

Estimating survival benefit for health technology assessment

New challenges presented by immuno-oncology treatments?

BBS / PSI 1-Day Scientific Meeting: Empower the immune system to fight cancer

June 15, 2017

Dr Nicholas Latimer, Senior Research Fellow, NIHR Post-doctoral Fellow,
University of Sheffield, Sheffield, UK

Acknowledgements: Andrew Briggs and Scott Ramsey have allowed me to use

Plan

1. Survival modelling for HTA
2. Issues raised by immuno-oncology
3. Possible solutions (and limitations)
 - Flexible parametric models
 - Mixture models
 - Response-based models
4. Summary

HTA – objectives

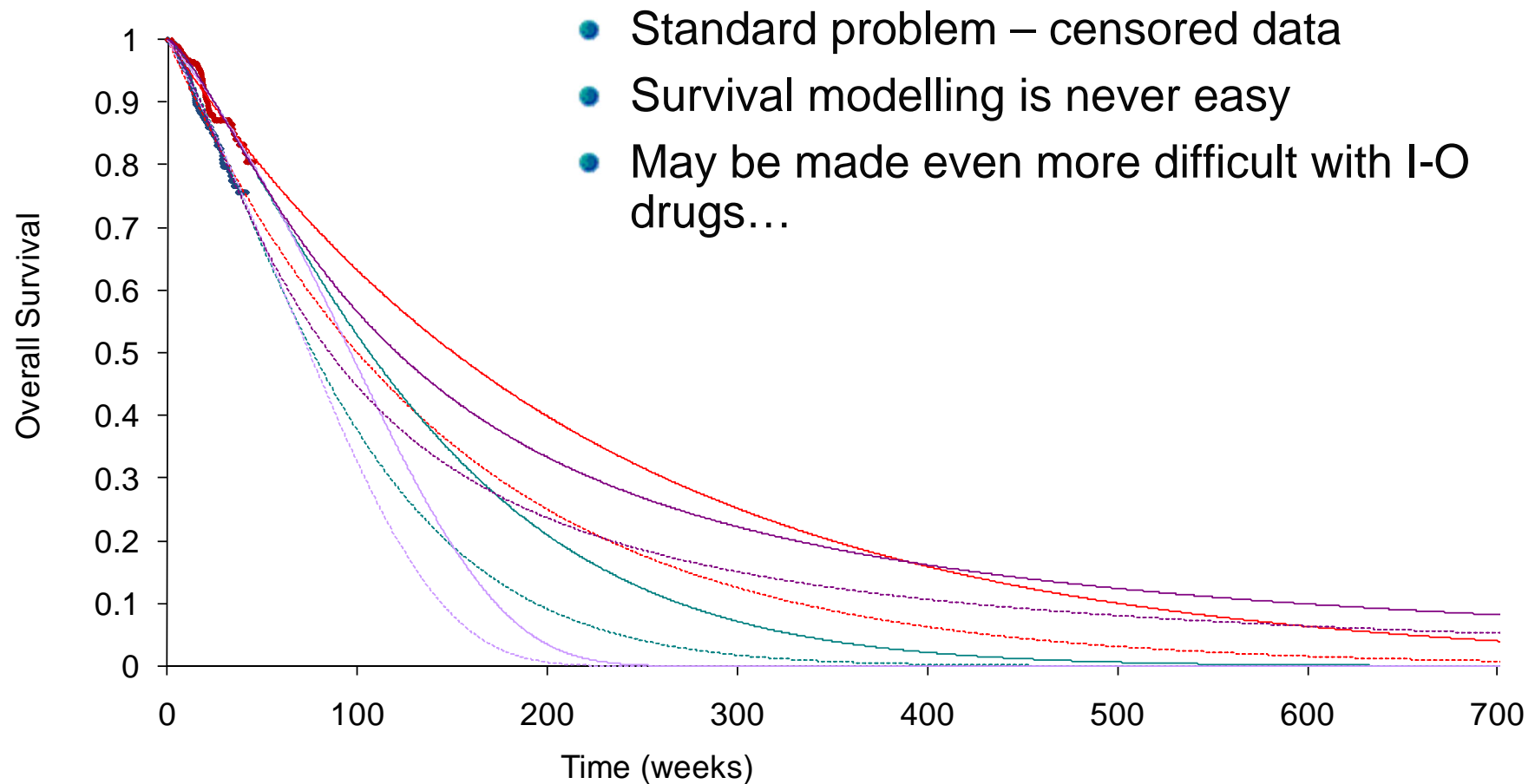
Allocation of scarce healthcare resources

Decisions need to be made based on treating the entire (eligible) disease population



Need to estimate mean survival advantage (not median)

HTA – objectives



The issue

- New I-O drug approvals increasingly characterised by:
 - Less mature data
 - Often without a control group
 - Intermediate endpoints rather than overall survival

AND

- Several agents appear to result in difficult-to-model survival curves
 - I-O drugs may be associated with a **delayed effect**, **long-term survivors** (a “**cure**” proportion) and therefore **complex hazard functions** with a **non-proportional treatment effect**

The issue

- Survival modelling is never straightforward, for any drug for any disease where we have to extrapolate into the future

Now we have fewer data and treatments that have increasingly complex effects



How do we deal with this?

