Flexible Parametric Survival Models as an alternative to Cox regression

Therese M-L Andersson

Department of Medical Epidemiology and Biostatistics,Karolinska Institutet, Stockholm, Sweden

About me

- Background in mathematical statistics.
- PhD from KI in 2013.
- Senior lecturer in biostatistics at MEB.
- Research interest in survival analysis and cancer epidemiology.
- Mostly population-based studies on cancer patient survival, both applications and method development.

Overview

- Short introduction to survival analysis and Cox proportional hazards model.
- Introduction to flexible parametric models in survival analysis as an alternative to the Cox model.
- Discuss the importance of modelling non-proportional hazards and how this can be done within the flexible parametric approach.
- Show the use of the models in Stata, using the `stpm2` command.
- Examples of ways of presenting results from survival data using flexible parametric survival models.

Why do we need survival analysis?

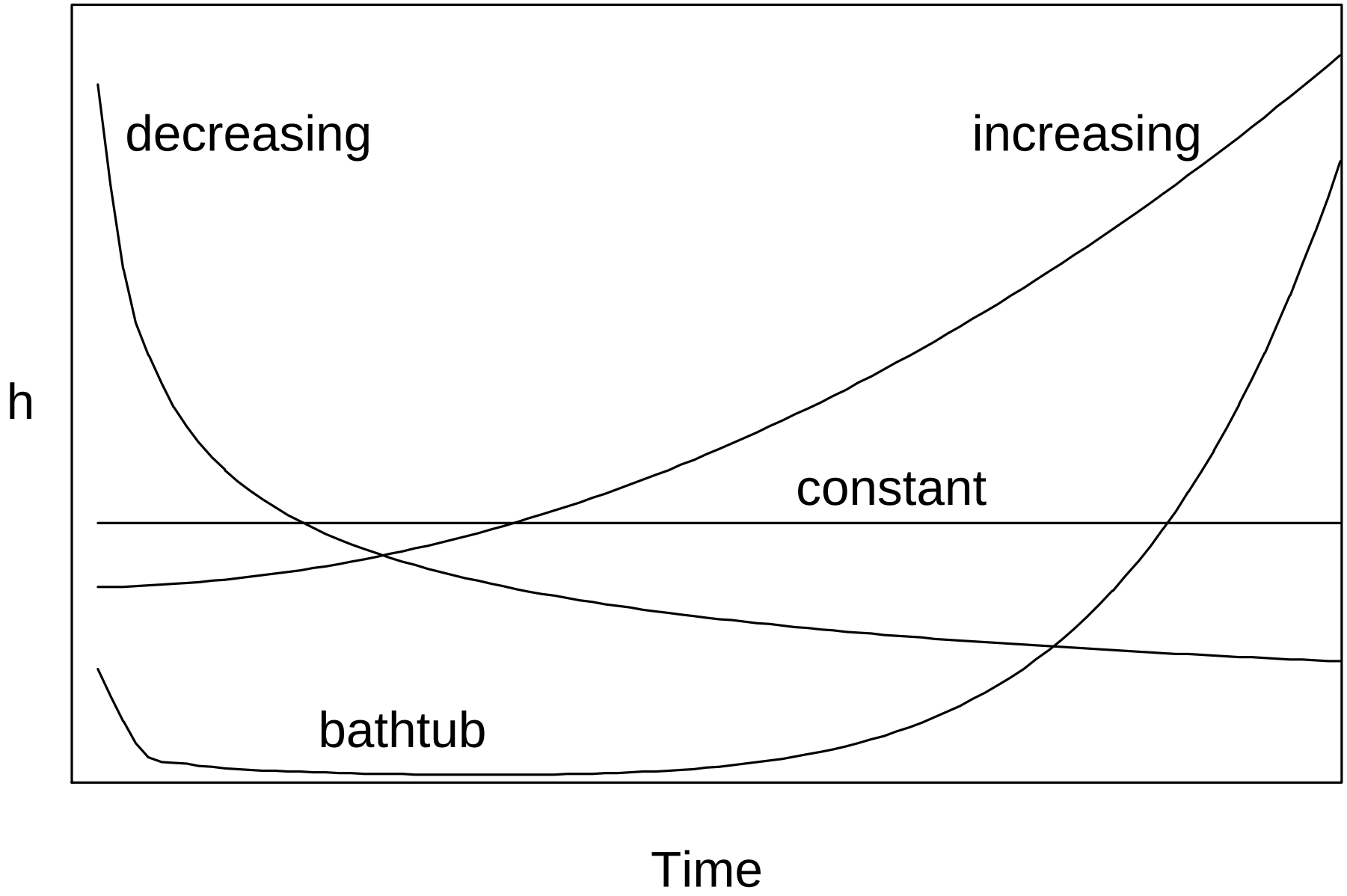
- The characteristic that complicates the use of standard statistical methods is censoring - unobserved values of the response measurement of interest.
- Censoring leads to differences in follow-up time between individuals.
- The main message is that, in survival analysis, the outcome has two dimensions – the event indicator and the time at risk.

Why do we need survival analysis?

- What we would like is some measure of the risk of an event adjusted for the fact that individuals were at risk for different lengths of time.
- Methods used for making inference about proportions (e.g., logistic regression) are only appropriate when all individuals have the same time at risk. This is typically not the case when we have survival data.
- If, however, individuals are at risk for differing lengths of time we use 'person-time' as the denominator and estimate the event rate (e.g. mortality rate).

$$\text{event rate} = \frac{\text{number of events}}{\text{person-time at risk}}$$

The hazard function



The Cox Model

- The Cox proportional hazards model is by far the most common regression model used in survival.

$$h(t) = h_0(t)\exp(\beta X)$$

$$\ln(h(t)) = \ln(h_0(t)) + \beta X$$

- The quantities estimated from a Cox model are *(log) hazard ratios*.
- The baseline hazard, $h_0(t)$ is not estimated from a Cox model
- The advantage of this is that one does not need to make (parametric) assumptions about the shape of the baseline hazard function. The disadvantage is that the underlying shape of the hazard function is often ignored.

The Cox Model

- Let's look at an example. We are interested in comparing the cause-specific mortality for colon cancer patients diagnosed in two calendar periods, 1975-1984 and 1985-1994.

_t	Haz. ratio	Std. err.	P>|z|	95% CI	
1985-1994	.8487157	.0202711	0.000	.8099006	.8893909

- The Cox model automatically adjusts for the underlying time-scale, even though the baseline is not estimated.
- The mortality rate for those diagnosed in the later calendar period is 15% lower than for those diagnosed in earlier period.
- The HR is the same across time, although the underlying rates can differ across time.

The Cox Model

- We can also include other covariates.

_t	Haz. ratio	Std. err.	P>|z|	95% CI	
1985-1994	.8797365	.0210338	0.000	.839462	.9219432
stage					
Regional	2.25628	.0933739	0.000	2.080496	2.446916
Distant	8.022359	.2377227	0.000	7.569703	8.502082

- The Cox model includes the covariates calendar period and stage at diagnosis (localised/regional metastasis/distant metastasis).
- However, also adjusts for the underlying time-scale.
- The HR of 0.88 is the same within each stage and across time.

The Cox Model

- The crucial assumption of the Cox model is that the estimated parameters are not associated with time. In other words, any hazard ratio estimated from the model is assumed not to vary over follow-up time, i.e we assume proportional hazards.
- If you are only interested in the relative effect of a covariate on the hazard rate, the assumption of proportional hazards is reasonable then the Cox model is probably the most appropriate model.
- However, whenever we estimate a relative effect we should ask "relative to what?"
- Also, the proportional hazards assumption is often not valid.

