Phases of development of the Multivariable Fractional Polynomial Interaction (MFPI) procedure

Willi Sauerbrei
IMBI, Medical Center – University of Freiburg, Germany
Patrick Royston
MRC Clinical Trials Unit, London, UK

28.08.2025 – ISCB 2025 Basel

Overview

- Some background of MFPI
- Phases of development
- Further development – meta-analysis

MFPI

- MFPI is an extension of MFP to investigate for interactions of a continuous with a categorical variable

- Experiences with a binary treatment variable in RCTs – investigations with or without adjustment for other variables

Fractional polynomial models

- Conventional polynomial of degree m with powers p = (1,..., m)

$$\beta_1 X^1 + \beta_2 X^2 + \ldots + \beta_m X^m$$

- Fractional polynomial of degree m with powers p = (p_1,..., p_m)

$$FPm = \beta_1 X^{p_1} + \beta_2 X^{p_2} + \ldots + \beta_m X^{p_m}$$

- Powers p are taken from a predefined set S
- S = {-2, -1, -0.5, 0, 0.5, 1, 2, 3} 0 means log X
- 'Repeated powers' are included e.g. (-2,-2)

$$FP2(-2,-2) = \beta_1 X^{-2} + \beta_2 X^{-2} \ln X$$

➢ M = 2 is sufficient for most analyses

Example: Metastatic renal cancer

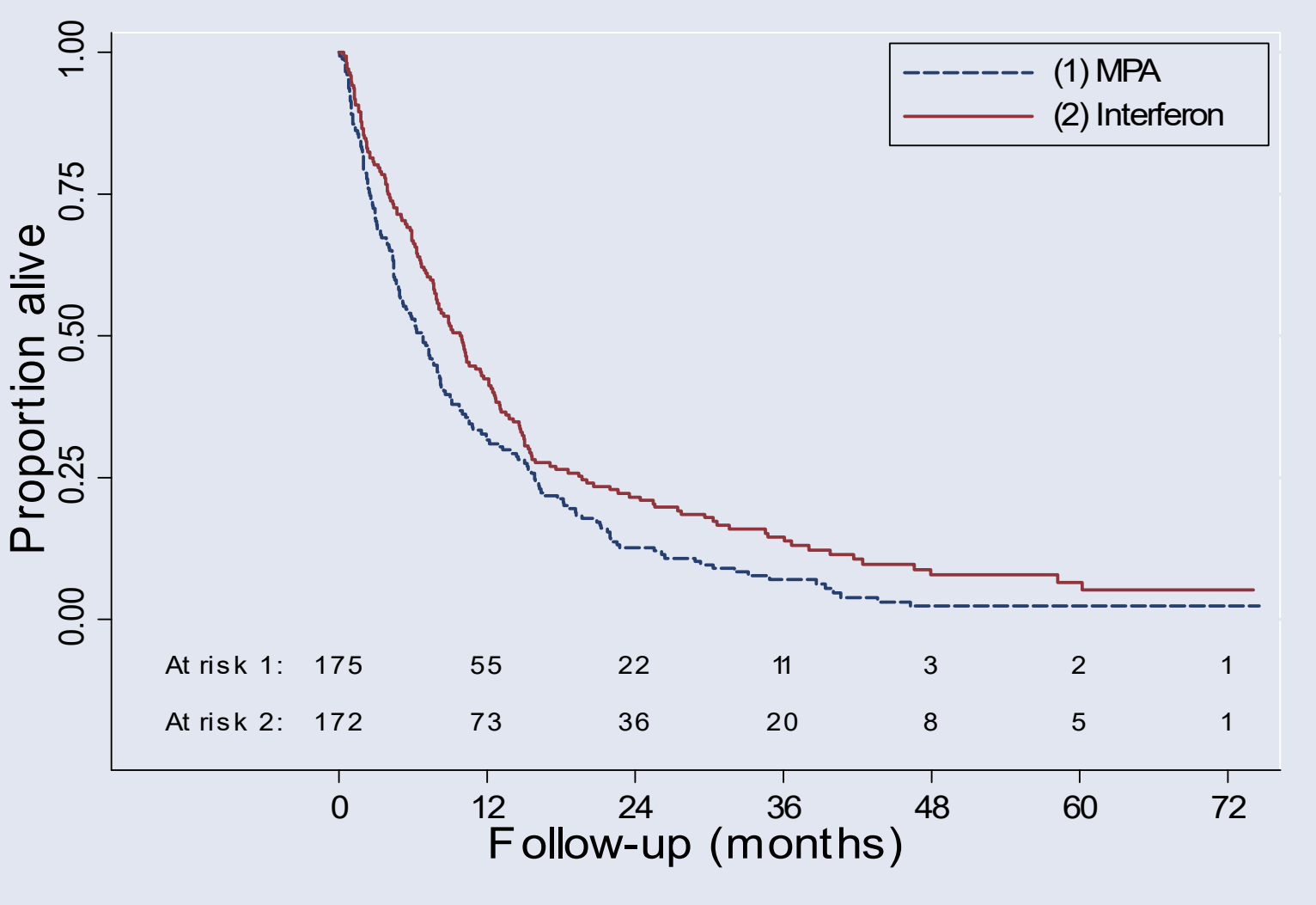
RCT in UK to compare interferon-α with MPA

