

---

# Use of multistate models to jointly model progression-free and overall survival and improve decision-making in clinical trials

*Kaspar Rufibach*

*Methods, Collaboration & Outreach Group, Department of Biostatistics, Roche Basel  
MRC Biostatistics unit*

*14th November 2019, Cambridge, UK*

Generated 2019-11-14 at 23:27:30.



# Who

Meller *et al.* (2019):



Beyer *et al.* (2019):



# Multistate model for PFS and OS

# Introduction

**Oncology** endpoints:

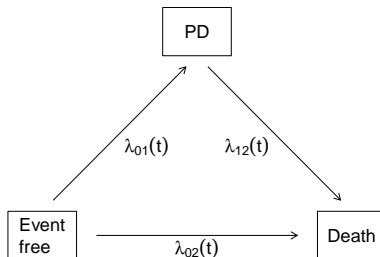
- Progression-free survival (PFS): Time from randomization to earlier of progression or death.
- Overall survival (OS): Time from randomization to death.

PFS common **surrogate** for OS in clinical trials.

Sophisticated methods to quantify amount of surrogacy, e.g. [Buyse et al. \(2016\)](#).

**Correlation** between PFS and OS important aspect, [Li and Zhang \(2015\)](#).

# Multistate model for PFS and OS

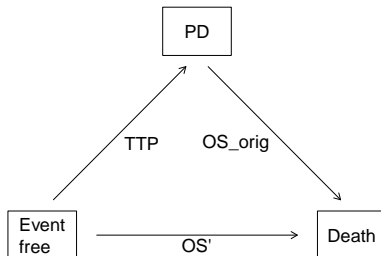


Standard **illness-death model without recovery**:

- Process  $X(t) \in \{0, 1, 2\}$ ,  $t \geq 0$  models the state occupied at time  $t$ .
- All patients in state 0 at time 0:  $P(X(0) = 0) = 1$ .
- PFS: waiting time in initial state 0,  **$\text{PFS} = \inf\{t : X(t) \neq 0\}$** .
- OS: time until reaching state 2,  **$\text{OS} = \inf\{t : X(t) = 2\}$** .

# Alternatives

# Latent failure time model



**Latent failure time model** (LFTM), [Fleischer et al. \(2009\)](#), [Li and Zhang \(2015\)](#):

- $PFS = \min(TTP, OS_{orig})$ .
- $OS = PFS$  if  $PFS \neq TTP$ ,  $TTP + OS'$  else.

Challenges:

- **Impossible sampling space:**  $TTP > OS_{orig} \Rightarrow$  progression after death  $\Rightarrow$  awkward idea.
- Issue with assumptions for estimation.

# Copulas

**Copulas**, e.g. Burzykowski *et al.* (2001), Fu *et al.* (2013), Emura *et al.* (2017):

- Model general bivariate survival data (lifetimes of twins).

Challenges:

- PFS - OS structure more specific.
- $\text{PFS} \leq \text{OS} \Rightarrow$  copulas do not place such a restriction on pair of event times.
- Reality: death without progression  $\Rightarrow P(\text{PFS} = \text{OS}) > 0$ . Copula model with **continuous** “marginal” survival functions for PFS and OS:  $P(\text{PFS} = \text{OS}) = 0$ .



# Multistate model formulation

Transition probabilities:

- **Full description** of multistate model by only assuming existence of intensities  $\alpha_{01}$ ,  $\alpha_{02}$  and  $\alpha_{12}$ .
- Formulas, even for **non-Markov** case: [Aalen et al. \(2008\)](#).

[Meller et al. \(2019\)](#):

- Embed PFS and OS in multistate model framework,
- formulas for  $P_{lm}$ 's assuming **Weibull** transition hazards for time-inhomogeneous Markov and semi-Markov (explicit),
- inference via **counting process likelihood**,
- $P(\text{PFS} \leq u, \text{OS} \leq v)$  for  $X$  non-Markov (generic).

Allows derivation of any functional of PFS and OS.

Exemplary application: **Pearson correlation**.

# Multistate model for PFS and OS

**Marginal** distributions:

$$\begin{aligned}S_{PFS}(t) &= P(\text{PFS} > t) = P_{00}(0, t), \\S_{OS}(t) &= P(\text{OS} > t) = P_{00}(0, t) + P_{01}(0, t),\end{aligned}$$

**Joint** distribution:

$$\begin{aligned}P(\text{PFS} \leq u, \text{OS} \leq v) &= P(X(u) \in \{1, 2\}, X(v) = 2) \\&= P(X(v) = 2 | X(u) = 1) \cdot P_{01}(0, u) + P_{02}(0, u).\end{aligned}$$

$X$  **inhomogeneous Markov**:  $P(X(v) = 2 | X(u) = 1) = P_{12}(u, v)$  independent of progression time  $t_1 \leq u$ .

$X$  **non-Markov**:

- Integrate  $P_{12}(u, v; t_1)$  over conditional distribution of all possible progression times  $t_1 \leq u$ .
- Formula tedious (see [Meller et al. \(2019\)](#))  $\Rightarrow$  **simulate** in applications.

# Advantages of multistate model

Multistate model:

- Assumptions on  $X(t)$  induce properties of transition intensities, (joint) probabilities, and thus PFS and OS.
- No progression after death.
- PFS = OS easily possible.
- Estimation: No assumption about “in-/dependence” of PFS and OS.

Multistate = (most?) parsimonious model

# Correlation coefficient

## Correlation coefficient

$$\text{Corr}(\text{PFS}, \text{OS}) = \frac{\text{Cov}(\text{PFS}, \text{OS})}{\sqrt{\text{Var}(\text{PFS}) \text{Var}(\text{OS})}} = \frac{\mathbb{E}(\text{PFS} \cdot \text{OS}) - \mathbb{E}(\text{PFS}) \mathbb{E}(\text{OS})}{\sqrt{\text{Var}(\text{PFS}) \text{Var}(\text{OS})}}.$$

Mean, variance of PFS and OS: via survival functions.

$\mathbb{E}(\text{PFS} \cdot \text{OS})$ : Use

$$P(\text{PFS} \cdot \text{OS} > t) = P(\text{PFS} > \sqrt{t}) + \int_{(0, \sqrt{t}]} P_{11}(u, t/u; u) P(\text{PFS} > u-) \alpha_{01}(u) \, du.$$

Proof: manipulations using law of total probability.

# Estimation and inference for Markov models

# Estimation and inference for Markov models

## Parametric:

- Plug parametric assumption in formulas for  $P_{lm}(s, t)$ ,  $S_{PFS}$ ,  $S_{OS}$ ,  $\text{Corr}(\text{PFS}, \text{OS})$ .
- Estimate parameters using **Counting Process Likelihood**, Andersen *et al.* (1993).  
Product of patient-specific likelihood-contributions to each state transition.
- Inference via delta method or bootstrap (results comparable).

## Nonparametric:

- Transition probabilities: Aalen-Johansen estimator, Aalen and Johansen (1978).
- Plug in estimates into formulas for PFS, OS,  $\text{Corr}(\text{PFS}, \text{OS})$ .
- Challenge: need to **extrapolate tail beyond where we have data**.
- Inference via bootstrap.

# Estimation and inference for Markov models

LFTM in [Fleischer et al. \(2009\)](#) and [Li and Zhang \(2015\)](#):

- Group patients depending on their path from 0 to 1 or 2, or censored.
- Likelihood uses **assumption of independence** of TTP,  $OS_{\text{orig}}$ . Cannot tell from (even uncensored!) data! [Aalen \(1987\)](#): “artificial problem”, as LFTM not needed, see also [Beyersmann et al. \(2012\)](#).

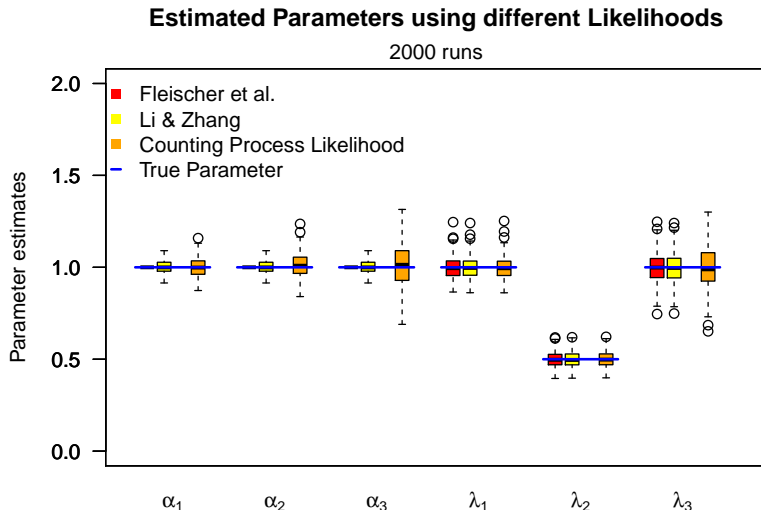
[Weber and Titman \(2019\)](#):

- Kendall's  $\tau$ , based on multistate, nonparametric, and copula models.
- Use again LFTM for estimation.