How to allow for non-proportional hazards

- There are several ways, but I think the easiest way is to include interaction between the covariate and the time-scale.
- Non-proportional hazards means that the effect of a covariate differs across time.
- This is the same as interaction, or effect modification, with time.
- Compare to what you would do if the effect of calendar period would be different for different stages at diagnosis. You would include an interaction.

Why Parametric Models

- We have the Cox model so why use parametric models?
- Parametric Models have advantages for
  - Prediction.
  - Extrapolation.
  - Quantification (e.g., absolute and relative differences in risk).
  - Modelling time-dependent effects.
  - Understanding.
  - Complex models in large datasets (time-dependent effects / multiple time-scales)
- The estimates we get from flexible parametric survival models are incredibly similar to those obtained from a Cox model.

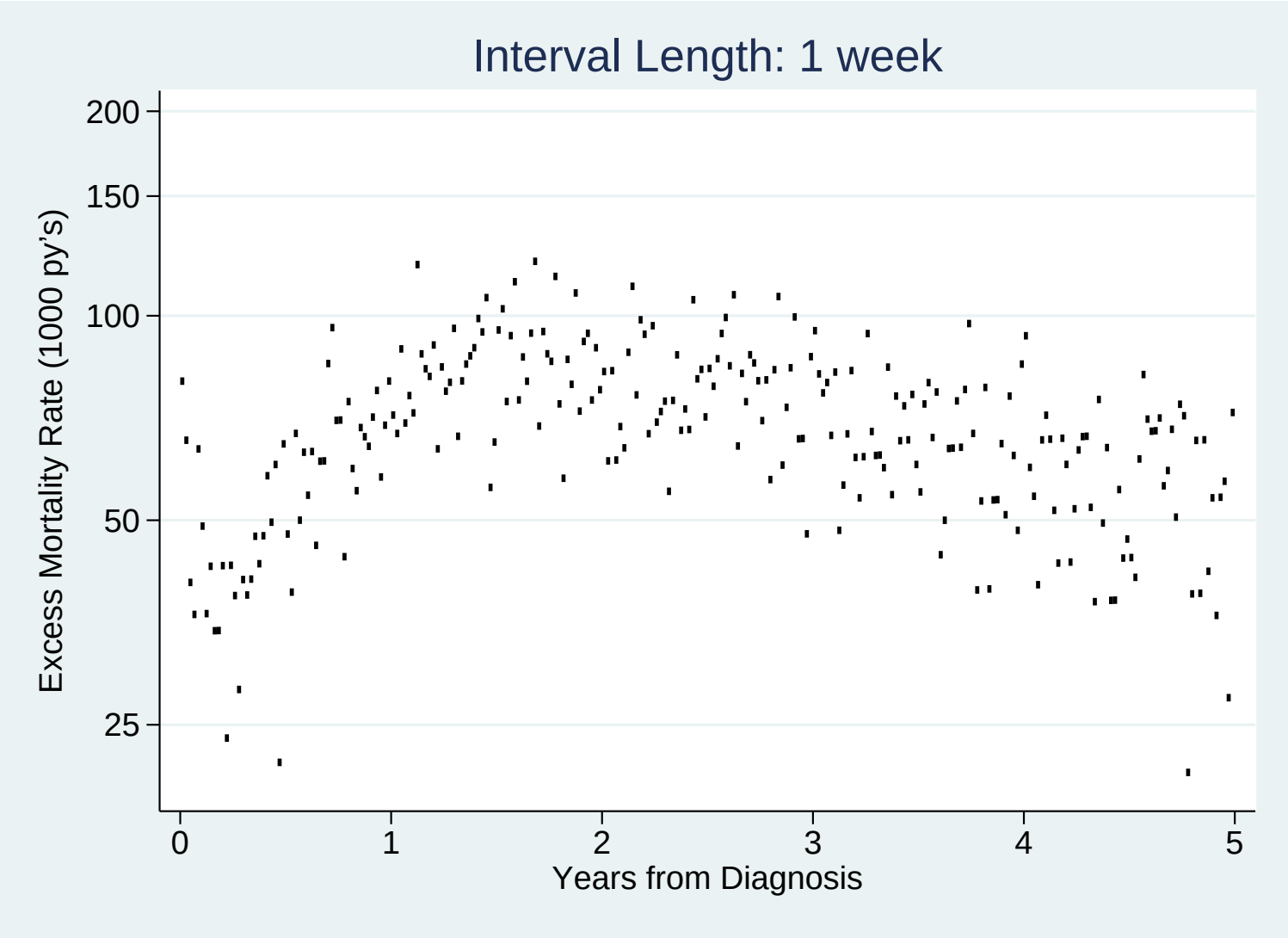
Flexible parametric survival models

- In Cox regression the baseline hazard is not estimated.
- Many parametric models make strong assumptions about the baseline which might not be plausible in many settings.
- Flexible parametric models are an alternative, using splines to explicitly model the baseline hazard.
- Splines are a way of modeling continuous variables in a flexible way.
- By modelling the baseline it is easier to plot and make different types of predictions.
- Also easier to allow for non-proportional hazards, since it is an interaction.

Splines

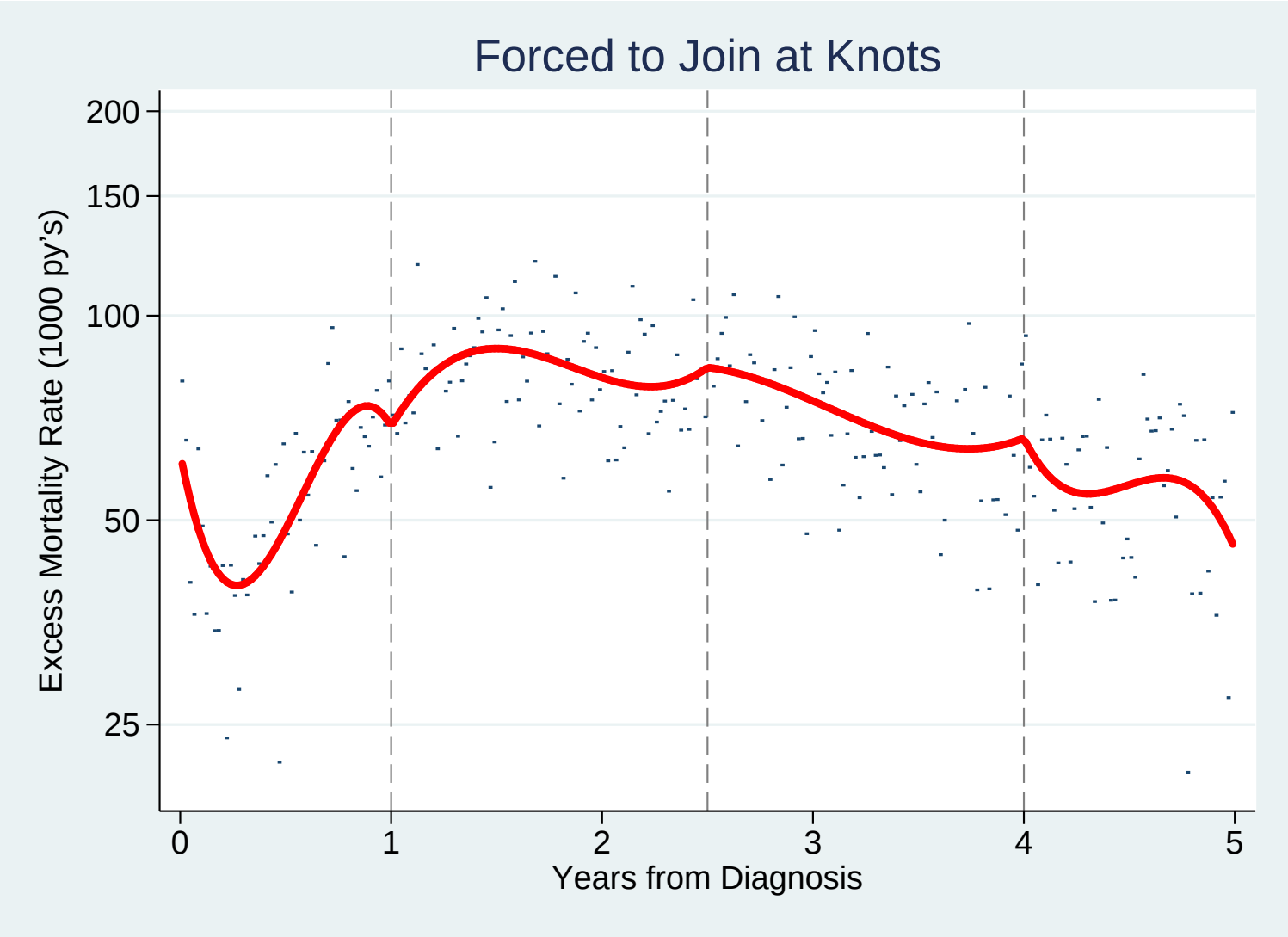
- Flexible mathematical functions defined by piecewise polynomials.
- Used in regression models for non-linear effects
- The points at which the polynomials join are called knots.
- Constraints ensure the function is smooth.
- The most common splines used in practice are cubic splines.
- However, splines can be of any degree, n.
- Function is forced to have continuous 0^{th}, 1^{st} and 2^{nd} derivatives.

