smoking, it cannot hold marginally (and vice versa)

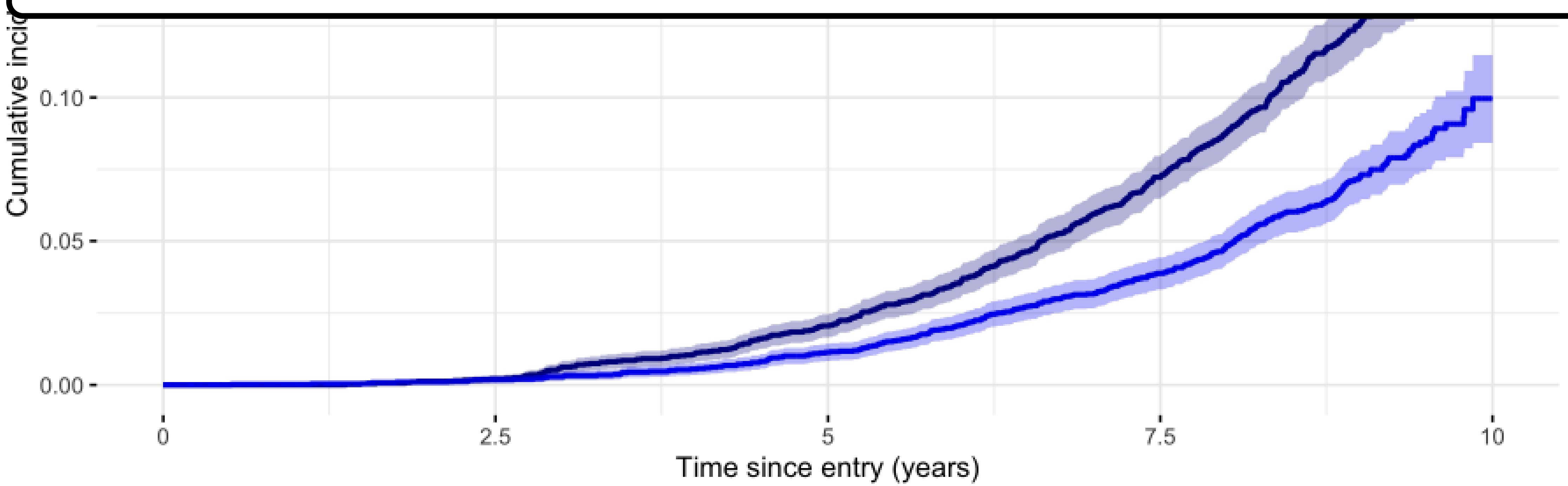
Can check by comparing with Kaplan–Meier