N = 347, 322 Death

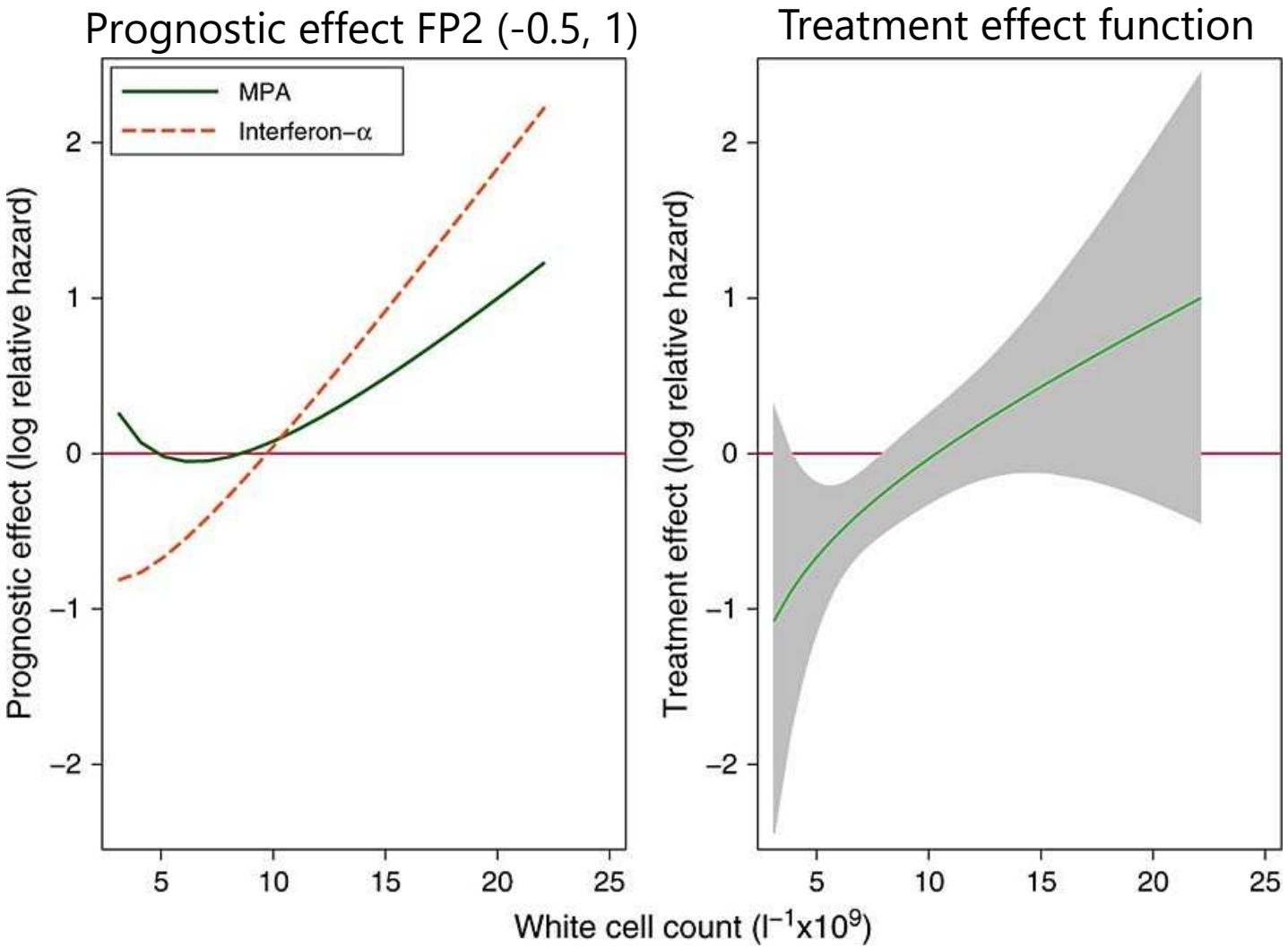
14 potential prognostic factors

Main analysis:
Interferon improves survival
HR: 0.75 (0.60 - 0.93), p = 0.009

Main effect of treatment



MFPI - Treatment Effect Function (TEF) dependent on WCC?



Prognostic effect in subgroups

TEF – difference of effect in two groups depends on WCC

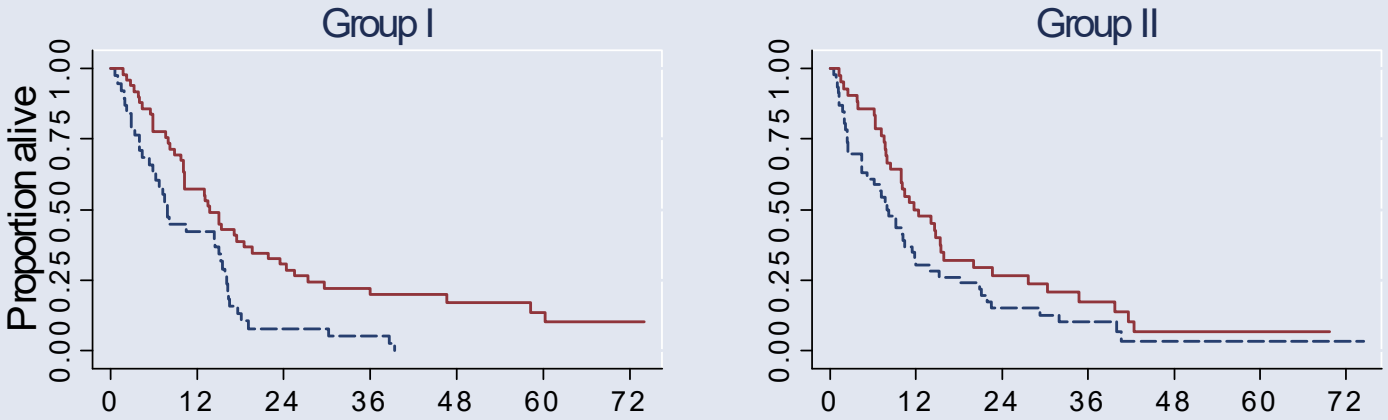
About 25% of patients
with WCC > 10 seem
not to benefit from interferon