Piecewise hazard function

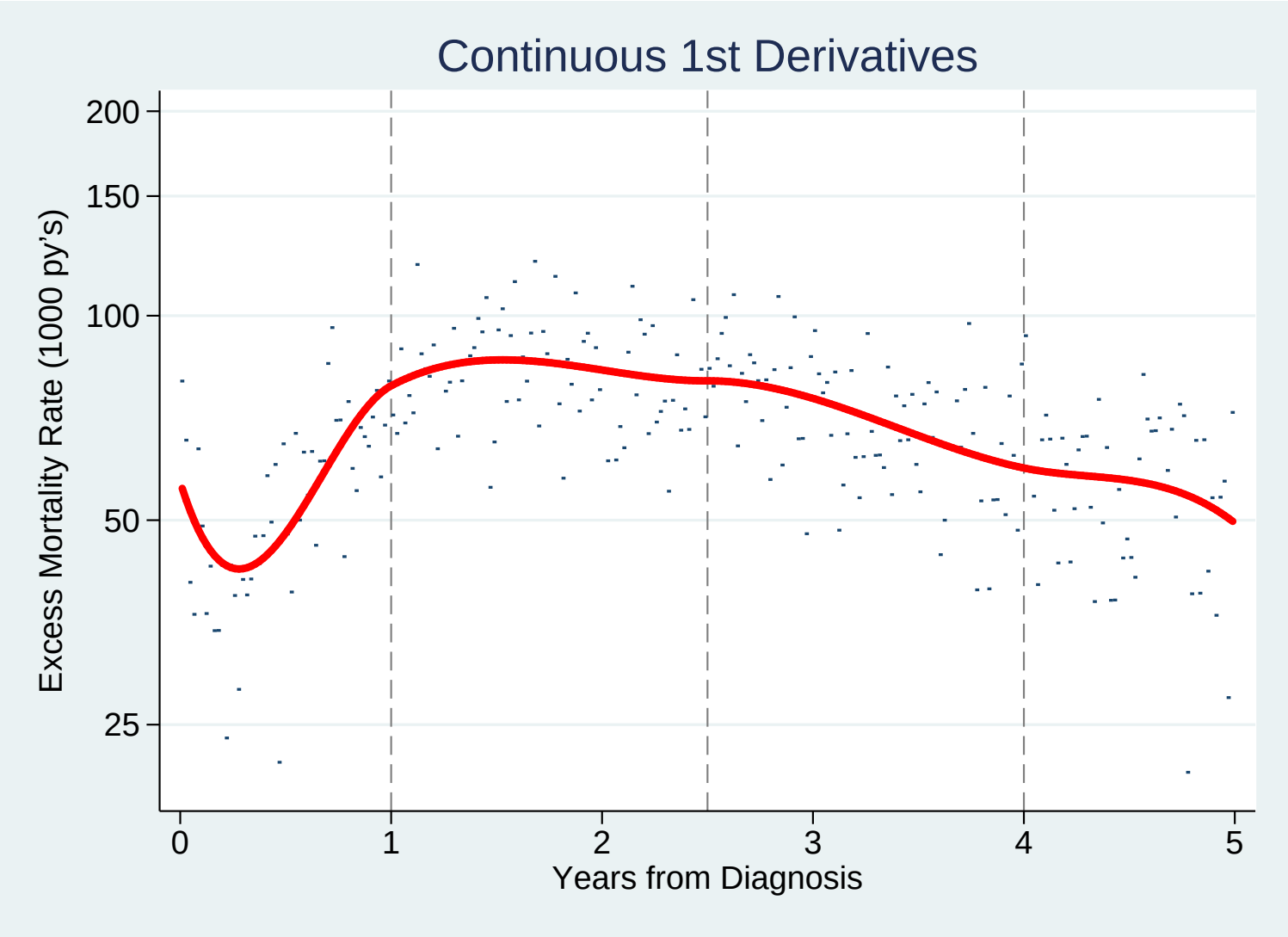


No Continuity Corrections

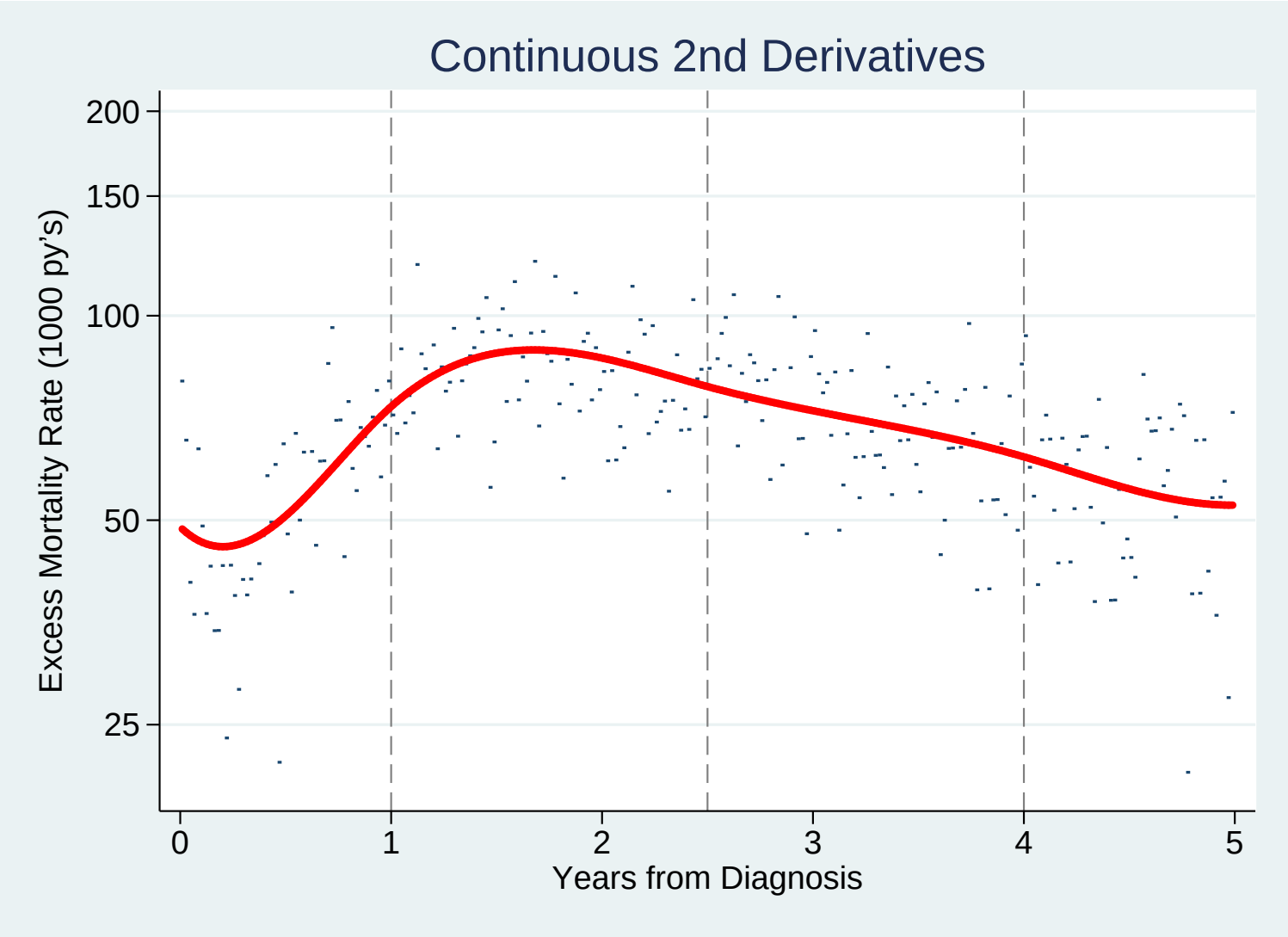
Function forced to join at knots



Continuous first derivative



Continuous second derivative



Flexible parametric survival model

- The model can be written as:

$$\ln(H(t)) = s(\ln t; \gamma_0, K) + \beta \mathbf{x}$$

  where K is the number of knots.
- Modelling on the hazard scale requires computationally intensive numerical integration
- Another potential problem is that many parameters might be needed, since the hazard function can have take any shape. The flexible parametric survival model is instead using the cumulative hazard, which is an increasing function.
- Parameter estimates are still interpreted as hazard ratios (if a PH model).
- Easy to transform to the survival or hazard scale.

Flexible parametric survival model

- The model shown on the previous slide is a proportional hazards model.
- Non-proportional hazards models (time-dependent effects) can be modeled by including interactions between covariates and splines for time.

$$\ln(H(t)) = s(\ln t; \gamma_0, K) + \beta \mathbf{x} + \sum_{j=1}^{D} s(\ln t; \gamma_j, K_j) * x_j$$

where D is the number of covariates with non-proportional hazards and K_j is the number of knots for each interaction.

Flexible parametric survival model

- Let's again revisit the example of colon cancer.
- We will focus on the HR of cancer-specific death, comparing the two calendar periods. Adjusting for stage at diagnosis.