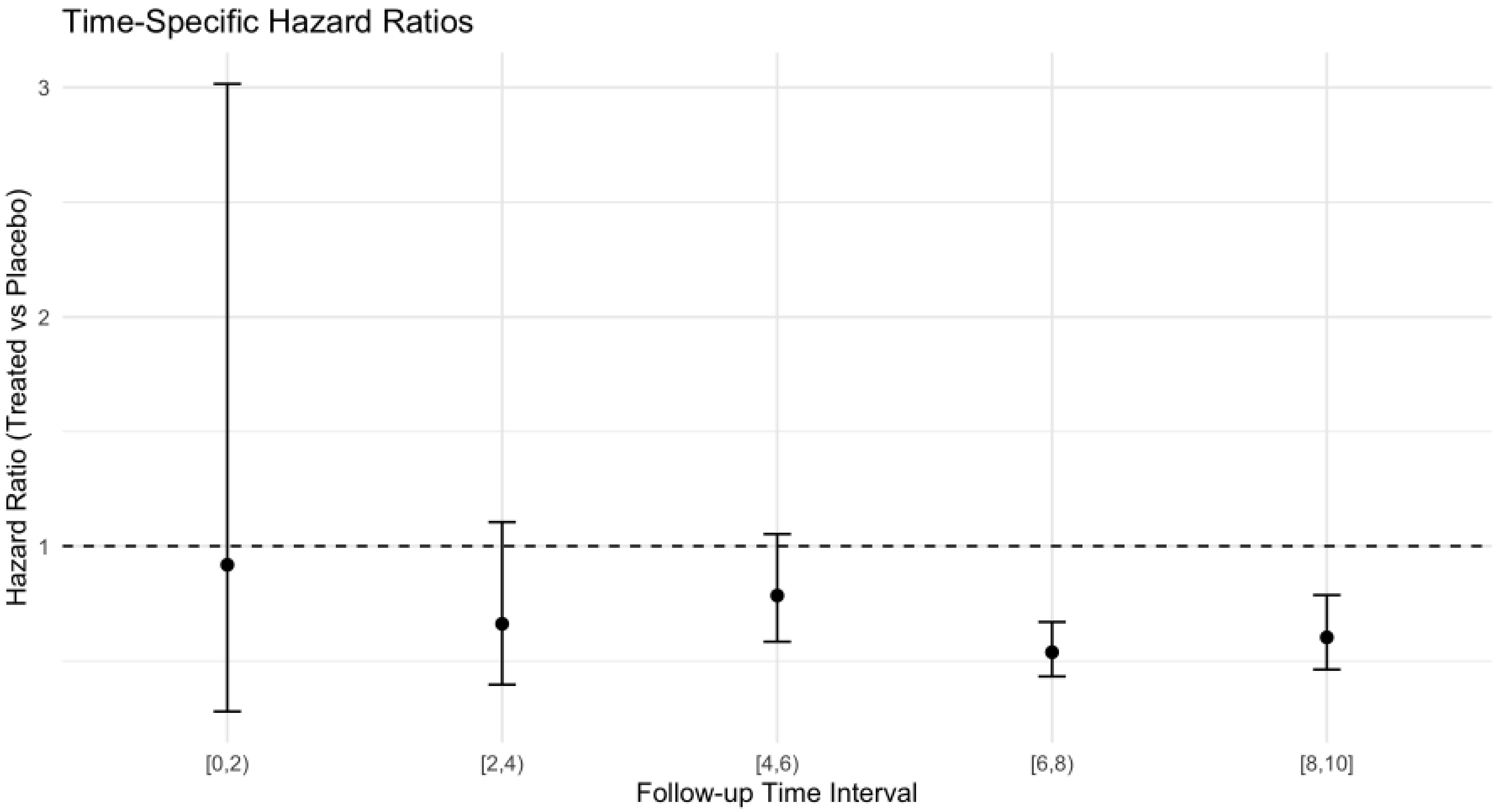
Can check by comparing with Kaplan–Meier

Comparison

—As it happens, there is very little noticeable difference!
—Confirmed by inspecting the period-specific HRs:



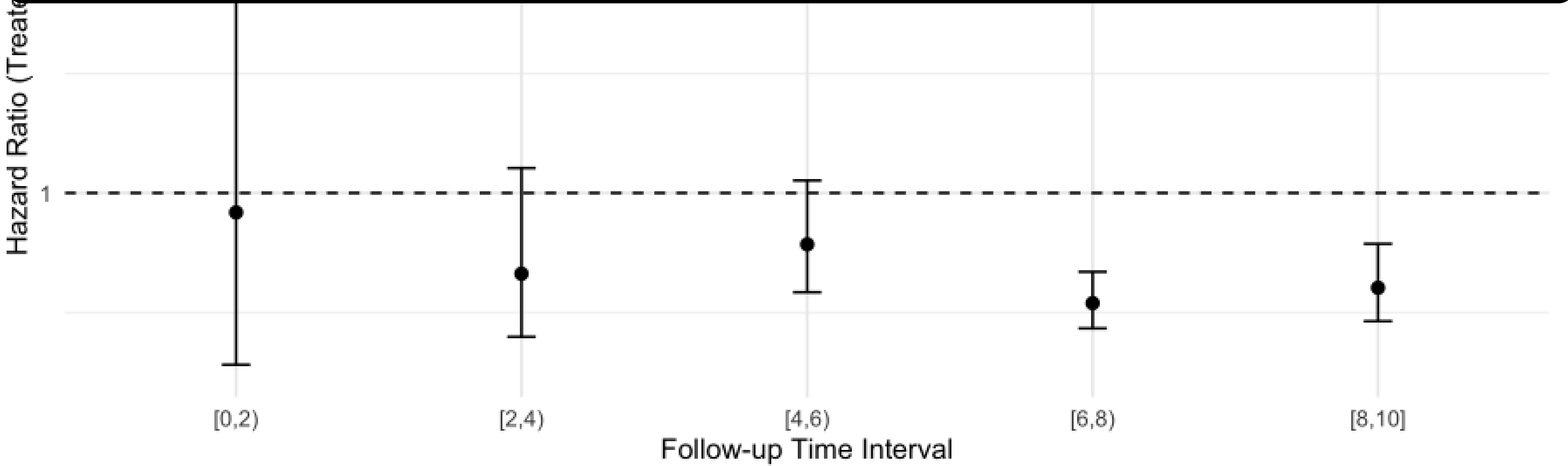
Can check by comparing with Kaplan–Meier



