

A new strategy for meta-analysis of continuous covariates in observational studies with IPD

Willi Sauerbrei & Patrick Royston

Overview

- Motivation
- Continuous variables – functional form
- Fractional polynomials
- Brief description of traditional meta-analysis of trials data
- Extension to continuous covariates
- Meta-analysis of continuous covariates – linear approach
- Our proposed strategy based on fractional polynomials
- Discussion

MA – standard approaches to analysis

- ... in randomized controlled trials
- “Treatment effect” in each trial is often binary
 - one effect size (ES) parameter to estimate per trial
- *Mechanically*, one combines ES values as a weighted average over the n studies and computes its 95% CI
- *Implicitly*, one has a model – usually “fixed effects” or “random effects”

Extension to observational studies

- In principle, no different to RCTs, but must adjust effect estimate of interest for confounding
- Adjustment for confounding is problematic for several reasons
 - Different confounders per study, different measurement techniques, ...
- Doing MA in this setting is virtually impossible without individual participant data (IPD)

Extension of MA for continuous covariates

- Continuous covariates are problematic for researchers
- Some categorize continuous covariates into groups and proceed as for standard MA
 - Analyzing one or more dummy variables
 - (But different cutpoints may be used across studies)
- Alternatively, some assume the dose/response function to be **linear**, and meta-analyze the regression slopes
 - with or without adjustment for confounders
- None of this allows for **heterogeneity in functional form** between studies
- What to do if **confounders differ** across studies?

Continuous variables – what functional form?

Traditional approaches

a) Linear function

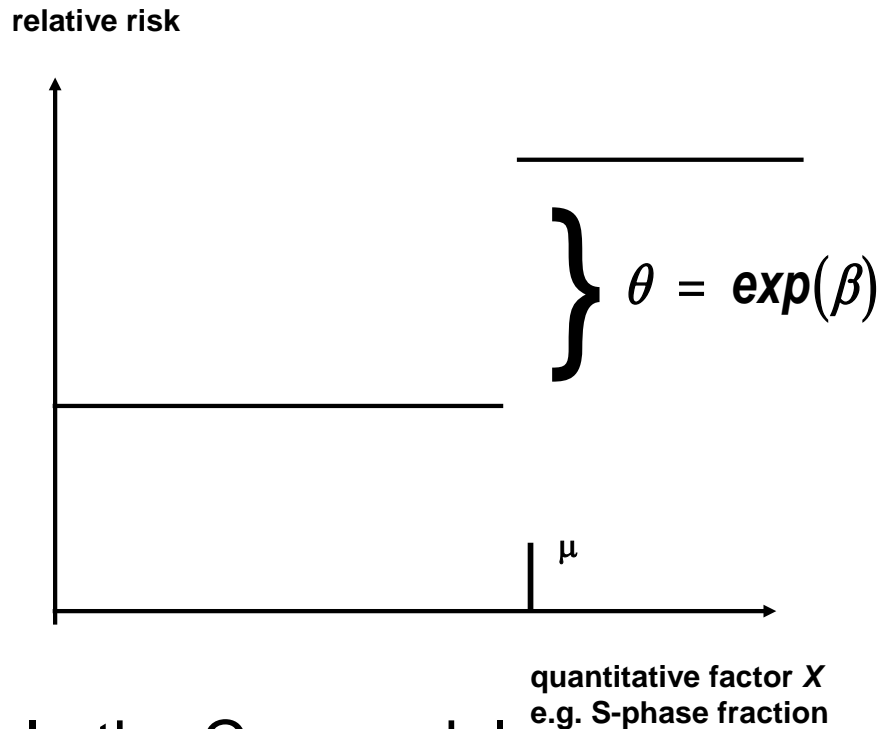
- may be inadequate functional form
- misspecification of functional form may lead to wrong conclusions

b) 'best' 'standard' transformation

c) Step function (categorical data)

- Loss of information
- How many cutpoints?
- Which cutpoints?
- Bias introduced by outcome-dependent choice

Step function - the cutpoint problem



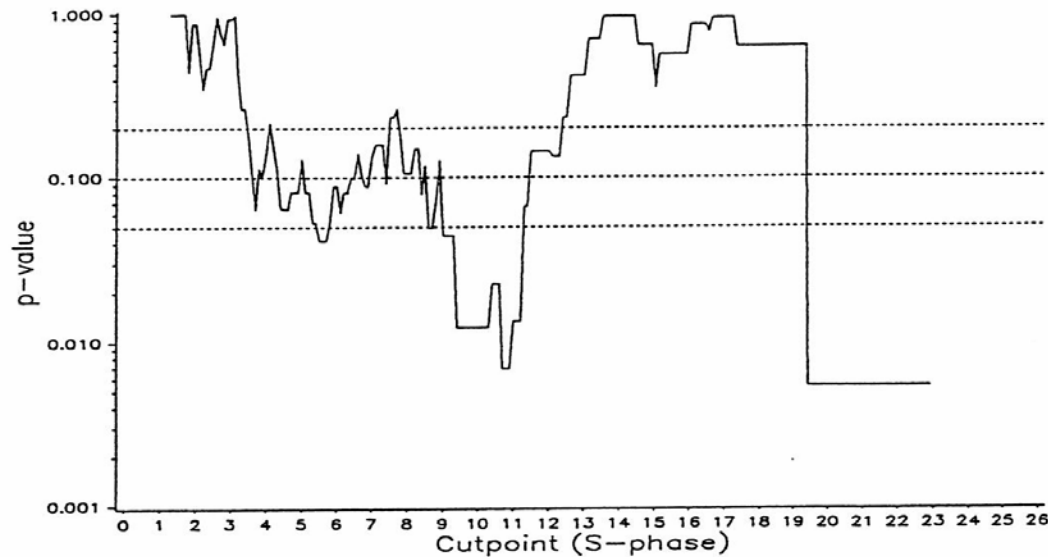
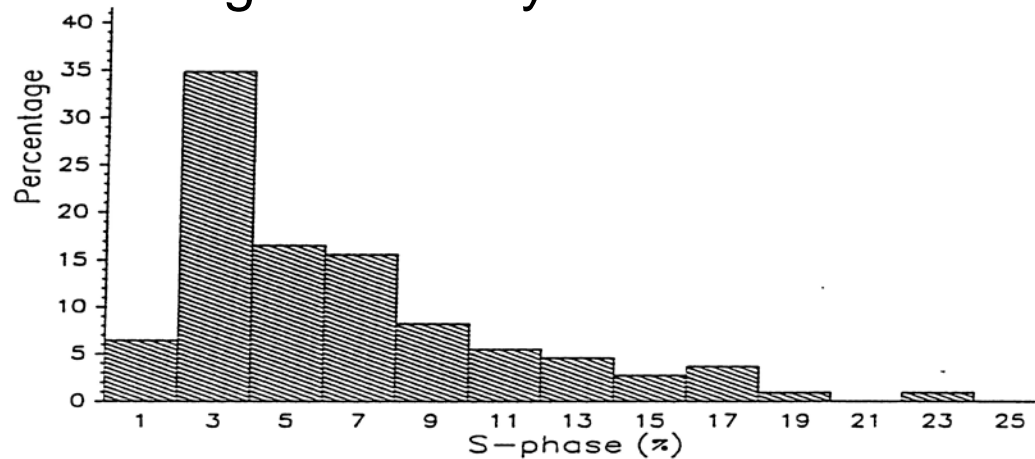
In the Cox model

$$\lambda(t|X > \mu) = \exp \beta \lambda(t|X \leq \mu)$$

$\hat{\mu}$: estimated cutpoint for the comparison
of patients with X above and below μ .