- First a flexible parametric model with proportional hazards.
- Then a flexible parametric model allowing for non-proportional hazards for stage, i.e. including an interaction between time and stage.

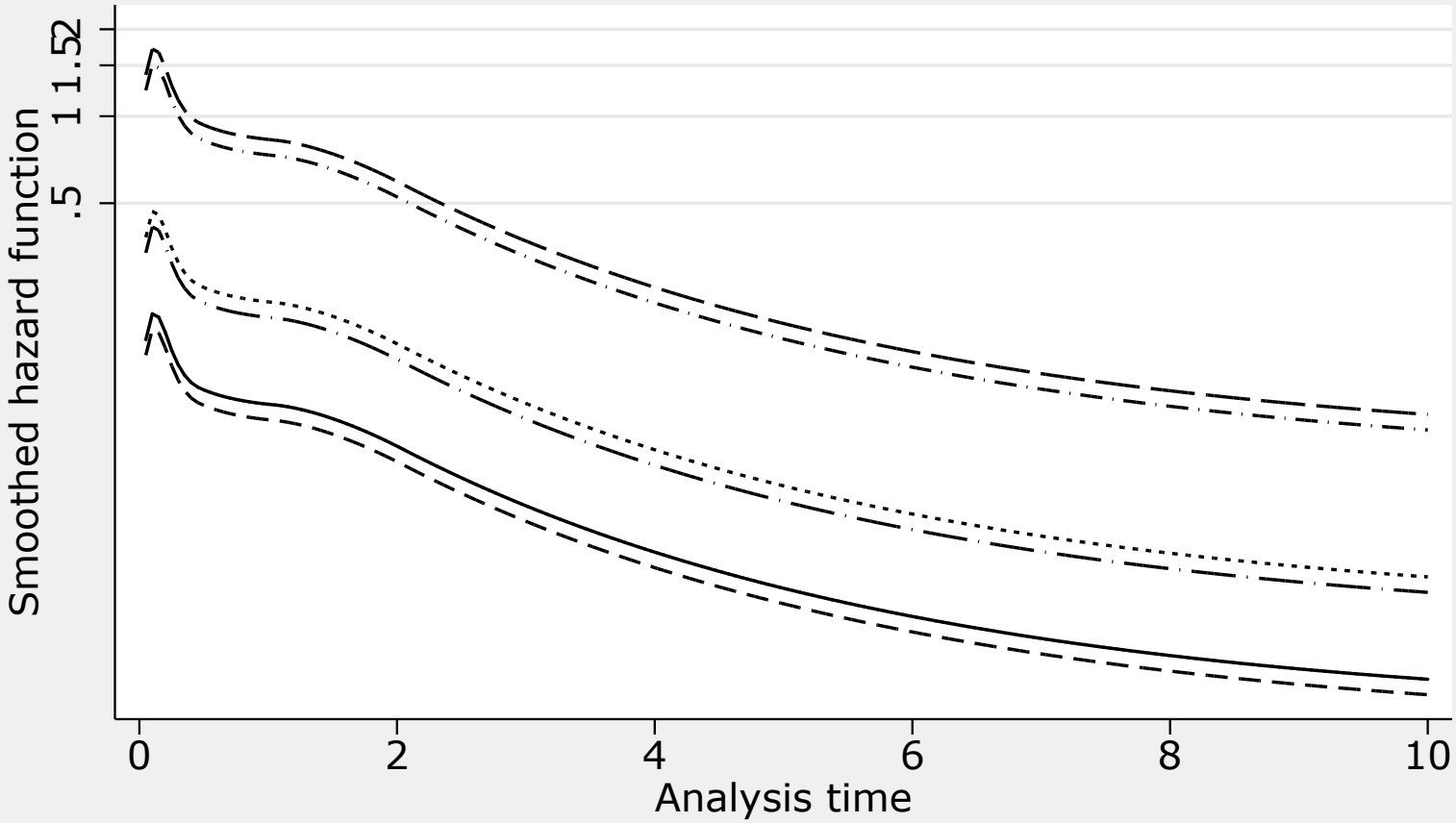
Flexible parametric survival model

```
. stpm2 year8594  i.stage, scale(hazard) df(5) eform

Log likelihood = -17317.704     Number of obs     =     13,208

------------------------------------------------------------------------------
             |     exp(b)   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
xb           |
1985-1994    |   .8838128   .0211523    -5.16   0.000     .8433125    .9262582
             |
       stage |
Regional     |   2.260989   .0935674     19.71   0.000     2.084841    2.452021
Distant      |   8.249828   .2439361     71.37   0.000     7.785313     8.74206
             |
_rcs1        |   3.150462   .0371294     97.37   0.000     3.078524    3.224082
_rcs2        |   1.302572   .0120809     28.50   0.000     1.279108    1.326467
_rcs3        |   .9965837   .0059341     -0.57   0.565     .9850207    1.008282
_rcs4        |   1.048574   .0037925     13.11   0.000     1.041167    1.056034
_rcs5        |   1.022945   .0028764      8.07   0.000     1.017323    1.028598
_cons        |   .1395332   .0040286    -68.21   0.000     .1318566    .1476567
------------------------------------------------------------------------------
```