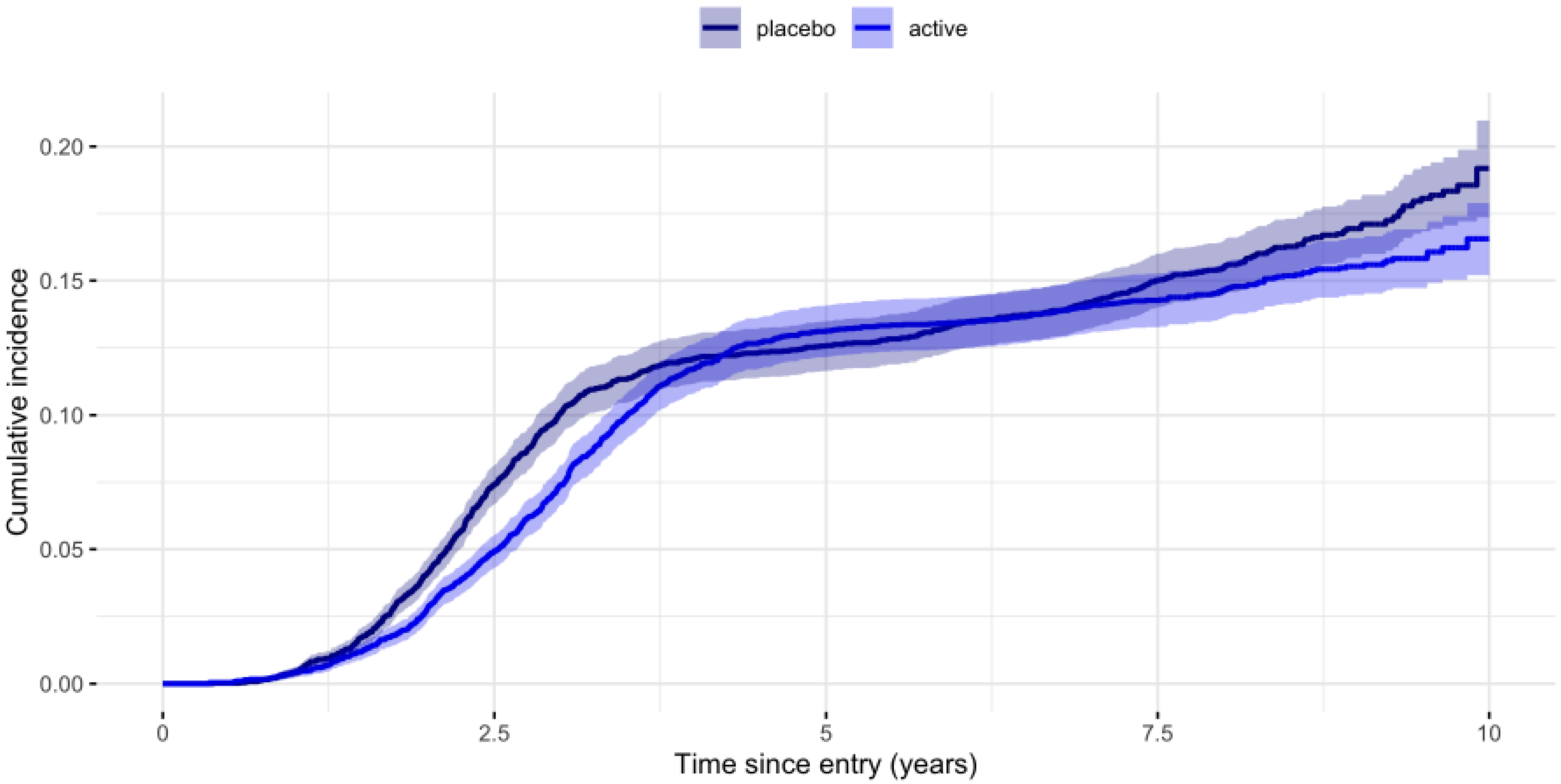
Can check by comparing with Kaplan–Meier