Searching for optimal cutpoint minimal p-value approach

SPF in Freiburg DNA study

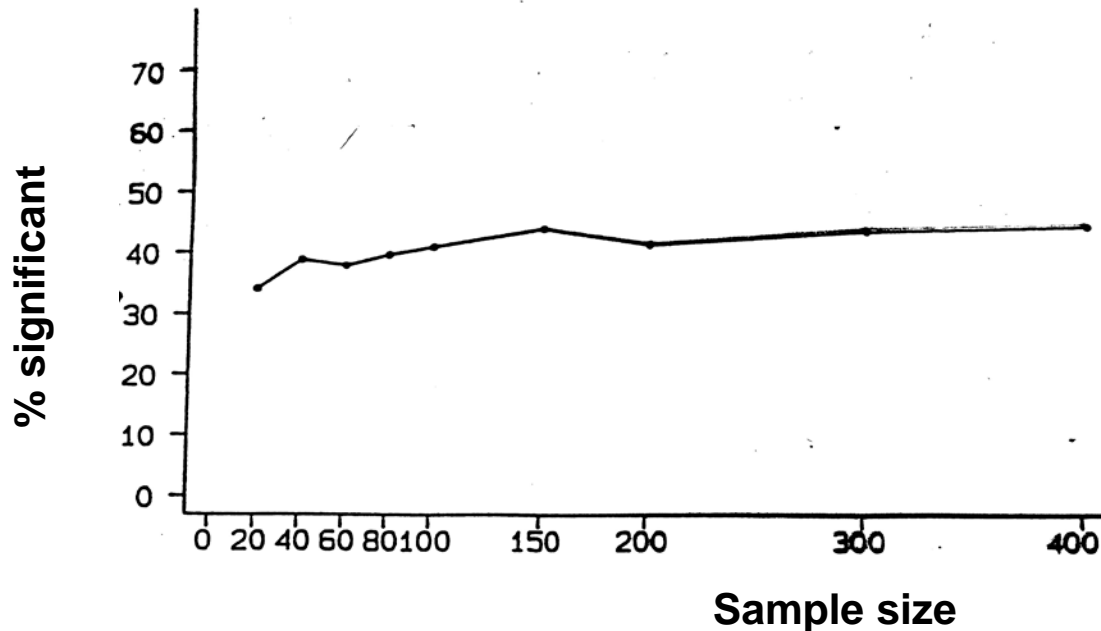


Problem

multiple testing => inflated type I error

Searching for optimal cutpoint

Inflation of type I errors



Simulation study

Type I error about 40% instead of 5%, does not disappear with increased sample size (in contrast to type II error)

Severe bias of estimated effect

Different cutpoints in each study

Step function – biologically plausible??

Example 1: Prognostic factors

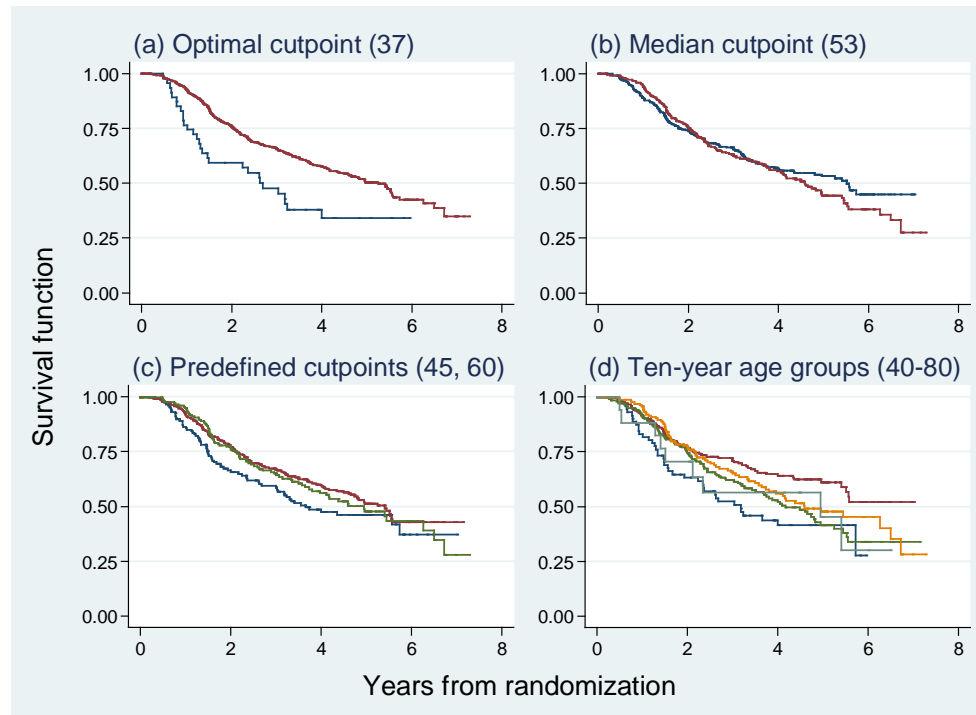
GBSG-study in node-positive breast cancer

299 events for recurrence-free survival time (RFS) in
686 patients with complete data

7 prognostic factors, of which **5** are continuous

Tamoxifen yes/no

Age as prognostic factor – cutpoint analyses



The **youngest group** is always in **blue**.

(a) 'Optimal' (37 years); HR (older vs younger) 0.54, $p = 0.004$

(b) median (53 years); HR (older vs younger) 1.1, $p = 0.4$

(c) predefined from earlier analyses (45, 60 years);

(d) popular (10-year groups)

Dichotomizing continuous predictors in multiple regression: a bad idea

Patrick Royston^{1,*†}, Douglas G. Altman² and Willi Sauerbrei³

StatMed 2006, 25:127-141

Fractional polynomial models

- *Fractional* polynomial of degree 2 with powers $p = (p_1, p_2)$ is defined as

$$FP2 = \beta_1 X^{p_1} + \beta_2 X^{p_2}$$

- Powers p are taken from a predefined set

$$S = \{-2, -1, -0.5, 0, 0.5, 1, 2, 3\}$$

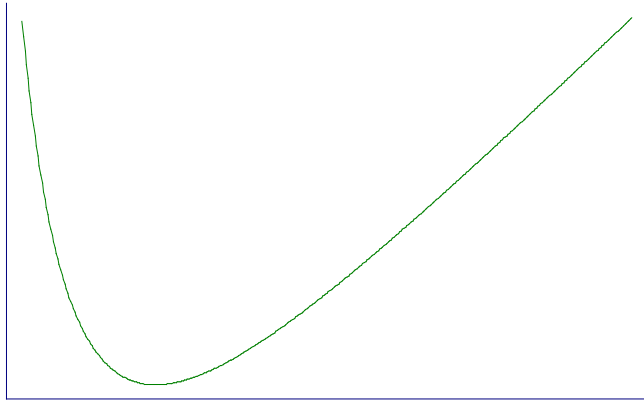
0- $\log X$

$$\text{Example } FP2 = \beta_1 X^{0.5} + \beta_2 X^3$$

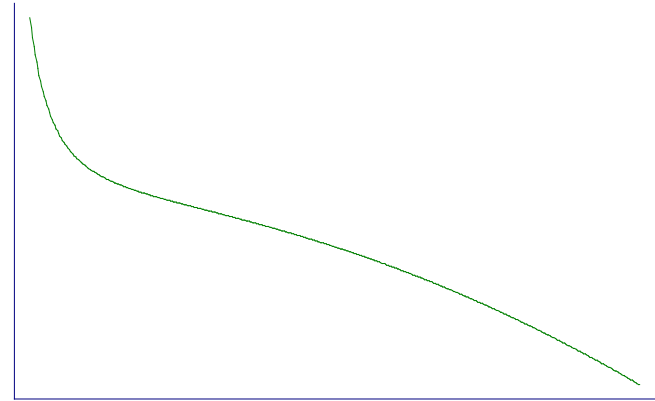


Examples of FP2 curves - varying powers

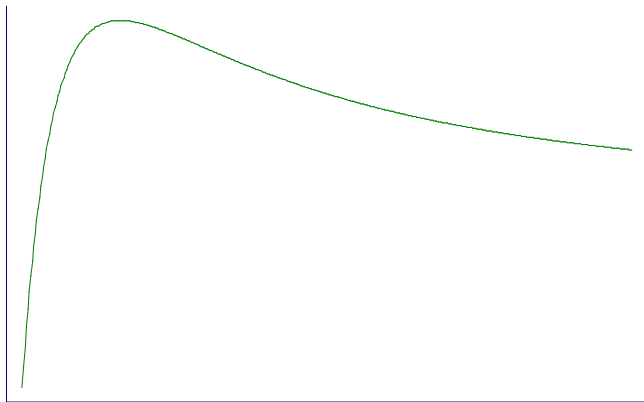
(-2, 1)



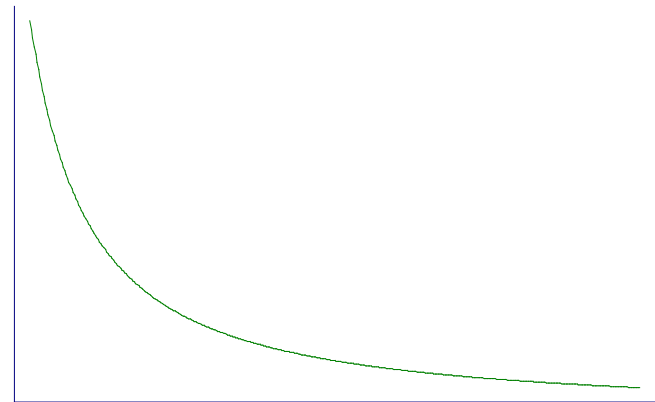
(-2, 2)



(-2, -2)