Flexible parametric survival model

- Patients diagnosed in the later calendar period have 12% lower cancer-specific mortality compared to earlier calendar period, after controlling for stage at diagnosis (and the underlying time scale), and this difference is assumed be the same for all stages.
- Patients with regional metastases have more than 2 times the mortality of patients with localised stage, after controlling for calendar period (and the underlying time scale), and the effect is assumed to be the same within both calendar periods.
- Patients with distant metastases have more than 8 times the mortality of patients with localised stage, after controlling for calendar period, and the effect is assumed to be the same within both calendar periods.
- The rest of the parameters are for the splines, and they are not interpreted one by one. However, together they give the function of the baseline.

Flexible parametric survival model



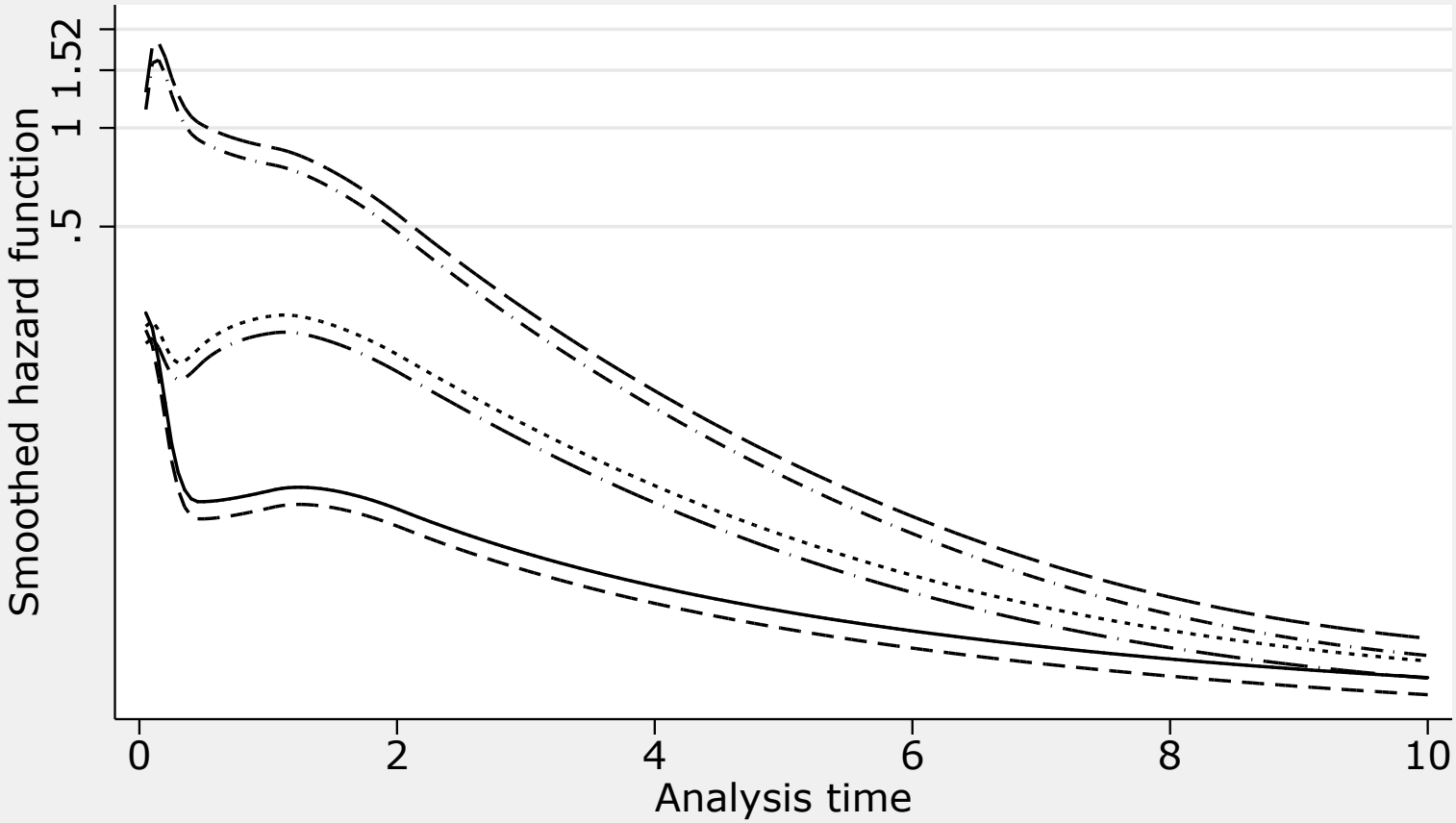
Flexible parametric survival model

```
. stpm2 year8594 stage2 stage3, scale(hazard) df(5) ///
tvc(stage2 stage3) dftvc(3) eform

Log likelihood = -17132.061                     Number of obs     =     13,208
------------------------------------------------------------------------------
             |     exp(b)   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
xb           |
1985-1994    |   .8864017   .0212132    -5.04   0.000     .8457845    .9289694
stage2       |   1.904373   .1029325    11.92   0.000     1.712947     2.11719
stage3       |   7.669099   .2661191    58.71   0.000     7.164857    8.208829
_rcs1        |   2.720272   .0693896    39.23   0.000     2.587615     2.85973
_rcs2        |   1.072641   .0178232     4.22   0.000     1.038271    1.108149
_rcs3        |   .9587595   .0098322    -4.11   0.000     .9396812    .9782252
_rcs4        |   1.046235   .0055504     8.52   0.000     1.035413    1.057171
_rcs5        |   1.024975   .0028437     8.89   0.000     1.019417    1.030564
_rcs_st21    |   1.369303   .0736305     5.85   0.000     1.232334    1.521495
_rcs_st22    |   1.130488   .0418797     3.31   0.001     1.051315    1.215625
_rcs_st23    |   1.142751   .0245353     6.21   0.000      1.09566    1.191865
_rcs_st31    |   1.126492    .032725     4.10   0.000     1.064144    1.192493
_rcs_st32    |   1.337784   .0264739    14.71   0.000      1.28689    1.390691
_rcs_st33    |   1.034256   .0130236     2.67   0.007     1.009043      1.0601
_cons        |   .1445927   .0048611   -57.52   0.000     .1353723     .154441
------------------------------------------------------------------------------
```

Flexible parametric survival model

- Patients diagnosed in the later calendar period have 11% lower cancer-specific mortality compared to earlier calendar period, after controlling for stage at diagnosis with non-proportional hazards (and the underlying time scale), and this difference is assumed be the same for all stages.
- Since stage is allowed to have non-proportional hazards, i.e. an interaction between stage and the time-scale, the HR changes over time, and is not one number found in the output.
- However, the HR for stage can be plotted as a function of time (see later example).

