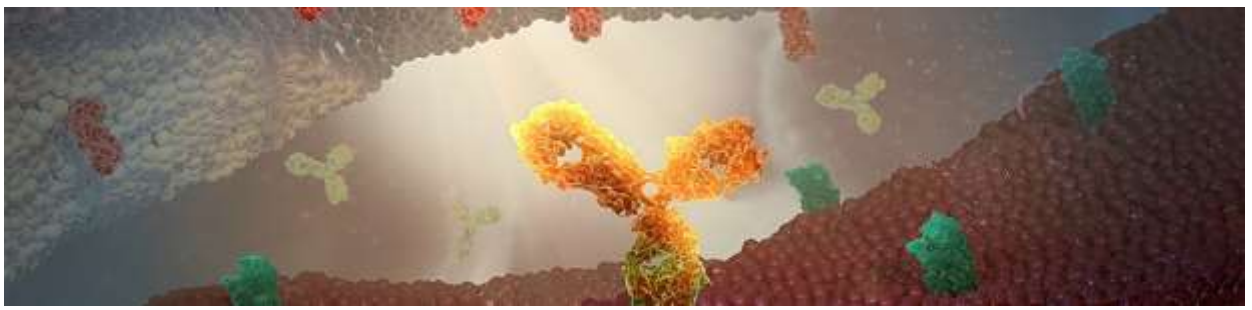
NOVEL APPROACHES TO COMMON CHALLENGES IN MODELLING OF SURVIVAL OUTCOMES FOR COST-EFFECTIVENESS ANALYSES IN ONCOLOGY

ISPOR 2018 Barcelona: W8 Advanced Workshop

Discussion Leaders: Richard Birnie, Mario Ouwens, Bart Heeg, Matthew Dyer

Tues Nov 13, 2018



Today's Workshop Agenda

1) Indirect Treatment Comparison in the presence of time varying relative treatment effects

- **Richard Birnie, PhD,** Principal Statistician, BresMed Health Solutions Ltd, Sheffield, UK

2) Use of clinical expert opinion in the estimation of survival extrapolation distributions

- **Mario Ouwens, PhD**, Statistical Science Director, AstraZeneca, Mölndal, Sweden

3) Subsequent treatment adjustment in oncology trials

- **Bart Heeg, PhD,** Managing Partner, Ingress-Health, Rotterdam, Netherlands

Our People.
Your Success.

Indirect Treatment Comparison in the presence of time varying relative treatment effects

Richard Birnie
ISPOR Europe
Barcelona 2018

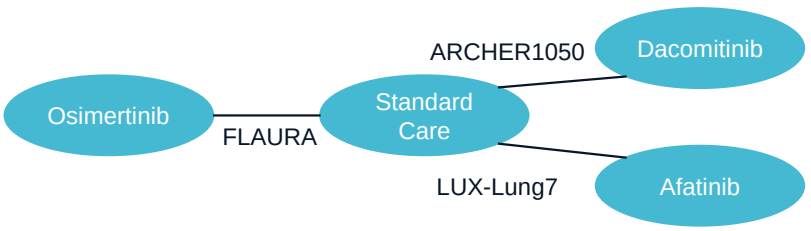
BresMed

Introduction

- To inform cost-effectiveness analyses and health technology assessments it is often necessary to make comparisons between treatments that have never been directly compared in clinical trials
- It is also common to extrapolate outcomes beyond the observed period of the clinical trial to estimate the effects of treatment over the whole lifetime of the patient
- Common to have access to individual patient data (IPD) for studies of a new intervention but only aggregate data for comparator studies.
- Commonly achieved as a two step process
  - Parametric survival modelling applied to IPD from the intervention study
  - Indirect treatment comparison method estimates relative effect of treatments

Estimating the relative effect of TKIs for the treatment of EGFR mutation positive NSCLC

- One study per treatment comparison
- All trials compare TKI (osimertinib, afatinib, dacomitinib) relative to standard care (gefitinib or erlotinib)
- Typical approach in this scenario would be network meta-analysis (NMA) pooling published hazard ratios from each study
  - Assumes that relative treatment effects remain constant over time