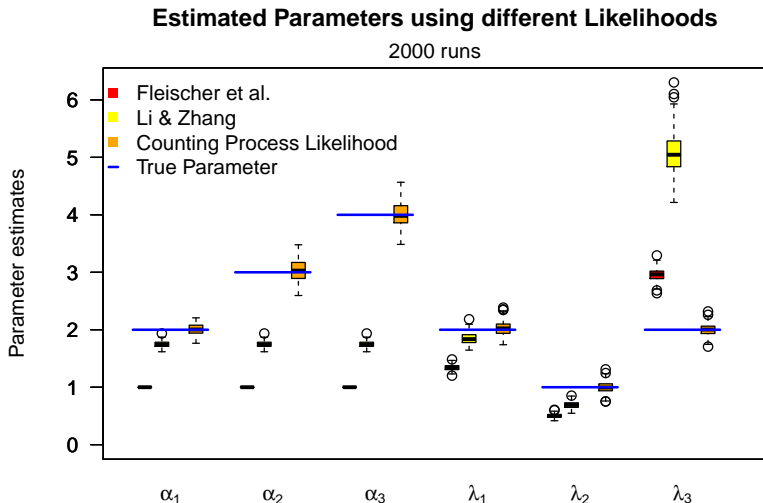


# Results

## Results: estimated parameters Exponential



# Results: estimated parameters time-inhomogeneous Weibull



# Conclusions

# Conclusions & outlook

Model PFS and OS within **illness-death without recovery multistate model**.

Advantages:

- Compared to LFTM avoids **questionable** and **uncheckable** assumptions,
- properties of PFS and OS induced through **transparent assumption on  $X(t)$** ,
- allows for **straightforward derivation** of survival functions and correlation for parametric models, e.g. no need to assume common Weibull shape parameter as in [Li and Zhang \(2015\)](#) to get tractable formulas,
- allows for **parametric** and **nonparametric** estimation and **inference** (at least) in Markov models using standard multistate modelling tools,
- engine to simulate PFS and OS times.

Outlook:

- How to best extrapolate tail of nonparametric survival function estimates?
- Shorten time for computation of bootstrap confidence intervals?
- R package?

# Multistate models for early decision-making

**How do we typically decide whether  
to move an oncology molecule  
into Phase 3?**

# Decision-making in early oncology development

- ① Small single-arm trial for **experimental** drug (e.g.  $n = 40$ ).
- ② Response proportion, duration of response.
- ③ Compare to “corresponding” quantities from literature for **control** treatment.

But:

- **P(wrong decision)** may be high.
- Primary endpoint in Phase 3: **Overall survival**.



**Proposal:**

**Decide in early phase based  
on OS prediction.**

**Decrease  $P(\text{wrong decision})$ .**

# Challenges and proposal

## Challenges:

- ① Response-type endpoint?
- ② Surrogacy? **Poor** in many indications.
- ③ Immunotherapy (CIT): no effect on response, relevant OS effect.
- ④ **Non-randomized** comparison  $\Rightarrow$  confounding.

Proposal: Base decision-making on **OS prediction from multistate model**.

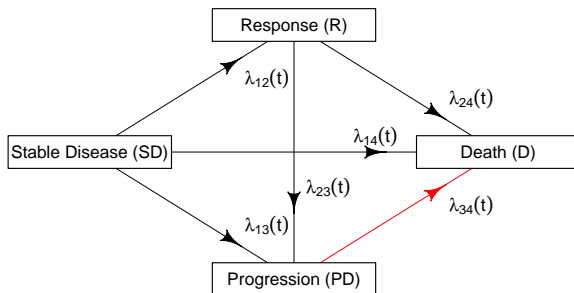
- ① **Predicted survival function for experimental arm.**
- ② Combine  $S_{\text{exp}}$  with  $S_{\text{control}}$  to get **predicted OS HR**.
- ③ Experimental drug might act on certain transitions only  $\Rightarrow$  not captured through simple modelling of OS. Potential **efficiency gain!**
- ④ **Propensity scoring.**

**Idealized scenario: Retrospective data  
from Phase 3 RCTs.**

**Long-term follow-up in both arms.**

**Randomization  $\Rightarrow$  no confounding.**

# Multistate model for early decision-making



- Follow-up of patient until **PD or death without PD**.
- Post-progression hazard  $\lambda_{34}$ : **borrowing** from historical data.
- Transitions  $SD \rightarrow D$ ,  $R \rightarrow D$  rare, hazards  $\approx$  same in both arms.
- Markov assumption.

## Predicted survival function in experimental arm, $S_{\text{exp}}$

Compute transition probabilities for each transition.

$$S_{\text{exp}}(t) = 1 - \left( P_{SD \rightarrow D}(0, t) + P_{SD \rightarrow \text{PD} \rightarrow \text{D}}(0, t) + P_{SD \rightarrow R \rightarrow D}(0, t) + P_{SD \rightarrow R \rightarrow \text{PD} \rightarrow \text{D}}(0, t) \right).$$

$\lambda_{34}$  corresponding to  $\text{PD} \rightarrow \text{D}$  transition borrowed from historical data.

# Historical borrowing for $\lambda_{34}$

Experimental treatment expected to provide benefit **beyond PD**?

**No:**

- E.g. chemotherapy or antibody-dependent cellular cytotoxicity.
- **Plug-in** hazard function estimate from historical control.
- No post-PD information required for experimental arm.

**Yes:**

- E.g. chemoimmunotherapy.
- Estimate post-PD hazard ratio assuming **proportionality**.
- How much post-PD deaths needed in experimental arm to reliably **estimate post-PD HR**?

## Benefit beyond PD: Oak

Previously treated non-small-cell lung cancer.

Rittmeyer *et al.* (2017).

	Atezolizumab	Chemotherapy	Hazard ratio
Effect post-PD	expected	not expected	
Objective Response	58 (13.6%)	57 (13.4%)	
Duration of Response	26.3 (10 - $\infty$ )	6.2 (4.9 - 7.6)	
Overall Survival			0.73 (0.62, 0.87)



**If this were early phase data -  
would you initiate Phase 3?**

**Competitors used this  
mechanism of action.**

# OS prediction when post-PD hazards assumed proportional

Random variable:

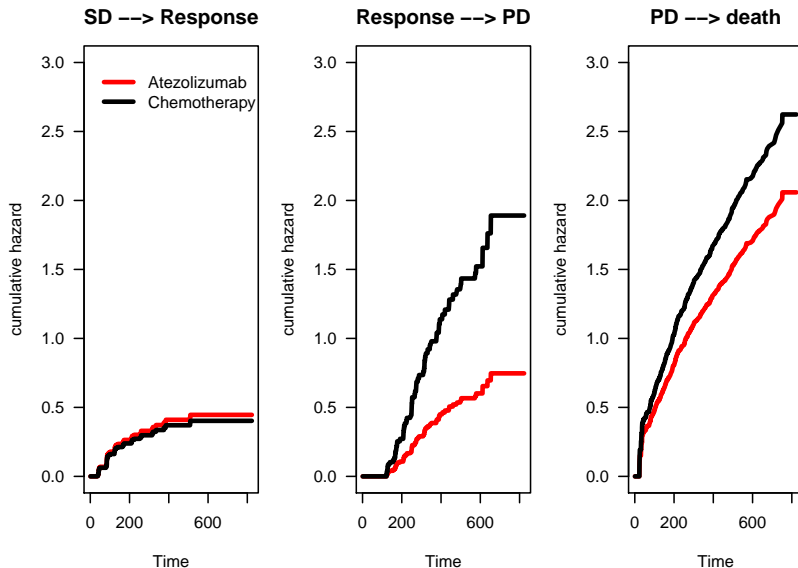
$$Z = \begin{cases} 0 & \text{if patient in control,} \\ 1 & \text{if in experimental group.} \end{cases}$$

$$\lambda_{34}(t | Z) = \lambda_{34,0}(t) \exp(\beta_{34}Z)$$

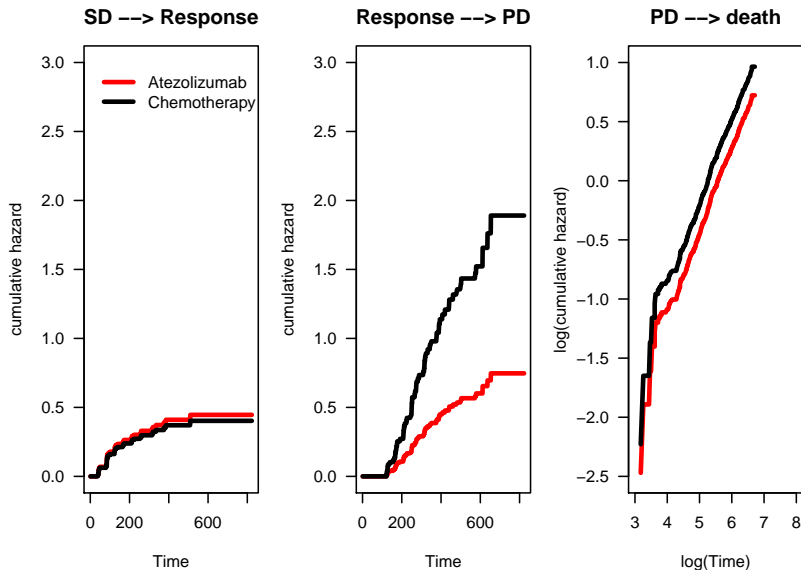
Baseline hazard  $\lambda_{34,0}$  **estimated from both arms combined.**