Modelling predictive factors using fractional polynomials – the MFPI procedure

- Have one continuous factor X of interest (pre-specified hypothesis or investigate whether interactions exist)
- Find best FP2 transformation of X with same powers in each treatment group
- LRT of equality of reg coefficients
- Test against main effects model (no interaction) based on with 2df
- Modifications available

- Metastatic renal cancer
  We investigated 14 variables, only WCC was significant at 0.01

Check result of FP modelling
Treatment effect in subgroups of WCC



HR (Interferon to MPA;
overall: 0.75 (0.60 – 0.93)
adjusted values similar)

I : 0.53 (0.34 – 0.83)
II : 0.69 (0.44 – 1.07)
III : 0.89 (0.57 – 1.37)
IV : 1.32 (0.85 –2.05)

Phases of MFPI – Phase I

.. new idea ...valid or invalid from a theoretical point of view

- *All RCTs have several continuous variables. A suitable approach to investigate for interactions is needed. Dichotomization seems to be the standard*

- Royston, P., & Sauerbrei, W. (2004). A new approach to modelling interactions between treatment and continuous covariates in clinical trials by using fractional polynomials. Statistics in medicine.
- Sauerbrei, W., & Royston, P. (2007). Modelling to extract more information from clinical trials data: On some roles for the bootstrap. Statistics in Medicine.
  - *Investigation of function stability*

Phases of MFPI – Phase II

Use of methods with real data, small simulations, limited comparison with other methods

- Royston, P., Sauerbrei, W., & Ritchie, A. (2004). Is treatment with interferon-α effective in all patients with metastatic renal carcinoma? A new approach to the investigation of interactions. British journal of cancer.
  - ➢ *Check for an interaction in an RCT*
- Royston, P., & Sauerbrei, W. (2008). Multivariable model-building: a pragmatic approach to regression anaylsis based on fractional polynomials for modelling continuous variables. John Wiley & Sons. Sections: 7.4, 7.5, 7.6.
  - ➢ *More examples, present MFPI to a broader audience*

Phases of MFPI – Phase II continued

Comparison with STEPP (Subpopulation Treatment Effect Pattern Plot), which motivated MFPI

- Bonetti, M., & Gelber, R. D. (2000). A graphical method to assess treatment–covariate interactions using the Cox model on subsets of the data. Statistics in medicine.
- Bonetti, M., & Gelber, R. D. (2004). Patterns of treatment effects in subsets of patients in clinical trials. *Biostatistics*, *5*(3), 465-481.
  - *STEPP is focused on producing a non parametric estimate of the treatment effect, expressed graphically*
- Sauerbrei, W., Royston, P., & Zapien, K. (2007). Detecting an interaction between treatment and a continuous covariate: A comparison of two approaches. Computational statistics & data analysis.
  - *Comparison with STEPP, stability investigations, small simulation of type I error MFPI*
- Royston, P., & Sauerbrei, W. (2008). Interactions between treatment and continuous covariates: a step toward individualizing therapy. Journal of Clinical Oncology.
  - *Editorial STEPP vs MFPI.*
- Royston, P., & Sauerbrei, W. (2009). Two techniques for investigating interactions between treatment and continuous covariates in clinical trials. The Stata Journal.
  - *Stata programs for MFPI and STEPP described*

Phases of MFPI – Phase III

..comparison with competitors ...simulations with wide range of scenarios (ideally neutral) ..realistic comparative example data analyses, ...when can the method be used?

- Royston, P., & Sauerbrei, W. (2013). Interaction of treatment with a continuous variable: simulation study of **significance level** for several methods of analysis. Statistics in medicine.

- Royston, P., & Sauerbrei, W. (2014). Interaction of treatment with a continuous variable: simulation study of **power** for several methods of analysis. Statistics in medicine.

  - *Simulations to assess properties and compare our fractional polynomial approach (FP1 (4 flexibility), FP2 (4)) with linear, categorization, scores, and splines (altogether **13 competitors**).*
  - *Based on results, an important MFPI default from R&S 2004 was changed.*

Phases of MFPI – Phase IV

..review of existing evidence, extended simulations, when preferred method, pitfalls in analysis

- *Pitfalls – effect of influential points (see discussion in R&S 2013, 2014)*

– Schandelmaier, S., Briel, M., Varadhan, R., …, Sauerbrei, W, … Guyatt, G 2020. Development of the **Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN)** in randomized controlled trials and meta-analyses. *Cmaj*, *192*(32), pp.E901-E906.