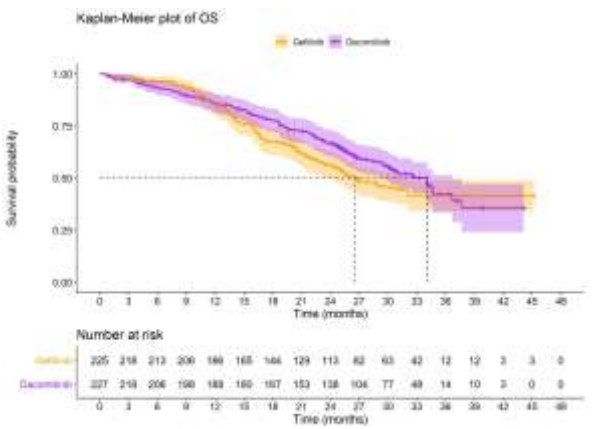


5

1. FLAURA, NCT02296125; 2. LUX-Lung7, NCT01466660; 3. ARCHER1050, NCT01774721

Time varying relative treatment effects

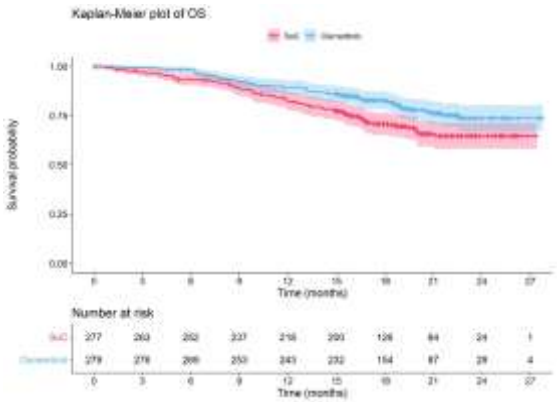
- Kaplan-Meier curves cross in ARCHER 1050 at ~12 months and ~36 months
  - Typically indicates violation of the proportional hazards assumption
  - Direction of treatment effect changes at different points in time



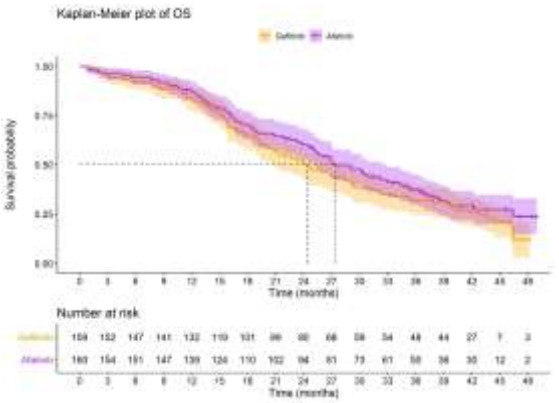
6

OS summary – FLAURA and LUX-Lung 7

FLAURA



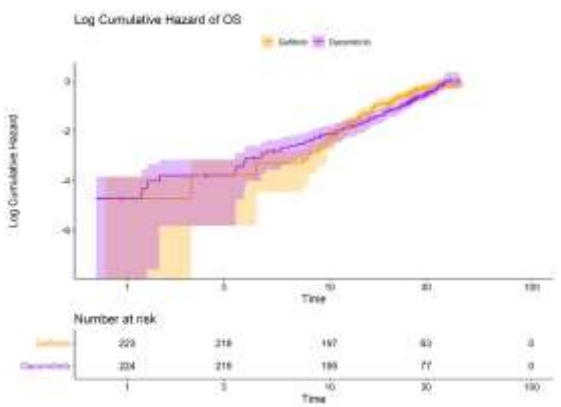
LUX-Lung7

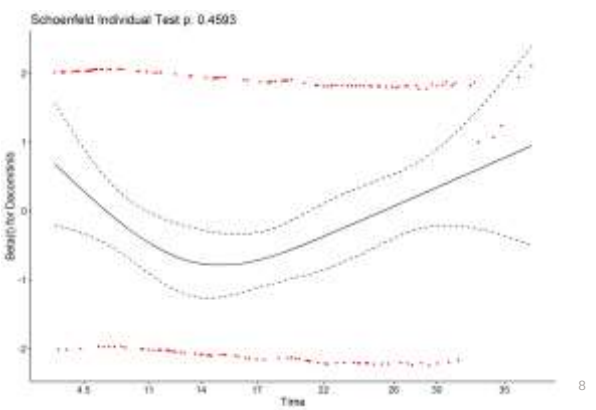


7

Assessment of proportional hazards – ARCHER1050

- If proportional hazards holds the log cumulative hazard (LCH) curves are expected to be parallel
  - Crossing indicates a reversal of the relative treatment effect
- Schoenfeld plot shows variation in HR over time





8

NMA of relative effects on multiple parameters of the survival model 1

- Synthesis of relative treatment effects that influence multiple parameters of the parametric survival curve to reflect a time varying treatment effect
- Relative treatment effects are captured in a survival regression model by including terms for study and for treatment arm
  - Parametric survival curve and relative treatment effects estimated in a single analysis
  - Combined ITC and extrapolation in one step

Parameter	Scenario 1	Scenario 2	Scenario 3
Location (e.g. scale)	• Vary between treatments • Vary between studies	• Vary between treatments • Vary between studies	• Vary between treatments • Vary between studies
Non-location (e.g. shape)	• Vary between treatments • Vary between studies	• Vary between treatments • Constant between studies	• Constant between treatments • Constant between studies

Ouwens MJ, Philips Z and Jansen JP. Network meta-analysis of parametric survival curves. *Res Synth Methods*. 2010; 1(3-4):258-71