Do we need new methods?

(Should we be using better methods anyway)?

Standard methods

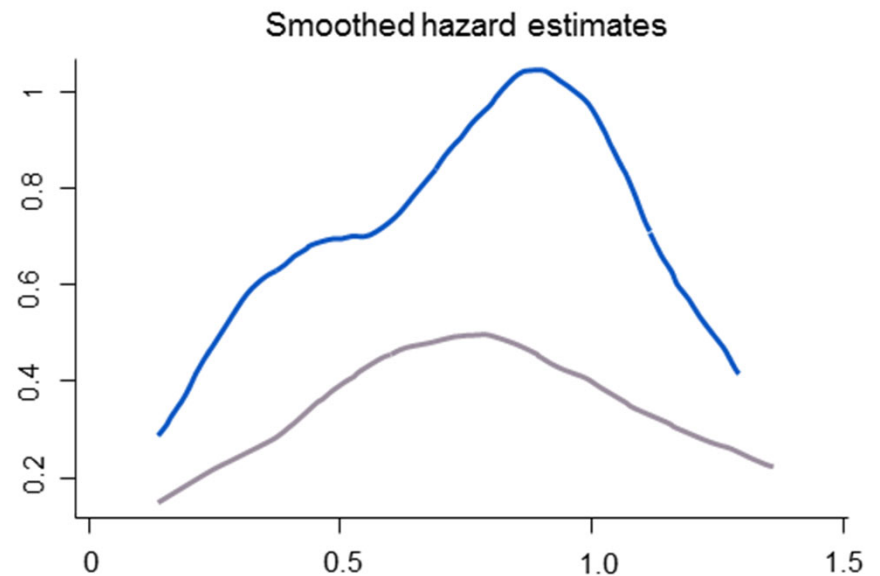
- In oncology HTAs standard parametric models are usually used to estimate long-term survival (e.g. Weibull, exponential, Gompertz...)
- These can be fitted separately to treatment arms to address non-PH
- But, they are also limited with regards to the hazards that they can represent (constant, monotonically increasing, monotonically decreasing...)

Standard methods

- I-O drugs may be associated with a complex hazard function

- Standard parametric models may not provide a good fit
- Survival estimates may be poor

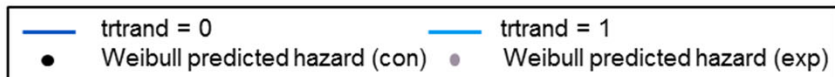
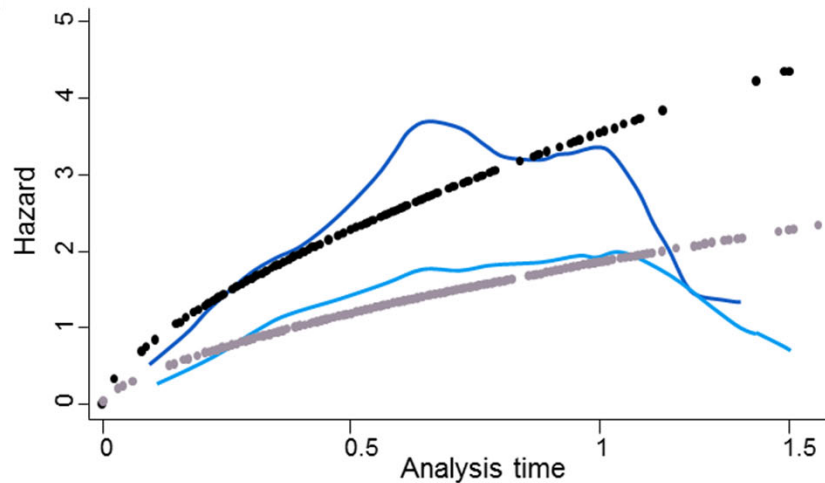
➔ What can we do?