- *9 core questions (only 4 relevant for both RCT and MA)*
- *If the effect modifier is a continuous variable, were arbitrary cutpoints avoided?*

Unfortunately, MFPI is hardly used …. R package missing

Further development – Meta-analysis of functions

Meta-analysis of (treatment effect) functions

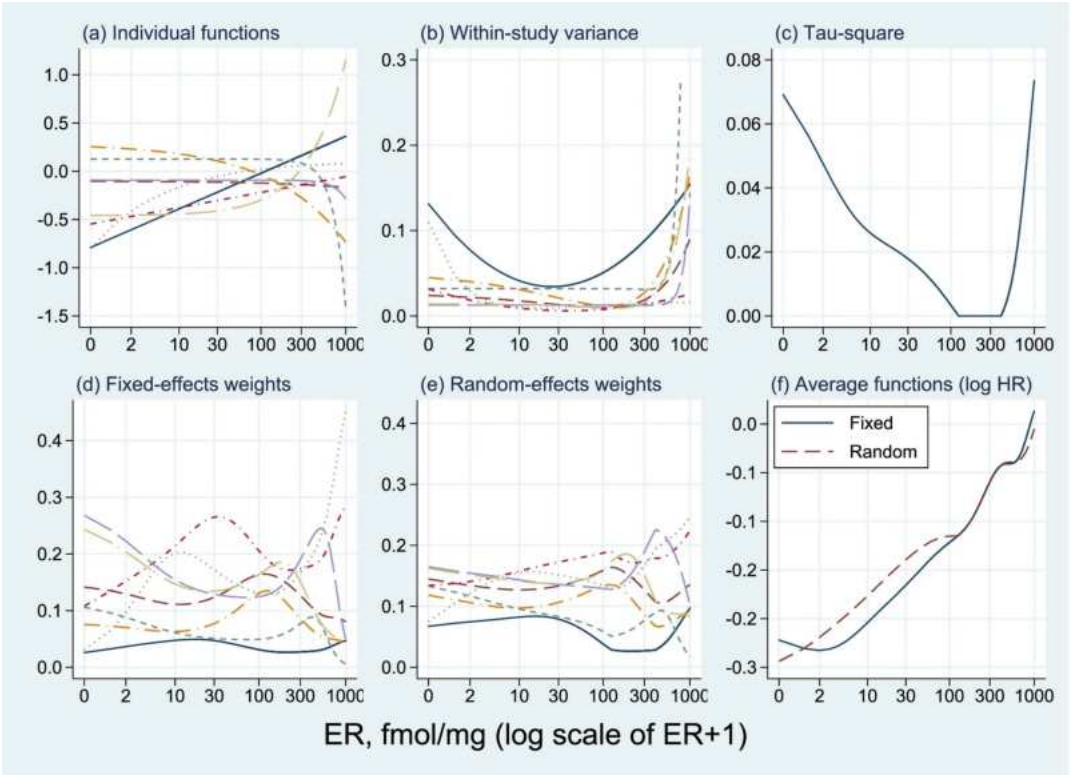
- Sauerbrei, W., & Royston, P. (2011). A new strategy for meta-analysis of continuous covariates in observational studies. Statistics in medicine.
- Kasenda, B., Sauerbrei, W., Royston, P., & Briel, M. (2014). Investigation of continuous effect modifiers in a meta-analysis on higher versus lower PEEP in patients requiring mechanical ventilation-**protocol** of the ICEM study. Systematic Reviews.
- Kasenda, B., Sauerbrei, W., Royston, P., et al. (2016). Multivariable fractional polynomial interaction to investigate continuous effect modifiers in a meta-analysis on higher versus lower PEEP for patients with ARDS. BMJ open.

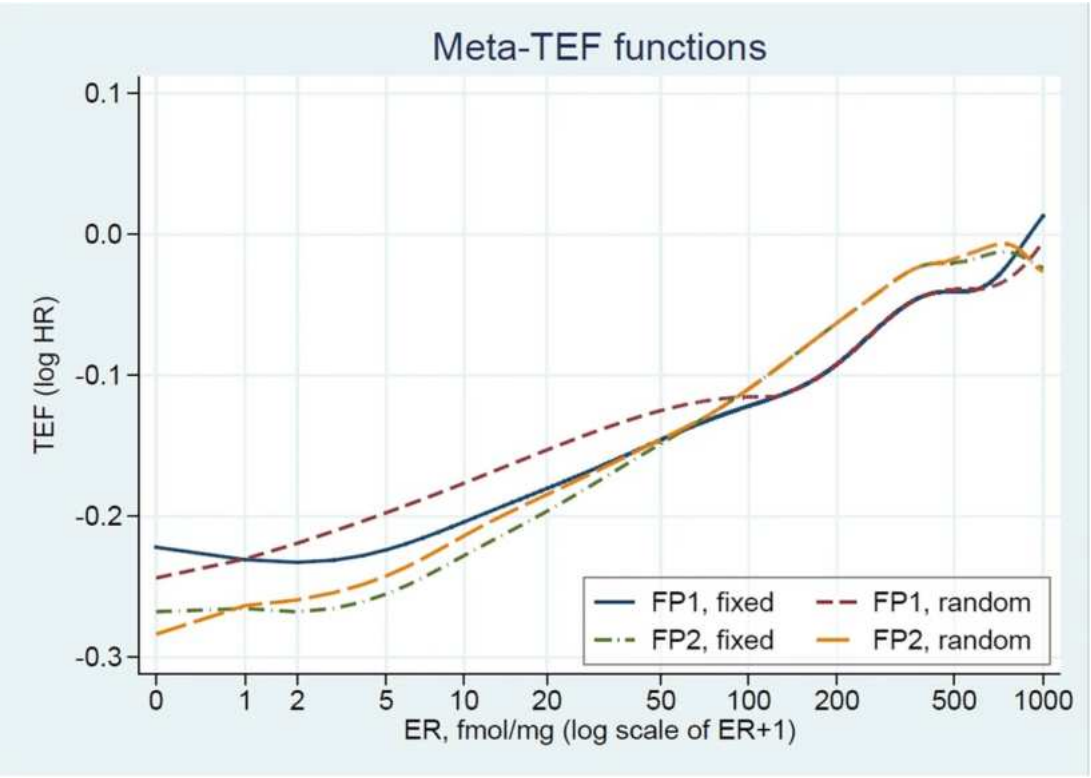
- Wang, X. V., Cole, B., Bonetti, M., & Gelber, R. D. (2016). Meta-STEPP: subpopulation treatment effect pattern plot for individual patient data meta-analysis. *Statistics in medicine*, *35*(21), 3704-3716.
- Wang, X. V., Cole, B., Bonetti, M., & Gelber, R. D. (2018). Meta-STEPP with random effects. *Research Synthesis Methods*, *9*(2), 312-317.