NMA of relative effects on multiple parameters of the survival model 2

$$\ln(h_{jkt}) = \upsilon_{jk} + \theta_{jk}\ln(t)$$

Weibull model

υ_{jk} = scale parameter for treatment *k* in trial *j*

θ_{jk} = shape parameter for treatment *k* in trial *j*

h_{jkt} = hazard rate for treatment *k* in trial *j* time t

- If we assume proportional hazards, only υ_{jk} varies between treatments
- θ_{jk} remains constant between treatments
  - Hazard functions have the same shape over time
  - Ratio of the hazards will be constant over time – proportional hazards
- Model can be extended by allowing both υ_{jk} and θ_{jk} to vary between treatments
  - Hazard functions will have the different shapes for each treatment over time over time
  - Ratio of the hazards will vary over time – non-proportional hazards – relative treatment effect varies over time
- Further extend the model by allowing both υ_{jk} and θ_{jk} to be influenced by other covariates
  - In these analyses, study acts as a proxy for IPD on other covariates

Model Comparison

- AIC indicates that best fitting models were:
  - Weibull with treatment and study effects on the location parameter only (Scenario 3)
  - Generalised gamma with treatment and study effects on the location parameter only (Scenario 3)
- Generalised gamma model has three parameters; mean, sigma, Q
  - Generalised gamma includes Weibull as a special case when Q = 1
  - In this analysis Q = 0.98 (95%CI 0.70 to 1.25)
  - Effectively the same as the Weibull in this example

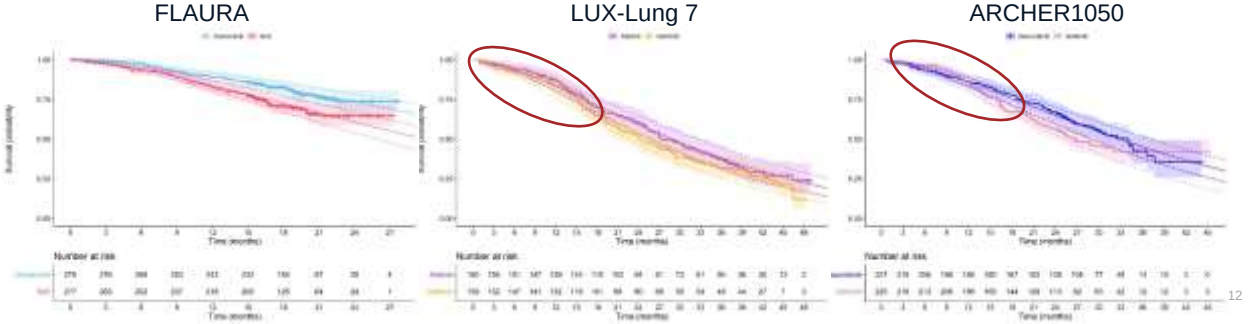
Model	Scenario 1	Scenario 2	Scenario 3
Weibull	5648.1	5643.7	**5639.9**
Generalised Gamma - Mean + Sigma	5649.9	5645.6	**5641.9**
Log logistic	5653.0	5649.5	5649.8
Gompertz	5674.9	5669.7	5667.1
Log normal	5699.7	5698.7	5702.3

11

Weibull model: Treatment and study effects applied to scale parameter only

- Equivalent to assuming proportional hazards – unlikely to be plausible based on earlier work
- Poor fit to the gefitinib arm of the ARCHER 1050 study (pink) and to both arms of LUX-Lung 7
- Underestimates survival up 12 months in both studies
- Similar results observed with generalised gamma model

FLAURA LUX-Lung 7 ARCHER1050



12

Exploring more complex models – Generalised F

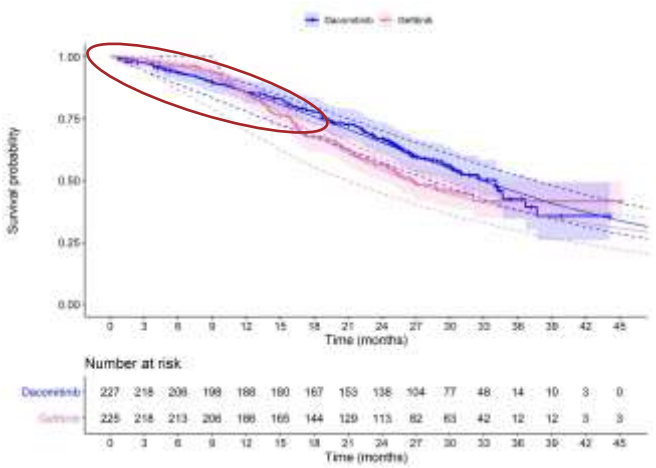
- Generalised F distribution has four parameters: mean, sigma, Q, P
  - More parameters = greater flexibility
- Lowest AIC (best fit) was observed for models with
  - Treatment and study effects applied to both mean and Q (Scenario 1)
  - Treatment and study effects applied to mean only (Scenario 3)
- Slightly higher AIC for the generalized F with treatment and study effects on mean and Q
  - Increased penalty for model complexity
  - Minimal impact compared to the improvement in visual fit to the observed Kaplan-Meier curve

