

Why meta-analysis should be based on individual patient data

Marc Buyse

*IDDI, Louvain-la-Neuve, and
Hasselt University, Diepenbeek, Belgium*

*ROeS Seminar 2007, 10 September 2007
Bern, Switzerland*

Types of meta-analyses

- Based on the literature
- Based on summary statistics
- Based on individual patient data (IPD)

Disadvantages of IPD meta-analyses

- Time
- Cost
- Feasibility

Advantages of IPD meta-analyses

- Reduction of bias
- Maximization of power
- Quality checks
- Investigations of heterogeneity
- Further analyses

Advantages of IPD meta-analyses

- Reduction of bias
 - Publication bias
 - Published statistics are often biased
- Maximization of power
- Quality checks
- Investigations of heterogeneity
- Further analyses

Advantages of IPD meta-analyses

- Reduction of bias
- Maximization of power
 - Power is lost in time to event analyses if event is frequent and time is ignored
 - Methods to extract information from Kaplan-Meier curves require assumptions
- Quality checks
- Investigations of heterogeneity
- Further analyses

Advantages of IPD meta-analyses

- Reduction of bias
- Maximization of power
- Quality checks
 - Validity of randomization can partially be checked
 - Patient characteristics can be explored
- Investigations of heterogeneity
- Further analyses

Advantages of IPD meta-analyses

Dates of entry	Experimental	Control
November 1981 to March 1982	12	0
April 1982 to July 1985	56	59
August 1985 to October 1985	0	8
Total	68	67

Advantages of IPD meta-analyses

- Reduction of bias
- Maximization of power
- Quality checks
- Investigations of heterogeneity
 - Heterogeneity can be due to design and/or patient mix, or to random variation
 - Meta-regression potentially subject to ecologic and/or baseline risk bias
- Further analyses

Ecologic bias

Meta-regression (in which the units of analysis are the trials) may be subject to ecologic bias and imply incorrect relationship between treatment effect and prognostic factors at the individual level

Ref: Buyse and Piedbois, Stat in Med 1996;15:2797.

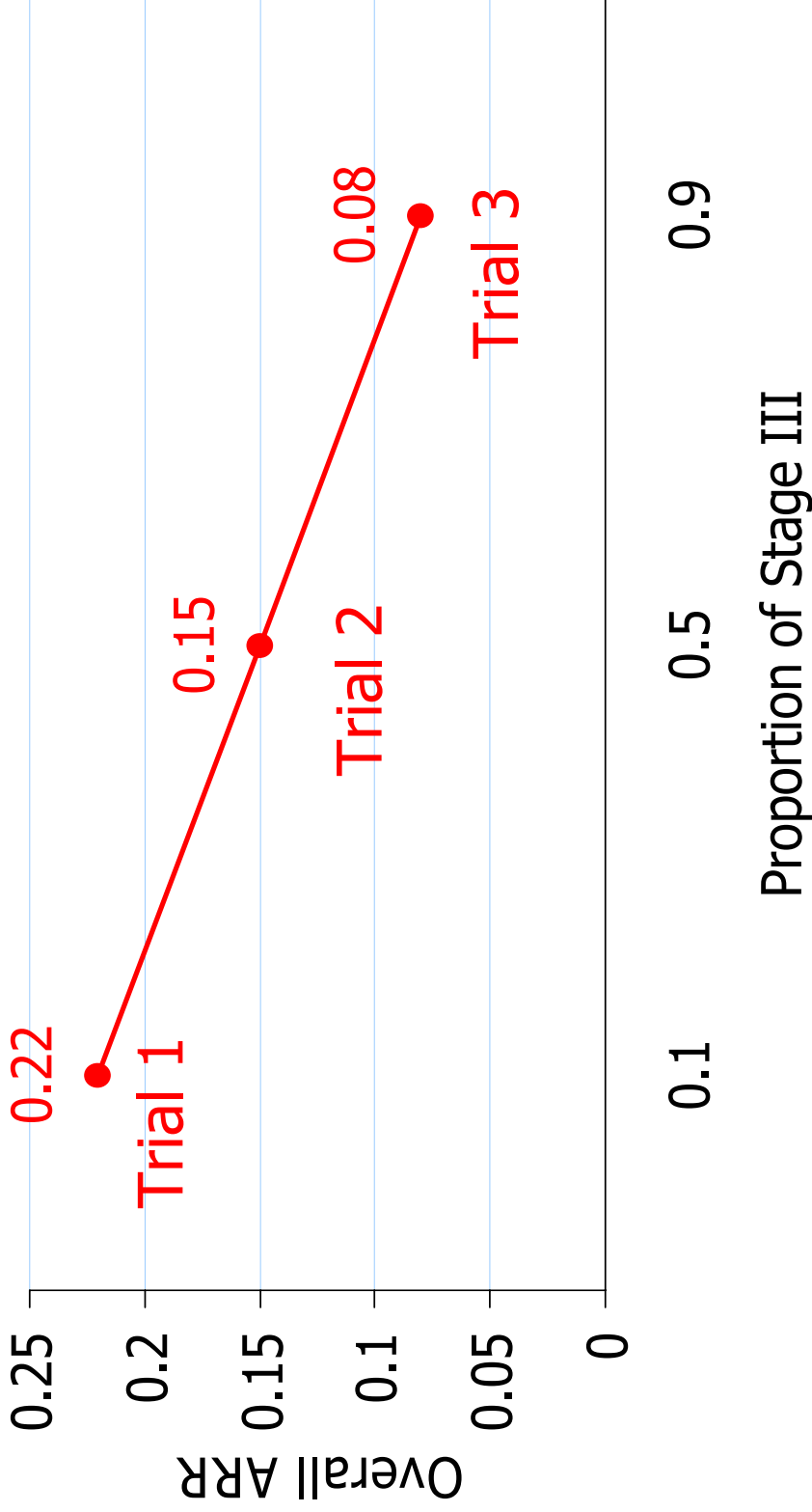
Ecologic bias

Example:

Absolute risk reductions (ARR) with adjuvant treatment for early colorectal cancer in three trials including patients with stage II and stage III disease

Ref: Buyse and Piedbois, Seminars Oncol 2001;S1:20.

Ecologic bias



Ecologic bias

	ARR _{III}	ARR _{II}	Prop _{III}	Prop _{II}	ARR
Trial 1	.4	.2	.1	.9	.22
Trial 2	.2	.1	.5	.5	.15
Trial 3	.08	.04	.9	.1	.08



$$ARR_{III} = 2 \text{ } ARR_{II}$$

Ecologic bias

- *at trial level:*

Overall absolute risk reduction decreases as the proportion of stage III patients increases, suggesting that treatment effect is larger in lower risk patients:
 $ARR_{III} < ARR_{II}$
- *at individual level:*

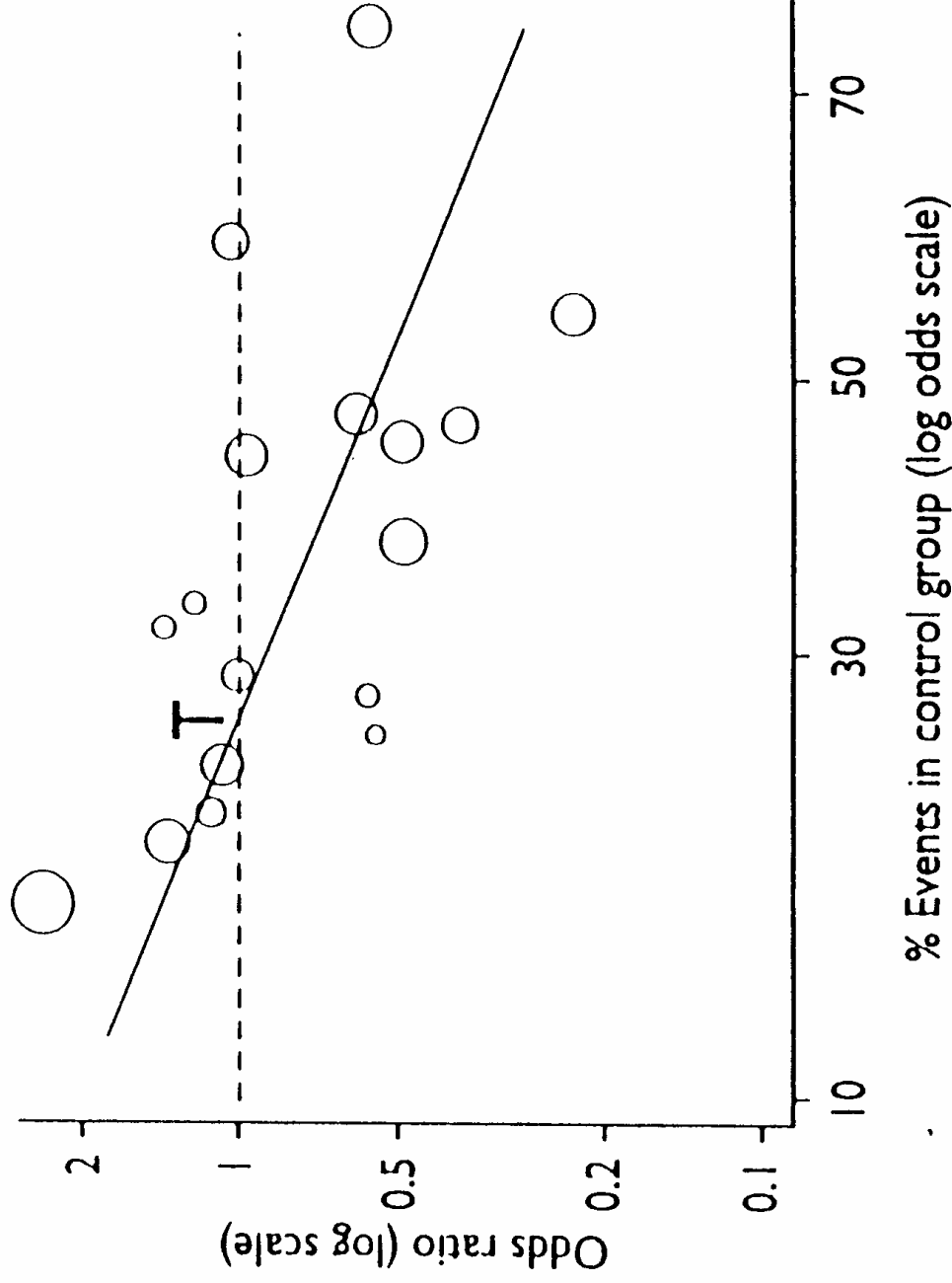
Absolute risk reduction is in fact twice larger in stage III patients than in stage II!
 $ARR_{III} = 2 \text{ } ARR_{II}$

Baseline risk bias

Meta-regression on baseline risk (in which the units of analysis are trials) is *always* subject to bias and may imply incorrect relationship between treatment effect and prognostic factors

Ref: Sharp et al. *Brit Med J* 1996;**313**:735.