- Riley, R. D., Debray, T. P., Fisher, D., Hattle, M., Marlin, N., Hoogland, J., ... & Ensor, J. (2020). Individual participant data meta-analysis to examine interactions between treatment effect and participant-level covariates: statistical recommendations for conduct and planning. *Statistics in medicine*, *39*(15), 2115-2137.

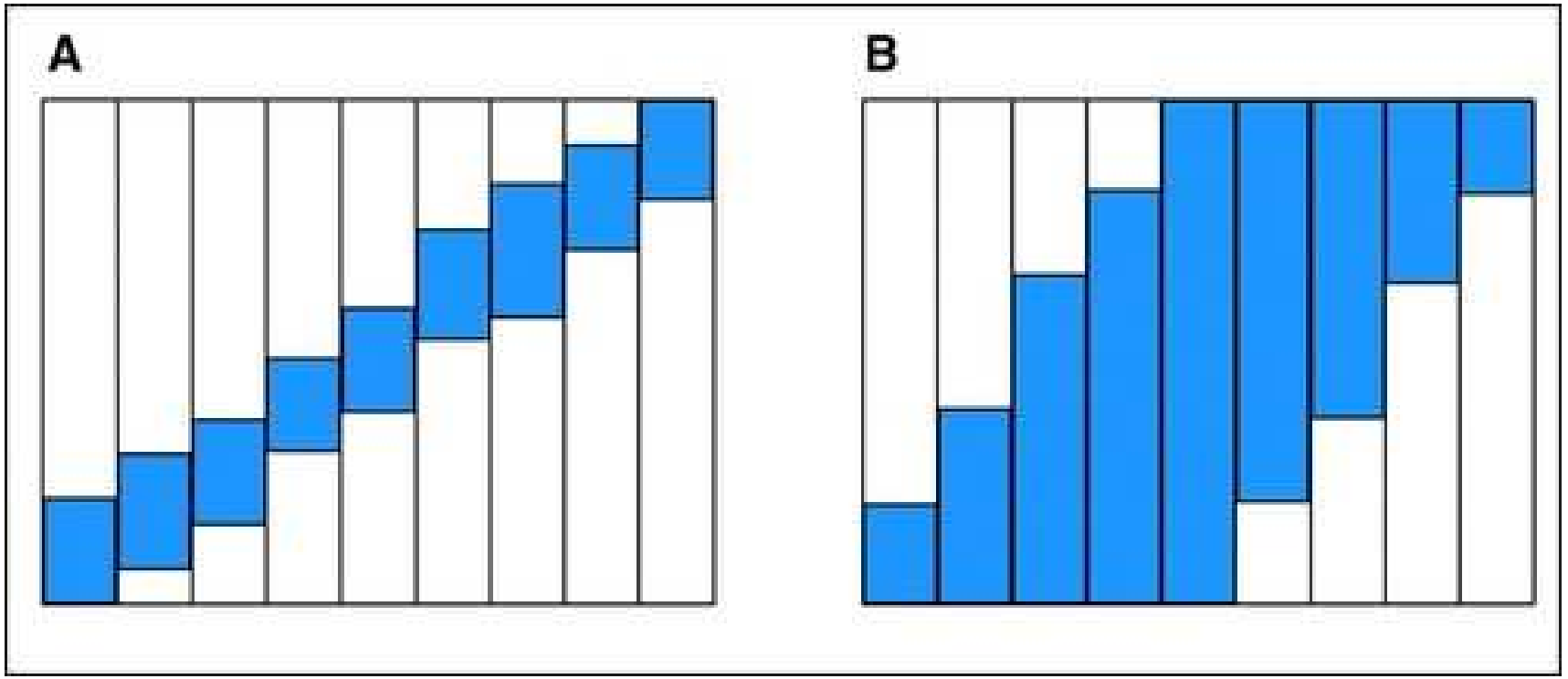
Meta-analysis of functions
…continued

- Sauerbrei, W., & Royston, P. (2022). Investigating treatment-effect modification by a continuous covariate in IPD meta-analysis: an approach using fractional polynomials. BMC medical research methodology.
  - ➢ *Illustrate various issues in an example with IPD. Eight breast cancer studies*