**Post-progression** hazard ratio  $\beta_{34}$ ?

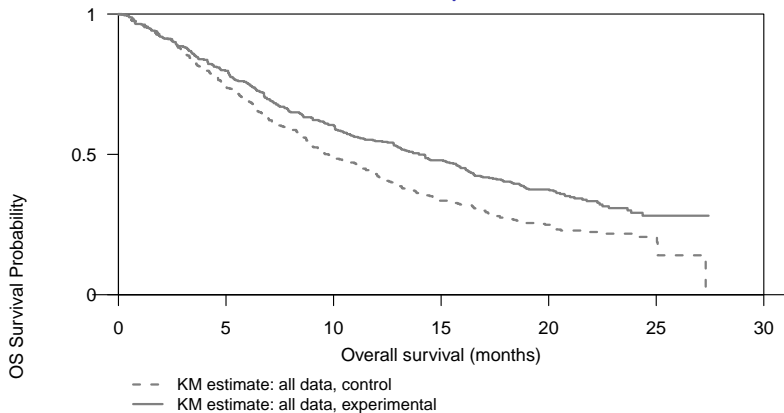
# Oak: raw cumulative hazard estimates (of interest)



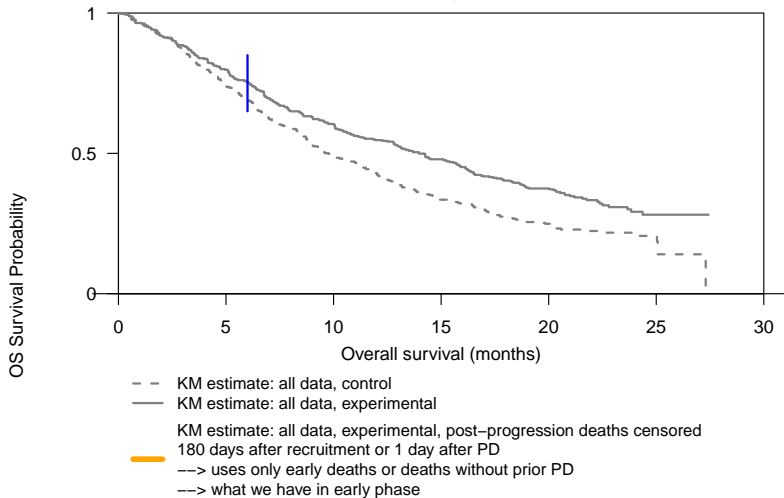
# Oak: raw cumulative hazard estimates (of interest)



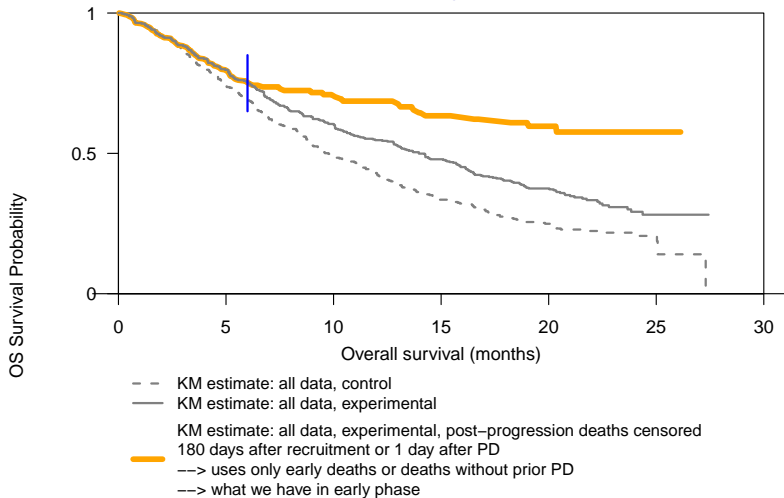
## Oak: estimates / predictions of $S_{\text{exp}}$



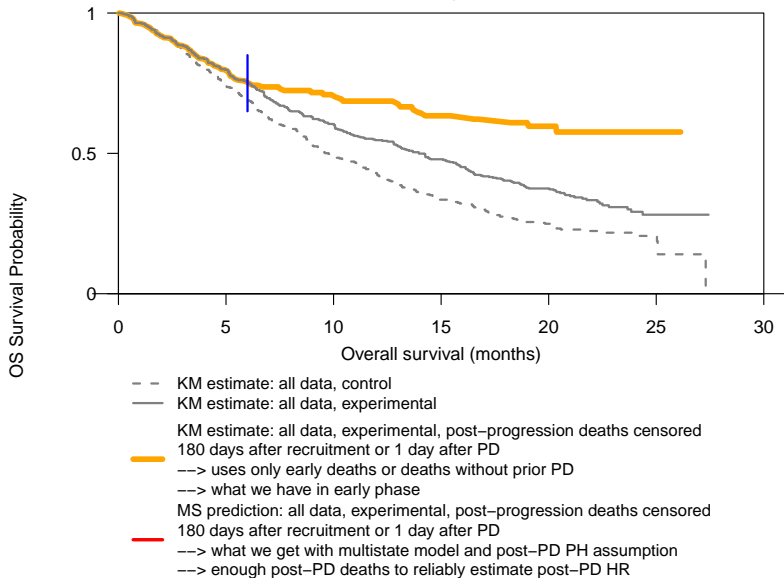
## Oak: estimates / predictions of $S_{\text{exp}}$



## Oak: estimates / predictions of $S_{\text{exp}}$

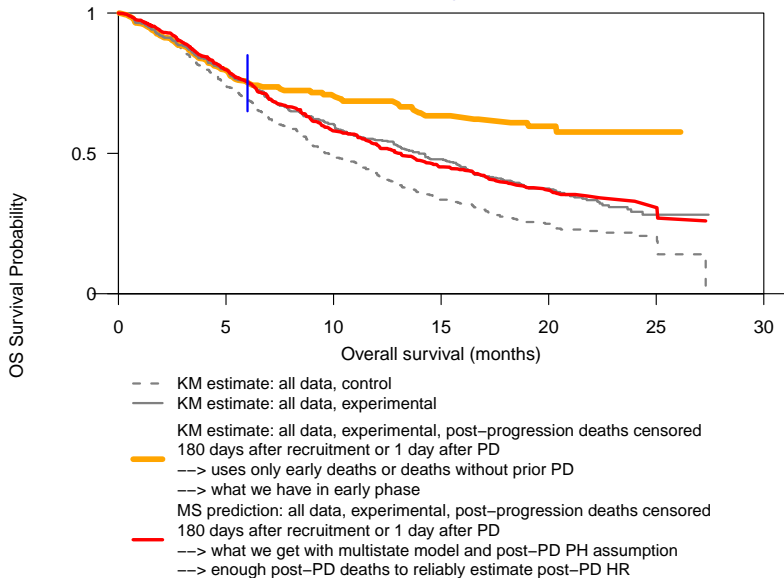


# Oak: estimates / predictions of $S_{\text{exp}}$





## Oak: estimates / predictions of $S_{\text{exp}}$



**Early phase decision based on  
multistate prediction:**

**$P(\text{wrong decision})?$**

# OS prediction from mimicked early phase data

Historical control: Oak control arm data.

**False-positive** decision: Sample early phase trial from Oak control arm.

**False-negative** decision: Sample early phase trial from Oak experimental arm.

Sample early phase trial:

- 40 patients,
- 6 months uniform recruitment,
- analysis 15 months after first patient entered,
- censor post-PD follow-up **one day after PD**,
- estimate  $\lambda_{12}, \lambda_{13}, \lambda_{14}, \lambda_{23}, \lambda_{24}$  from this data.

Cox regression for post-PD transition  $\Rightarrow \hat{\lambda}_{34}(t|Z)$ .

Compute prediction of  $S_{\text{exp}}$ .

# OS HR prediction based on early phase trial

Approximate HR by fitting exponential distribution to both arms  $\Rightarrow \widehat{HR}$ .

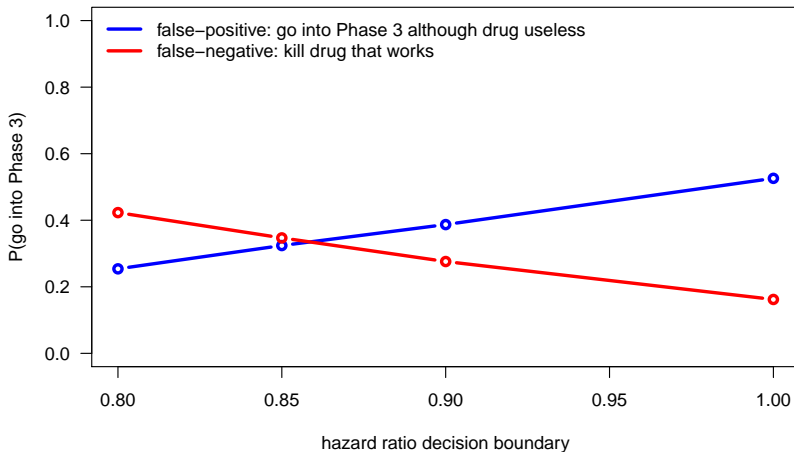
Decision to move to Phase 3:  $\widehat{HR} \leq \text{boundary} \in \{0.80, 0.85, 0.90, 1.00\}$ .

Repeat 1000 times.

Resampling  $\Rightarrow$  **quantification of uncertainty**.

## Oak: P(wrong decision)

**P(go into Phase 3) = P(approximated HR  $\leq$  boundary)**



**How many post-PD deaths to  
estimate HR of PD  $\rightarrow$  death transition?**

**How many post-PD deaths to  
estimate HR of PD  $\rightarrow$  death transition?**

**Ask during Q&A.**

## Conclusions for early-decision making proposal



# Conclusions

Early phase decision-making based on **multistate OS prediction**:

- Assumption on  $\lambda_{34} \Rightarrow$  need to understand **disease and treatment**.
- **Avoids difficulty in interpretation of response-type endpoints**.
- Feasibility assessed in **idealized scenario**.
- Recommendation **how much post-PD follow-up** needed to estimate  $\beta_{34}$ .
- Needs **long-term individual-patient** data in control arm!