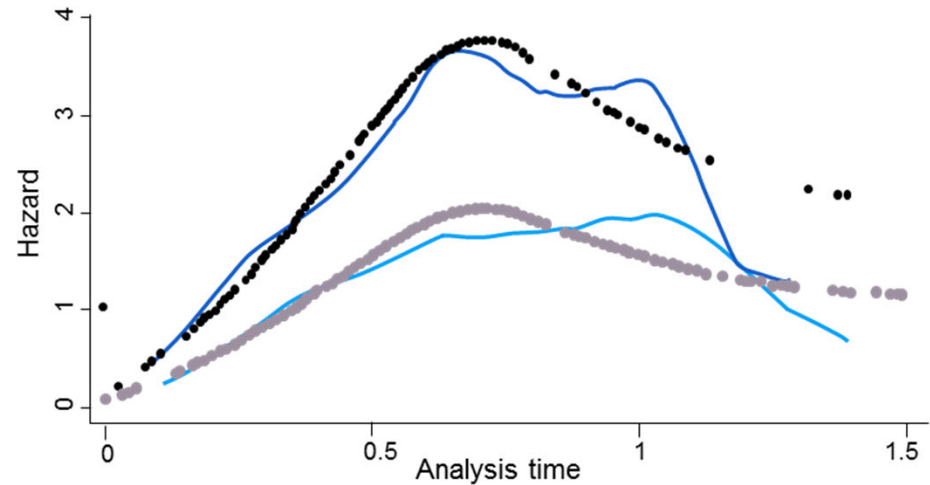
Solutions - FPMs

- Flexible parametric models use restricted cubic splines to estimate the shape of the log-cumulative hazard function
- Knots are positioned, usually placed at centiles of the distribution of log survival times, and sections of the curve separated by these knots are fitted
- FPMs can accurately reflect complex hazard functions, with turning points (Royston and Parmar, 2002; Rutherford, Crowther and Lambert, 2015)...

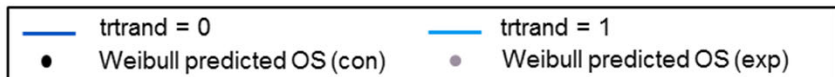
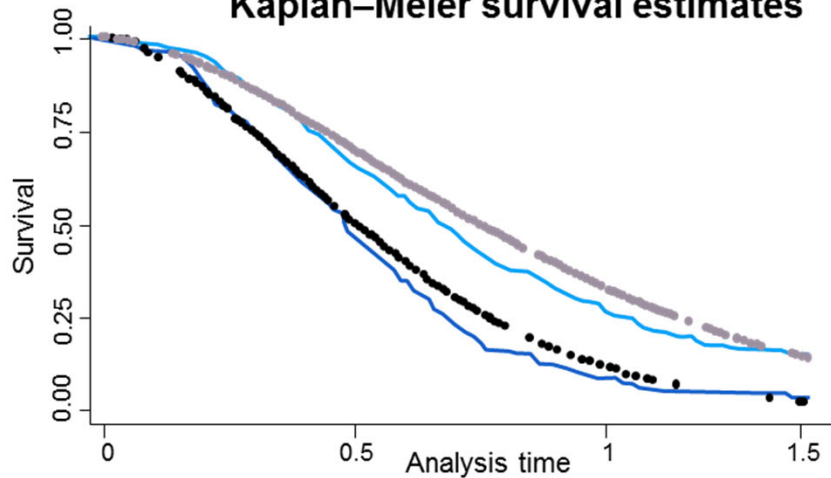
Smoothed hazard estimates