Model	Scenario 1	Scenario 2	Scenario 3
Generalised F - Mean + Q	**5641.8**	5644.0	NA
Generalised F - Mean + P	5646.9	5643.9	NA
Generalised F - Mean + Sigma	5647.4	5643.3	**5640.3**
Weibull	5648.1	5643.7	**5639.9**

13

Generalised F distribution - Treatment and study coefficients applied to both mean and Q

- Provides good visual fit to ARCHER1050
- Closest fit to the gefitinib arm of ARCHER1050 among the models tested

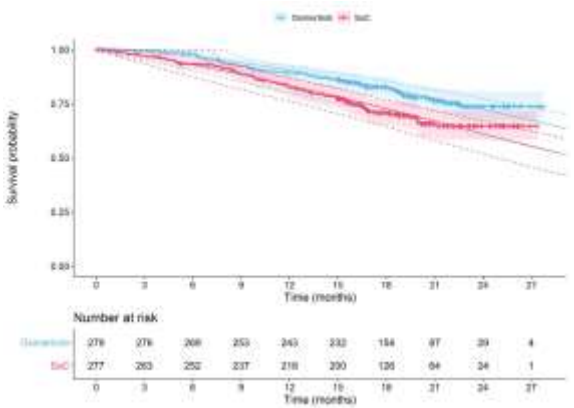


14

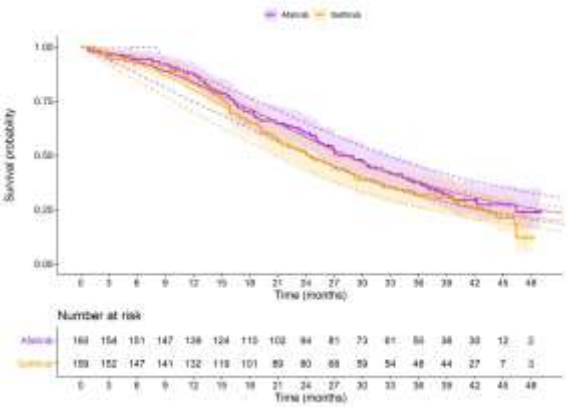
Generalised F distribution - Treatment and study coefficients applied to both mean and Q

- Provides reasonable visual fit to FLAURA and LUX-Lung 7

FLAURA



LUX-Lung 7



15

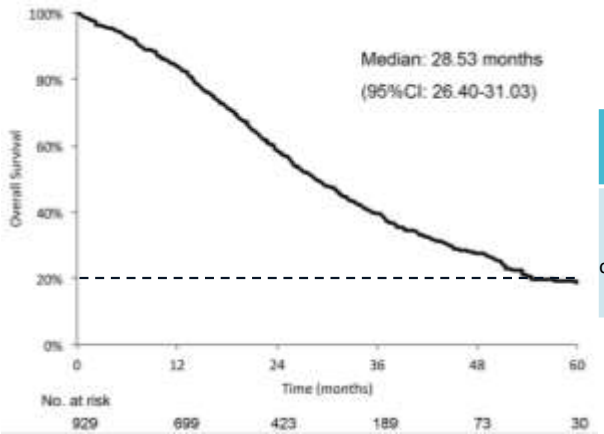
Extrapolation of overall survival

- Weibull and generalised gamma models give similar estimates of 10 year survival
  - Consistent with previous observations. These models are basically equivalent
- Generalised F models give much higher estimates of 10 year survival
- Models in scenario 3 rely on assumptions of constant relative treatment effect (Weibull, Gen Gamma, Gen F)
  - Unlikely to be plausible based on LCH and Schoenfeld plots presented earlier
  - Does not capture variation in relative effect over time

Treatment	Model	Estimated 5 year % survival (95% CI)	Estimated 10 year % survival (95% CI)
Osimertinib	S3: Weibull	29.2 (18.8, 38.7)	3.0 (0.7, 7.6)
	S3: Gen F	29.7 (22.5, 37.8)	10.3 (5.1, 18.3)
	S1: Gen F - Mean + Q	29.4 (23.4, 36.9)	12.8 (7.1, 19.5)
	S3: Gen Gamma	29.4 (19.5, 39.3)	3.3 (0.6, 9.3)
Afatinib	S3: Weibull	18.8 (8.8, 31.5)	0.9 (0.1, 3.9)
	S3: Gen F	23.2 (14.8, 34.2)	7.9 (2.7, 16.5)
	S1: Gen F - Mean + Q	23.8 (16.6, 36.8)	8.3 (3.8, 17.4)
	S3: Gen Gamma	19.1 (8.8, 30.5)	1.0 (0.0, 4.8)
Dacomitinib	S3: Weibull	20.0 (9.4, 31.6)	1.0 (0.1, 4.2)
	S3: Gen F	24.8 (16.0, 35.1)	8.5 (3.3, 16.6)
	S1: Gen F - Mean + Q	27.5 (18.8, 42.5)	6.0 (3.5, 18.2)
	S3: Gen Gamma	20.3 (10.0, 32.7)	1.2 (0.1, 5.5)
Standard care	S3: Weibull	13.9 (8.1, 21.4)	0.4 (0.1, 1.5)
	S3: Gen F	19.0 (13.3, 26.6)	6.5 (2.3, 13.3)
	S1: Gen F - Mean + Q	19.5 (15.3, 27.0)	7.0 (3.6, 12.7)
	S3: Gen Gamma	14.2 (7.2, 22.2)	0.4 (0.0, 2.6)