Comparison

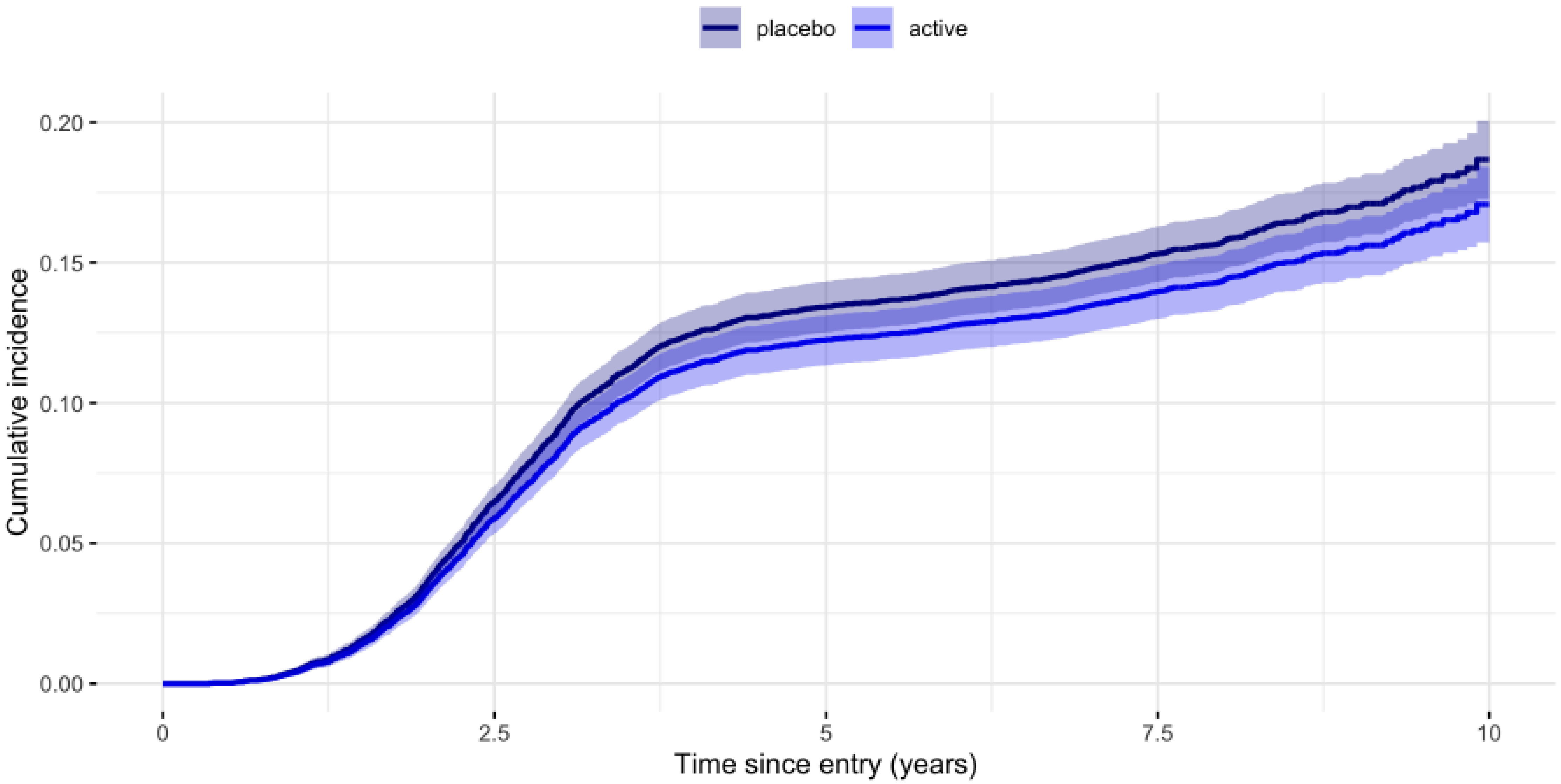
—As it happens, there is very little noticeable difference!
—Confirmed by inspecting the period-specific HRs:
—Here is a more extreme example (with a baseline hazard that varies more over time)
—But still (1) no confounding, (2) proportional hazards conditional on smoking, (3) no heterogeneity by smoking