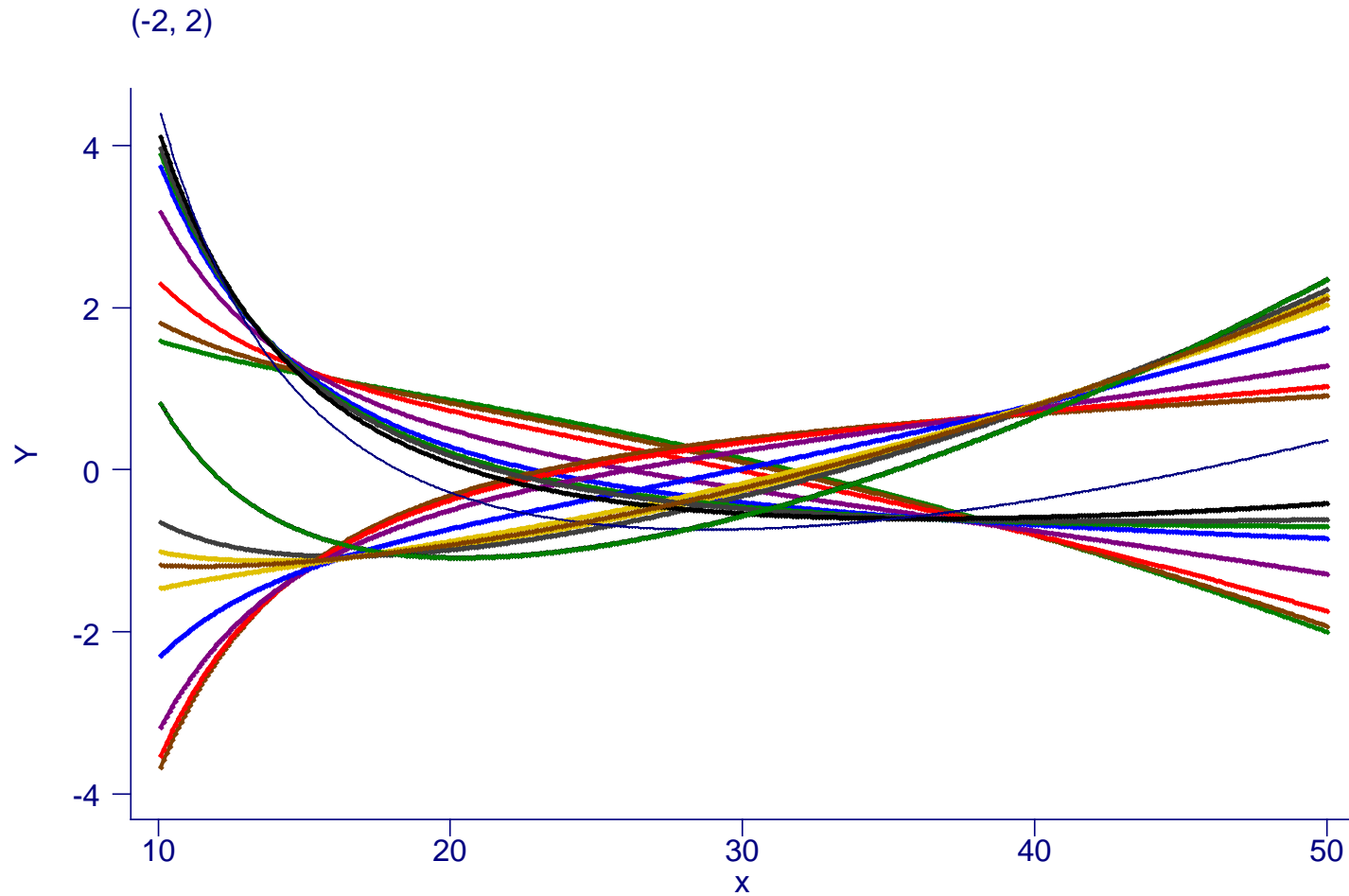


(-2, -1)



Examples of FP2 curves

- single power, different coefficients



Our philosophy of function selection

- Prefer simple (linear) model
- Use more complex (non-linear) FP1 or FP2 model if indicated by the data
- Contrasts to more local regression modelling
 - Already starts with a complex model

FP analysis for the effect of age

Degree 1		Degree 2								
Power	Model	Powers		Model	Powers		Model	Powers		Model
	chi-square			chi-square			chi-square			chi-square
<u>-2</u>	<u>6.41</u>	-2	-2	17.09	-1	1	15.56	0	2	11.45
-1	3.39	-2	-1	17.57	-1	2	13.99	0	3	9.61
-0.5	2.32	<u>-2</u>	<u>-0.5</u>	<u>17.61</u>	-1	3	12.37	0.5	0.5	13.37
0	1.53	-2	0	17.52	-0.5	-0.5	16.82	0.5	1	12.29
0.5	0.97	-2	0.5	17.30	-0.5	0	16.18	0.5	2	10.19
<u>1</u>	<u>0.58</u>	-2	1	16.97	-0.5	0.5	15.41	0.5	3	8.32
2	0.17	-2	2	16.04	-0.5	1	14.55	1	1	11.14
3	0.03	-2	3	14.91	-0.5	2	12.74	1	2	8.99
		-1	-1	17.58	-0.5	3	10.98	1	3	7.15
		-1	-0.5	17.30	0	0	15.36	2	2	6.87
		-1	0	16.85	0	0.5	14.43	2	3	5.17
		-1	0.5	16.25	0	1	13.44	3	3	3.67

Function selection procedure (FSP)

Effect of age at 5% level?

	χ^2	df	p-value
Any effect?			
Best FP2 versus null	17.61	4	0.0015
Linear function suitable?			
Best FP2 versus linear	17.03	3	0.0007
FP1 sufficient?			
Best FP2 vs. best FP1	11.20	2	0.0037

Many predictors – MFP

With many continuous predictors selection of best FP for each becomes more difficult → MFP algorithm as a standardized way to variable and function selection

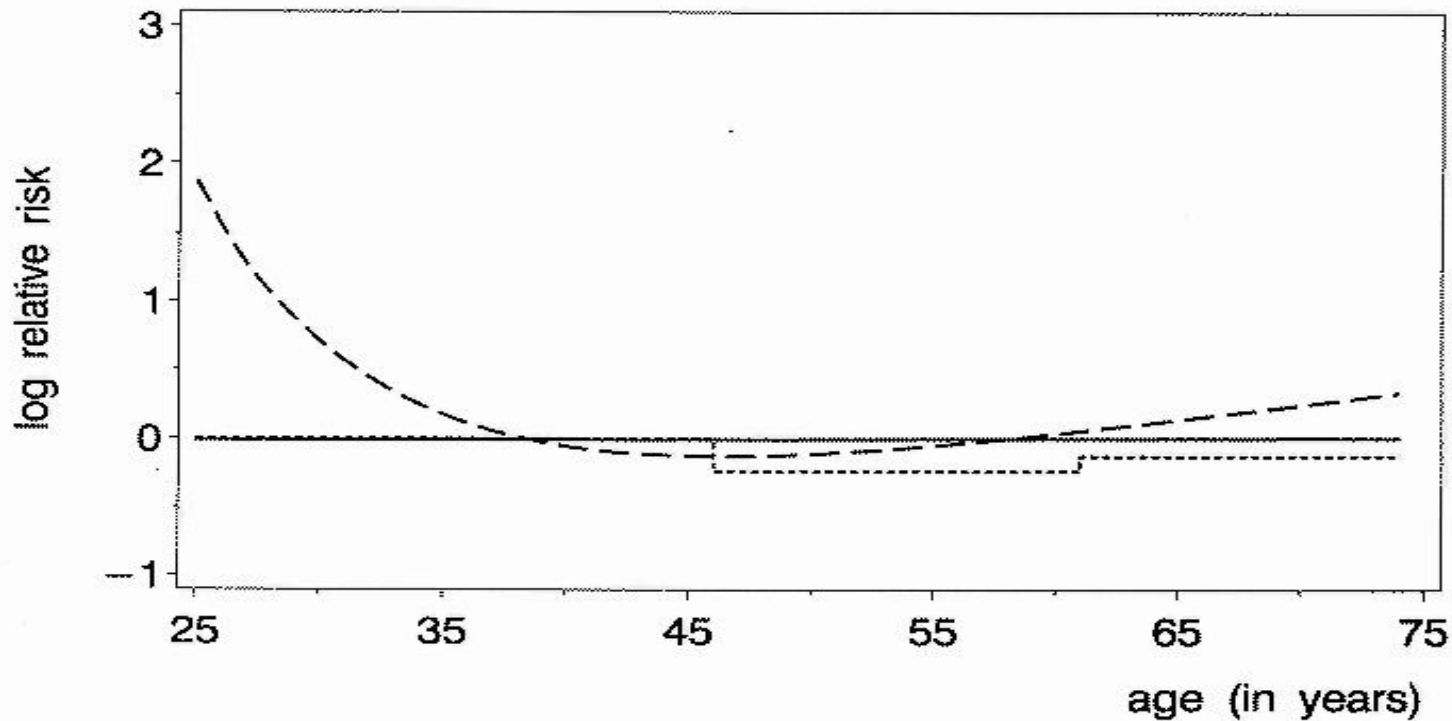
(usually binary and categorical variables are also available)

MFP algorithm combines
backward elimination with
FP function selection procedures

Continuous factors

Different results with different analyses

Age as prognostic factor in breast cancer (adjusted)



— linear function step function - - - fract. polyn.