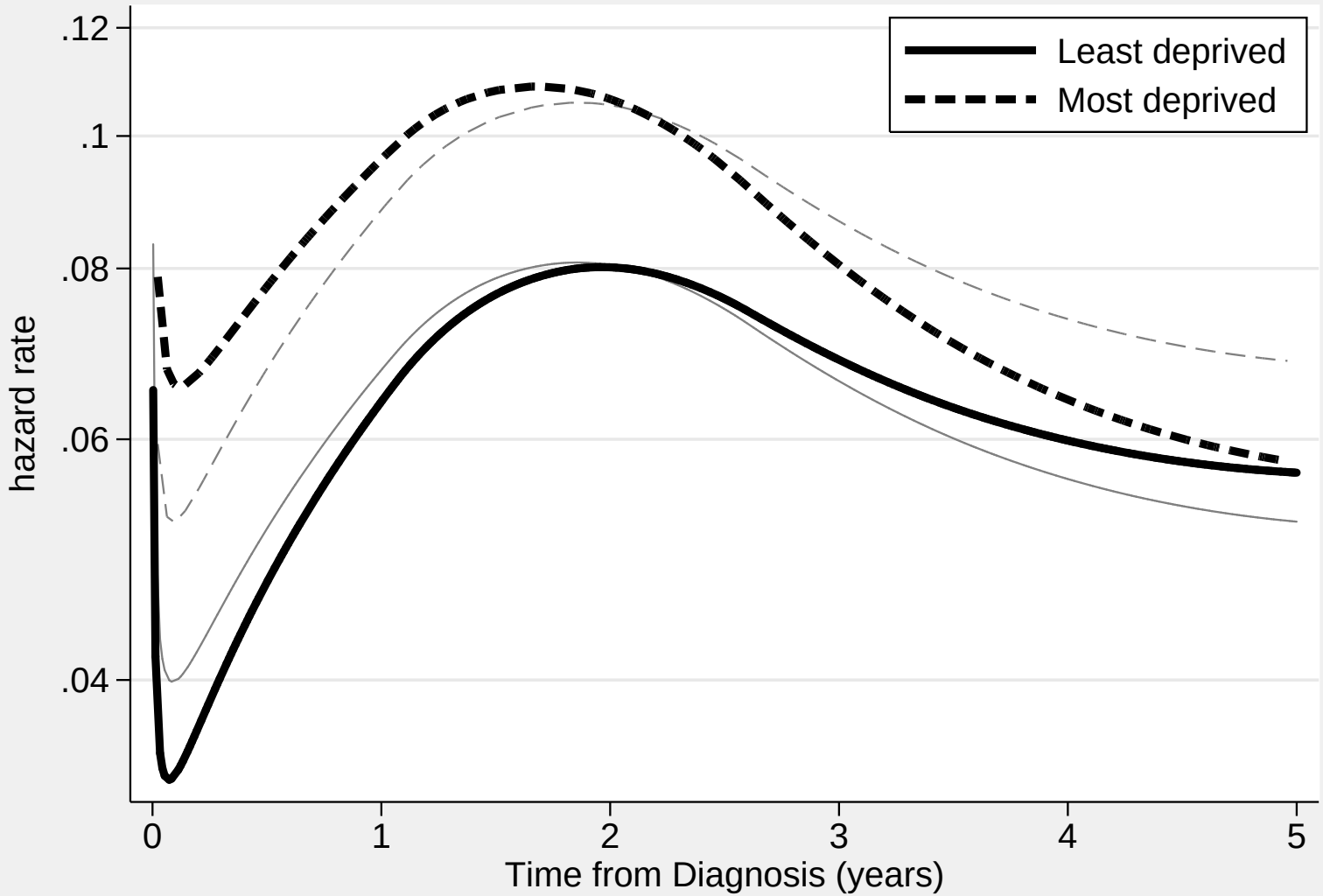
Flexible parametric survival model



Flexible parametric survival model

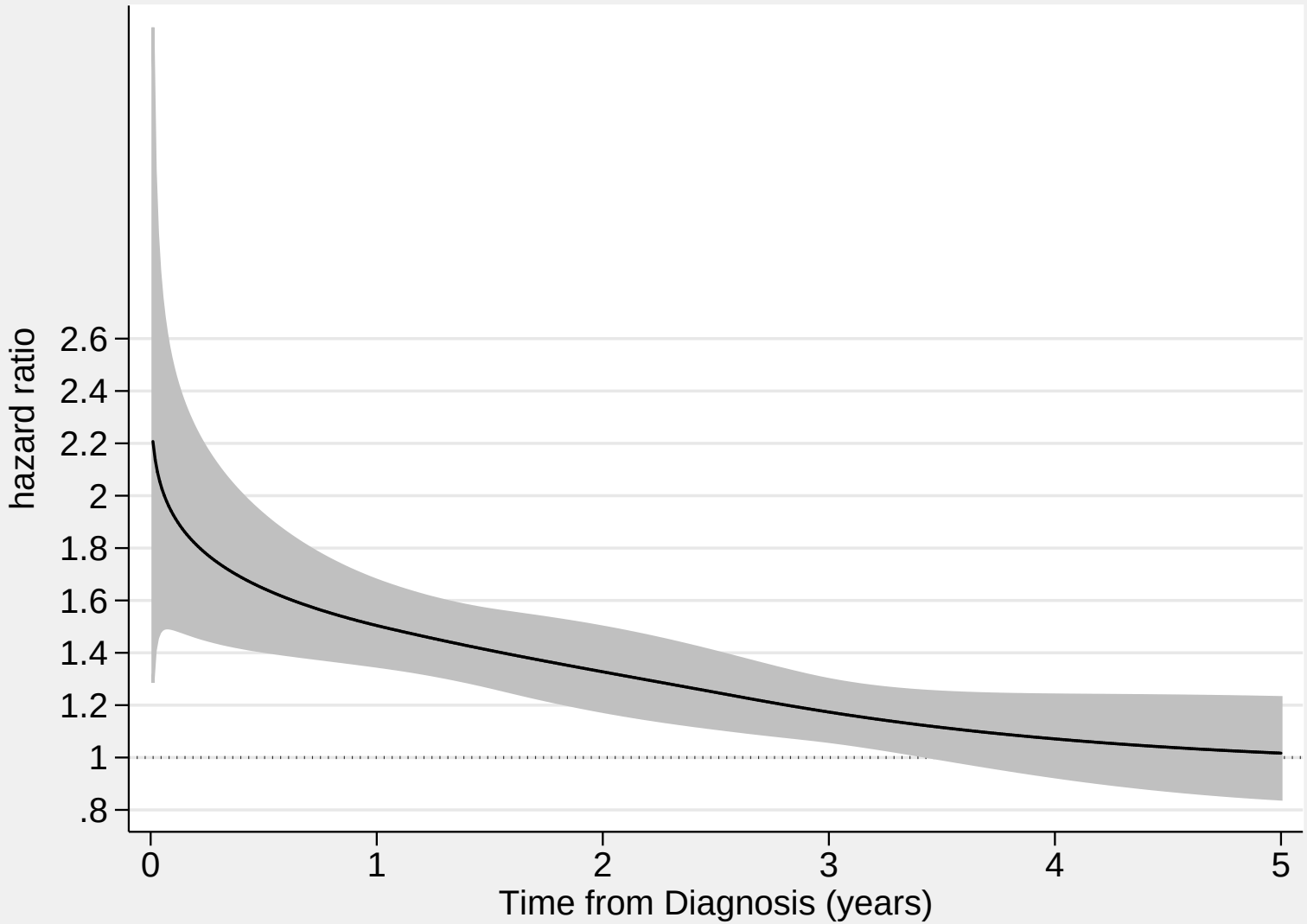
- Since the baseline hazard is estimated as a continuous function in the flexible parametric survival model it is easy to present results using graphs, and to present results on the hazard scale, as hazard ratios, or the survival scale.
- This is illustrated in the following graphs.
- A flexible parametric survival model fitted to data on breast cancer patients in England, with breast cancer death as the outcome.
- The variable of interest is deprivation status, and results are shown for the lowest and highest group.