Baseline risk bias



Baseline risk bias

$ARR = R_T - R_C$, absolute risk reduction, where

R_C , risk in control group (= baseline risk)

R_T , risk in treated group

$$ARR = \alpha + \beta R_C$$

Even when $\beta = 0$, $E(\beta) = -\sigma_w^2 / (\sigma_b^2 + \sigma_w^2)$

where σ_w^2 is (common) variance within trials
and σ_b^2 is variance between trials

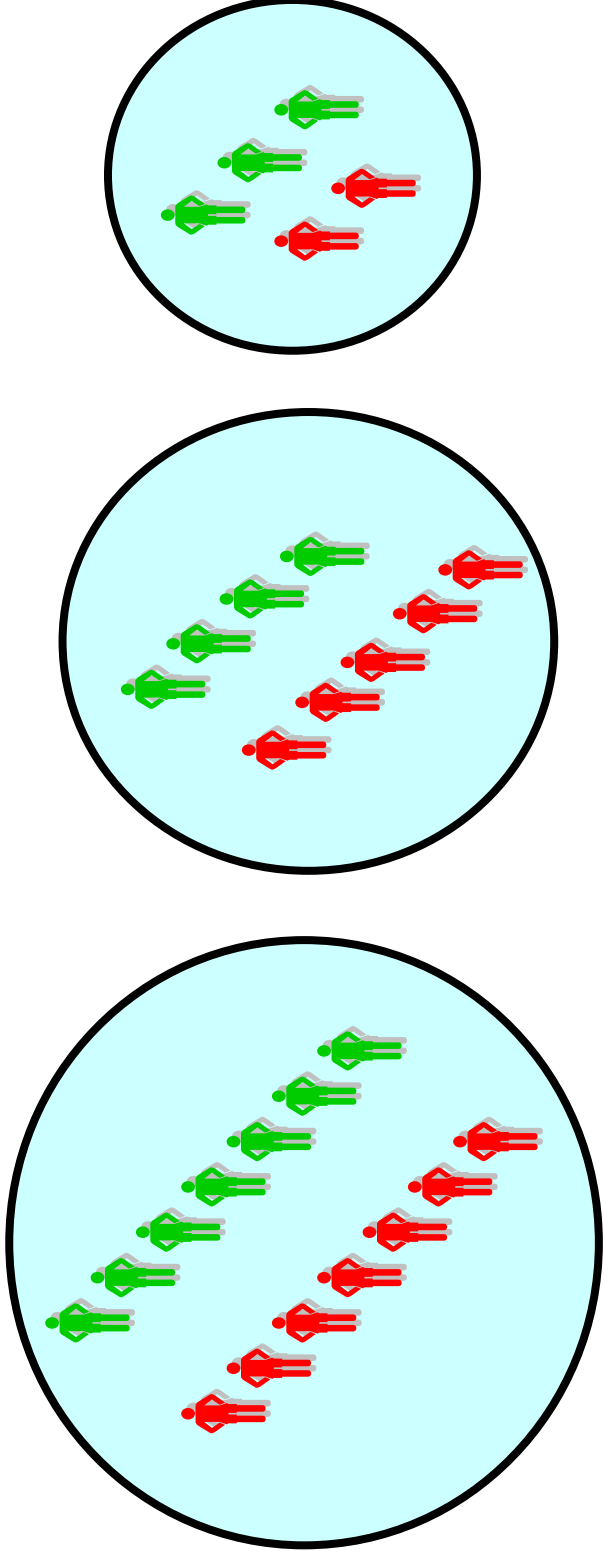
Baseline risk bias

- *at trial level:*
Treatment benefit increases as baseline risk increases, suggesting that treatment effect is larger in higher risk patients
- *at individual level:*
Treatment benefit is in fact exactly the same in high risk as in low risk patients

Multi-level information

Trial level ($i = 1 \dots I$):

Summary statistics on endpoints and treatment effects



Individual level ($j = 1 \dots J_i$)