**What about confounding?**

**Real-world data as historical control.**

**Combine proposal with propensity scoring.**

# Conclusions

# Multistate models

**Multistate models** useful:

- Canonical **extension of survival analysis**.
- Get more **insight** in how disease and drug work.
- **Prediction** in well-specified, as opposed to black-box, model.
- **Jointly** model three key oncology endpoints: response, PFS, OS.
- Applications by no means restricted to oncology!

Many potential applications:

- Improved **early stage decision-making**  $\Rightarrow$  [Beyer et al. \(2019\)](#).
- Improved **communication** of effect and optimized **sample size** computation.
- Bivariate modelling of PFS and OS to help inform **surrogacy** questions  $\Rightarrow$  [Meller et al. \(2019\)](#).

**Thank you for your attention.**

**[kaspar.rufibach@roche.com](mailto:kaspar.rufibach@roche.com)**

**<http://www.kasparrufibach.ch>**

** [numbersman77](#)**

** [numbersman77](#)**

# References I

- ▶ Aalen, O., Borgan, Ø., and Gjessing, H. (2008). *Survival and event history analysis: a process point of view*. Springer Science & Business Media.
- ▶ Aalen, O. O. (1987). Dynamic modelling and causality. *Scandinavian Actuarial Journal*, **1987**(3-4), 177–190.
- ▶ Aalen, O. O. and Johansen, S. (1978). An empirical transition matrix for non-homogeneous markov chains based on censored observations. *Scandinavian Journal of Statistics*, **5**(3), 141–150.
- ▶ Andersen, P. K., Borgan, Ø., Gill, R. D., and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer.
- ▶ Baselga, J. and Cortes, J. et. al. (2012). Pertuzumab plus trastuzumab plus docetaxel for metastatic breast cancer. *N. Engl. J. Med.*, **366**(2), 109–119.
- ▶ Beyer, U., Dejardin, D., Meller, M., Rufibach, K., and Burger, H. U. (2019). A multistate model for early decision making in oncology. *Biom J*, to appear.
- ▶ Beyersmann, J., Allignol, A., and Schumacher, M. (2012). *Competing Risks and Multistate Models with R*. Springer.
- ▶ Burzykowski, T., Molenberghs, G., Buyse, M., Geys, H., and Renard, D. (2001). Validation of surrogate end points in multiple randomized clinical trials with failure time end points. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **50**(4), 405–422.

# References II

- ▶ Buyse, M., Molenberghs, G., Paoletti, X., Oba, K., Alonso, A., Van der Elst, W., and Burzykowski, T. (2016). Statistical evaluation of surrogate endpoints with examples from cancer clinical trials. *Biom J*, **58**(1), 104–132.
- ▶ Emura, T., Nakatochi, M., Murotani, K., and Rondeau, V. (2017). A joint frailty-copula model between tumour progression and death for meta-analysis. *Stat Methods Med Res*, **26**(6), 2649–2666.
- ▶ Fleischer, F., Gaschler-Markefski, B., and Bluhmki, E. (2009). A statistical model for the dependence between progression-free survival and overall survival. *Stat. Med.*, **28**(21), 2669–2686.
- ▶ Fu, H., Wang, Y., Liu, J., Kulkarni, P. M., and Melemed, A. S. (2013). Joint modeling of progression-free survival and overall survival by a Bayesian normal induced copula estimation model. *Stat Med*, **32**(2), 240–254.
- ▶ Gaschler-Markefski, B., Schiefele, K., Hocke, J., and Fleischer, F. (2014). *Multi-state Models Used in Oncology Trials*, pages 283–304. Springer Berlin Heidelberg, Berlin, Heidelberg.
- ▶ Li, Y. and Zhang, Q. (2015). A Weibull multi-state model for the dependence of progression-free survival and overall survival. *Stat Med*, **34**(17), 2497–2513.
- ▶ Meller, M., Beyersmann, J., and Rufibach, K. (2019). Joint modeling of progression-free and overall survival and computation of correlation measures. *Statistics in medicine*, **38**, 4270–4289.

# References III

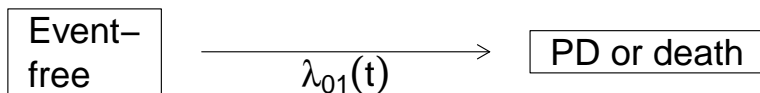
- ▶ Rittmeyer, A., Barlesi, F., Waterkamp, D., Park, K., Ciardiello, F., Von Pawel, J., Gadgeel, S. M., Hida, T., Kowalski, D. M., Dols, M. C., *et al.* (2017). Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (OAK): a phase 3, open-label, multicentre randomised controlled trial. *The Lancet*, **389**(10066), 255–265.
- ▶ Swain, S. M. and Baselga, J. (2015). Pertuzumab, trastuzumab, and docetaxel in HER2-positive metastatic breast cancer. *N. Engl. J. Med.*, **372**(8), 724–734.
- ▶ Weber, E. M. and Titman, A. C. (2019). Quantifying the association between progression-free survival and overall survival in oncology trials using kendall's  $\tau$ . *Statistics in medicine*, **38**, 703–719.



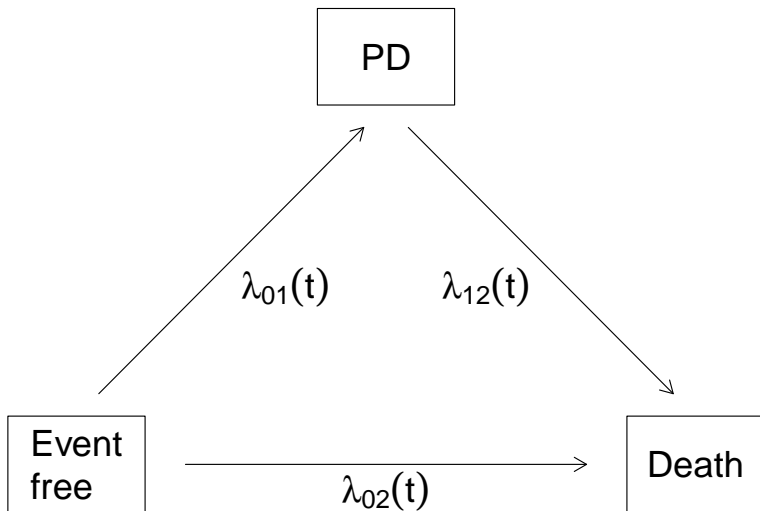
# Backup

# Multistate models

# Canonical extension of survival analysis



## Canonical extension of survival analysis



# Multistate models

Multistate model:

- 1-1 correspondence **hazard - probability** breaks down.
- Transition probabilities: (Markov) process  $X(t)_{t \geq 0}$  with state space  $\{0, 1, 2\} = \{\text{event-free, progression, death}\}$ . Then,

$$P_{lj}(s, t) := P(X_t = j | X_s = l, \text{Past}).$$

- Estimate  $P_{lj}$ 's **nonparametrically** by **Aalen-Johansen** estimator.
- PFS: Kaplan-Meier of time-to-progression simply censoring death is **biased**!
- OS: Aalen-Johansen offers **higher precision** compared to simple Kaplan-Meier estimate, [Andersen et al. \(1993\)](#) (p. 315 and Fig. IV.4.16).
- Markov assumption **stronger** than what is needed for Kaplan-Meier though.

# Prediction in multistate models

Rates (hazards, intensities):

- Modelling of effects of covariates on **transition hazards**.
- Hazard ratios (HR) from Cox regression.

Transition probabilities look at **cumulative effects**:

- Effects on transition probabilities may be different from what HRs suggest.
- **Intermediate** events in multistate model also contribute to cumulative effects.
- How to estimate such cumulative effects?

**Prediction from multistate model!**

# Prediction in multistate models

General problem: estimate **conditional probability of some future clinical event**, given

- event history,
- set of values for prognostic factors of a patient.

Derive formulas for these conditional probabilities, or simulate.

Final result: survival function for OS, as function of

- covariates and
- relevant **cumulative hazards**.

# PFS - OS



# Multistate vs. latent failure time model

Fleischer *et al.* (2009), Li and Zhang (2015): LFTM with **uncheckable** and **questionable** (**unrealistic?**) independence assumption.

Connections to multistate model? We are still figuring that out, work in progress.

Parametric models: formula for  $S_{PFS}$  identical for all three models below, and

- **Time-homogeneous Markov, Exponential:** model so simple that  $\nexists$  time-inhomogeneous Markov process.  $S_{OS}$  identical to Exponential LFTM.
- **Time-homogeneous Markov, Weibull:** formula for  $S_{OS}$  identical to Weibull LFTM  $\Rightarrow$  are model assumptions equivalent? **No!**
- **Time-inhomogeneous Markov, Weibull:** formulas for  $S_{OS}$  are **different**.

**BUT:** values of estimated parameters differ between LFTM and multistate for **all three parametric models**, as **likelihoods differ!**

Not clear (?) how to nonparametrically estimate LFTM  $\Rightarrow$  possible for (Markov) multistate.