Expected overall survival

- Inoue 2016 reported overall survival for a Japanese cohort receiving first line gefitinib, N = 929
- Reported OS at 5 years was ~20%, similar to generalised F model estimates for standard care



Treatment	Model	Estimated 5 year % survival (95% CI)	Estimated 10 year % survival (95% CI)
Standard care	S3: Weibull	13.9 (8.1, 21.4)	0.4 (0.1, 1.5)
	S3: Gen F	19.0 (13.3, 26.6)	6.5 (2.3, 13.3)
	S1: Gen F - Mean + Q	19.5 (15.3, 27.0)	7.0 (3.6, 12.7)
	S3: Gen Gamma	14.2 (7.2, 22.2)	0.4 (0.0, 2.6)

Inoue 2016 Japanese Journal of Clinical Oncology, 46(5) 462–467

Conclusions

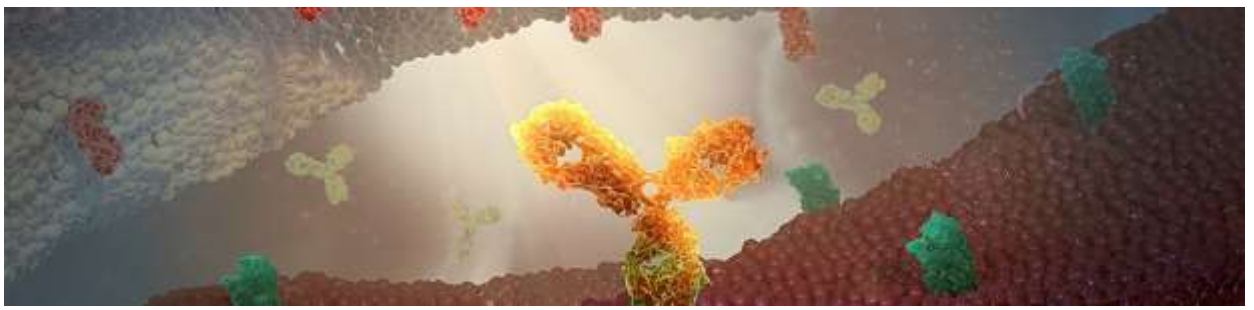
- Standard approach to NMA assumes that relative treatment effects are constant over time
- Kaplan-Meier curves from the ARCHER1050 study showed evidence of time varying relative treatment effects
  - Likely due to the combination of study design and treatment switching observed in this study
- Several models gave similar goodness of fit statistics
  - Three out of four best fitting models assumed constant relative treatment effects
    - Unlikely to be plausible
- Four parameter generalised F model with treatment and study effects on two parameters (mean and Q) gave a better fit to the observed Kaplan-Meier data
  - Slight penalty in goodness of fit statistics due to increased complexity
  - Increased flexibility leads to higher estimates of long term survival compared to alternative models
  - Estimated 5 year OS was comparable to external sources for the standard care arm

Thank you

Questions?

ISPOR EU 2018 - Use of clinical opinion in the estimation of survival extrapolation distributions

Mario Ouwens
Statistical Science Director, Advanced Analytics Centre, AstraZeneca R&D



Disclaimer

The views and opinions expressed in the following PowerPoint slides are those of the individual presenter and should not be attributed to AstraZeneca.

Can extrapolation be more clinical information based?