STEPP – analyses in subpopulations (motivated MFPI)

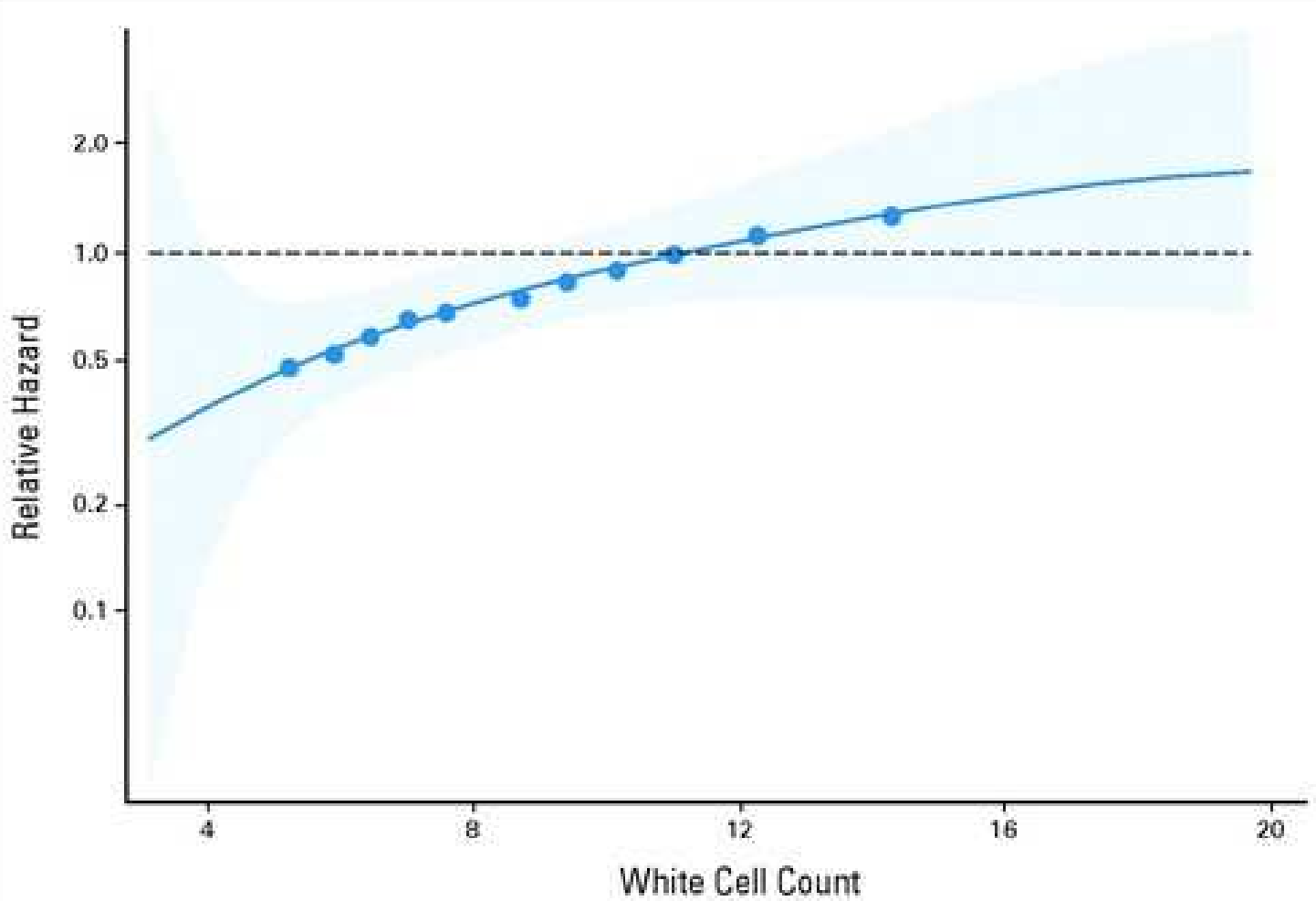


Sliding window

Tail-oriented

How many subpopulations**?**

Comparison of MFPI and STEPP



Metastatic renal cancer

MFPI treatment effect function (TEF)
STEPP tail-oriented, with 11 subgroups

Simulation study - methods

Table III. Methods of analysis (MICVs) used in the simulation studies.

MICV no.	MICV name	Class	Description
1	`lin`	Linear	Linear function at each level of t
2	`cat2`	Categorical	Two equal classes (one dummy variable)
3	`cat3a`	Categorical	Three equal classes (two dummy variables)
4	`cat3b`	Categorical	Three unequal classes ('Cox' cut-points: 27 and 73 centiles)
5	`cat4a`	Categorical	Four equal classes (three dummy variables)
6	`cat4b`	Categorical	Four unequal classes ('Cox' cut-points: 16.3, 50, and 83.7 centiles)
7	`score3a`	Categorical	Linear on `cat3a` scores
8	`score3b`	Categorical	Linear on `cat3b` scores
9	`score4a`	Categorical	Linear on `cat4a` scores
10	`score4b`	Categorical	Linear on `cat4b` scores
11	`fp1`	FP	FP1 function at each level of t (with four levels of flexibility)
12	`fp2`	FP	FP2 function at each level of t (with four levels of flexibility)
13	`spline`	Splines	Regression splines with 2, 3, or 4 DOF; automatic knot placement

See text for details of terminology.
FP, fractional polynomial; MICV, method of investigating interactions with continuous variables.