# Assumptions for multistate model

# Assumptions for multistate model

Multistate model **sufficiently smooth** so that following intensities exist:

$$\begin{aligned}\alpha_{0j}(t) &= \lim_{\Delta t \searrow 0} \frac{P(\text{PFS} \in [t, t + \Delta t), X(\text{PFS}) = j \mid \text{PFS} \geq t)}{\Delta t}, j = 1, 2, \\ \alpha_{12}(t; t_1) &= \lim_{\Delta t \searrow 0} \frac{P(X(t + \Delta t) = 2 \mid X(t-) = 1, \text{PFS} = t_1)}{\Delta t} \\ &= \lim_{\Delta t \searrow 0} \frac{P(\text{OS} - \text{PFS} \in [t - t_1, t - t_1 + \Delta t) \mid \text{OS} \geq t, \text{PFS} = t_1)}{\Delta t} \quad \text{for } \mathbf{t_1} < \mathbf{t}.\end{aligned}$$

$t_1$ : observed PFS time, i.e. time when leaving state 0.

# Assumptions for multistate model

$X(t)$  **Markov**:

- **Time-inhomogeneous**: intensity of death after progression does not depend on time of progression,  $\alpha_{12}(t; t_1) = \alpha_{12}(t)$  for all  $t_1 < t$ .
- **Homogeneous**: intensities are time-constant, i.e. **Exponential**,  $\alpha_{ij}(t) = \alpha_{ij}$ ,  $i, j = 0, 1, 2$ .

$X(t)$  **non-Markov** (= semi-Markov for illness-death model without recovery):

- Intensities depend on state patient is in at  $s$  and entire history  $\leq s$ , i.e. all transitions.
- Relevant for  $1 \rightarrow 2$  transition only, as  $0 \rightarrow 1, 2$  are rooted in initial state 0.

As soon as a quantity depends on  **$1 \rightarrow 2$  transition** we need to be specific about assumption on  $X(t)$ .

# Illness-death multistate model for PFS and OS

Transition probabilities to move from state  $l$  at time  $s$  to state  $m$  at time  $t$ :

$$P_{lm}(s, t) := P(X(t) = m | X(s) = l, \text{history}).$$

Illness-death model w/o recovery,  $P_{lm}$  as functions of transition intensities, Aalen *et al.* (2008):

$$\begin{aligned}P_{00}(s, t) &= \exp\left(-\int_s^t \alpha_{01}(u) + \alpha_{02}(u) \, du\right), \\P_{11}(s, t; \mathbf{t}_1) &= \exp\left(-\int_s^t \alpha_{12}(u; \mathbf{t}_1) \, du\right), \\P_{22}(s, t) &= 1, \\P_{01}(s, t) &= \int_s^t P_{00}(s, u_-) \alpha_{01}(u) P_{11}(u, t; u) \, du, \\P_{12}(s, t; \mathbf{t}_1) &= 1 - P_{11}(s, t; \mathbf{t}_1), \\P_{02}(s, t) &= 1 - \left(P_{00}(s, t) + P_{01}(s, t)\right).\end{aligned}$$

If  $X(t)$  non-Markov:

- $P_{11}$  and  $P_{12}$  depend on **PFS time  $t_1$** .
- Although  $P_{01}, P_{02}$  depend on  $\alpha_{12}$  they **do not depend on  $t_1$** .

# Intuition behind transition probabilities

$P_{00}(s, t)$ ,  $P_{11}(s, t; t_1)$ : exp of cumulative hazards  $\Rightarrow$  standard survival functions.

$P_{01}(s, t) = \int_s^t P_{00}(s, u_-) \alpha_{01}(u) P_{11}(u, t; u) du$ : integral of

- $P_{00}(s, u_-) \alpha_{01}(u)$ : “infinitesimal probabilities” to move from 0 to 1 at time  $u$ ,  $u \in (s, t]$ ,
- $P_{11}(u, t; u)$ : subsequently stay in state 1 until at least time  $t$ , with progression happened in  $u$ .

# Illness-death multistate model for PFS and OS

**Marginal** distributions:

$$\begin{aligned}S_{PFS}(t) &= P(\text{PFS} > t) = P_{00}(0, t), \\S_{OS}(t) &= P(\text{OS} > t) = P_{00}(0, t) + P_{01}(0, t),\end{aligned}$$

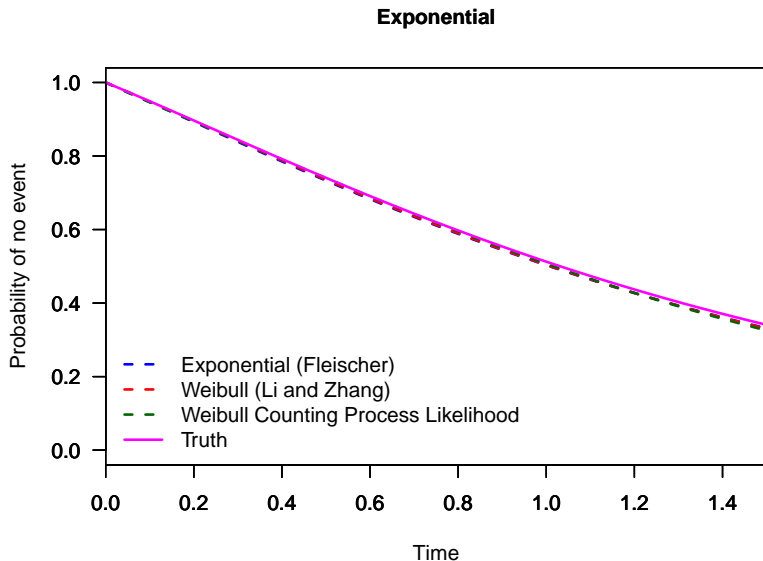
**Joint** distribution:

$$\begin{aligned}P(\text{PFS} \leq u, \text{OS} \leq v) &= P(X(u) \in \{1, 2\}, X(v) = 2) \\&= P(X(u) = 1, X(v) = 2) + P(X(u) = 2) \\&= P(X(v) = 2 | X(u) = 1) \cdot P(X(u) = 1 | X(0) = 0) \\&\quad + P(X(u) = 2 | X(0) = 0) \\&= P(X(v) = 2 | X(u) = 1) \cdot P_{01}(0, u) + P_{02}(0, u).\end{aligned}$$

**X inhomogeneous Markov:**  $P(X(v) = 2 | X(u) = 1) = P_{12}(u, v)$  independent of progression time  $t_1 \leq u$ .

**X non-Markov:** integrate  $P_{12}(u, v; t_1)$  over conditional distribution of all possible progression times  $t_1 \leq u \Rightarrow$  final formula tedious.