Smoothed hazard estimates

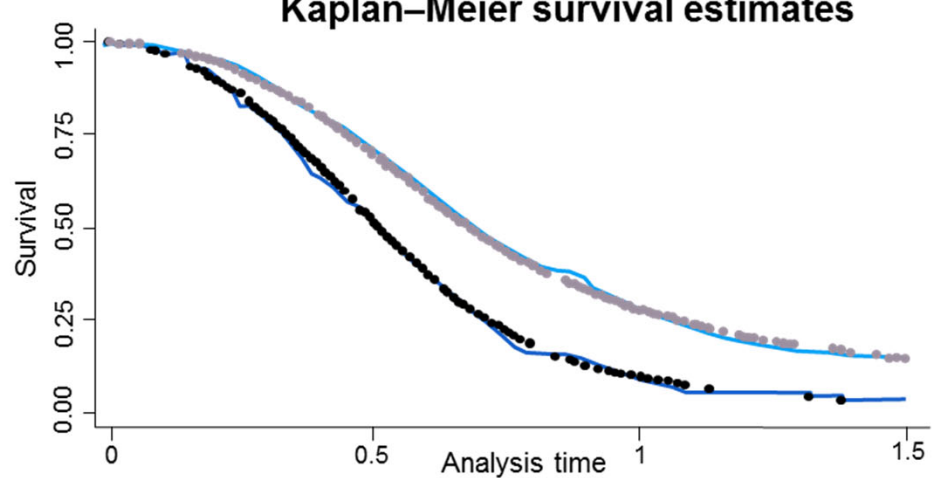


Kaplan–Meier survival estimates



Speaker's own data

Kaplan–Meier survival estimates



Limitations - FPMs

- The FPM extrapolates beyond the data using only the final segment of the curve. This may or may not be appropriate for achieving accurate projections
- How many knots to choose?
- “Joining the dots”

Solutions – Cure models

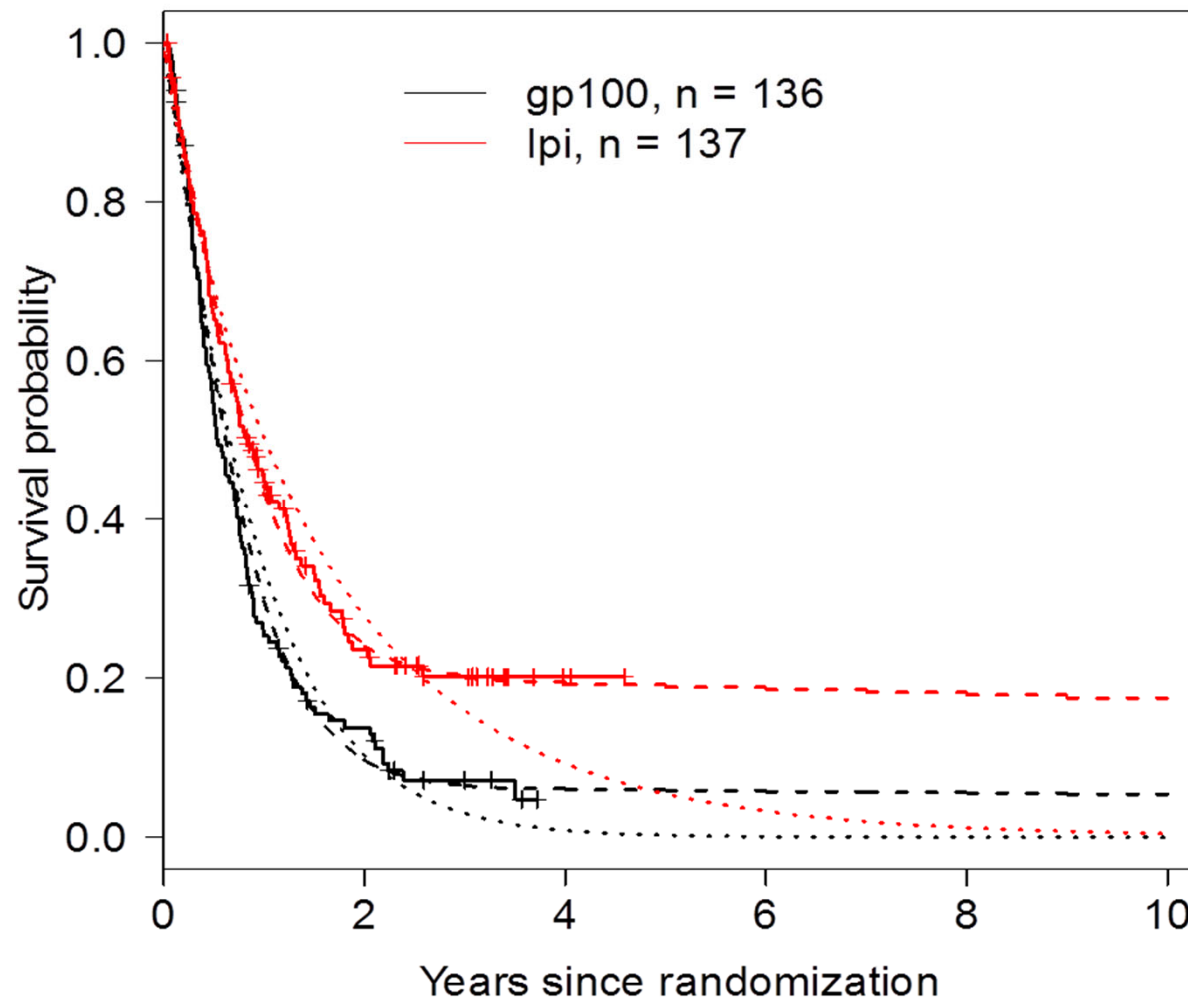
Parametric cure models

- Sometimes it might appear that a % of patients have been “cured”
- Model is used to:
 - Estimate **probability** that a patient is cured
 - Predict **survival** of patients who are not cured
- Survival distribution for cured patients is based on background mortality from external data

$$\text{Population survival} = p_{\text{cured}} * \text{survival}_{\text{cured}} + (1 - p_{\text{cured}}) * \text{survival}_{\text{uncured}}$$