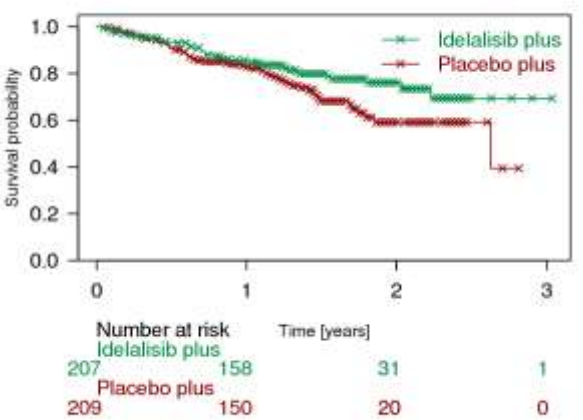
- Immature survival data may need to be extrapolated for HTA purposes
- Current process: Clinicians expectations asked after extrapolation
  - But what if all curves cross/all estimated curves are not clinically plausible?
- **Research question: Why not asking clinicians first and using clinical opinion in estimation thereafter?**

exp(-λ*t*)

Decision problem: 3 year idelalisib trial in chronic lymphocytic leukaemia

exp(-λ*t*)

- Decision problem: Zelenetz et al (2017):
  - Idelalisib compared to placebo
  - Relapsed or refractory chronic lymphocytic leukaemia



→ Clinical opinion:**10%-15%** 10 year survival for placebo arm

 Elicitation: In line with elicitation work from Kate Ren at PSI
https://www.psiweb.org/docs/default-source/default-document-library/kate-ren-slides.pdf?sfvrsn=1f2ededb_0

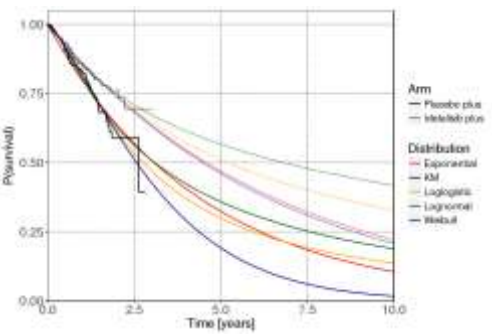
Large difference in estimated 10yr benefit, while having equivalent fit to KM

	AIC	BIC	ΔRMST benefit KM end	ΔRMST benefit 10 years
Weibull	1032	1045	0.22	1.97
Log logistic	1031	1045	0.21	1.68
Lognormal	1031	1045	0.20	1.74
Exponential	1032	1039	0.20	1.19

Difference in 10yr restricted mean survival benefit idelalisib versus placebo purely due to post-trial period:

- Benefit lognormal: 1.74
- Benefit exponential 1.19
- Benefit lognormal 46% larger

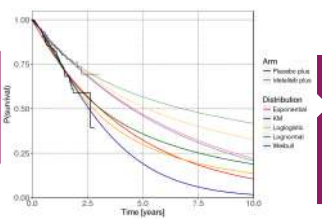
exp(-λ*t*)



25

So, what do we do about it?

Current approach

exp(-λ*t*) → Estimate exp(-λ*t*) →  → Go to clinician → 

New proposal

 → Go to clinician first → Estimate exp(-λ*t*) Using clinical opinion → exp(-λ*t*)

26 Additional option: when clinician rejects distribution, ask him to provide reasonable percentage range and re-estimate

How does it work? Bayesian estimation

Construction of a priori distribution for exponential distribution:

Clinician: ”At 10 years, placebo survival percentage is between 10 and 15%”

exp(-λ*t*)

Statistician: exp(-λ*t*)

”I can do my magic with the exponential”

S = exp(-λ*t*)

10% = exp(-10 $\lambda_{\text{upper bound}}$)
15% = exp(-10 $\lambda_{\text{lower bound}}$)

$\lambda_{\text{lower bound}}$ = 0.19, $\lambda_{\text{upper bound}}$ = 0.23

27 To be very specific: Our analyses used lognormal λ with log of bounds of confidence intervals

How does it work? Bayesian estimation; Challenge: What to do with multiple parameters

But what to do with distributions with more than 1 parameter? exp???

Distribution	Functional form	Rewritten for S=10% at t = 10
Exponential	S = Exp(-λ*t*)	λ = -ln(10%) / 10
Weibull	S = Exp(-λ*t* $^{\varphi}$)	λ = -ln(10%) / 10 $^{\varphi}$
Lognormal	$S = 1 - \Phi\left(\frac{\log(t)-\mu}{\sigma}\right)$	$\mu = \log(10) - \sigma\,\Phi^{-1}(1-10\%)$
Loglogistic	$S = \frac{1}{1+\exp\left(\frac{t}{\beta}\right)^{\varphi}}$	$\beta = 10 - (\log((1\text{-}10\%)/10\%))^{1/\varphi}$
Gompertz	$exp(-\varphi\,(\exp(\beta t) - 1))$	$\varphi = -\log(10\%)/(\exp(10\beta) - 1)$

28

How does it work? Bayesian estimation; Challenge: What to do with multiple parameters

Solution: Sample φ first and compute $\lambda_{\text{upper bound}}$ and $\lambda_{\text{lower bound}}$ using S and φ