Observations of individual endpoints

Validation of surrogate endpoints

Randomized
treatment

Trt

Intermediate endpoint,
potential surrogate

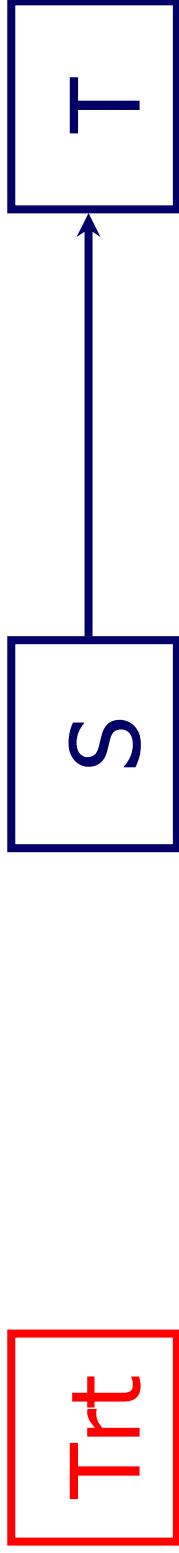
S

True endpoint

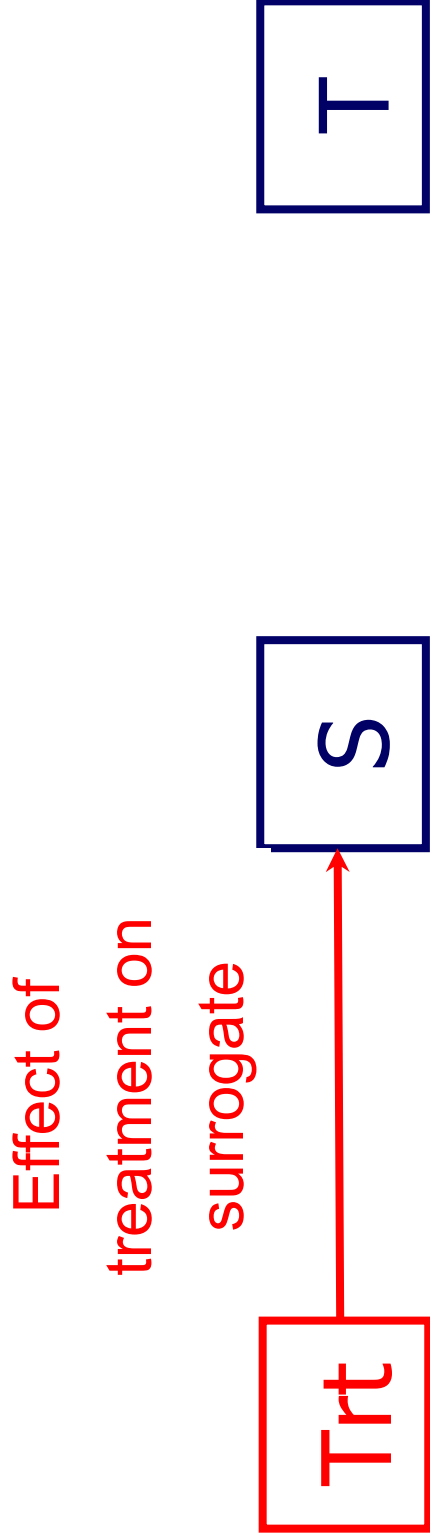
T

Validation of surrogate endpoints

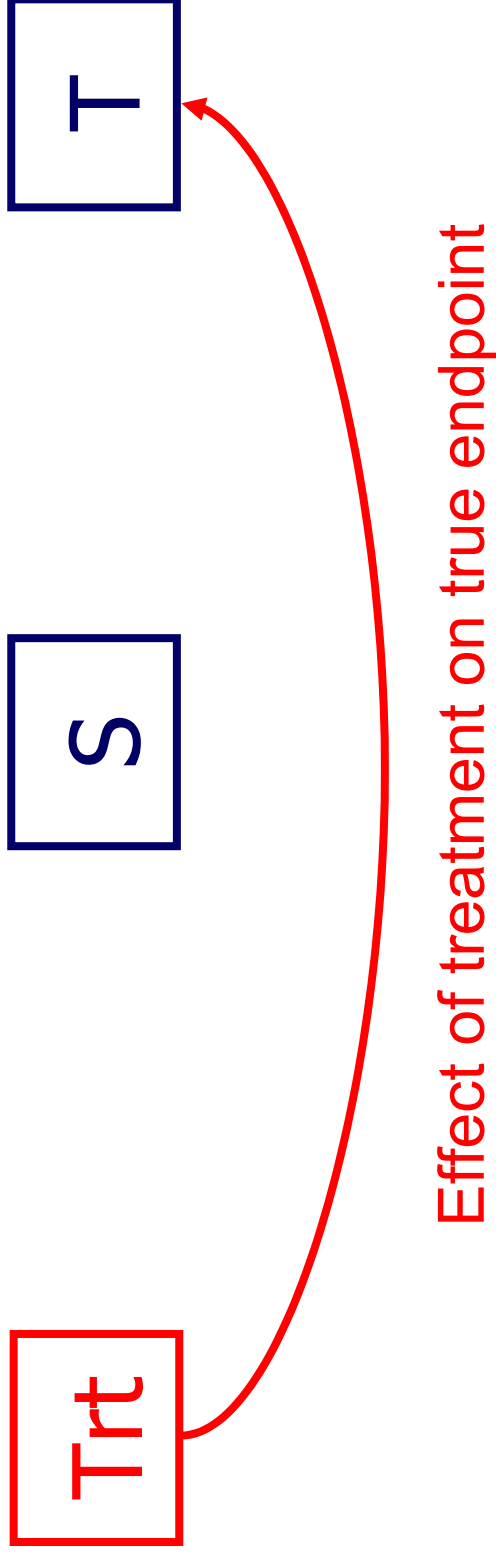
Effect of
surrogate on
true endpoint



Validation of surrogate endpoints



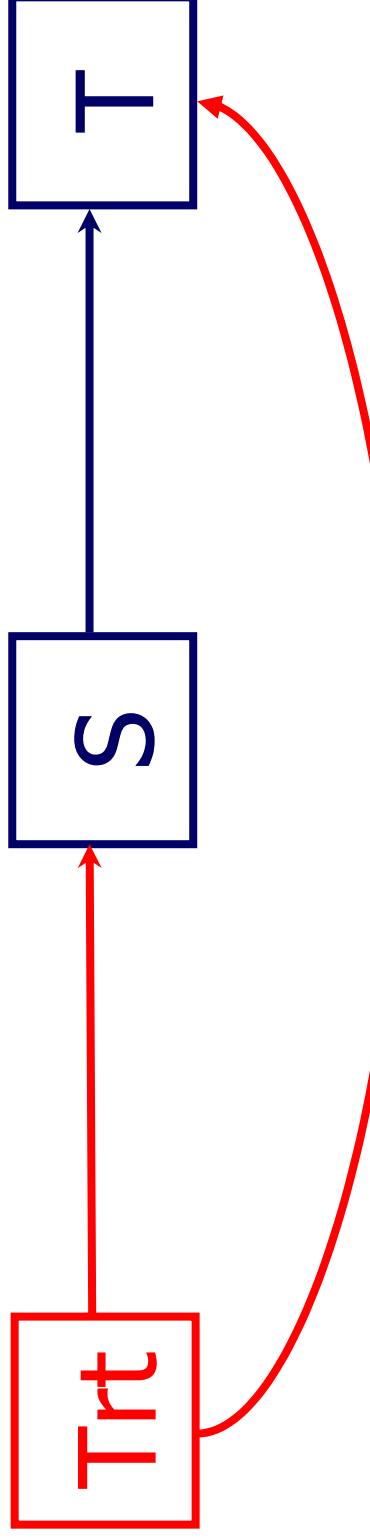
Validation of surrogate endpoints



Full capture of effect

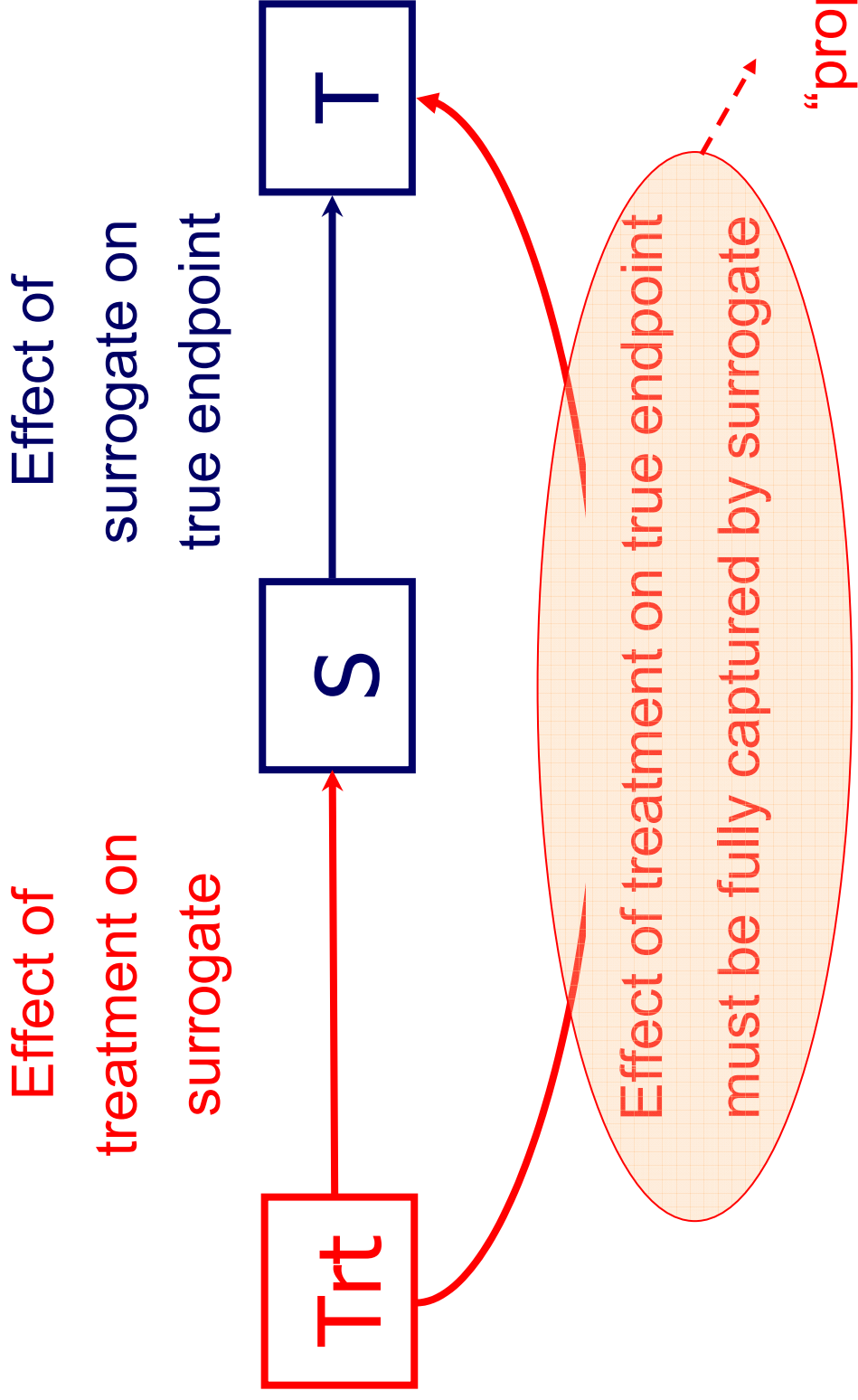