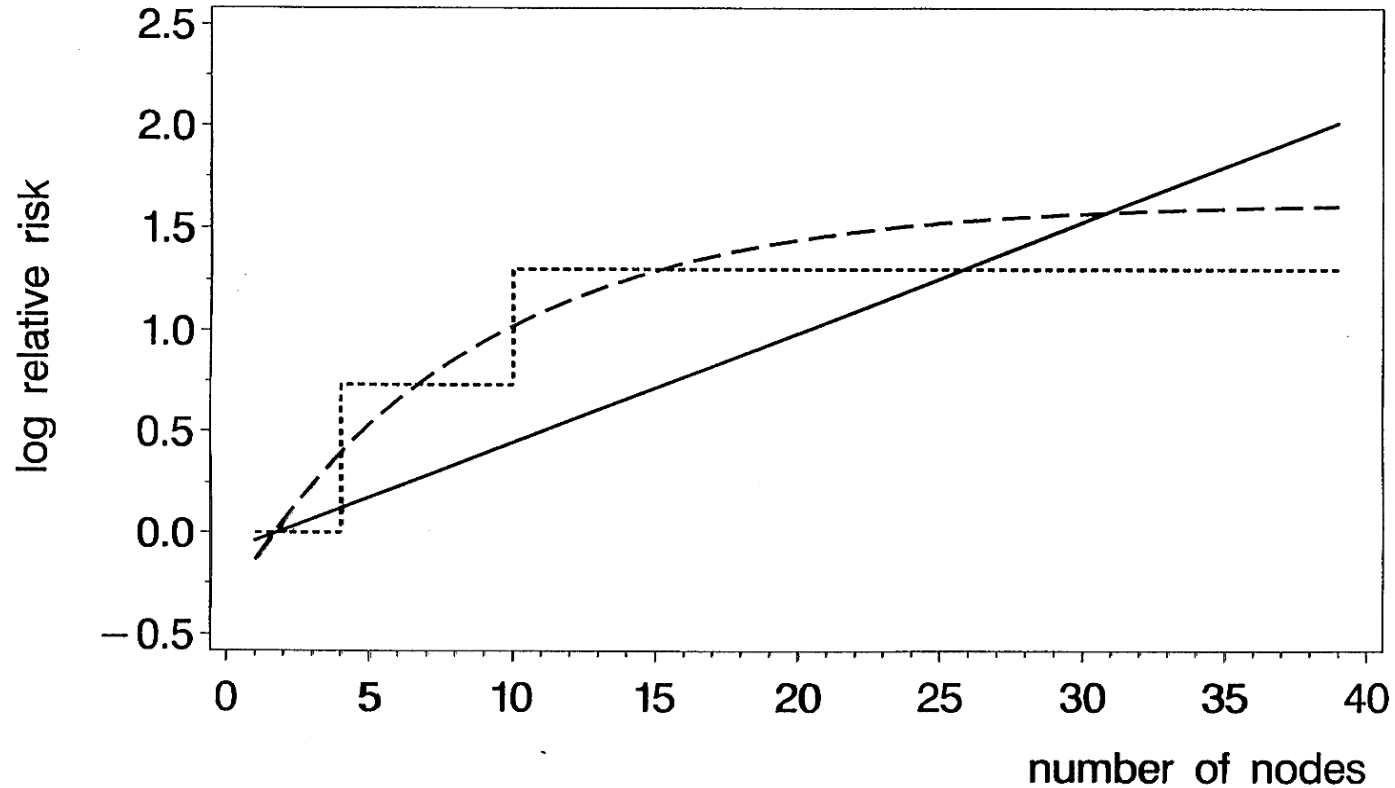
P-value 0.9

0.2

0.001

Results similar?

Nodes as prognostic factor in breast cancer (adjusted)



P-value 0.001

0.001

0.001

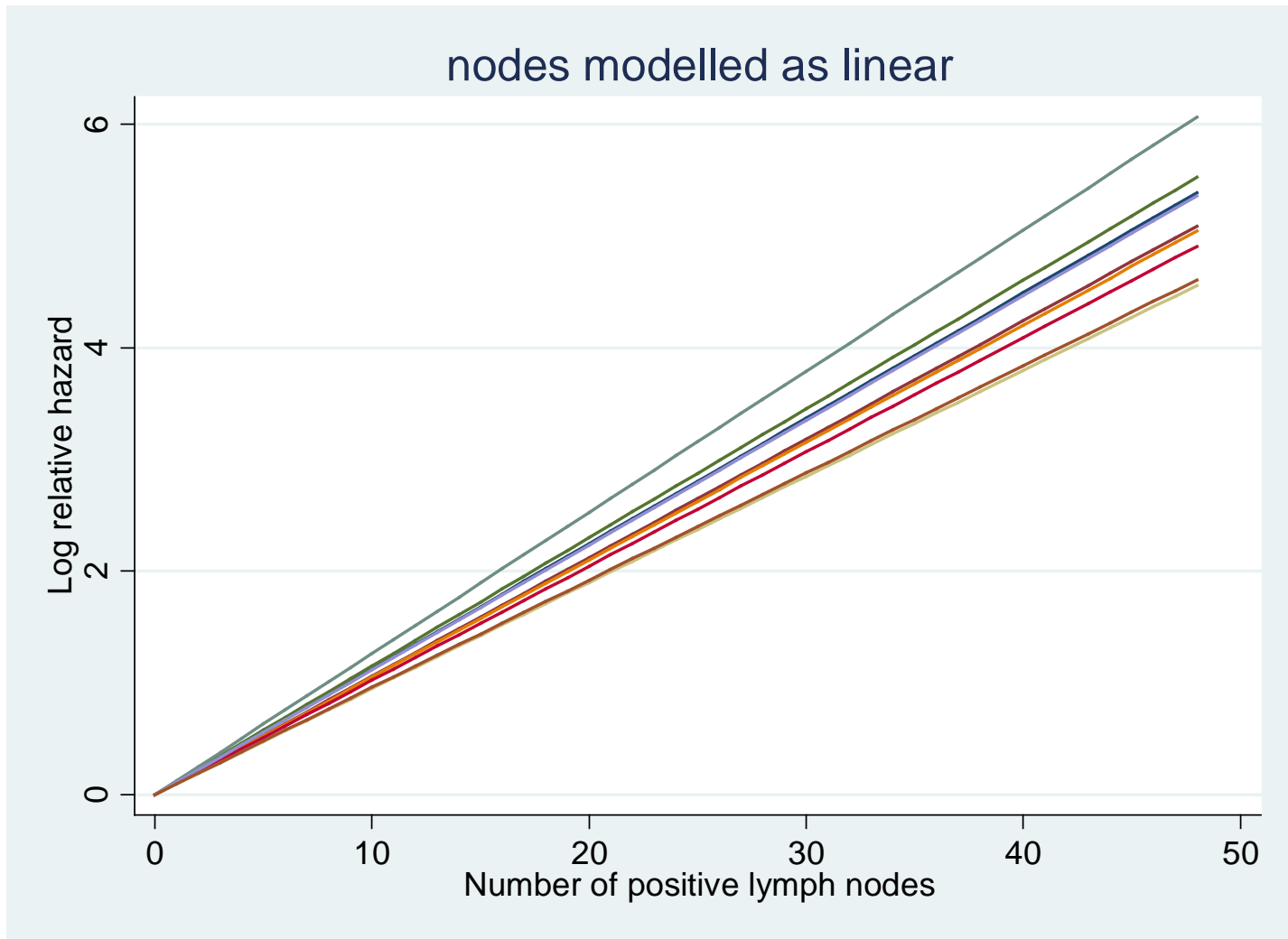
Meta-analysis: Simplest approach (for illustration)