Design of the simulation study - functions

	Group 0 ($t = 0$)	Group 1 ($t = 1$)	Main-effect model (no interaction)	
Scenario	$f(x, 0)$	$f(x, 1)$	$[f(x, 0) + f(x, 1)]/2$	Type
1	0	$0.25x$	$0.125x$	Linear
2	0	$0.1 \ln x$	$0.05 \ln x$	FP1
3	x	$0.6x$	$0.8x$	Linear
4	$\ln x$	$0.75 \ln x$	$0.875 \ln x$	FP1
5	0	$-0.1x^3$	$-0.05x^3$	FP1
6	0	$0.25x^{-2}$	$0.125x^{-2}$	FP1
7	$0.05x^{-1} + 0.15x^2$	$x^{0.5}$	$0.025x^{-1} + 0.075x^2 + 0.5x^{0.5}$	FP3
8	0	$0.075x^{-2} + 0.02x^3$	$0.0375x^{-2} + 0.01x^3$	FP2
9	$0.05x^2$	$\Phi[(x - 2)/0.6]$	$0.025x^2 + 0.5\Phi[(x - 2)/0.6]$	—

Simulation study – well and badly behaved cases

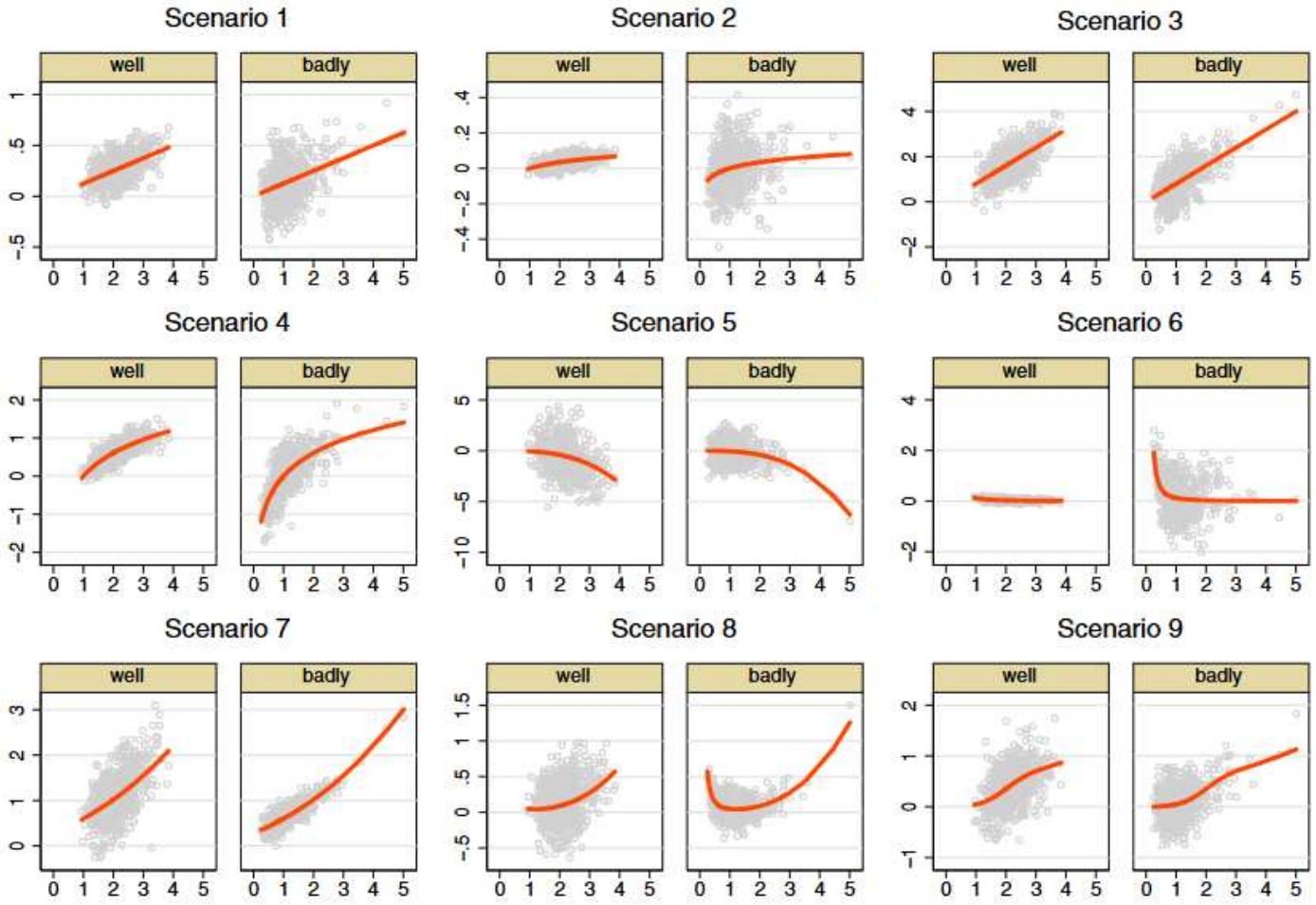


Figure 3. True functions and simulated y (one replication each) for the well-behaved and badly behaved cases in the nine scenarios.