Effect of
treatment on
surrogate

Effect of
surrogate on
true endpoint

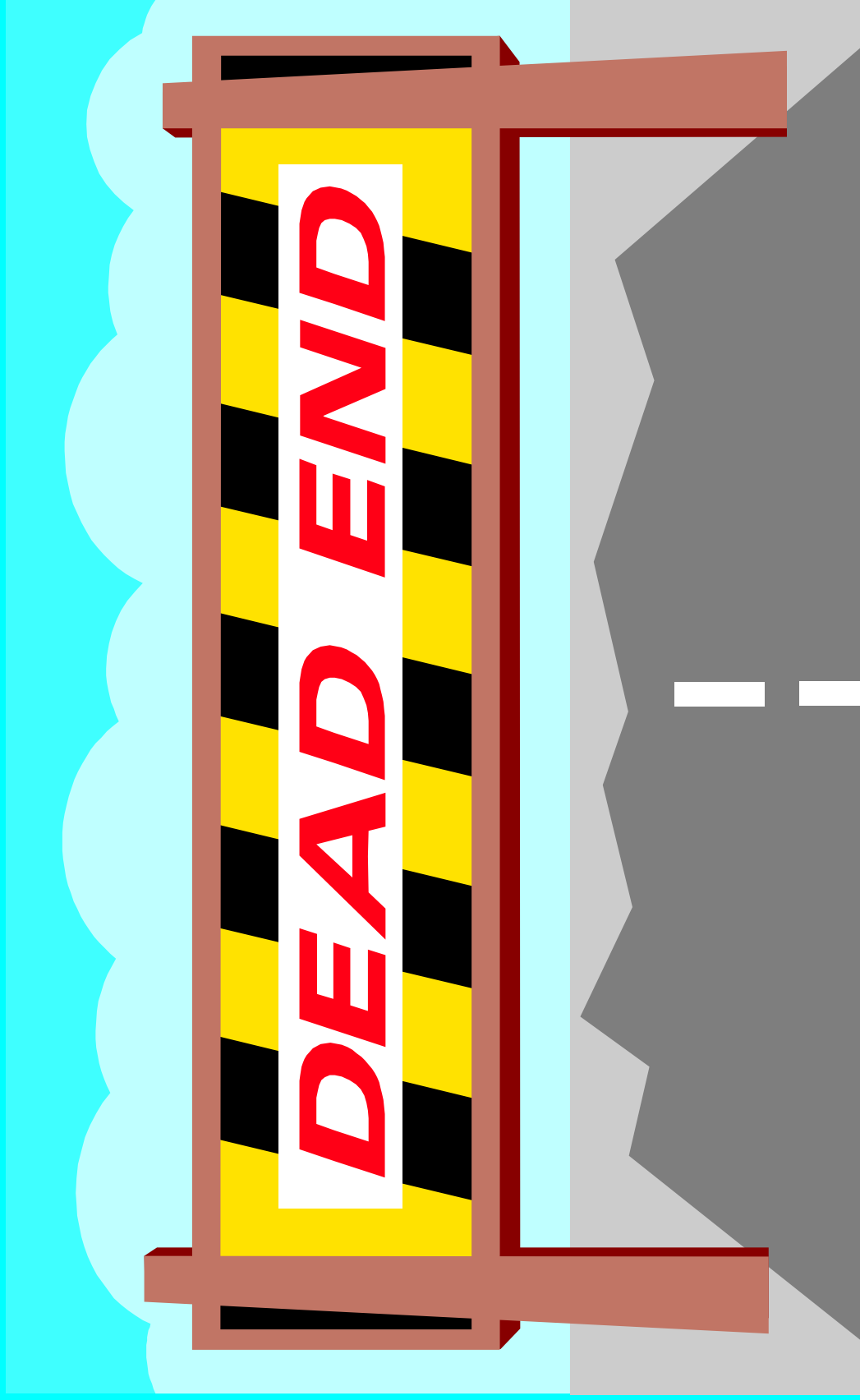


Effect of treatment on true endpoint
must be fully captured by surrogate

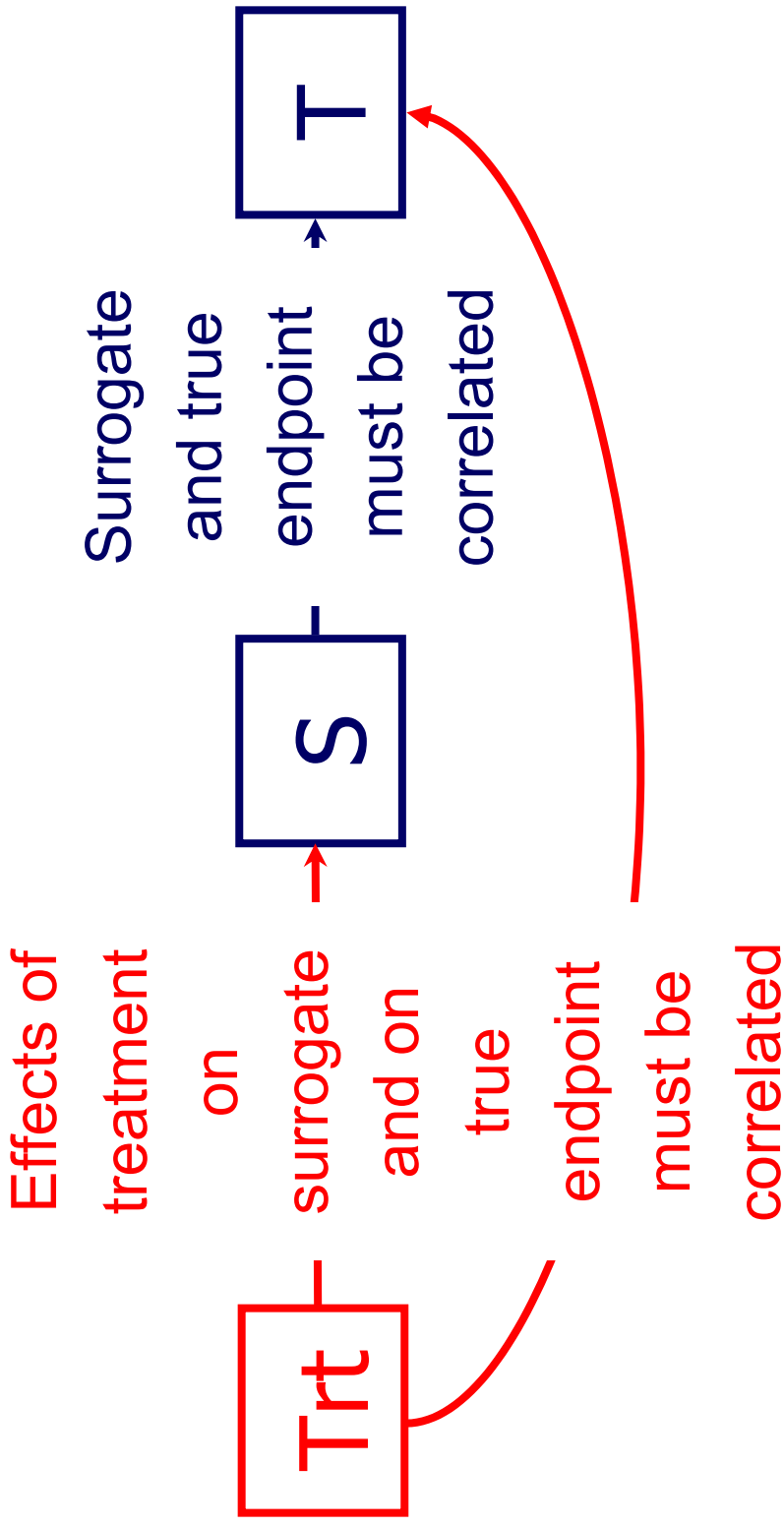
Proportion of effect explained



“Full capture of effect” in fact implies
perfect individual-level correlation



Multilevel approach



Ref: Buyse et al, *Biostatistics* 2000;1:49;

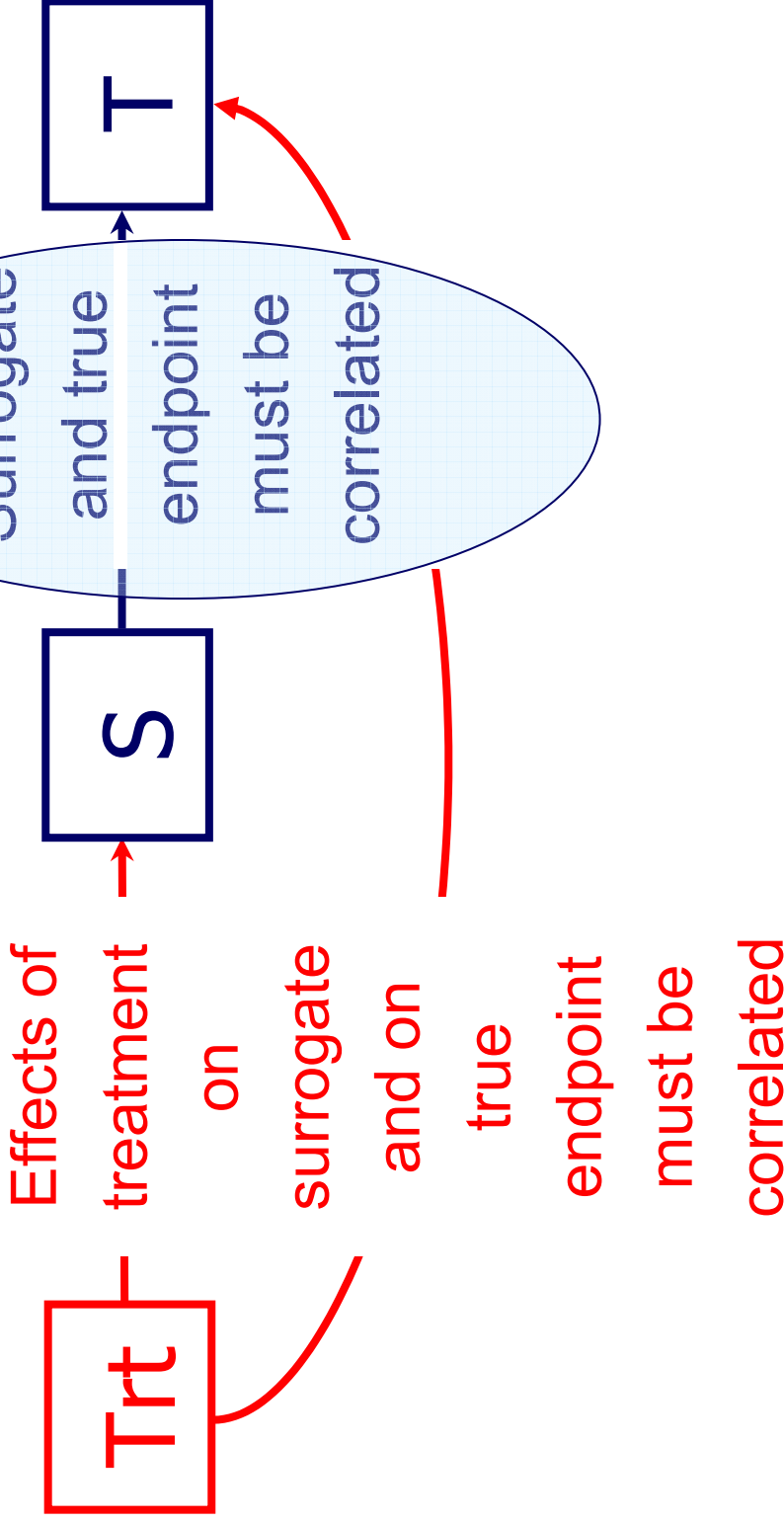
Gail, Pfeiffer and van Houwelingen, *Biostatistics* 2000;1:231.