- Univariate – not adjusted
- Assume linearity

Example with a “good” IPD dataset

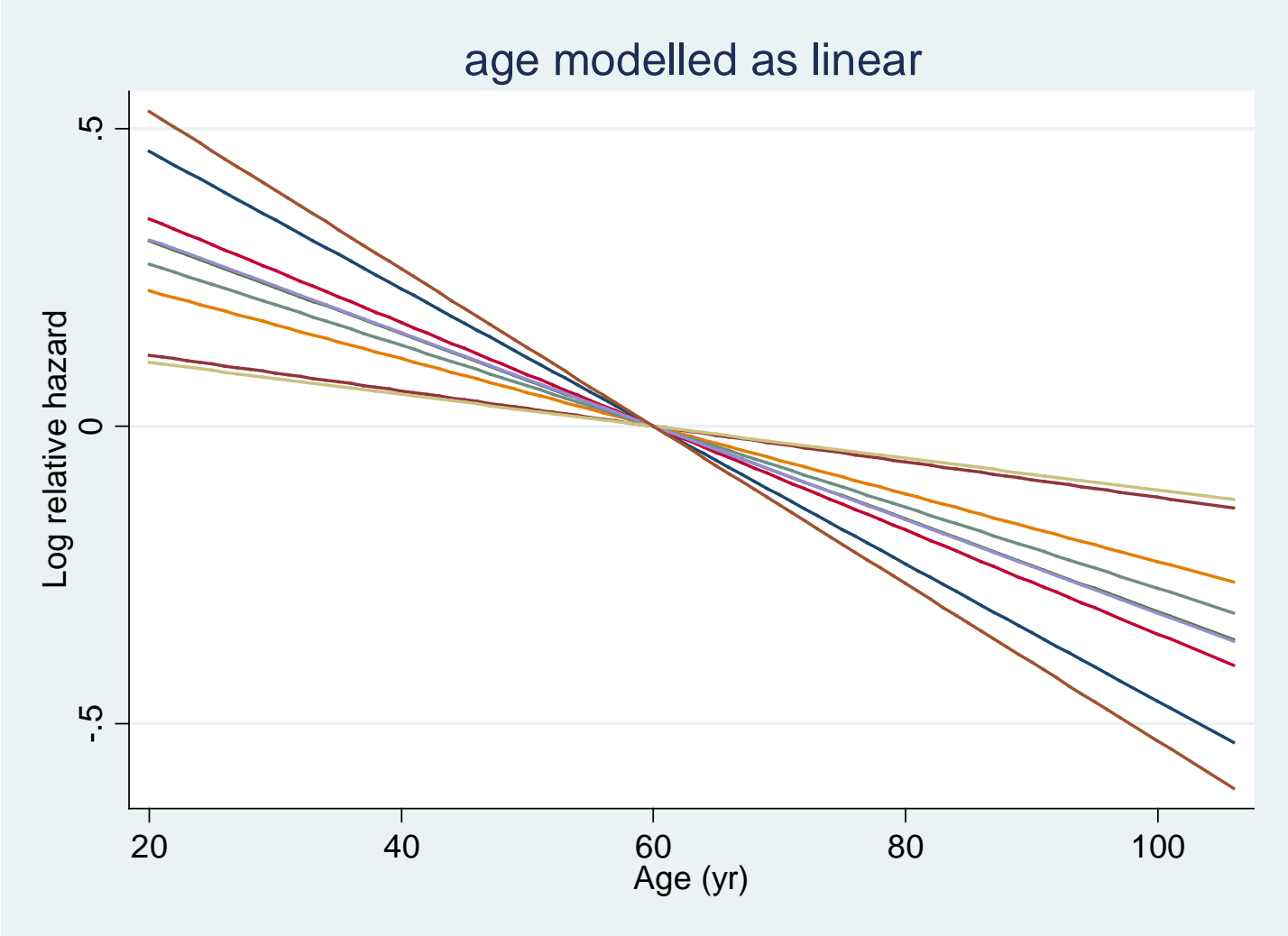
- Data: SEER registries in USA – we use the primary breast cancer database
- 9 registries; 83,804 patients; 8,099 events (deaths)
 - No registry (study) has a “small” sample size
- Time-to-event analysis using Cox regression
- Several continuous covariates of interest:
 - **nodes** (number of positive lymph nodes)
 - **age** (age at surgery for breast cancer)
 - **size** (tumour size)
- Several potential confounders
 - the above variables;
 - other clinical and demographic variables
- As examples, consider **nodes** - the strongest predictor and **age** – a weak predictor

SEER data: what do the linear **nodes** and **age** functions look like?



Node negative (0) chosen as reference point

Age modeled as linear

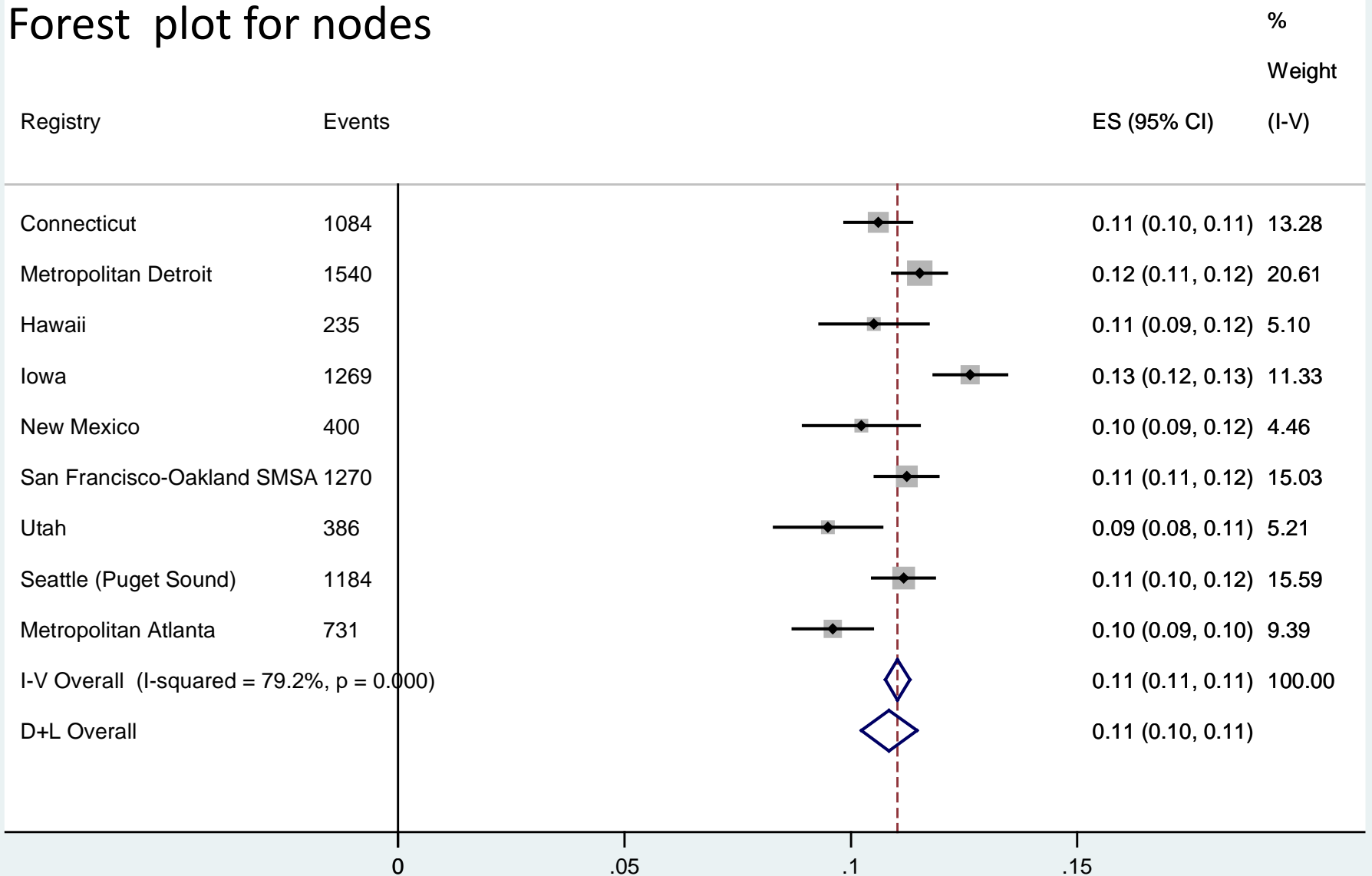


60 yrs chosen as reference point

Linear approach to MA – Cox model

- Do Cox regression on covariate in i^{th} study, estimate i^{th} slope and SE
- Get combined slope as weighted mean, and its SE