True functions and simulated y (one replication)

Significance level – well and badly behaved data

Table IV. Mean significance levels (%) for the well-behaved case.				for the badly behaved case.	
MICV		Sample size		Sample size	
No.	Name	250	500	250	500
1	lin	6:0	5:6	9.9	10.4
2	cat2	5:7	5:5	5.4	5.2
3	cat3a	5:1	5:1	5.6	5.1
4	cat3b	5:7	5:3	6.0	5.5
5	cat4a	5:5	5:1	6.3	5.6
6	cat4b	6:1	5:5	6.4	6.5
7	score3a	5:9	5:0	6.0	5.3
8	score3b	5:5	5:3	6.2	5.8
9	score4a	5:9	5:4	6.2	5.7
10	score4b	6:1	5:5	6.2	6.2
11a	fp1(flex1)	5:1	5:0	5.1	5.5
11b	fp1(flex2)	6:6	5:4	6.6	5.9
11c	fp1(flex3)	5:5	5:4	5.7	5.3
11d	fp1(flex4)	5:0	4:4	4.8	4.2
12a	fp2(flex1)	4:6	4:4	4.5	4.4
12b	fp2(flex2)	9:0	8:3	10.3	9.7
12c	fp2(flex3)	6:4	6:5	7.8	7.1
12d	fp2(flex4)	2:0	2:1	2.7	3.0
13a	spline(2 DOF)	5:6	5:3	6.7	7.7
13b	spline(3 DOF)	5:7	5:5	5.9	6.0
13c	spline(4 DOF)	5:9	5:4	6.1	6.1

Values are averages for each sample size: for all 9 scenarios for scenarios 1-6

<4.5

>6.5

>7.5

>8.5

Significance level too large:

1(lin)

12 b,c (FP2 with flex2, flex3)

13a (spline, 2df)

Lin remains candidate in FP1

Simulation study – Power - well behaved scenarios

significance level is unacceptable, no candidate for a sensible strategy

Table III. Summary of power results for Category 1 (well-behaved x, all nine scenarios). For each sample size, the mean of the nine scenarios and also the mean for the two sample sizes are shown.

MICV		Sample size			MICV		Sample size		
no.	Name	250	500	Average	no.	Name	250	500	Average
1	lin[a]	61	88	74.5	11a	fp1 (flex1)	61	88	74.5
2	cat2	40	69	54.5	11b	fp1 (flex2)	66	90	78
3	cat3a	41	71	56	11c	fp1 (flex3)	64	90	77
4	cat3b	43	73	58	11d	fp1 (flex4)	53	85	69
5	cat4a	40	70	55	12a	fp2 (flex1)	53	84	68.5
6	cat4b	42	74	58	12b	fp2 (flex2)[a]	61	88	74.9
7	score3a	49	79	64	12c	fp2 (flex3)[a]	58	87	72.5
8	score3b	51	80	65.5	12d	fp2 (flex4)[a]	35	71	53
9	score4a	53	82	67.5	13a	spline (2 d.f.)[a]	56	86	71
10	score4b	54	83	68.5	13b	spline (3 d.f.)	49	82	65.5
					13c	spline (4 d.f.)	45	78	61.5

Simulation study – Power - badly behaved scenarios

significance level is unacceptable, no candidate for a sensible strategy

Table IV. Summary of power for Category 2 (badly behaved x, six 'typical' scenarios).

MICV		Sample size			MICV		Sample size		
no.	Name	250	500	Average	no.	Name	250	500	Average
1	lin[a]	66	87	76.5	11a	fp1 (flex1)	68	93	80.5
2	cat2	40	66	53	11b	fp1 (flex2)	75	95	85
3	cat3a	43	71	57	11c	fp1 (flex3)	73	94	83.5
4	cat3b	46	74	60	11d	fp1 (flex4)	63	89	76
5	cat4a	43	72	57.5	12a	fp2 (flex1)	59	88	73.5
6	cat4b	46	77	61.5	12b	fp2 (flex2)[a]	69	92	80.5
7	score3a	50	77	63.5	12c	fp2 (flex3)[a]	66	90	78
8	score3b	52	80	66	12d	fp2 (flex4)[a]	48	80	64
9	score4a	54	82	68	13a	spline (2 d.f.)[a]	63	88	75.5
10	score4b	58	85	71.5	13b	spline (3 d.f.)	58	86	72
					13c	spline (4 d.f.)	54	84	69

Summary

- Investigations for an interaction with a continuous variable is a very important issue in any RCT. Despite well known weaknesses, dichotomization seems to be the standard.
- MFPI is a simple and well developed approach, which can be used to investigate for an interaction with a continuous variable.
- Several parts of Phases I to III have been completed by PR and WS with some support from colleagues. Unfortunately, MFPI was widely ignored by the research community.
- MFPI is discussed in the context of the very important Instrument to assess the Credibility of Effect Modification Analyses **(ICEMAN).**
- Evidence-based medicine requires reviews and meta-analyses. An approach for meta-analyses for functions is proposed.