Multilevel approach

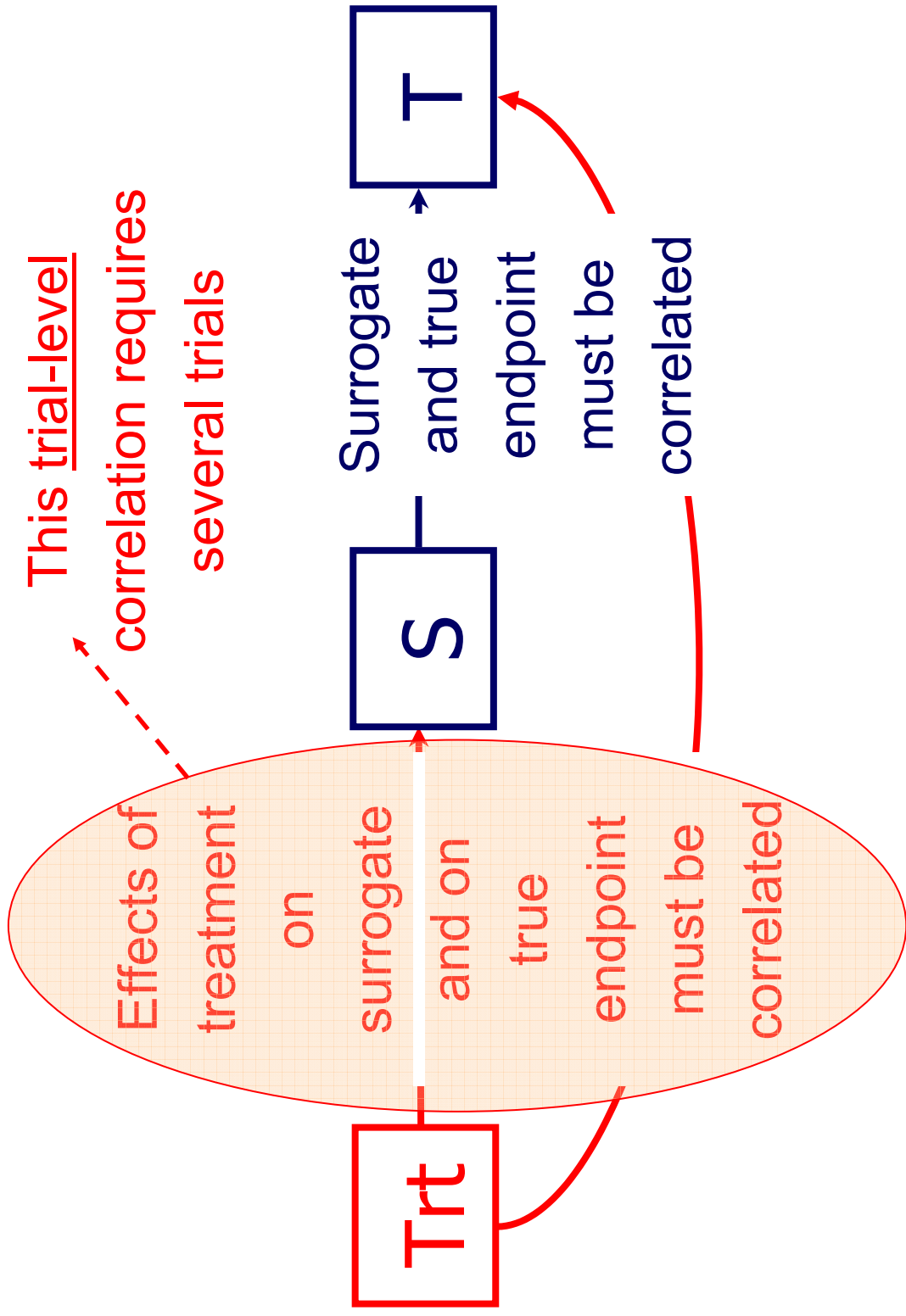
This individual-level

correlation can be

shown in a single trial



Multilevel approach



General model

Model (assuming S and T normally distributed):

$$S_{ij}|Z_{ij} = \mu_{Si} + \alpha_i Z_{ij} + \varepsilon_{Sij}.$$

$$T_{ij}|Z_{ij} = \mu_{Ti} + \beta_i Z_{ij} + \varepsilon_{Tij}$$

$$\Sigma = \begin{pmatrix} \sigma_{SS} & \sigma_{ST} \\ \sigma_{ST} & \sigma_{TT} \end{pmatrix}$$

Correlation between endpoints

Individual-level surrogacy:

$$R_{\text{indiv}}^2 = \frac{\sigma_{ST}^2}{\sigma_{SS} \sigma_{TT}}$$

Random effects model

$$\begin{pmatrix} \mu_{Si} \\ \mu_{Ti} \\ \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \mu_S \\ \mu_T \\ \alpha \\ \beta \end{pmatrix} + \begin{pmatrix} m_{Si} \\ m_{Ti} \\ a_i \\ b_i \end{pmatrix}$$

$$D = \begin{pmatrix} d_{SS} & d_{ST} & d_{Sa} & d_{Sb} \\ d_{TT} & d_{Ta} & d_{Tb} & \\ d_{aa} & d_{ab} & & \\ d_{bb} & & & \end{pmatrix}$$

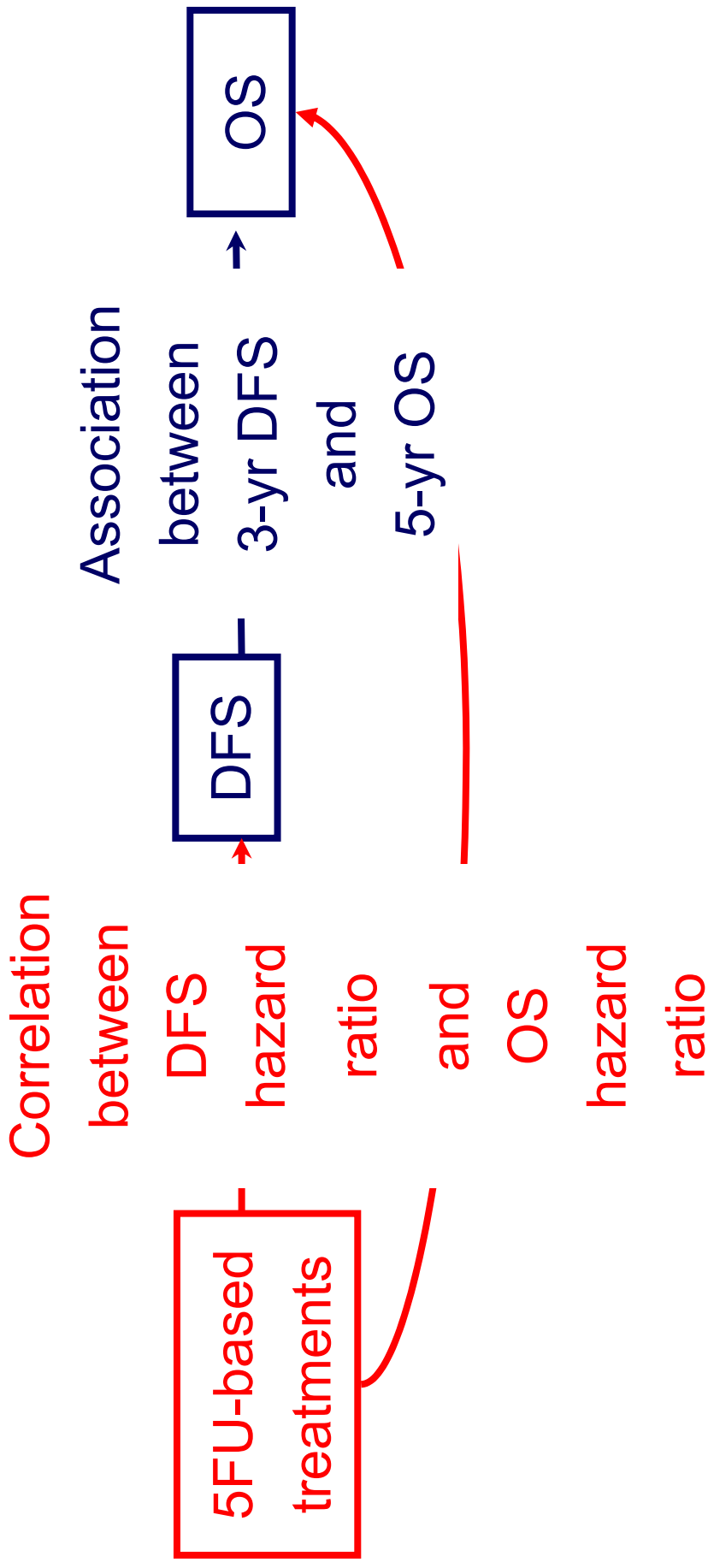
Correlation between treatment effects

Trial-level surrogacy:

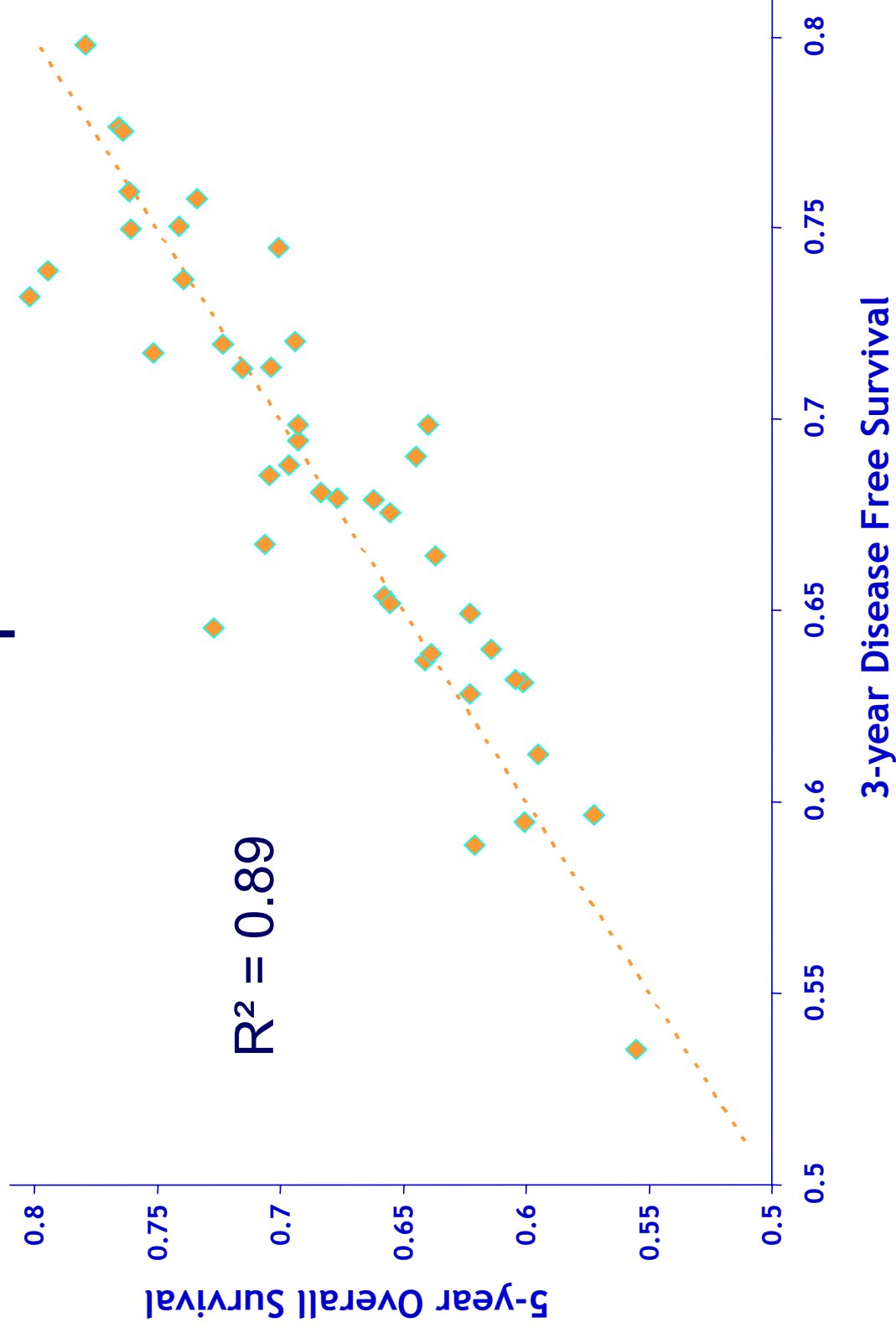
$$R_{\text{trial}(f)}^2 = \frac{\begin{pmatrix} d_{sb} \\ d_{ab} \end{pmatrix}^T \begin{pmatrix} d_{ss} & d_{sa} \\ d_{sa} & d_{aa} \end{pmatrix}^{-1} \begin{pmatrix} d_{sb} \\ d_{ab} \end{pmatrix}}{d_{bb}}$$