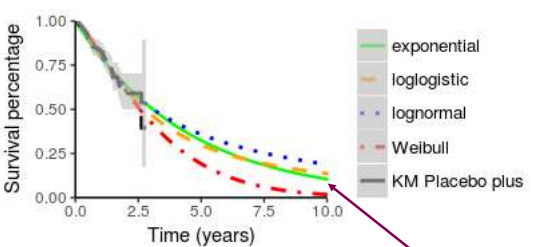
Distribution	Functional form	Rewritten for S=10% at t = 10
Exponential	S = Exp(-λ*t*)	λ = -ln(10%) / 10
Weibull	S = Exp(-λ*t* $^{\varphi}$)	λ = -ln(10%) / 10 $^{\varphi}$
Lognormal	$S = 1 - \Phi\left(\frac{\log(t)-\mu}{\sigma}\right)$	$\mu = \log(10) - \sigma\,\Phi^{-1}(1-10\%)$
Loglogistic	$S = \frac{1}{1+\exp\left(\frac{t}{\beta}\right)^{\varphi}}$	$\beta = 10 - (\log((1-10\%)/10\%))^{1/\varphi}$
Gompertz	$exp(-\varphi\,(\exp(\beta t)-1))$	$\varphi = -\log(10\%)/(\exp(10\beta)-1)$

29

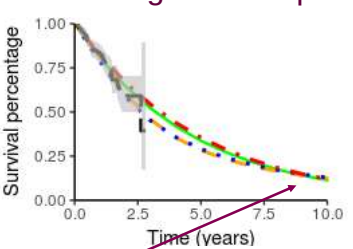
Did it help for comparator arm?

exp(-λ*t*)

Original fit



Fit using clinical opinion



Yes, it helped

30

Active arm; No 10 years experience

- **Active more difficult**
  - Less experience
  - Less convincing to HTA authorities

What could we at least do:

- Assume exp(-λ*t*)
  - proportional hazards (PH) or
  - constant accelerated failure time (AFT)

- Clinicians know that the 10 years survival percentage is between 0 and 100%
  - Thus: Use more uncertainty or larger ranges of survival percentages

- Run scenarios

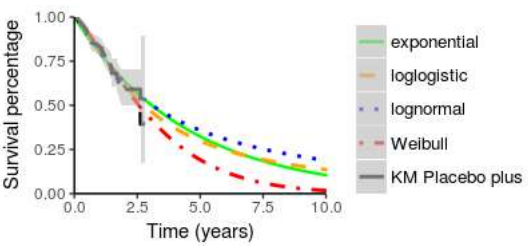
Did it help?

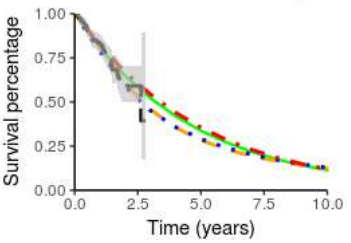
exp(-λ*t*)

Original fit

Fit using clinical opinion

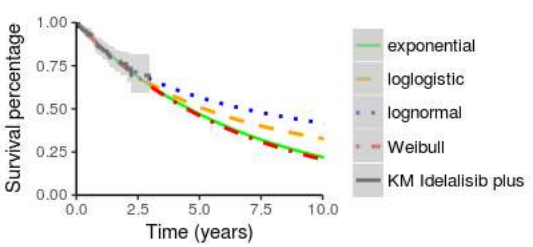
Placebo

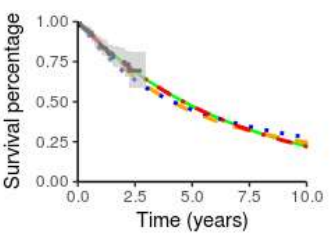




Aligned with clinical expectations and good fit to KM

Idelalisib





Subsequent treatment adjustment in Oncology trials

Bart Heeg, Managing Partner INGRESS Health

Background

- Cross-over/treatment switching might bias clinically reliability of reported (relative) survival
- Several methods exist that can adjust for treatment switching, e.g.
  - Rank preserving structural failure time models (RPSFTM)
  - Inverse probability of censoring weighing (IPCW)
  - Two stage method
- Traditionally these methods are used to generate counterfactual survival times/weights for patients in the placebo arm to reflect the situation that no one switched to the "novel treatment or to not reimbursed therapies in clinical practice".

Decision problem

- Novel oncology agents are usually developed first for relapsed refractory patients and subsequently for front-line patients
- In a front line clinical trial the novel agent can be prescribed as relapsed/refractory treatment (mostly in the placebo arm)
- Potential local decision problems/questions
  - In clinical practice the percentage of patients in the placebo arm being switched to the novel agent differs for the trial.
    - This impacts costs in the cost-effectiveness model but how does it impact expected relative survival ?
  - Isn't it "more cost-effective" to initiate novel treatment in a later line