→ Can represent hazard functions with turning points

Solutions – Cure models



- Othus *et al.* (2017)
- Standard Weibull model compared to mixture cure model

Solutions – Mixture models

Parametric mixture models

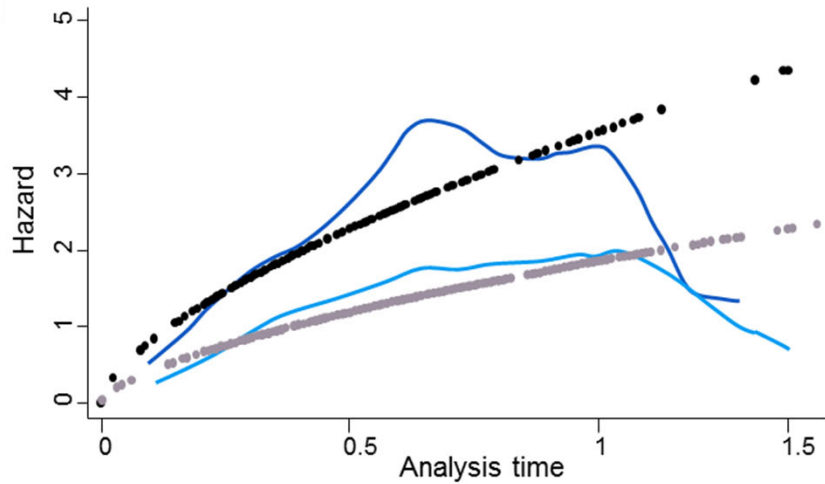
- May be some evidence of different survival distributions within data, but not necessarily a cure
- Parametric mixture models can be used to model with two (or more) distinct distributions (Lambert, 2007)
- E.g. mixture Weibull model:

$$s_0(t) = p \exp(-\lambda_1 t^{\gamma_1}) + (1 - p) \exp(-\lambda_2 t^{\gamma_2})$$

- p is the first mixture, $(1-p)$ is the second mixture

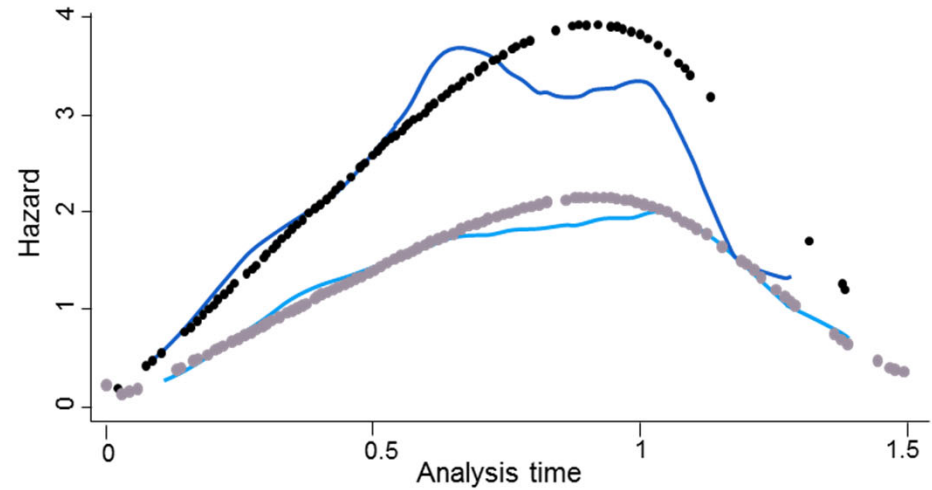
➔ Can represent hazard functions with turning points