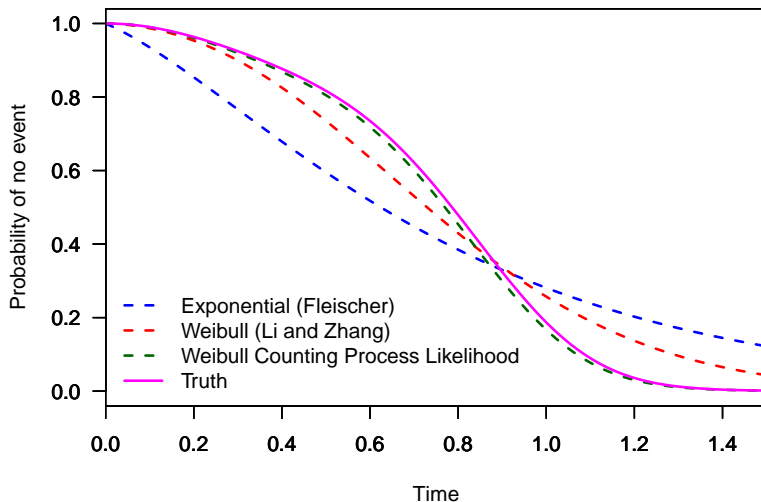
## Results: $S_{OS}$ for Exponential



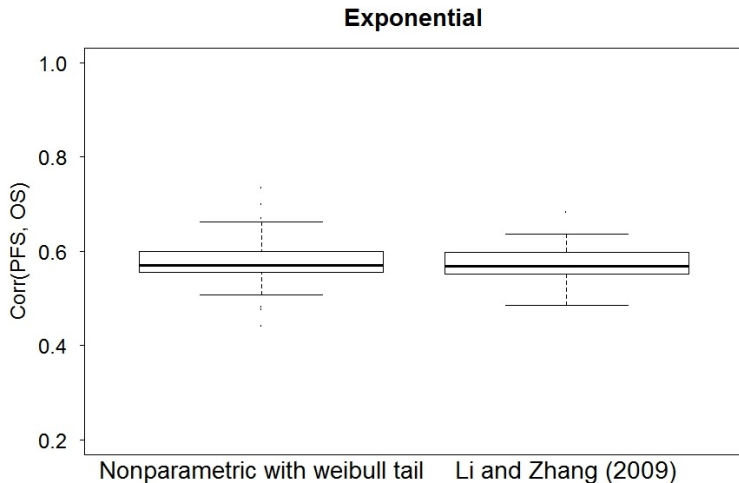


## Results: $S_{OS}$ for Weibull

Data from time-inhomogeneous Markov,  
Weibull with different shape,  $n = 500$

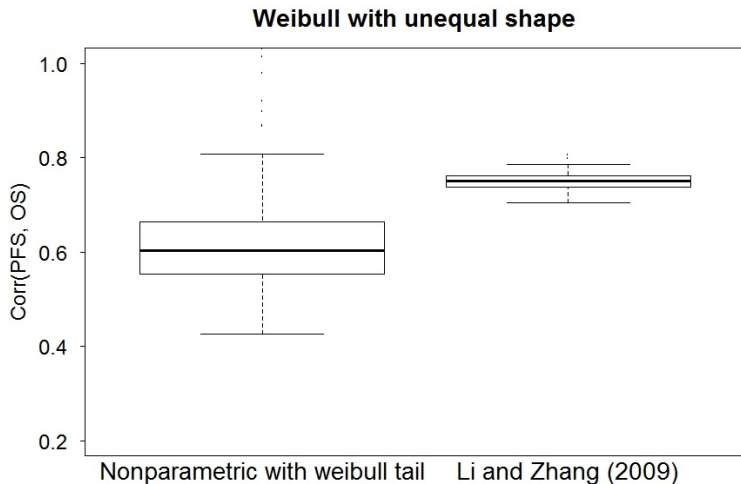


## Results: correlations Exponential



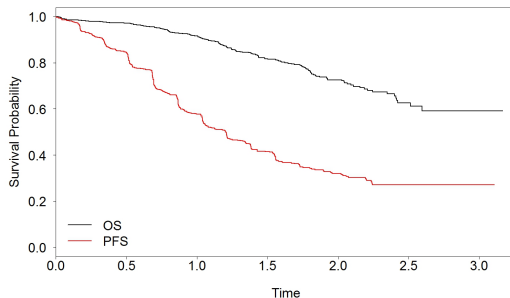
Corr(PFS, OS) for 200 simulated dataset from time-inhomogeneous Markov process.

## Results: correlations Weibull



Corr(PFS, OS) for 200 simulated dataset from time-inhomogeneous Markov process.

## Results: CLEOPATRA, Baselga and Cortes (2012).



	Exponential	Weibull	Weibull Markov	Nonparametric Markov
Corr(PFS, OS)	0.611	0.643	0.483	0.450
95% Bootstrap CI	[0.541; 0.673]	[0.584; 0.699]	[0.342; 0.643]	[0.297; 0.655]

**Table:** Correlation between PFS and OS in CLEOPATRA (1000 bootstrap samples).

# Early decision-making

# How many post-PD deaths needed?

Assumption:

$$\lambda_{34}(t | Z) = \lambda_{34,0}(t) \exp(\beta_{34}Z).$$

How many post-PD deaths needed in **experimental** arm to reliably estimate  $\lambda_{34}$ ?

Planning stage: only data for control arm are available.

- Fit multistate model to control data.
- Assume transition-specific hazard ratios corresponding to **clinically meaningful OS effect**.
- **Simulate**.

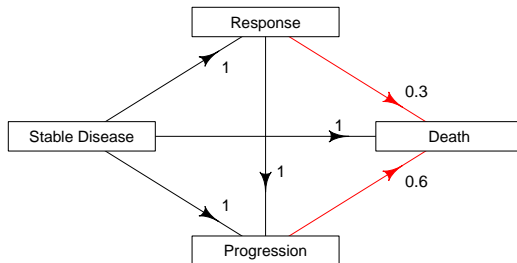
Various scenarios for post-PD follow-up time.

## Simulation details – mimick Oak

**NOT** power computation for hypothesis test – sample size too small anyway.

Rather: find cutoff timepoint from which on OS HR estimate remains **stable**.

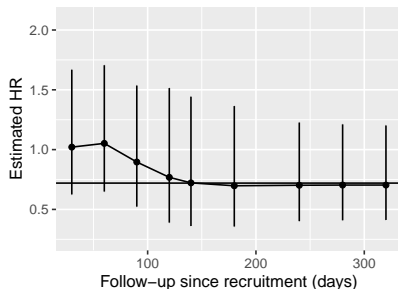
- Simulate 40 patient from experimental arm as before.



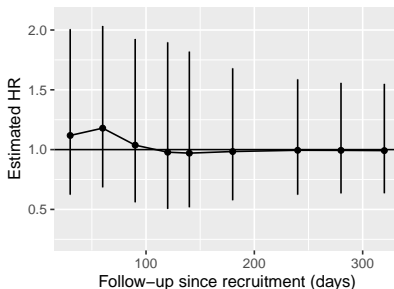
- Resulting **OS HR = 0.73**. Close to Oak OS HR.
- Follow-up post-PD for experimental arm truncated at 30, 60, 90, 120, 150, 180 and 240 days after recruitment.
- Repeat 1000 times.

# Stability of hazard ratio estimate

**A** Treatment Effect



**B** No Treatment Effect



**180-240 days sufficient** to obtain stable point estimate over time.

Typical early phase follow-up: Post-PD deaths censored **180 days after recruitment** in experimental arm.



## Example 1: Cleopatra

# Cleopatra

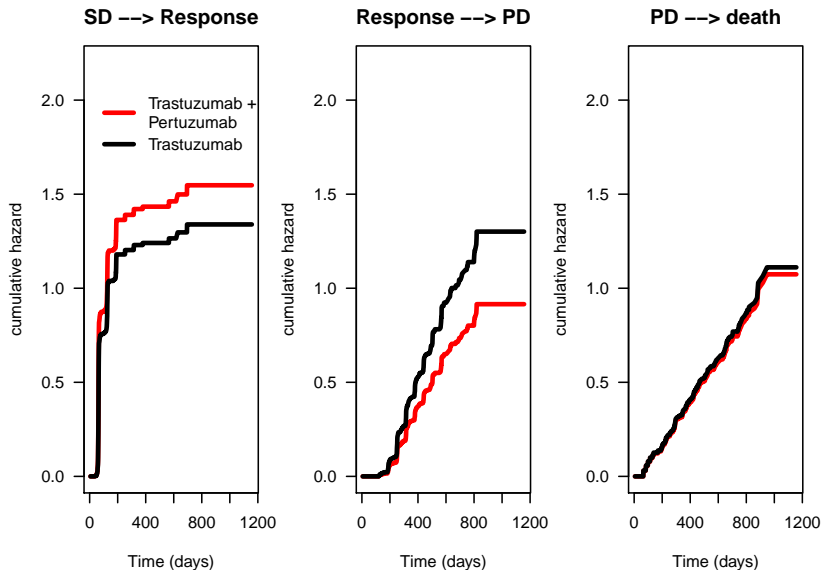
Baselga and Cortes (2012), Swain and Baselga (2015).

Previously untreated HER2-positive metastatic breast cancer patients.

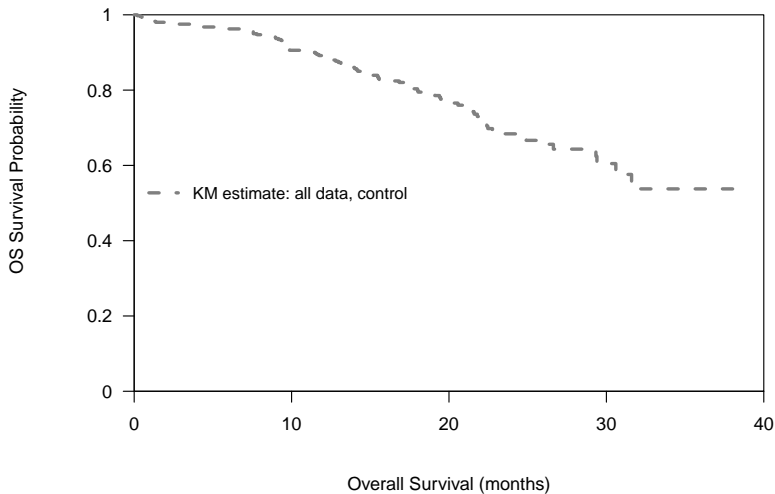
	Pertuzumab+Trastuzumab	Trastuzumab	HR (95% CI)
<b>Survival</b>	N=402	N=406	
Overall Survival			<b>0.64</b> <b>(0.47,0.88)</b>
Progression-free Survival			0.62 (0.51,0.75)
<b>Response</b>	N=343	N=336	
Objective Response	275 (80.2%)	233 (69.3%)	
Stable Disease	50 (14.6%)	70 (20.8%)	
Progressive Disease	13 (3.8%)	28 (8.3%)	
<b>Duration of Response</b>	N=275	N=233	
Median (months, 95% CI)	<b>20.2 (16.0,24.0)</b>	<b>12.5 (10.0-15.0)</b>	

- Moderate difference in response.
- Prolonged **duration of response** in experimental arm.
- Clear OS benefit.
- Experimental treatment induces antibody-dependent cellular cytotoxicity  $\Rightarrow$  no benefit beyond PD expected  $\Rightarrow$   $\lambda_{34}$  **same in both arms.**