- Early colorectal cancer :
DFS as a surrogate for survival**
- **43 treatment arms in 18 randomized trials (20,898 patients)**
 - **9 surgery alone control groups**
 - **34 5FU-based experimental treatment groups**
 - **Endpoints: disease-free survival and survival**

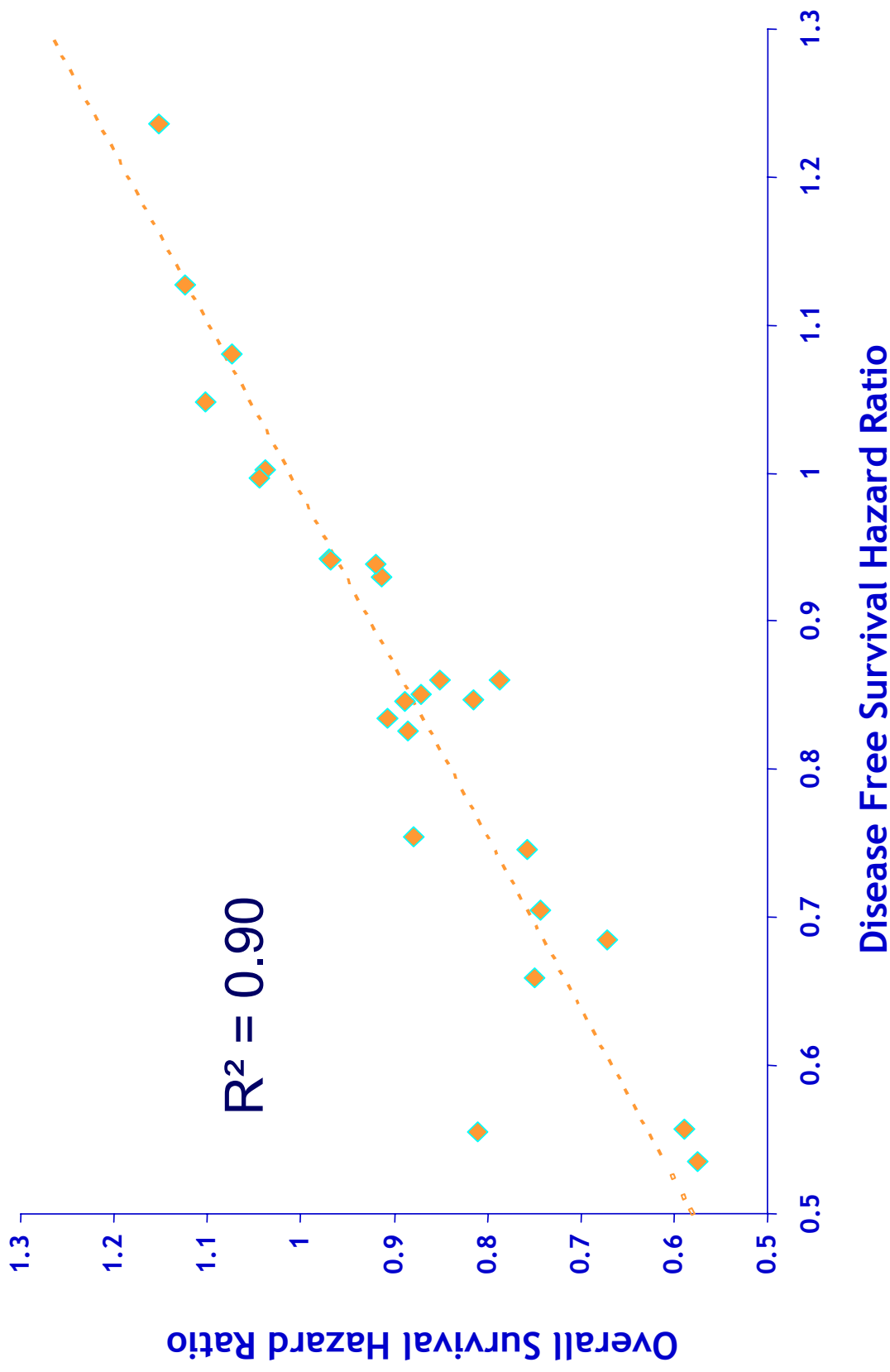
Early colorectal cancer : DFS as a surrogate for survival



Correlation between summary statistics on endpoints



Correlation between treatment effects



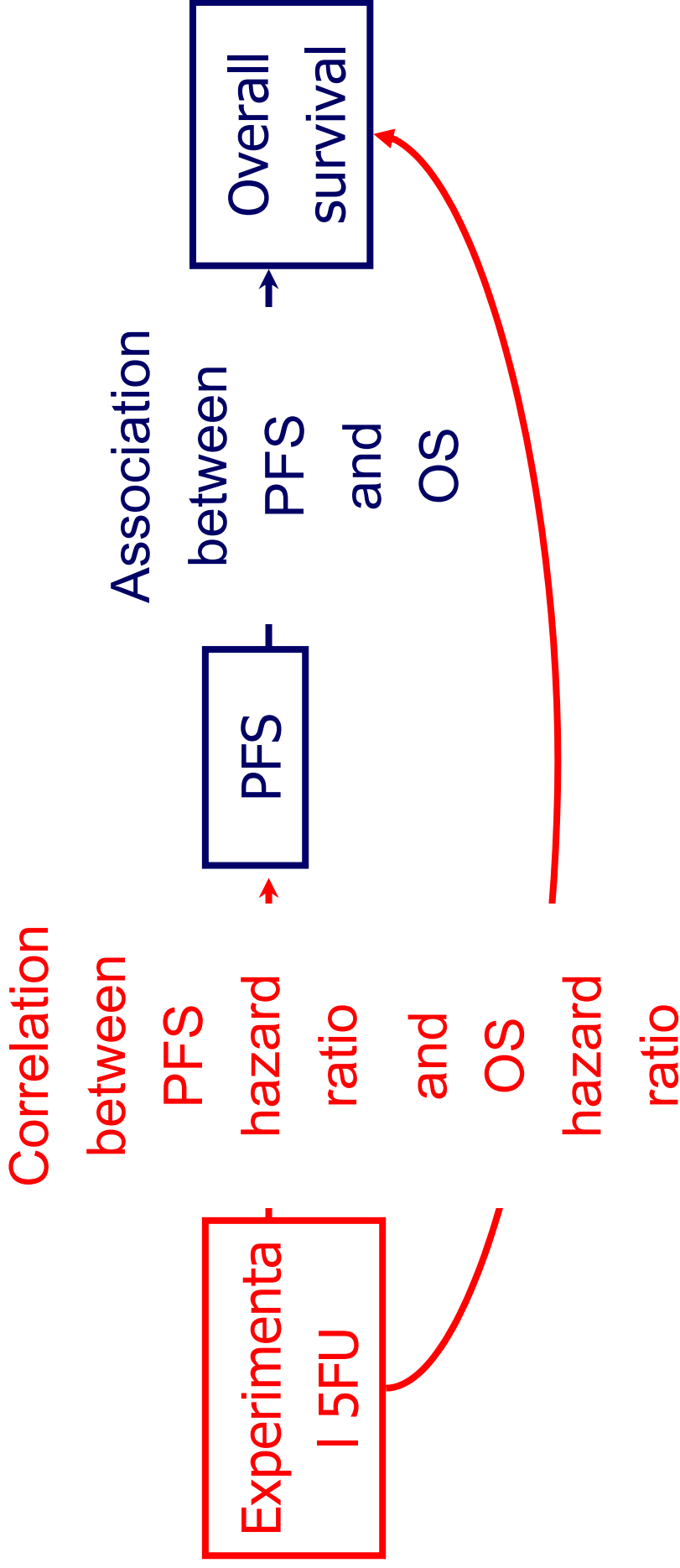
Early colorectal cancer : DFS as a surrogate for survival