Flexible parametric survival model

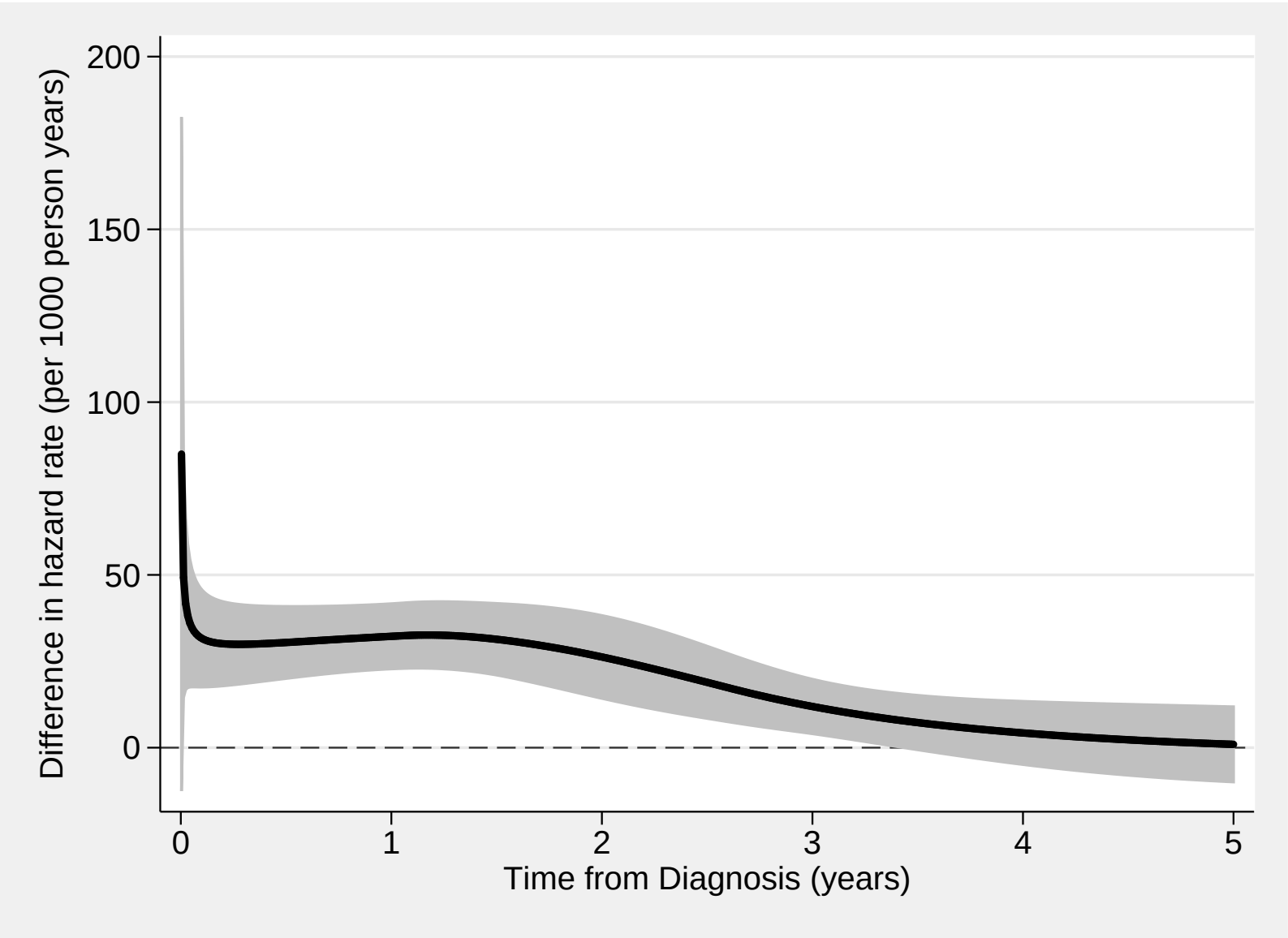


Thinner lines are predictions from proportional hazards model

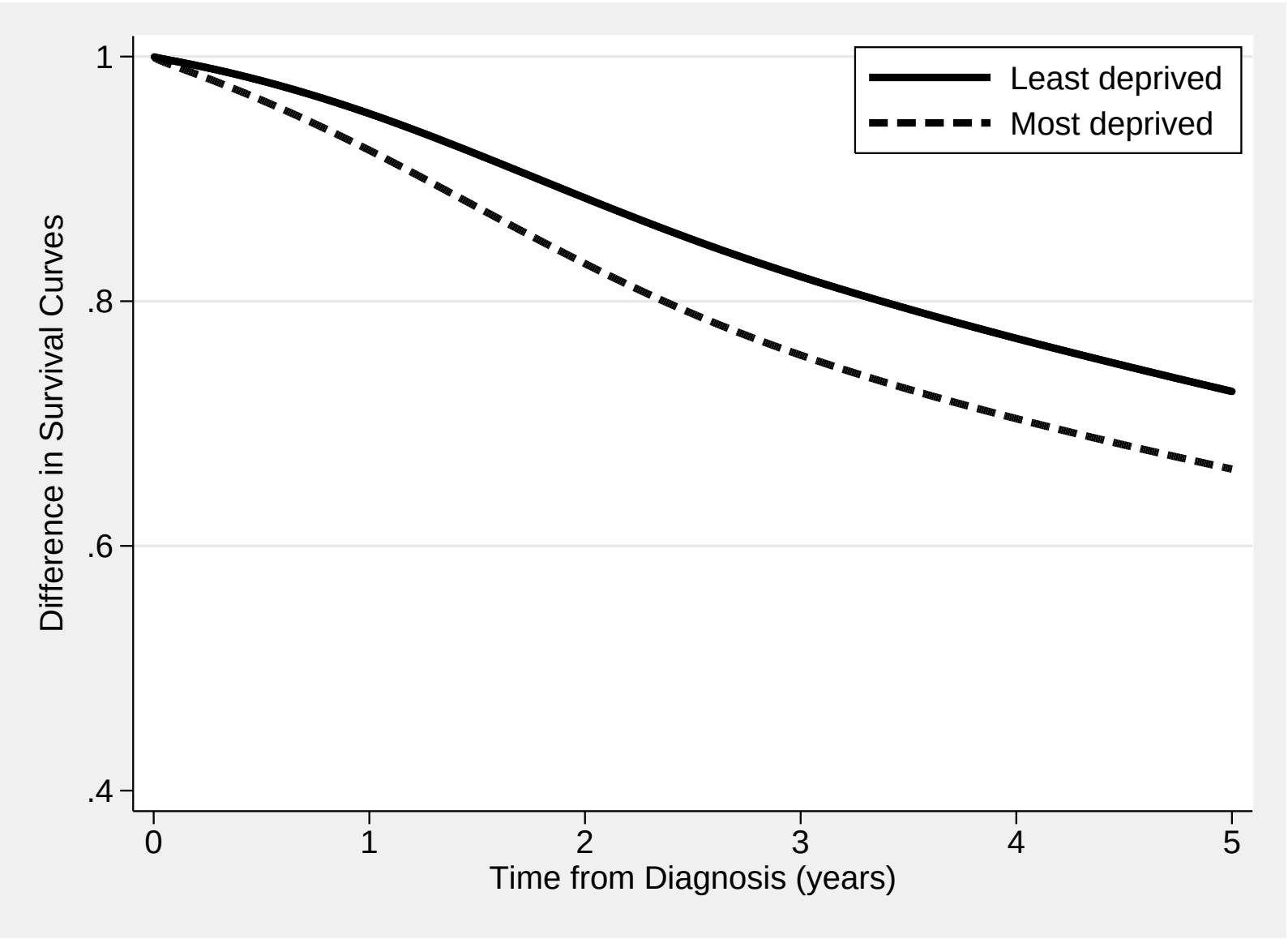
Flexible parametric survival model



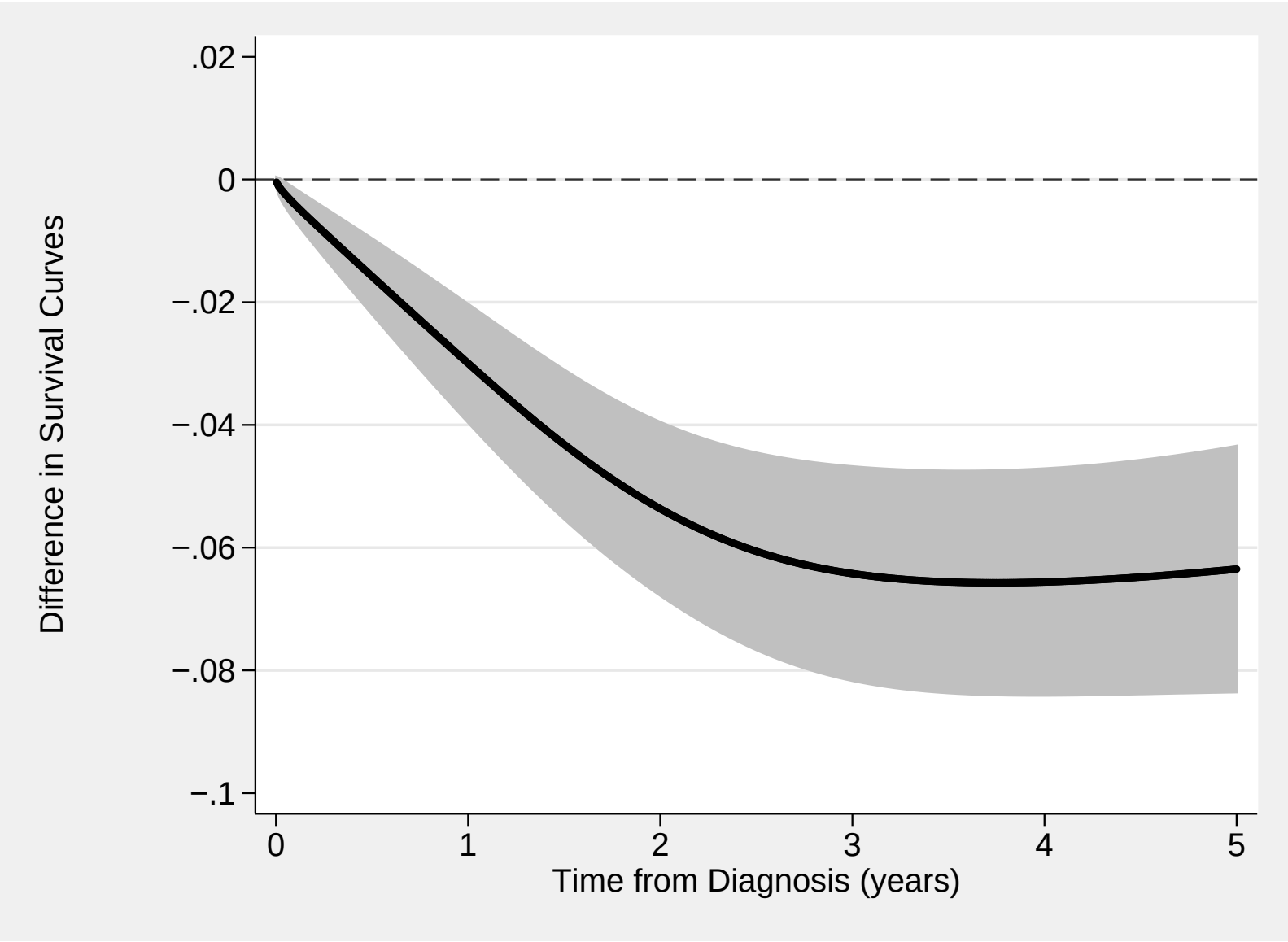
Flexible parametric survival model



Flexible parametric survival model



Flexible parametric survival model



Flexible parametric survival model

A few examples of other measures/extensions

- Cumulative incidence within a competing risks setting
- Marginal measures, e.g. marginal survival
- Restricted mean survival time
- Life expectancy and loss in life expectancy
- Relative survival models (excess mortality)

Summary, flexible parametric model

- Hazard ratios are very similar to hazard ratios from a Cox model.
- The time-scale is included as a continuous variable using splines.
- Since the baseline hazard is modelled it is easy to include non-PH by including interactions with time.
- Easy to present results, both on absolute and relative scales, using graphs.
- The parametric approach enables predictions and extrapolations.

Thank you

Acknowledgement

- Paul Lambert
- Paul Dickman
- Sandra Eloranta
- Anna Johansson
- Elisavet Syriopoulou
- Mark Rutherford
- Caroline Weibull

Some references

- Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. Stat Med. 2002
- Lambert PC, Royston P. Further development of fexible parametric models for survival analysis. The Stata Journal. 2009
- Royston P, Lambert PC. Flexible parametric survival analysis in Stata: Beyond the Cox model. Stata Press, 2011.
- https://pclambert.net/software/stpm2/
- https://pclambert.net/software/standsurv/
- Jackson CH. flexsurv: A Platform for Parametric Survival Modeling in R. J Stat Softw. 2016
- rstpm2 package - written by Mark Clements (MEB, KI)

Some references

- Hinchliffe RS, Lambert PC. Flexible parametric modelling of cause-specific hazards to estimate cumulative incidence functions. BMC Medical Research Methodology. 2013
- Syropoulou E, Wasterlid T, Lambert PC, Andersson TM. Standardised survival probabilities: a useful and informative tool for reporting regression models for survival data. Br J Cancer. 2022
- https://pclambert.net/software/standsurv/standardized_survival_rmst/
- Andersson TM, Dickman PW, Eloranta S, Lambe M, Lambert PC. Estimating the loss in expectation of life due to cancer using flexible parametric survival models. Stat Med. 2013
- Nelson CP, Lambert PC, Squire IB, Jones DR. Flexible parametric models for relative survival, with application in coronary heart disease. Stat Med. 2007