# Cleopatra: raw cumulative hazard estimates (of interest)



## Cleopatra: estimates / predictions of $S_{\text{exp}}$



406

347

150

28

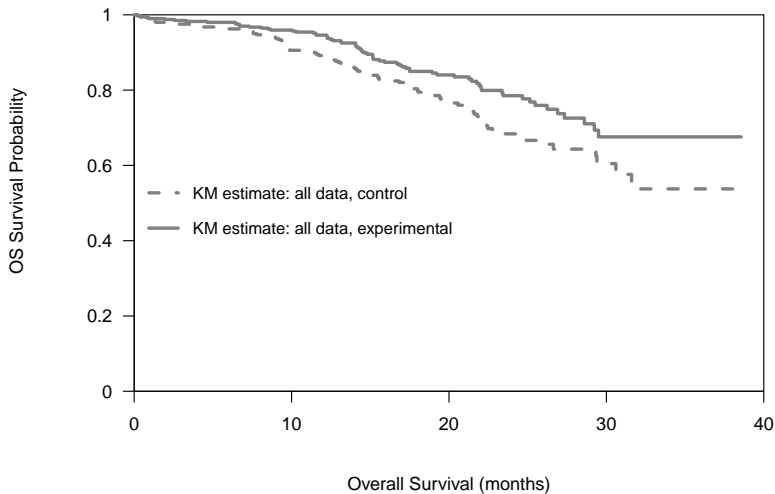
402

368

163

35

## Cleopatra: estimates / predictions of $S_{\text{exp}}$



406

347

150

28

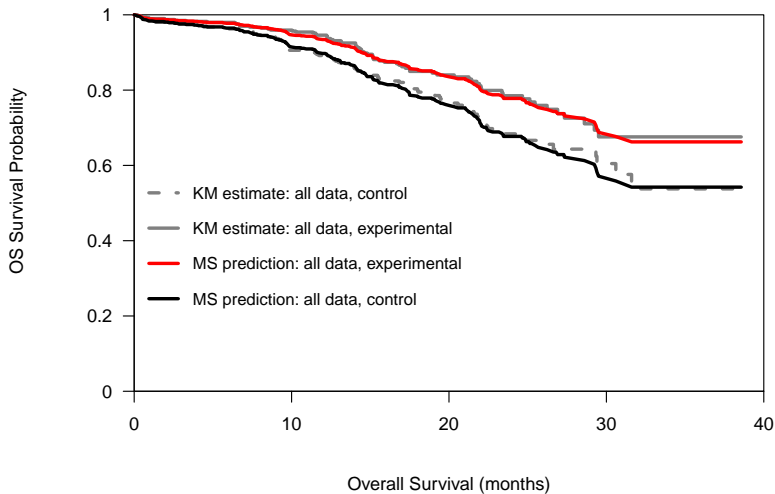
402

368

163

35

## Cleopatra: estimates / predictions of $S_{\text{exp}}$



406

347

150

28

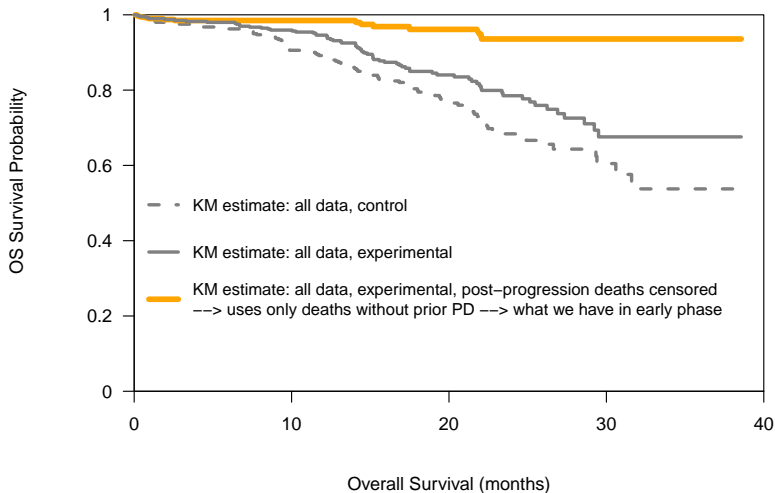
402

368

163

35

# Cleopatra: estimates / predictions of $S_{\text{exp}}$



406

347

150

28

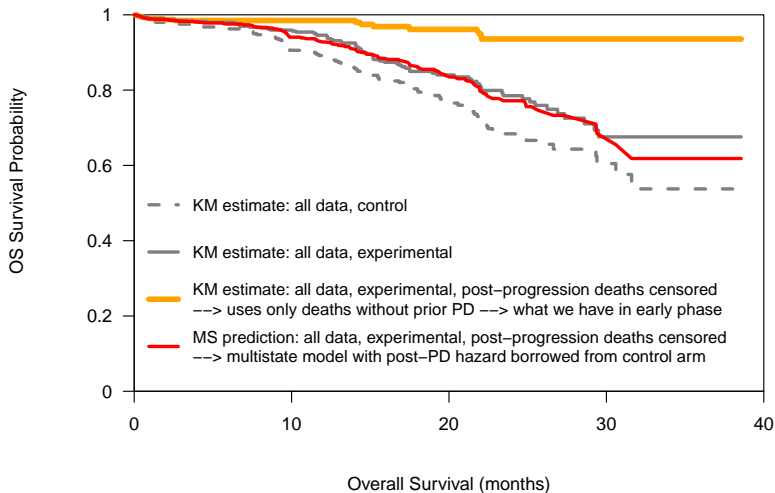
402

368

163

35

# Cleopatra: estimates / predictions of $S_{\text{exp}}$



406	347	150	28
402	368	163	35



# Conclusions for Cleopatra

For estimated / predicted survival function in experimental arm, based on **all data**:

- Majority of patients dies after observed PD.
- KM estimate of simply censoring post-PD deaths does not work  $\Rightarrow$  very **few deaths observed**.
- Multistate model prediction assuming post-PD hazards as in control provides good prediction.

**Early phase decision based on  
multistate prediction:**

**Operating characteristics?**

# OS prediction from mimicked early phase data

Sample early phase trial from **Cleopatra experimental arm**:

- 40 patients,
- 6 months uniform recruitment,
- analysis 15 months after first patient entered,
- censor post-PD follow-up **one day after PD**,
- estimate  $\lambda_{12}, \lambda_{13}, \lambda_{14}, \lambda_{23}, \lambda_{24}$  from this data,
- **borrow  $\hat{\lambda}_{34}$  from historical data** = Cleopatra control arm in idealized scenario,
- compute prediction of  $S_{\text{exp}}$  as described above.

# Resampling of operating characteristics

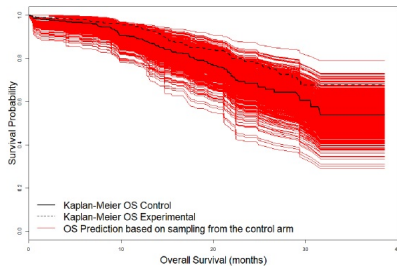
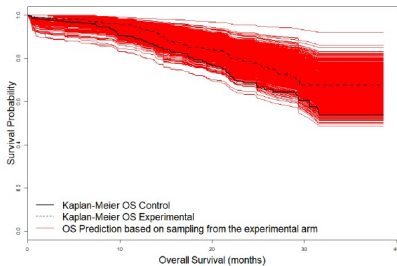
Setup:

- Use all data in control arm  $\Rightarrow$  corresponds to historical control.
- **False-positive** decision: Sample early phase trial from Cleopatra control arm.
- **False-negative** decision: Sample early phase trial from Cleopatra experimental arm.
- Approximate HR by fitting exponential distribution to both arms  $\Rightarrow \widehat{HR}$ .
- Decision to move to Phase 3:  $\widehat{HR} \leq \text{boundary} \in \{0.80, 0.85, 0.90, 1.00\}$ .
- Repeat 1000 times.

Resampling easily allows for **quantification of uncertainty**.

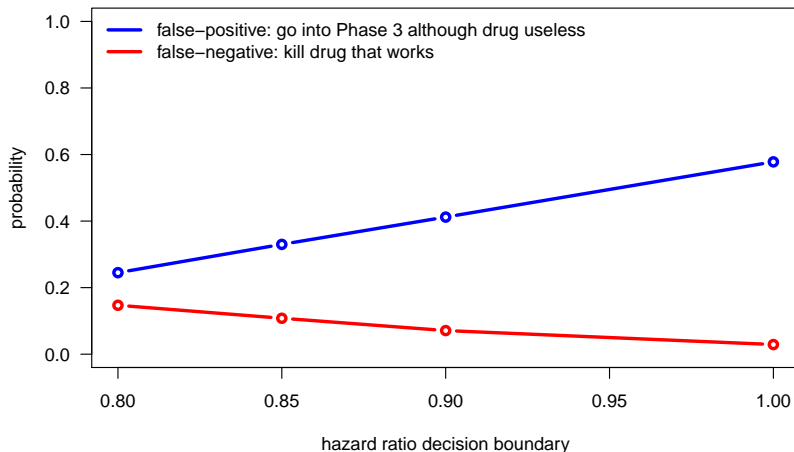
# Cleopatra: operating characteristics

Sampled from **experimental** and **control** arm.



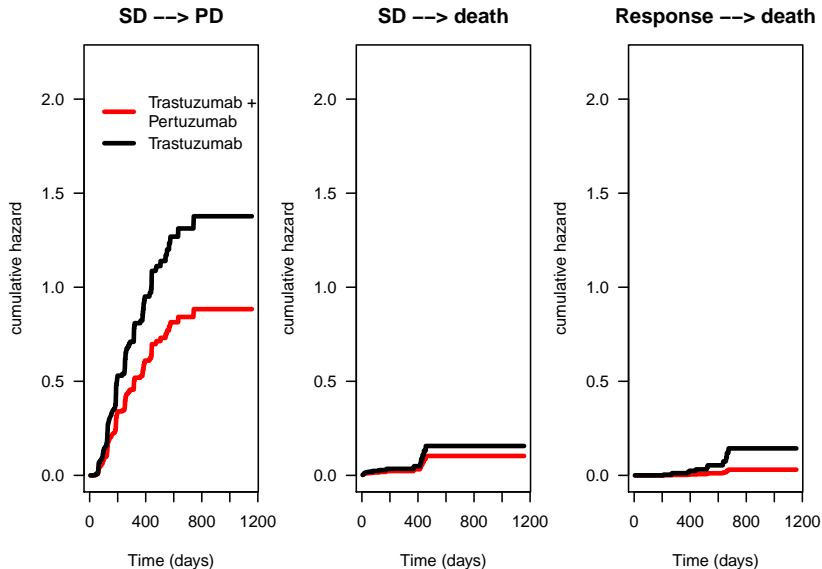
# Cleopatra: operating characteristics

probability to go into Phase 3:  $P(\text{approximated HR} \leq \text{boundary})$



Decision based on response:  $\approx 10\%$  difference, some prolongation of DOR  $\Rightarrow$  moved to Phase 3.