Smoothed hazard estimates



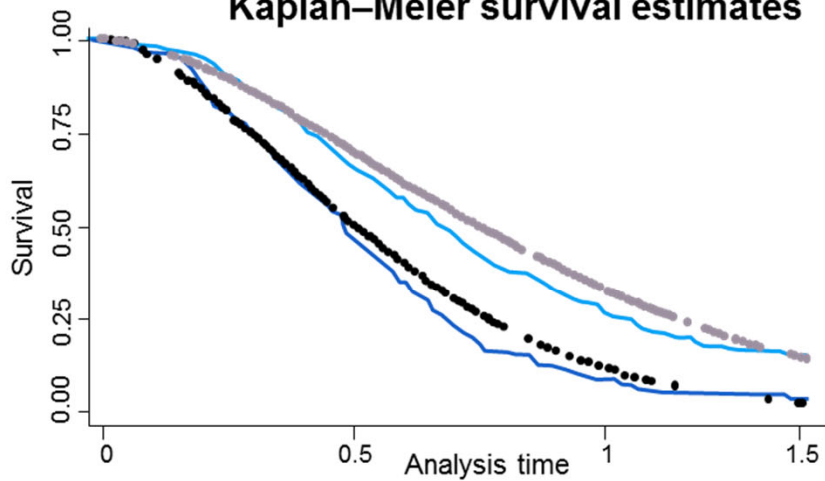
— ttrand = 0 — ttrand = 1
 • Weibull predicted hazard (con) • Weibull predicted hazard (exp)

Smoothed hazard estimates



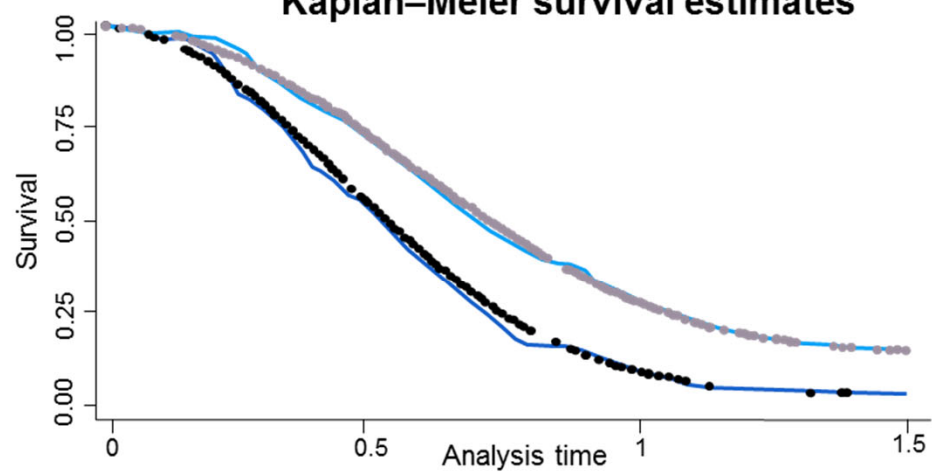
— ttrand = 0 — ttrand = 1
 • Mix Weibull predicted hazards (con) • Mix Weibull predicted hazards (exp)

Kaplan–Meier survival estimates



— ttrand = 0 — ttrand = 1
 • Weibull predicted OS (con) • Weibull predicted OS (exp)

Kaplan–Meier survival estimates



— ttrand = 0 — ttrand = 1
 • Mix Weibull predicted survival (con) • Mix Weibull predicted survival (exp)