Forest plot for nodes



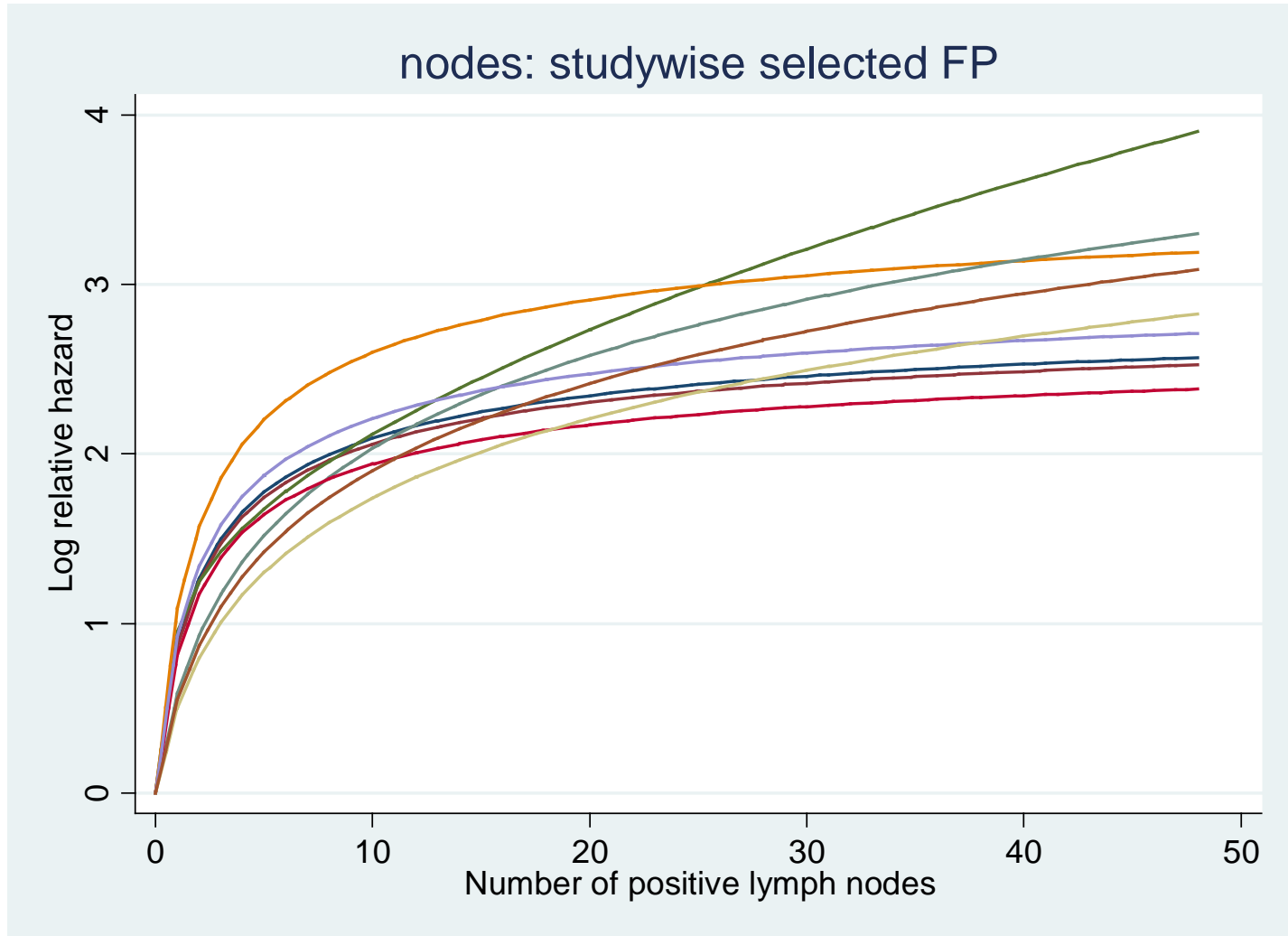
Comments on forest plot for **nodes**

- Some slope heterogeneity between “studies”
 - I-squared = 79% (see Higgins & Thompson 2002)
- Weights are not vastly different between studies – no one dataset dominates the results
- Fixed- and random-effects estimates of overall slope are similar (although CI's differ)

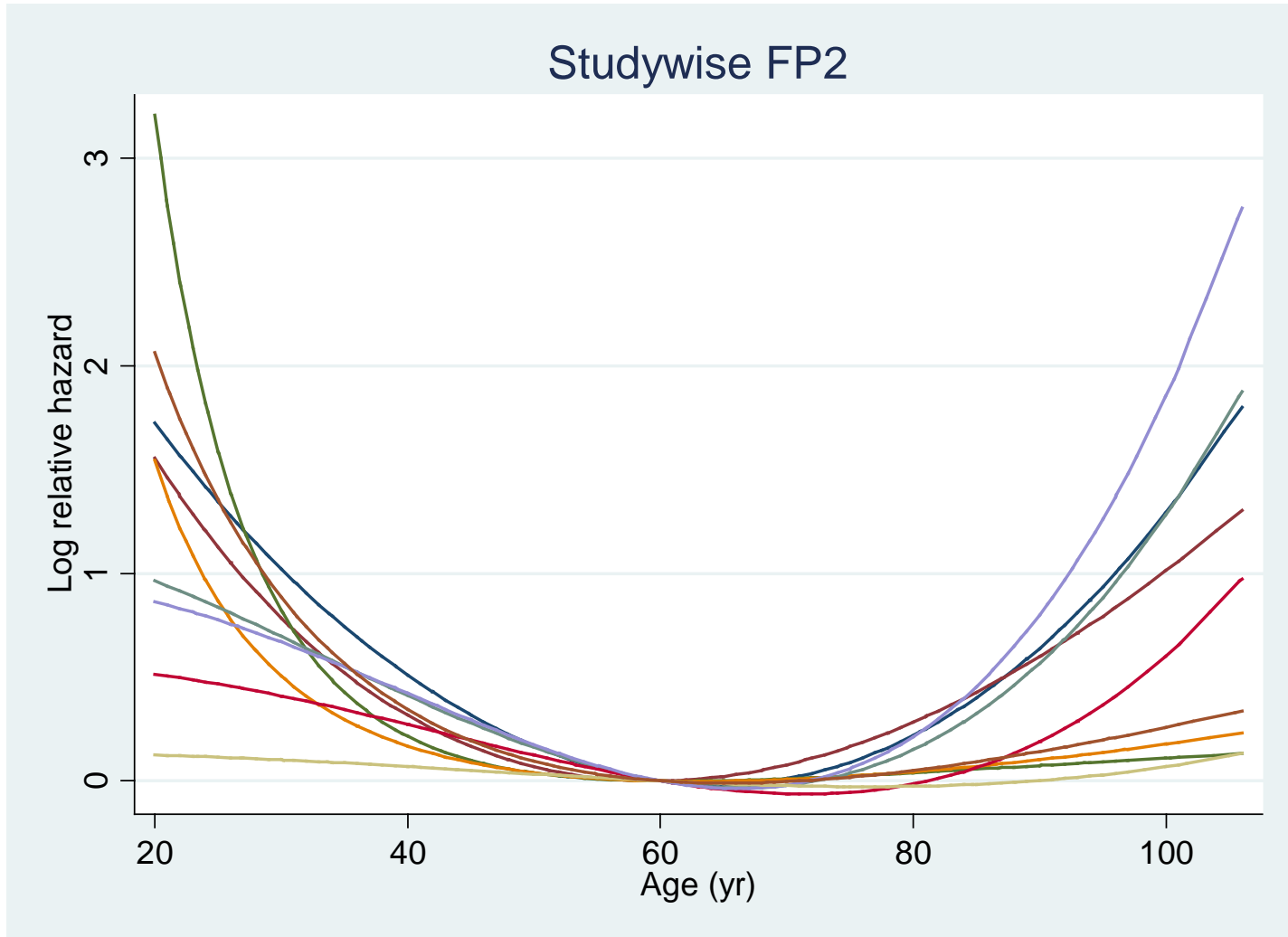
Going beyond linearity

- We have assumed functions for **age** and **nodes** are linear – is that sensible?
- If not, what do we do next?
- Relax linearity assumption – use fractional polynomials to model possible non-linearity

What does the **nodes** function “really” look like?



age (fitted by FP2 functions)



Comments

- Information on functional form within each study has clearly been lost with the linear approach
- Effect of **nodes** (strong predictor) is highly non-linear
- Effect of **age** (weak predictor) is variable but also appears non-linear
- We do need a more flexible approach to MA

Our proposal based on FPs

- Assume that the covariate of interest is **measured on the same scale** in each study
 - No problem with **nodes** and **age**
- FP approach offers a practical compromise between flexibility and stability for determining a dose/response function in a single study
 - With or without adjustment for confounders
- To obtain an overall (average) function from n studies, we suggest a **strategy with 3 main steps**

Our proposed strategy

- Step 1: Determine confounder model
 - use MFP, summarize selected model in a confounder index
- Step 2: Determine functional form
 - Adjusted for confounder index from step 1
 - 3 methods
 - Overall FP
 - Studywise FP2
 - Studywise selected FP
- Step 3: Pointwise average the fitted functions
 - Fixed or random effect weights

Results for SEER breast cancer data

- Step 1 (determine confounder model)
 - Treatment variables were included in all the confounder models
 - Somewhat different variables and FP transformations were selected across studies
- Step 2 (determine functional form)
 - Individual functions for **nodes** and **age**
- Step 3 (averaging the functions)
 - Average functions using **random-effects weights**

Step 1: Determine confounder model

- Each study can have **different potential confounders**
- Determine confounder models M_1, \dots, M_n per study
- Apply the MFP procedure* to confounders in study $i = 1, \dots, n$ to determine model M_i
 - Include “standard” confounders (e.g. age, sex, etc) in models
 - Don’t include continuous covariate of main interest
- Summarize confounder models as **index** (linear predictor) xb_i per study
- Adjust linearly for index values xb_i in further analyses