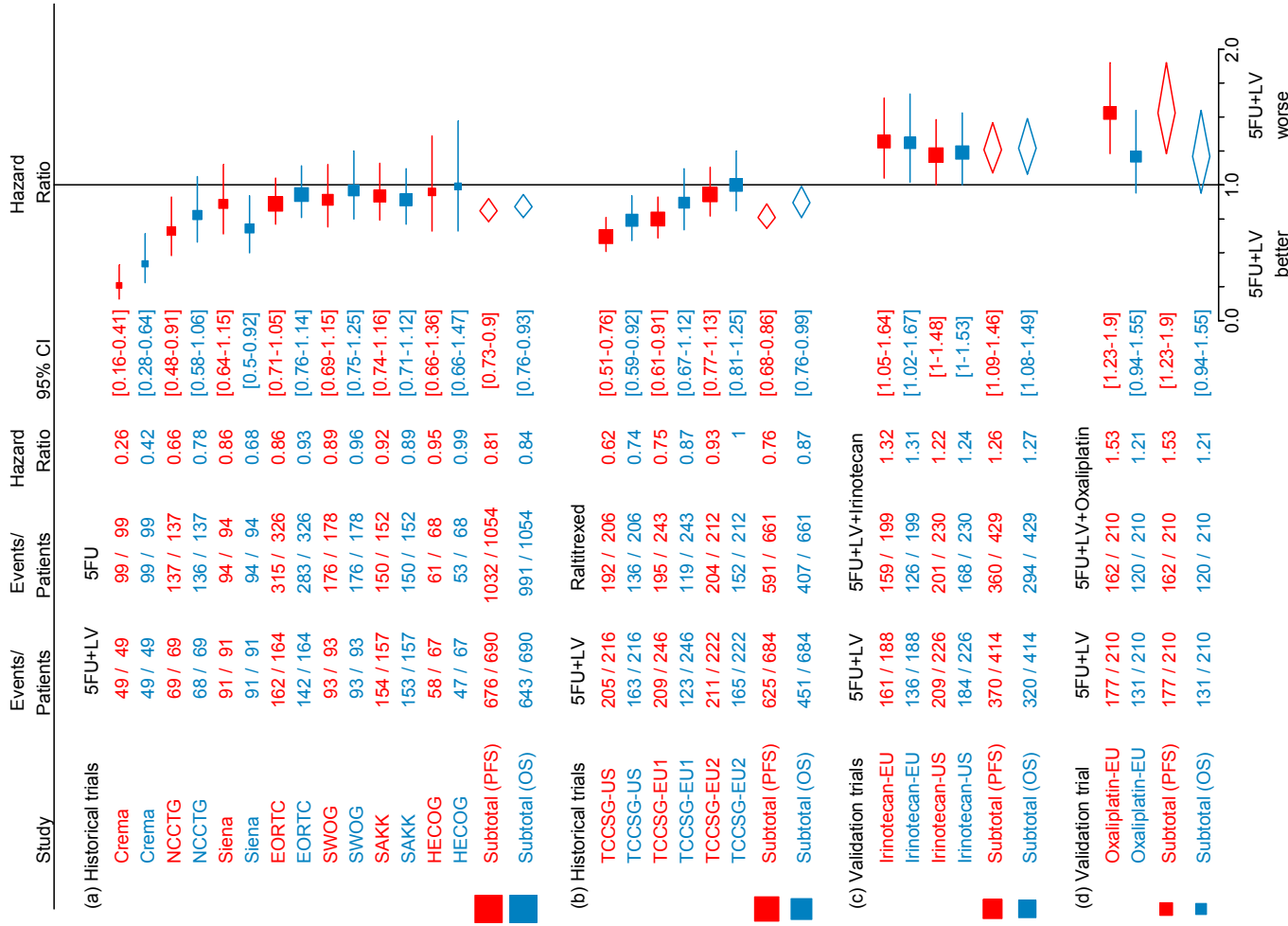
- Remarkably high correlations between endpoints as well as treatment effects
- However,
 - Individual-level surrogacy based on summary statistics → is loss of information important, and are results sensitive to choice?
 - Uncertainty in summary statistics is ignored:
 - is correlation over-estimated?
 - Only 5FU-based treatments

Advanced colorectal cancer: PFS as a surrogate for survival

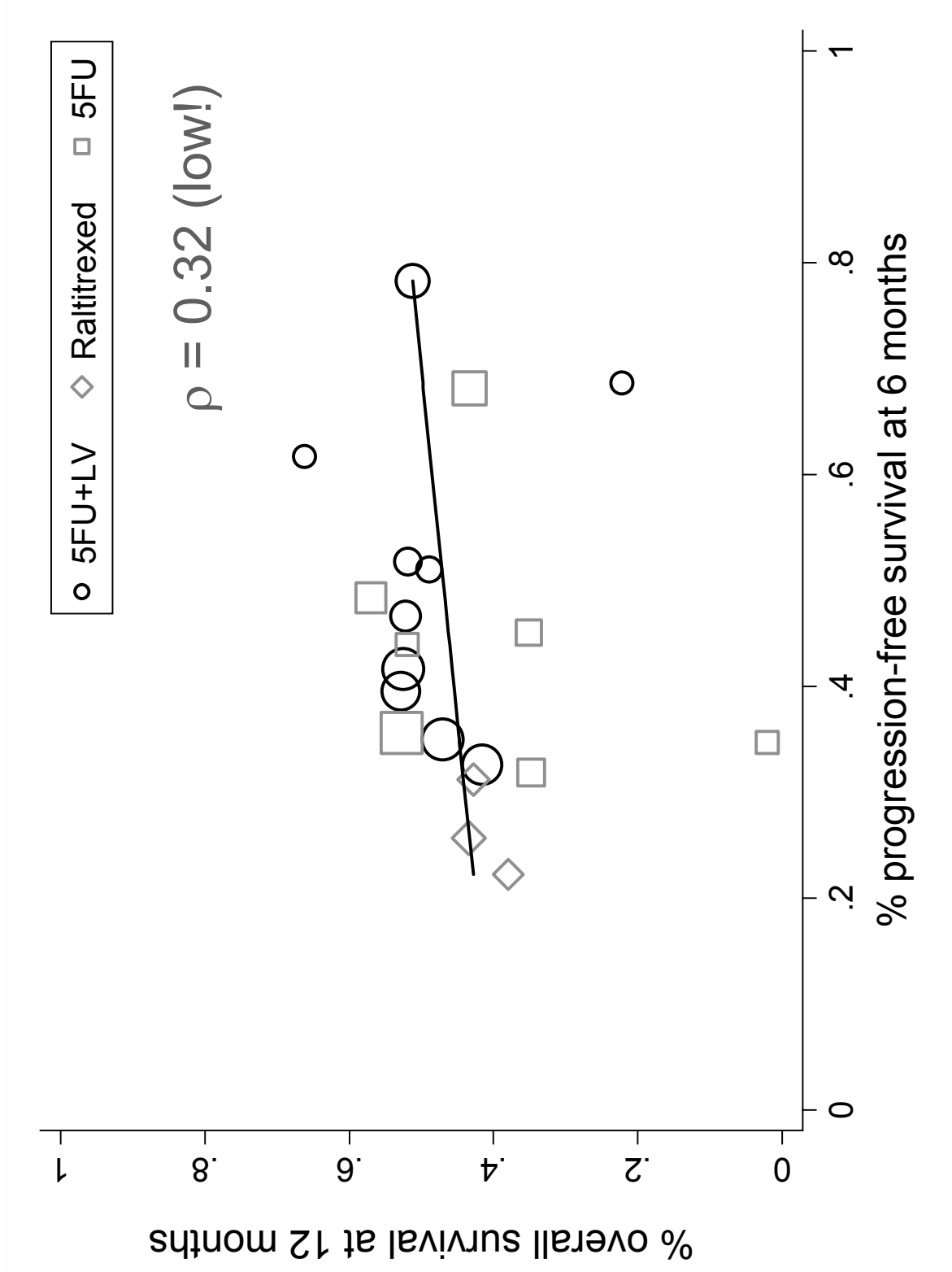
- 13 trials on 4,352 patients
- Treatments (5FU/LV common arm):
 - Training set (10 trials, 3,089 patients)
5FU Bolus or raltitrexed vs. 5FU/L
 - Validation set (3 trials, 1,263 patients)
oxaliplatin or irinotecan + 5FU/LV vs.
5FU/LV