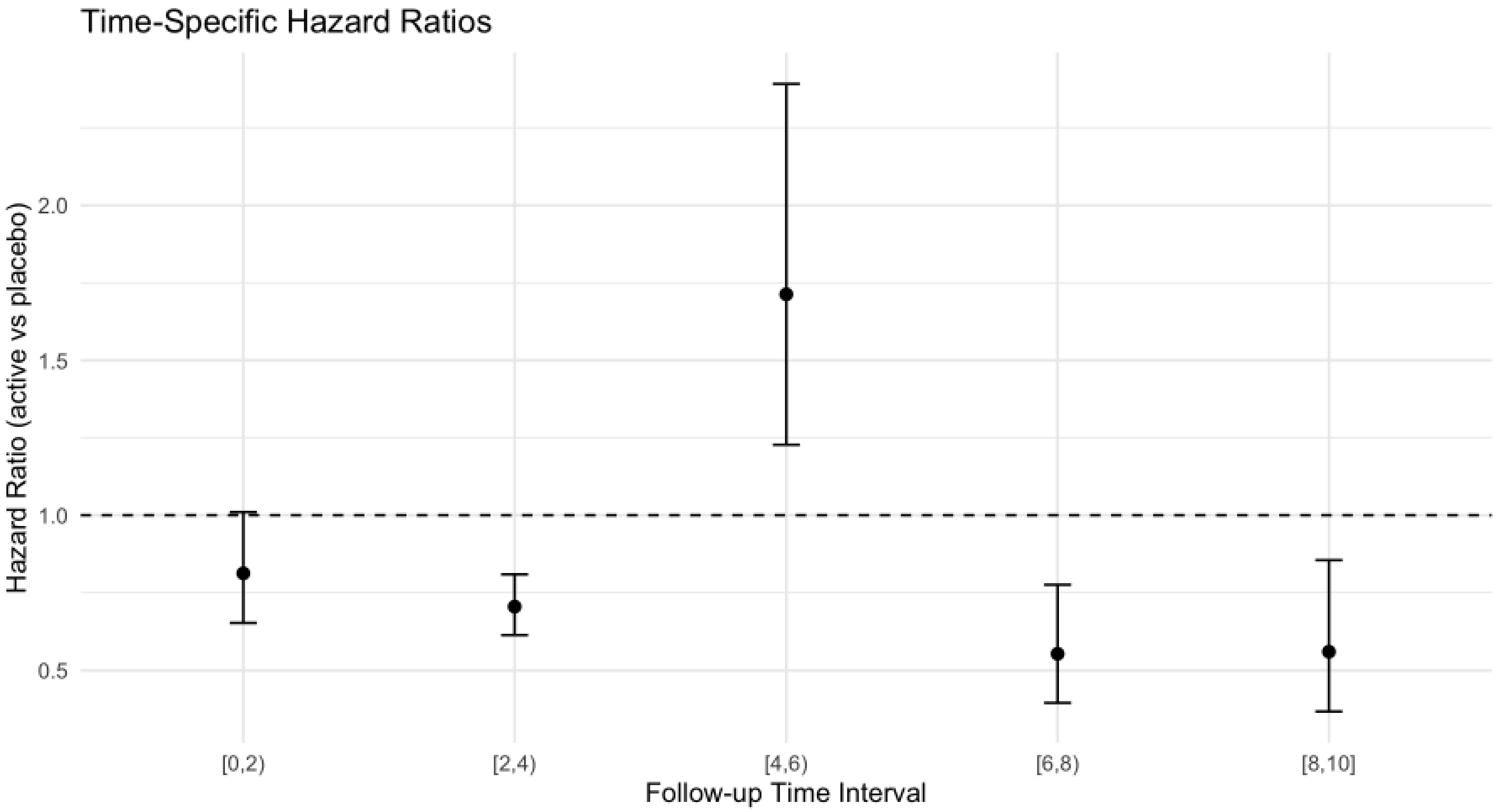
Unadjusted Kaplan–Meier



Unadjusted Cox PH fit



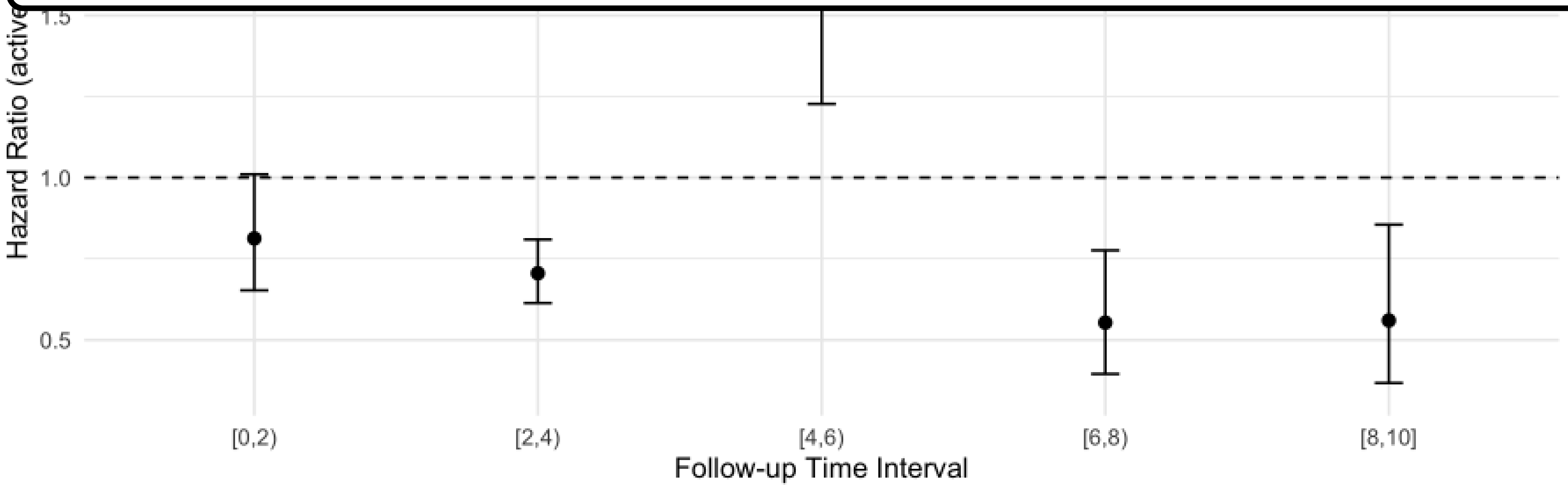
Estimated period-specific HRs



Estimated period-specific HRs

Remarks

—Unadjusted Cox PH model badly misspecified
—**Conditional HR** = 0.5, but estimated **marginal HR** = 0.9: ‘non-collapsibility’
—Even though conditional HR constant over time (and smoking status), estimated period-specific marginal HRs oscillate between ‘protective’ and ‘harmful’: **!HAZARDOUS!** to interpret these!



Outline

Conditioning vs. adjusting

- Traditionally conditional and adjusted are used interchangeably, likewise marginal and unadjusted.
- Useful suggestion from causal inference: use marginal/conditional for the estimand of interest and unadjusted/adjusted for the analysis performed.
- Can obtain an adjusted estimator of a marginal estimand.
- When comparing estimators of a marginal estimand, an adjusted estimator is usually more precise than an unadjusted estimator.
  - This is the case for binary and survival outcomes, not just for continuous outcomes, contrary to some discussions.
  - Confusion enters when people compare the SE of an (adjusted) estimator of a conditional estimand with the SE of an unadjusted estimator of a marginal estimand: apple vs. orange.
  - Williamson *et al.*, 2014; Daniel *et al.*, 2021

Outline

Key messages

- **All of this is highly confusing** (to me), but maybe this is an important lesson in itself!
- **The causal interpretation of hazard ratios is hazardous, esp. for time-varying HRs**
- **Proportional hazards is a fragile assumption**
  — if it holds conditional on a set of prognostic baseline covariates, it doesn't hold marginally, nor conditional on any other set. If it holds marginally, it doesn't hold conditionally
- **Hazard ratios are non-collapsible**
  — although the exact meaning of this is subtle given previous point
- **Conditional estimands are used to mean different things**
  — with and without treatment $\times$ covariate interactions
- **Adjusted $\neq$ conditional**: can obtain adjusted estimators of marginal estimands
  — and **efficiency/power are both increased** as long as the covariates adjusted for are prognostic: this is the case **for HRs too**!