*as explained in Willi Sauerbrei’s talk

Step 2: Determine functional form

- We suggest 3 methods for selecting a suitable function of a continuous covariate in each study
- Function estimates are always adjusted for the confounder indexes xb_1, \dots, xb_n

Three methods of function selection

- 1) **Overall FP**. Find the best FP transformation for the **pooled** dataset, **stratified** by study
 - 1) To avoid over-fitting when have a large dataset, select using a small significance level
 - 2) Across studies, functions have the *same overall FP transformation* but *different estimated regression coefficients*
- 2) **Studywise FP2**. Select the **best FP2** function for each study. These may differ substantially across studies
- 3) **Studywise selected FP**. Select the **best FP** function for each study. These may also differ substantially across studies, since power may particularly be an issue

Step 3: average the fitted functions

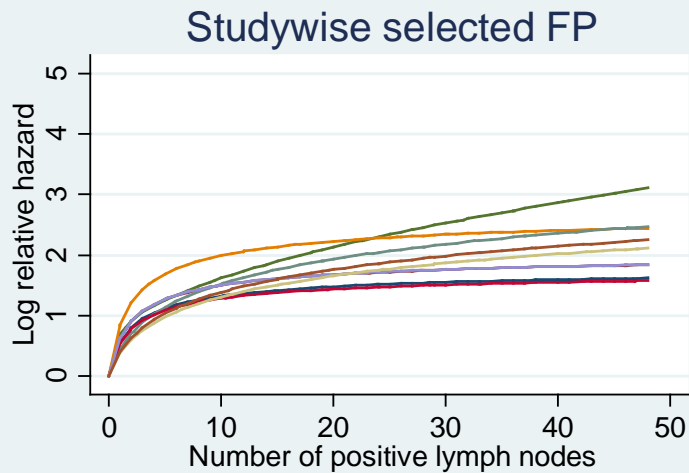
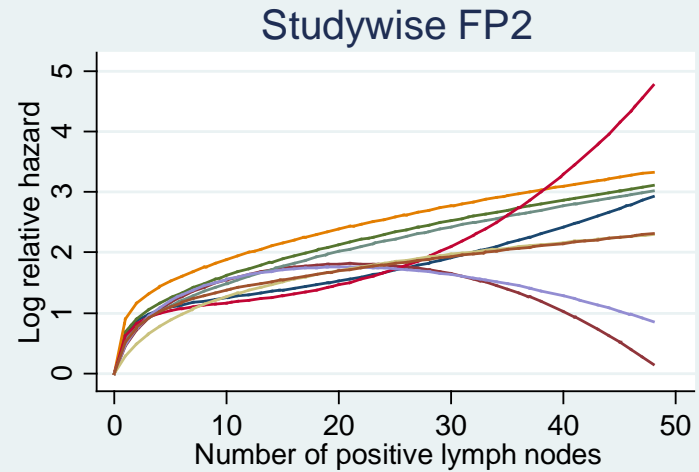
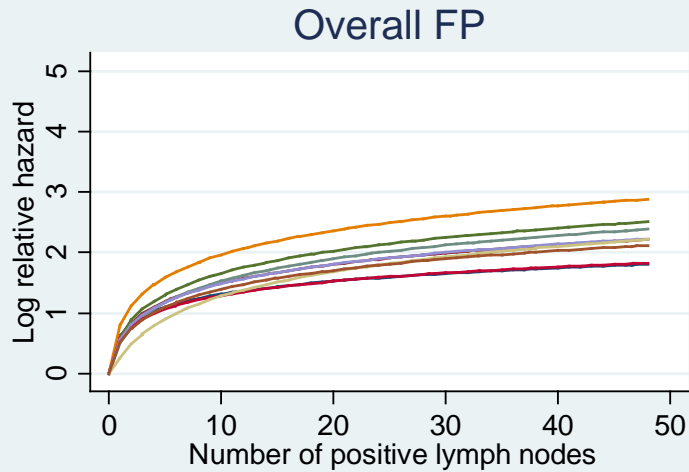
- Make an **informal assessment of heterogeneity**
 - E.g. plot of estimated functions against covariate
- Decide whether it is sensible to average functions across studies
- If it's sensible to average, we adopt standard methods for finding an average function, adjusted for the confounder indexes xb_1, \dots, xb_n
- The study-specific weights w_i are not constant any more
- Since the variance of effect size ES_i depends on Z , the **weights are also functions of Z** , i.e. $w_i = w_i(Z)$
- The approach can be used with all 3 FP function selection methods

Further details: Step 3 – averaging the functions

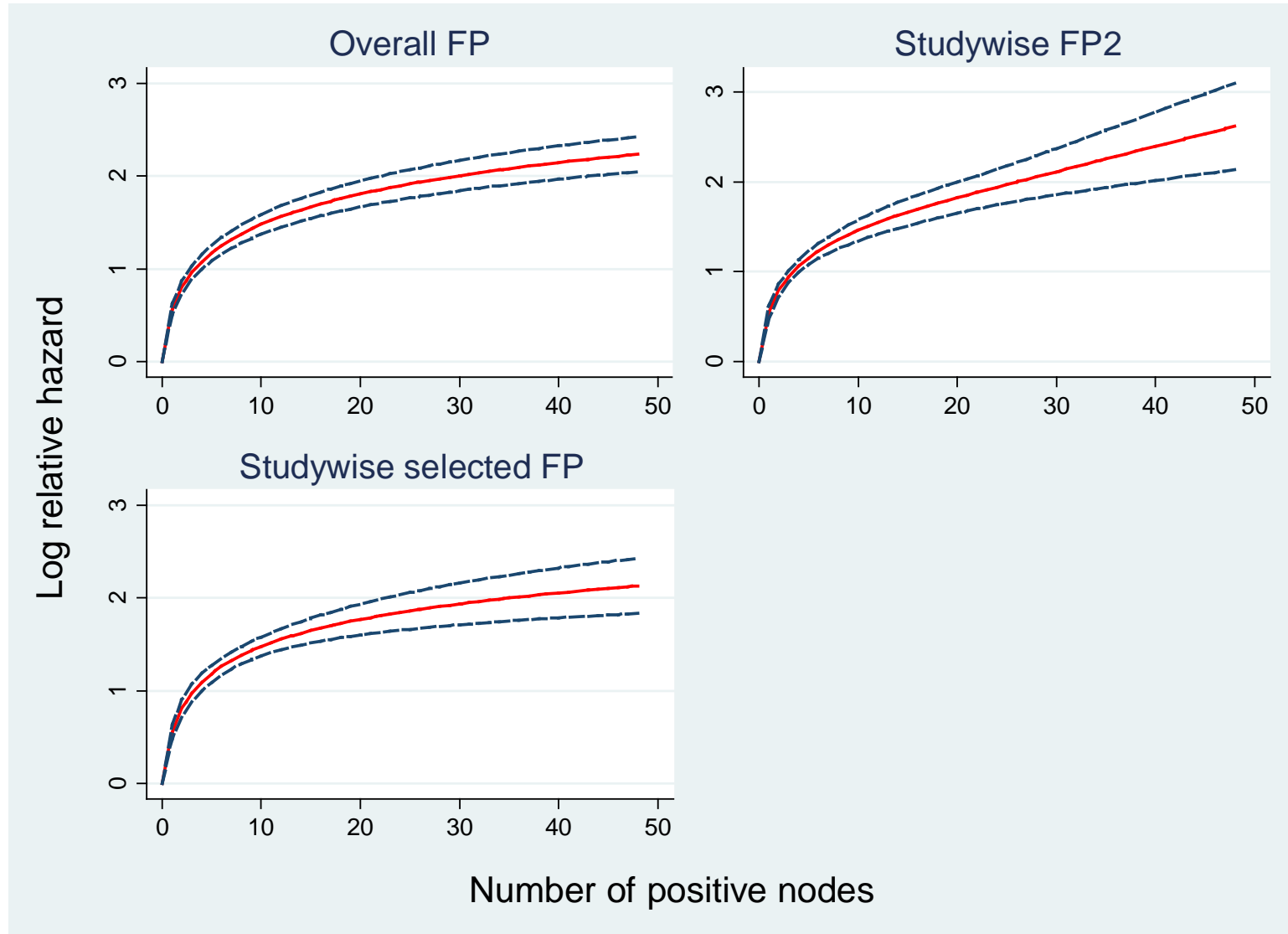
- Cox model has no intercept, so estimated function $ES_i(Z)$ in study **must first be standardized**
- Choose a suitable reference value Z_0 for all studies
- Subtract the function in each study i at that value
 - $ES_i^{\sim}(Z) = ES_i(Z) - ES_i(Z_0)$
- Weights and τ^2 are computed *pointwise* on Z
 - functions of Z – affects shape of mean function
- Compute overall mean function as pointwise weighted average of $ES_i^{\sim}(Z)$ across studies
- Pointwise fixed and random effects weights are calculated in the usual way