Advanced colorectal cancer

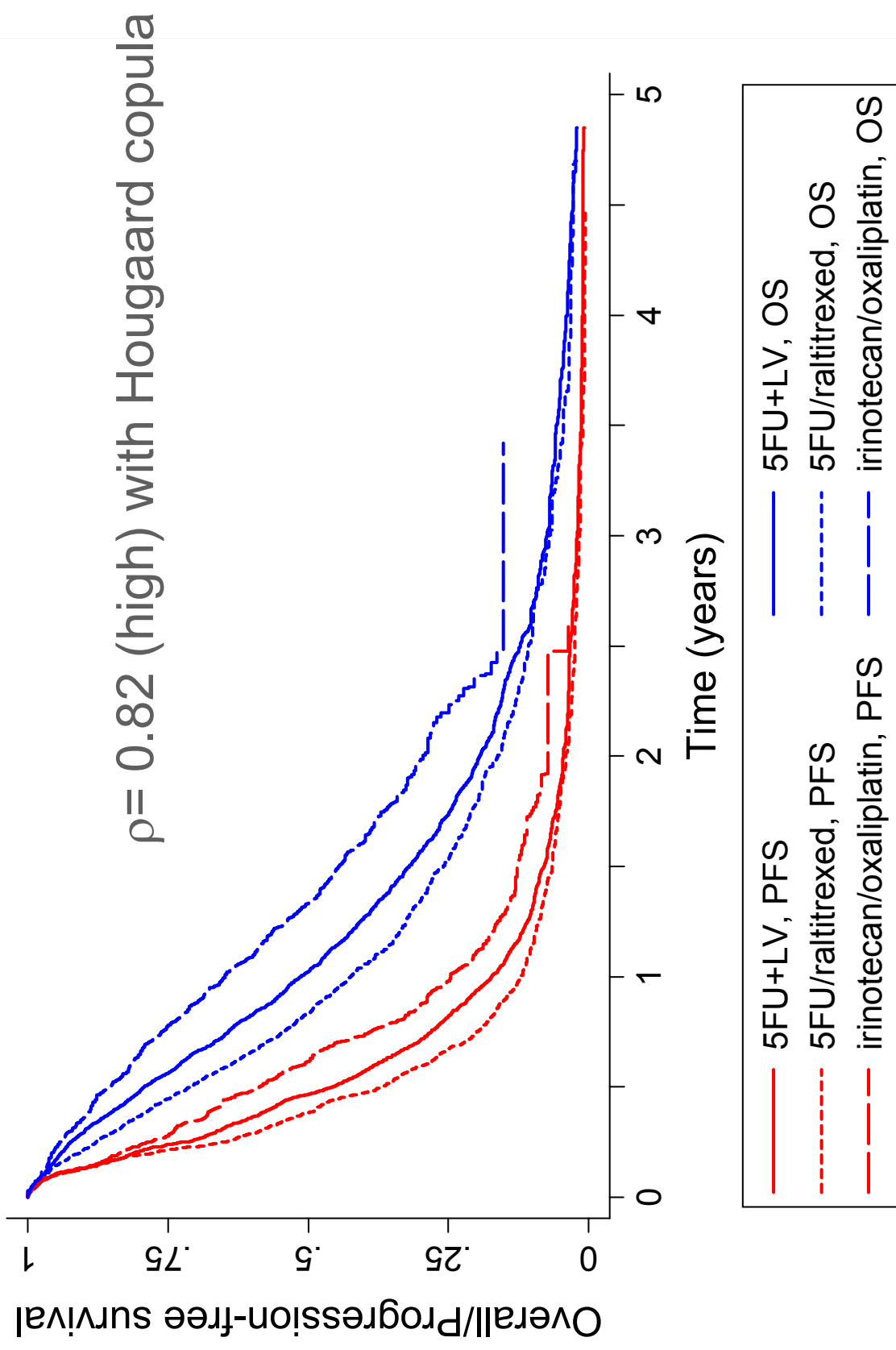




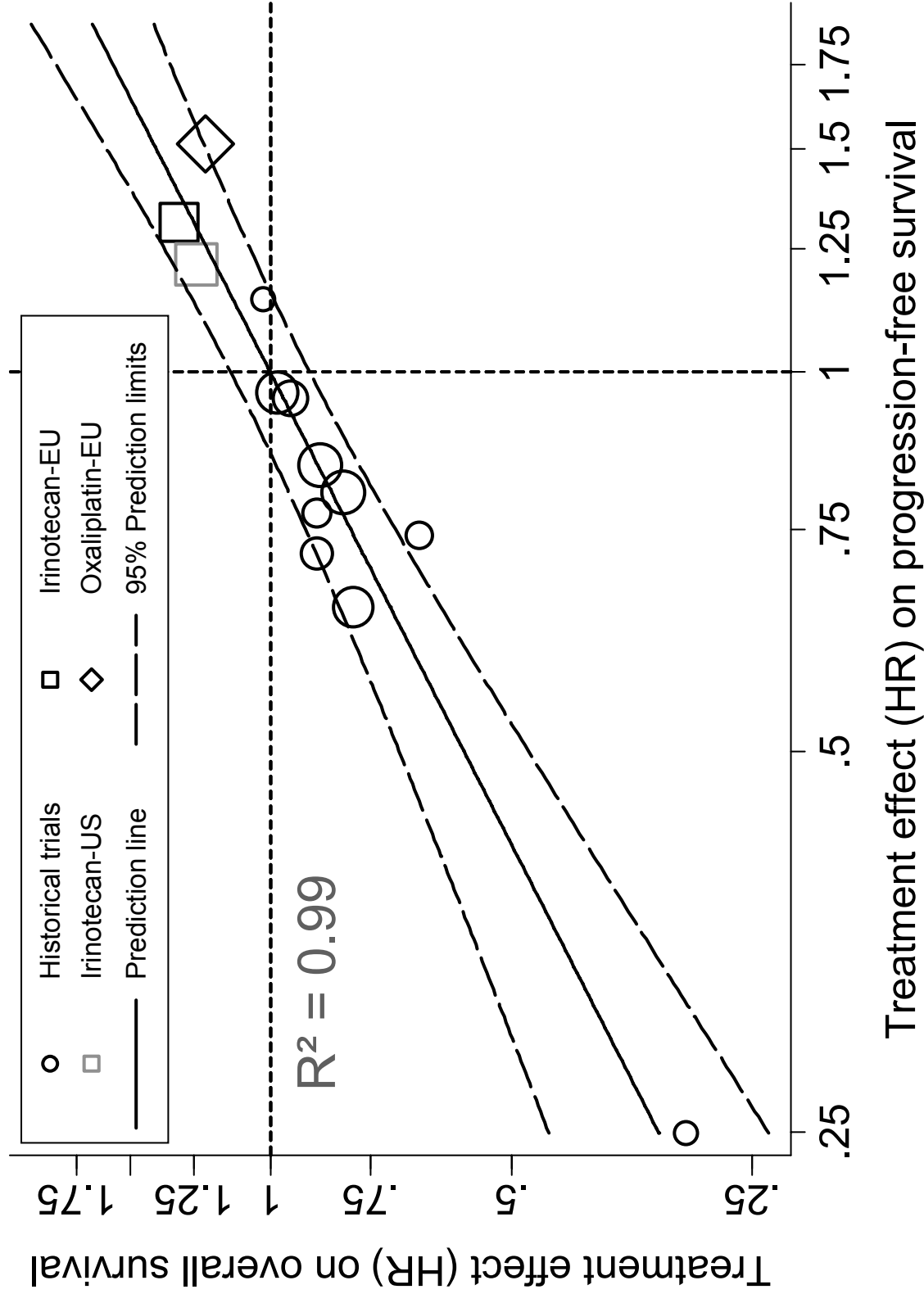
Individual level - summary statistics



Individual level - PFS and OS curves



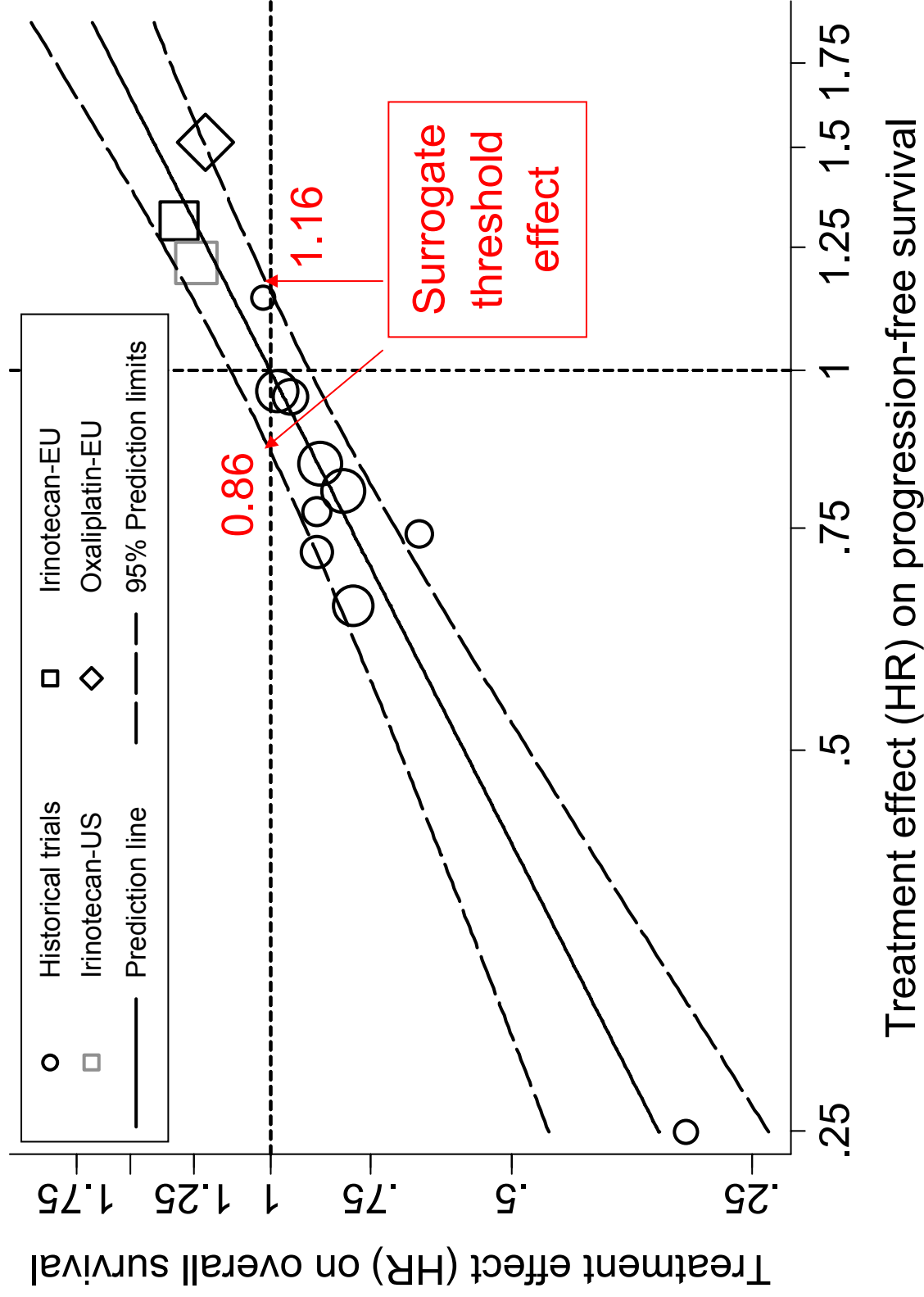
Trial level correlation between effects



Surrogate threshold effect

The “Surrogate Threshold Effect” is the treatment effect on the surrogate that would predict a statistically significant treatment effect on the true endpoint.

Surrogate threshold effect



Advanced colorectal cancer

- Remarkably high correlations between endpoints as well as treatment effects, including for new drugs oxaliplatin & irinotecan
- However,
 - Individual-level surrogacy based on summary statistics does not show a strong association → correlation based on individual patient data
 - Influential point for trial-level association → is trial-level association over-estimated?

Conclusion

“ If a problem is very important and needs an answer, statisticians should not just make do by calculating a weighted estimate based on published data (...) Presenting poor-quality estimates may mislead the reader and the public health practitioner into believing the estimates are based on good science. That is not the case.”

Ref: Sauerbrei et al, J Clin Epidemiol 2001;54:537.

Conclusion

If a meta-analysis is worth doing at all, it is worth doing well.

Conclusion

If a meta-analysis is worth doing at all, it is worth doing well.

But ...

Disadvantages of IPD meta-analyses

- Time
- Cost
- Feasibility

SLIDE # 3

The ONLY satisfactory solution (in the long run...)

Individual patient data from all randomized clinical trials should fall in the public domain after trial “completion”

- Re-analyses possible (meta-analyses / detection of low level signals)
- Precedent of clinical trial registration
- Feasible (Gene Expression Omnibus etc.)
- Technically easy (CDISC etc.)
- Ethical