Speaker's own data

Limitations – Mix/cure models

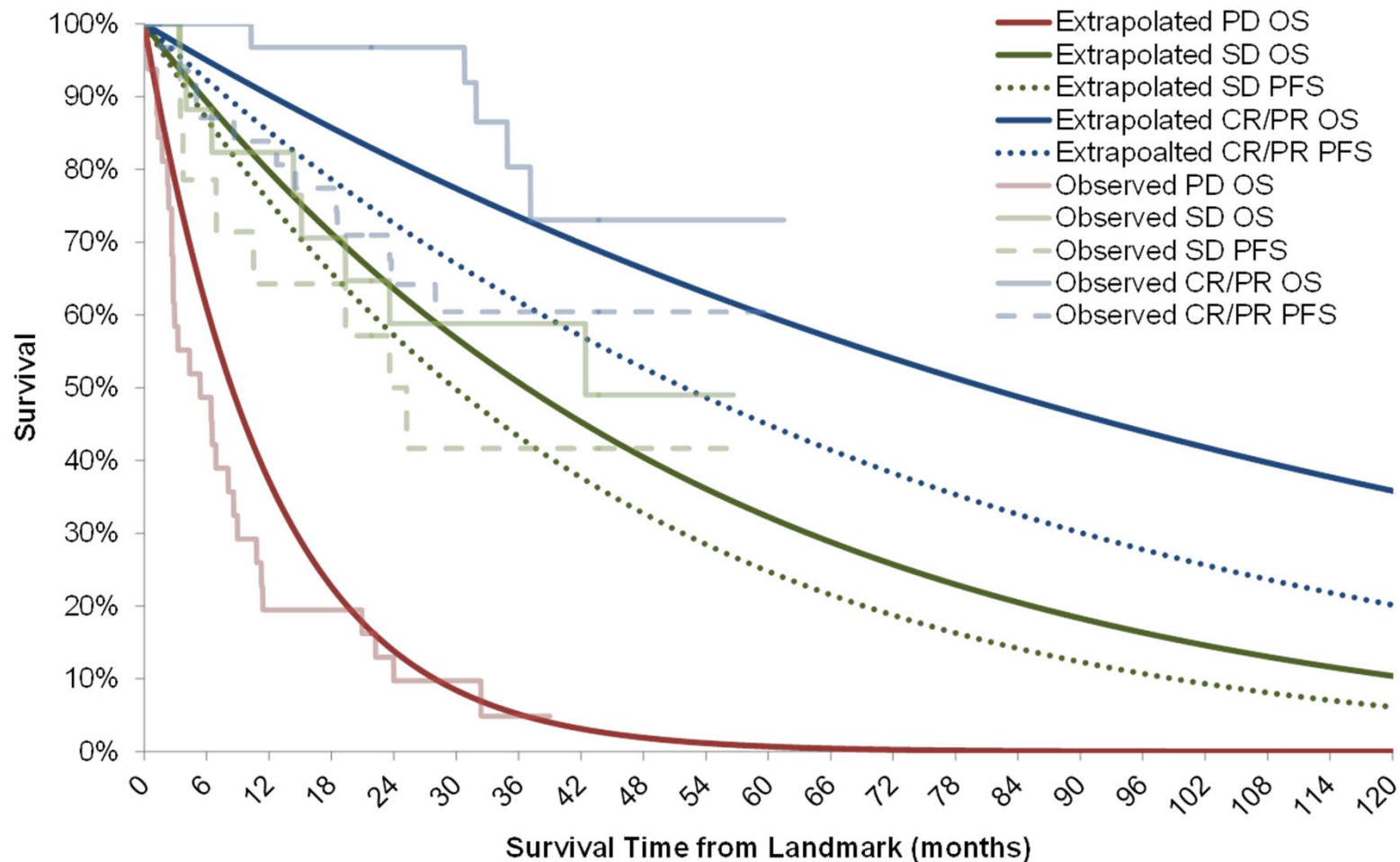
- Cure/mixture models have a nice rationale, but...
 - Can we prove that an assumption of a cure is reasonable?
 - Can we estimate the cure fraction based on short-term data?
 - How many mixes are there / do we need?
 - Do we fit cure models to PFS and OS? What if we get different cure fractions?
 - Do we fit from time 0? Cured at randomisation?
 - Are long-term hazards reasonable in the mixture?