# Cleopatra: cumulative hazards of secondary interest



Previously treated non-small-cell lung cancer. [Rittmeyer et al. \(2017\)](#).

- Control: no benefit post-PD expected.
- Experimental: CIT  $\Rightarrow$  benefit post-PD expected.

	Atezolizumab	Chemotherapy	HR (95% CI)
<b>Survival</b>	N=425	N=425	
Overall Survival			0.73 (0.62,0.87)
Progression-free Survival			0.95 (0.82,1.10)
<b>Response</b>	N=425	N=425	
<b>Objective Response</b>	<b>58 (13.6%)</b>	<b>57 (13.4%)</b>	
Stable Disease	150 (35%)	177 (42%)	
Progressive Disease	187 (44%)	117 (28%)	
Duration of Response	N=58	N=57	
Median (months, 95% CI)	26.3 (10,NE)	6.2 (4.9-7.6)	

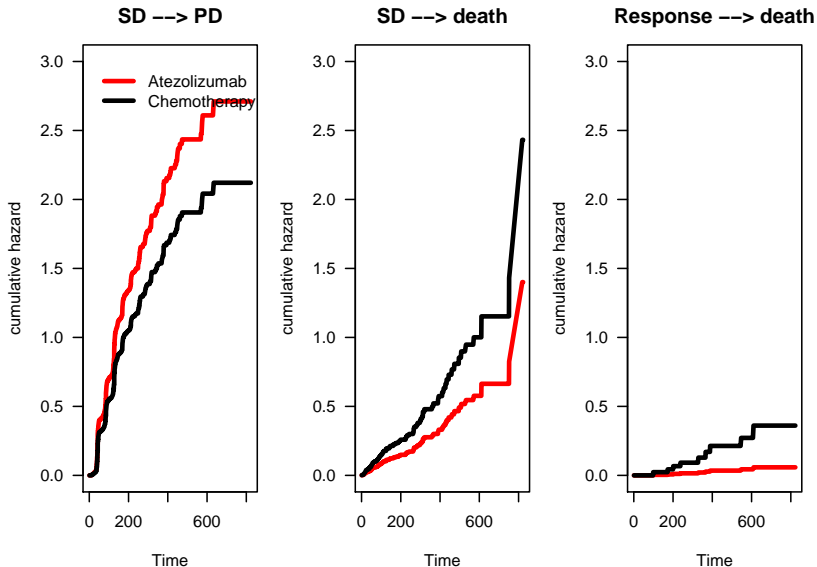
No observed difference in response.

Prolonged duration of response in experimental arm.

Clear survival benefit.

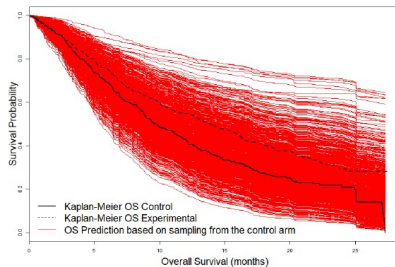
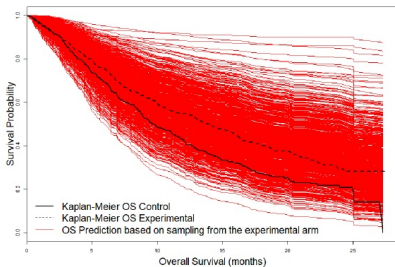


# Oak: cumulative hazards of secondary interest



# Oak: operating characteristics

Sampled from **experimental** and **control** arm.



# Non-proportional hazards via multistate model

**Immunotherapy:**  
**1) no difference in PFS,**  
**2) non-proportional hazards for OS.**

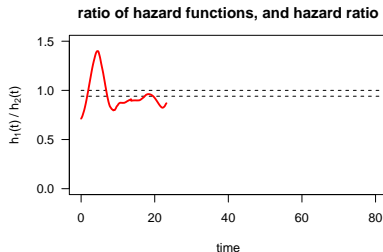
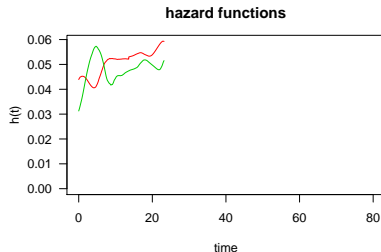
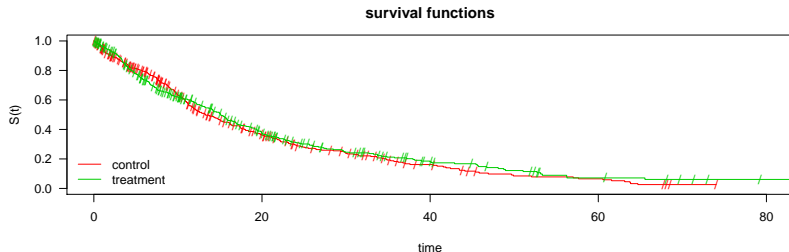
**How to quantify effect?**

# A fictional clinical trial

Simulated clinical trial:

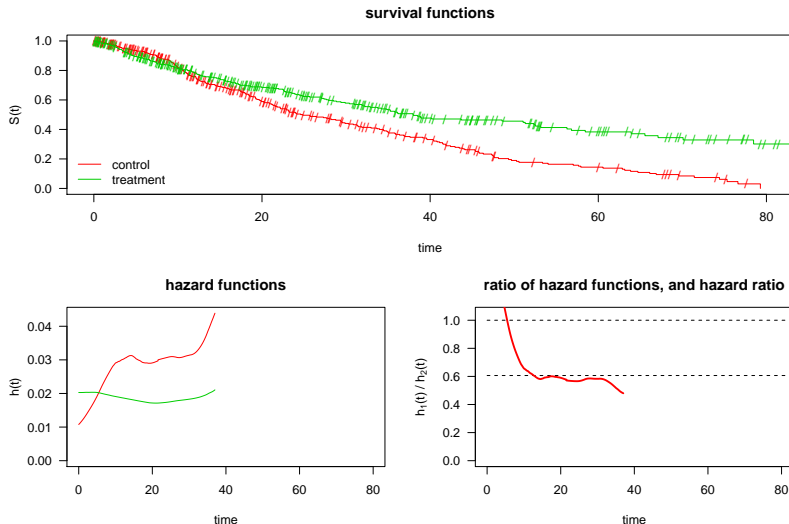
- 1:1 randomized, 400 and 400 patients per arm.
- No administrative censoring, but drop-out.

# PFS for simulated clinical trial



- Estimated hazard ratio: 0.94, 95% confidence interval [0.80, 1.11].
- Hypothesis test for PH:  $p = 0.24$ .

# OS for simulated clinical trial



- Estimated hazard ratio: 0.61, 95% confidence interval [0.50, 0.74].
- Hypothesis test for PH:  $p < 0.0001$ .

# Summarize treatment effect

**Non-proportional hazards** for OS. How to summarize effect of treatment?

Data was generated according to:

Transition	Control arm	Treatment arm
$0 \rightarrow 1$	$\lambda_{01}^c = \log(2)/25$	$\lambda_{01}^t = \lambda_{01}^c \cdot \mathbf{1}$
$0 \rightarrow 2$	$\lambda_{02}^c = \log(2)/30$	$\lambda_{02}^t = \lambda_{02}^c \cdot \mathbf{0.8}$
$1 \rightarrow 2$	$\lambda_{12}^c = \log(2)/15$	$\lambda_{12}^t = \lambda_{12}^c \cdot \mathbf{0.4}$

	coef	HR = exp(coef)	95% CI	p-value
transition event-free $\rightarrow$ PD	-0.04	0.96	[0.77, 1.19]	0.72
transition event-free $\rightarrow$ death	-0.09	0.91	[0.70, 1.18]	0.49
transition PD $\rightarrow$ death	-1.09	0.34	[0.24, 0.46]	< 0.0001

Gaschler-Markefski *et al.* (2014).



# Big vs. small data

# Big vs. small data

Often, information removed/alterd in **small data**:

- (Artificial) response categories instead of actual measurements: **dichotomization**,
- response proportions only: **ignoring the dynamics between states**,
- complicated subsets, e.g. those that respond only: **selection bias**,
- effect quantification in **one** number where biological process might suggest few numbers,
- ...

**Maximize** information from **small** data. AND look at **BIG** data.

Biostatisticians ideally placed to contribute to this!

# *Doing now what patients need next*

## **R version and packages used to generate these slides:**

R version: R version 3.6.0 (2019-04-26)

Base packages: stats / graphics / grDevices / utils / datasets / methods / base

Other packages: nls2 / proto / diagram / shape / ggplot2 / rocheBCE / muhaz / flexsurv / reporttools / xtable / mstate / etm / dplyr / mvna / prodlm / biostatKR / survival

This document was generated on 2019-11-14 at 23:27:30.