For **nodes**, 0 is obvious ref point; for **age**, we chose 60 yrs

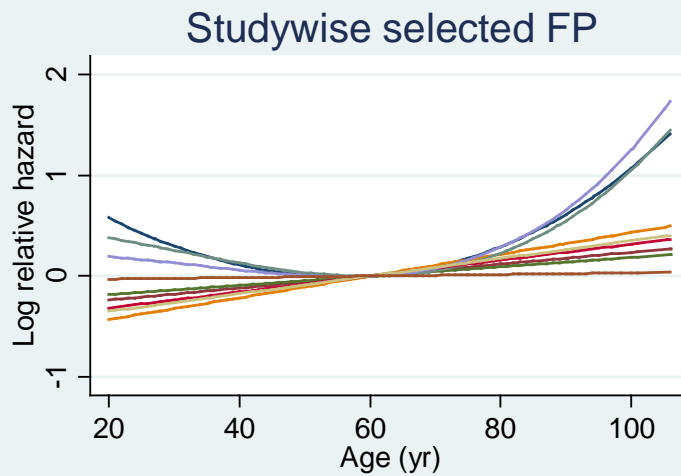
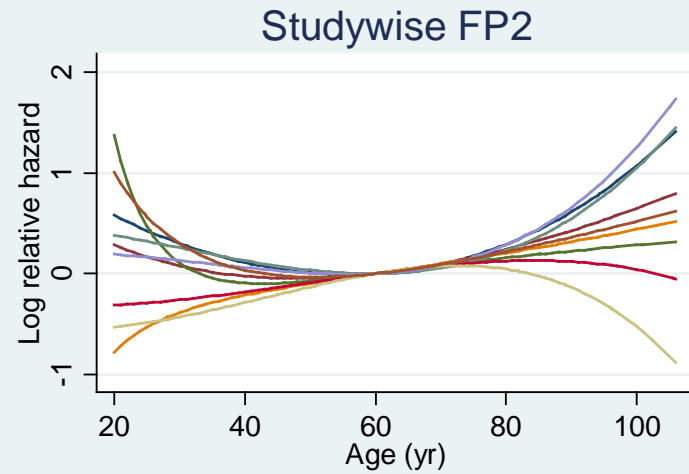
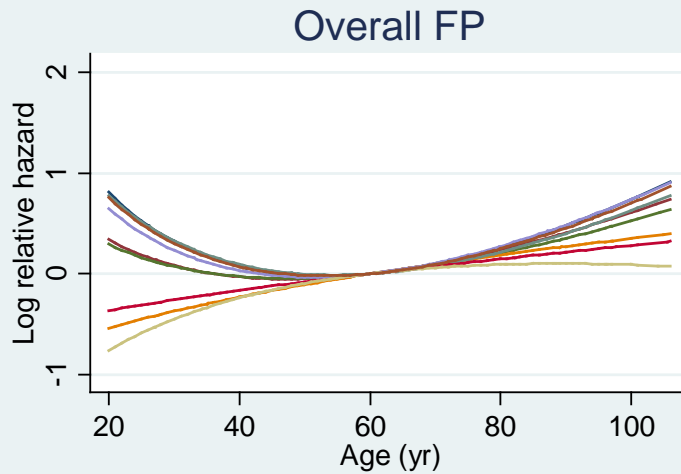
Individual functions for **nodes** [3 methods]



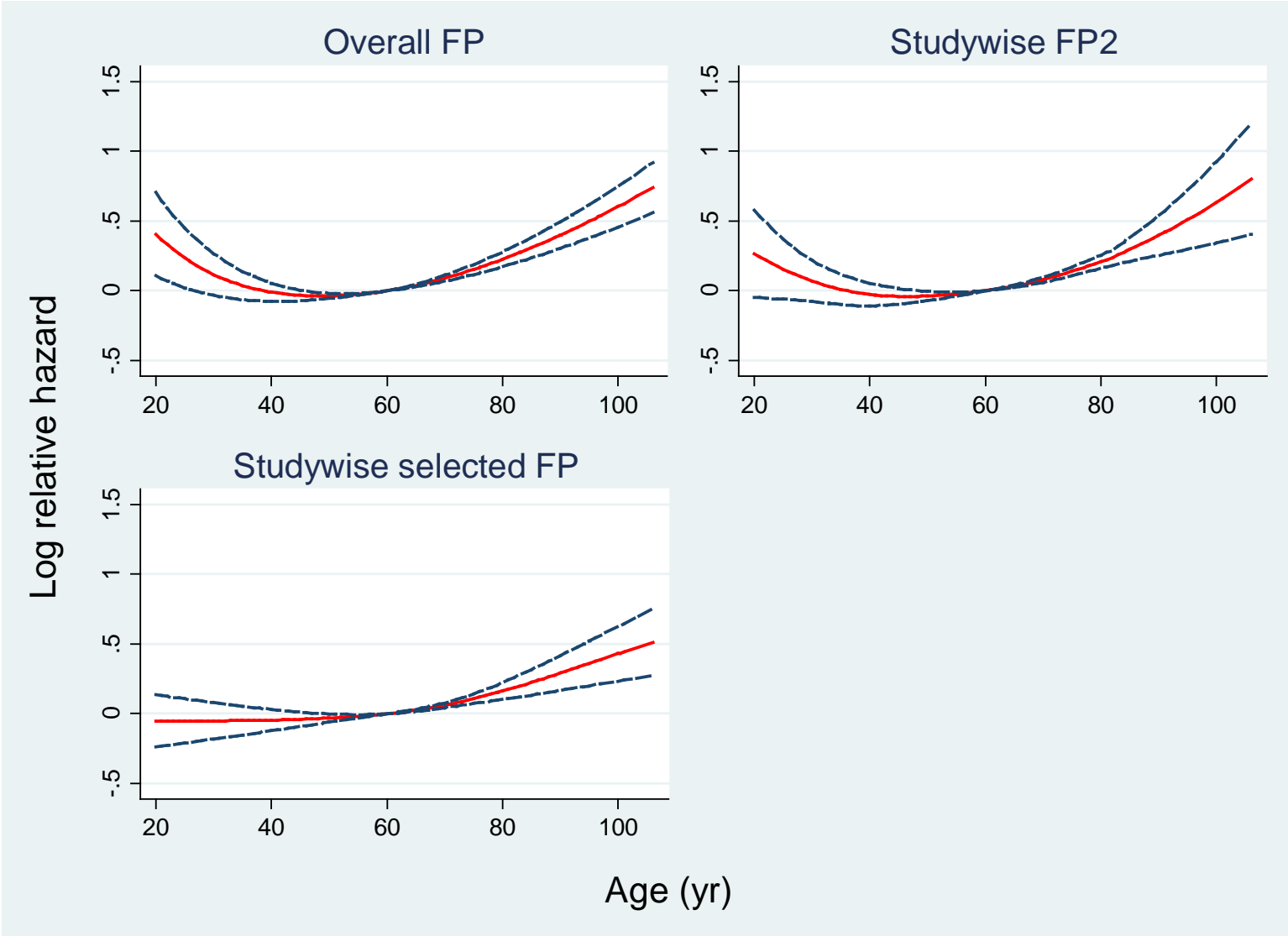
Average functions for **nodes** [RE weights]



Individual functions for **age** [3 methods]



Average functions for **age** [RE weights]



Comments on the results

- Linear approach is clearly **highly misleading**, both for **nodes** and **age**
- For age, the direction of the effect seems **reversed**
 - Linear functions suggest decreasing risk with age, whereas the FP2 functions are strongly non-monotonic

Conclusions

- Meta-analysis of continuous covariates based on **cutpoints** and dummy variables is **not sufficient**
- MA of regression coefficients for **linear functions** is **sometimes acceptable**
- When **non-linearity** is present, MA of FP functions **may offer a way forward**
- Several **critical points** need more consideration, experience and investigation

Some selected references

- Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. 2009. Introduction to meta-analysis. Wiley, Chichester.
- DerSimonian R, Laird N. 1986. Meta-analysis in clinical trials. *Controlled Clinical Trials* 7: 177-188.
- Higgins JPT, Thompson SG, Spiegelhalter DJ. 2009. A re-evaluation of random-effects meta-analysis. *JRSS(A)* 172: 137-159.
- Royston P., Altman D.G. (1994): Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling. *Applied Statistics*, 43(3): 429–467.
- Royston P, Sauerbrei W. 2008. Multivariable Model-building: A pragmatic approach to regression analysis based on fractional polynomials for modelling continuous variables. Wiley, Chichester.
- Sauerbrei W., Royston P., Binder H. (2007): Selection of important variables and determination of functional form for continuous predictors in multivariable model building. *Statistics in Medicine*, 26: 5512-5528.
- Sauerbrei W, Royston P. 2010. A new strategy for meta-analysis of continuous covariates in observational studies. *Statistics in Medicine* submitted.
- White IR. 2009. Multivariable random-effects meta-analysis. *Stata J* 9(1): 40-56