Solutions – Response models

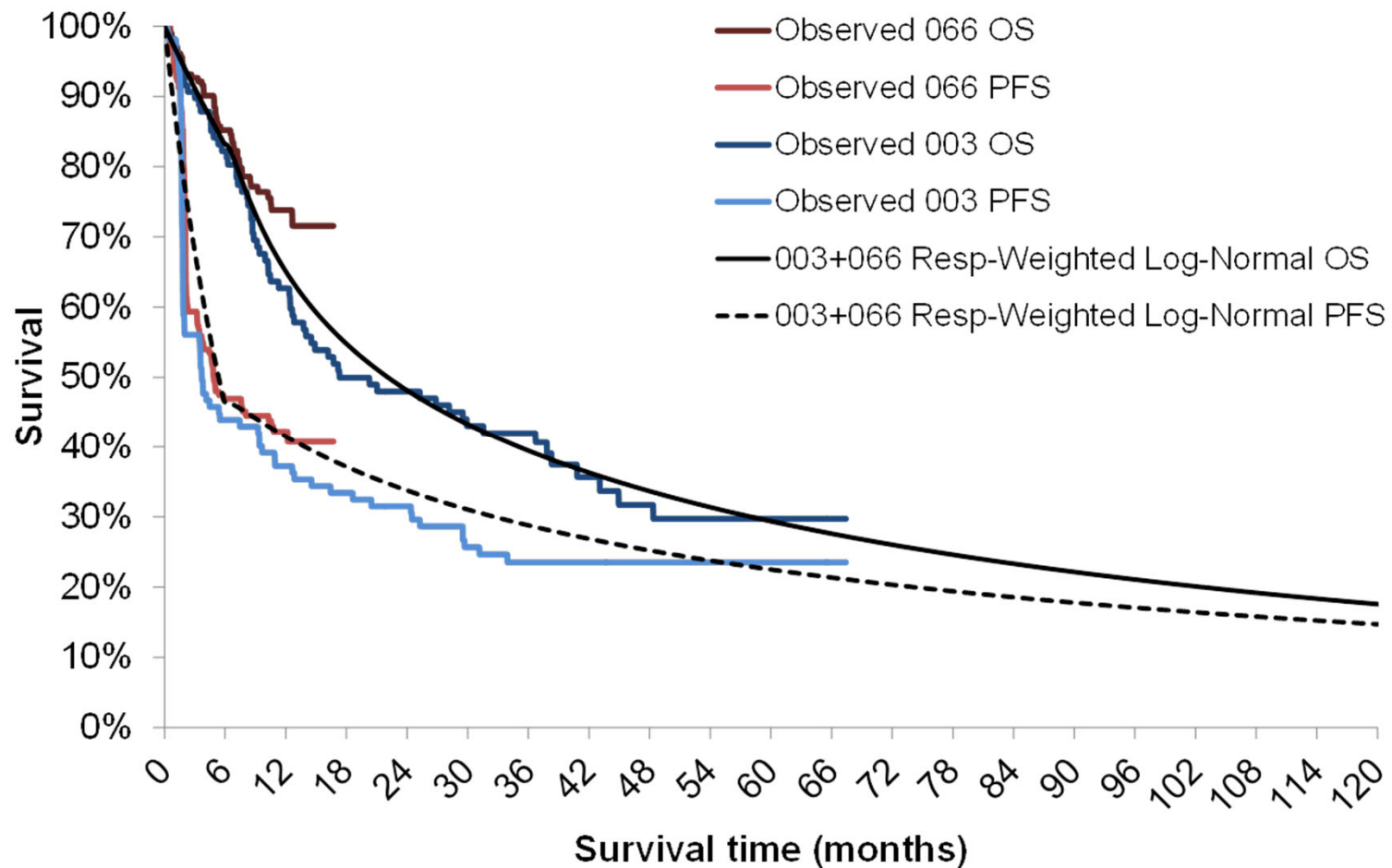
- Model based upon response categories:
 1. Select a landmark time-point, categorise patients into response groups
 2. Fit parametric survival models for response groups from landmark point
 3. Weight the response curves by the observed response distribution at the landmark time-point
- ➔ Can represent hazard functions with turning points



Solutions – Response models



Solutions – Response models



Limitations – Response models

- Fits the language used about I-O treatments: some patients don't benefit, but those that do benefit very substantially. But...
 - Are response measures adequate?
 - Pseudo-progression
 - Reliably distinguish patient prognosis, treatment effect only mediated through response
 - Which landmark time-point is suitable?
 - Delayed responses Vs reduced advantages if wait too long
 - Are standard parametric models appropriate within response groups – are long-term hazards appropriate?

Summary

- I-O drugs have encouraged increased attention on survival modelling techniques in HTA
 - This was probably needed anyway
- More complex methods are available – no need to stick to commonly used approaches
- Are decision makers equipped to review these methods?
- Can we assume that the “plateau” is there, without seeing it in an RCT?
- The more complex models have limitations – external validity remains crucial

References

- Royston, P. and Parmar, M.K.B. Flexible proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Statistics in Medicine* 2002; 21:2175-2197
- Rutherford MJ, Crowther MJ and Lambert PC. The use of restricted cubic splines to approximate complex hazard functions in the analysis of time-to-event data: a simulation study. *Journal of Statistical Computation and Simulation* 2015;85;4:777-793
- Lambert P. Modeling of the cure fraction in survival studies. *The Stata Journal* 2007 ;7;3:351-375
- Chen T. Statistical issues and challenges in immune-oncology. *Journal for ImmunoTherapy of Cancer* 2013;1:18
- Chen T. Predicting analysis times in randomized clinical trials with cancer immunotherapy. *BMC Medical Research Methodology* 2016;16;12
- Othus M, Bansal A, Koepl L, Wagner S, Ramsey S. Accounting for cured patients in cost-effectiveness analysis . *Value in Health* 2017;20:705–709
- Paly VF, Baker T, Gilloteau I, Orsini L, Briggs A. Long-term survival extrapolation for nivolumab (anti-PD-1) in advanced melanoma from trial data: A response-stratified approach. 11th EADO Congress and 8th World meeting of interdisciplinary melanoma/skin cancer centers
- Latimer N, Ramsey S, Briggs A. Cost–effectiveness models for innovative oncology treatments: How different methodological approaches can be used to estimate the value of novel therapies. ISPOR, Boston, US, 2017