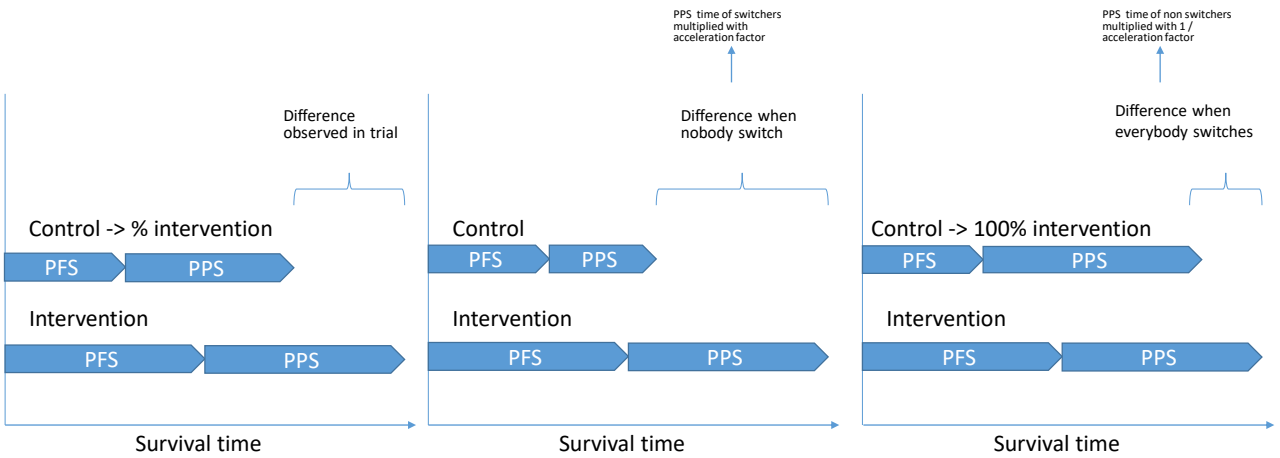
Methods

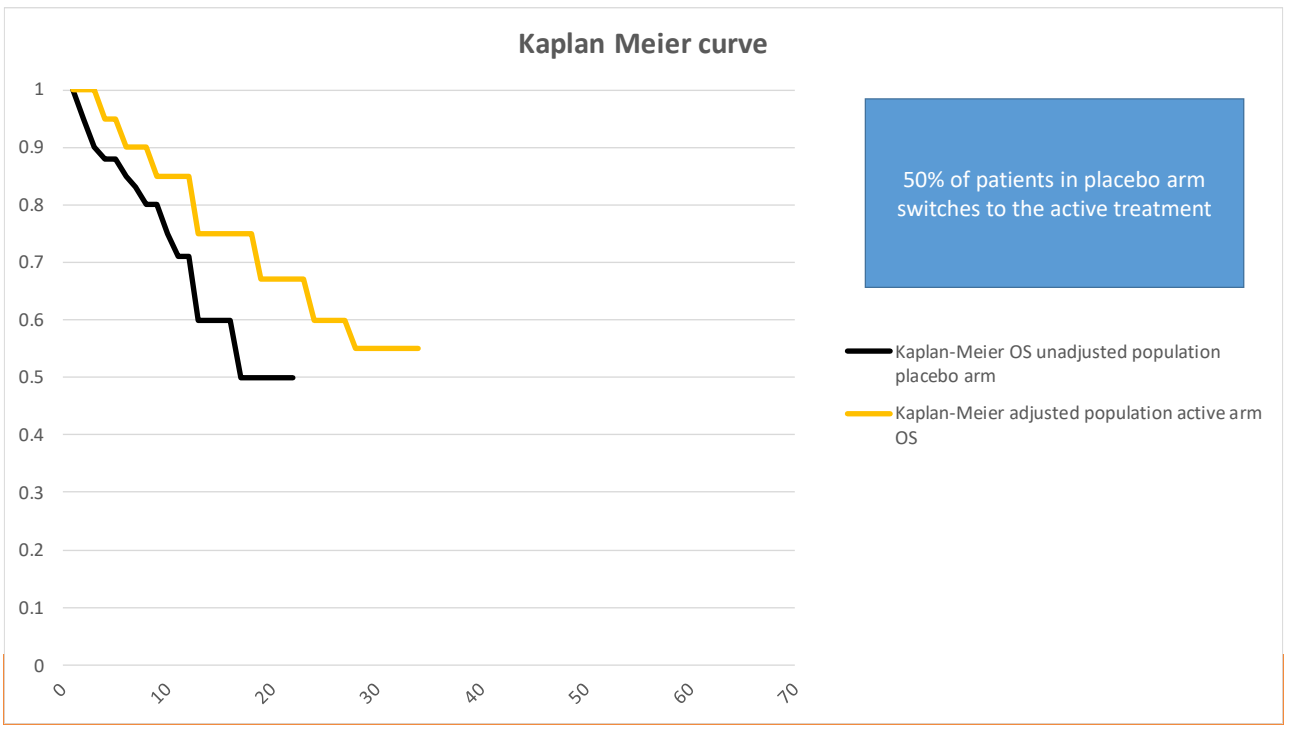
The acceleration factor reflects the amount by which an individual's expected survival time is increased by treatment.

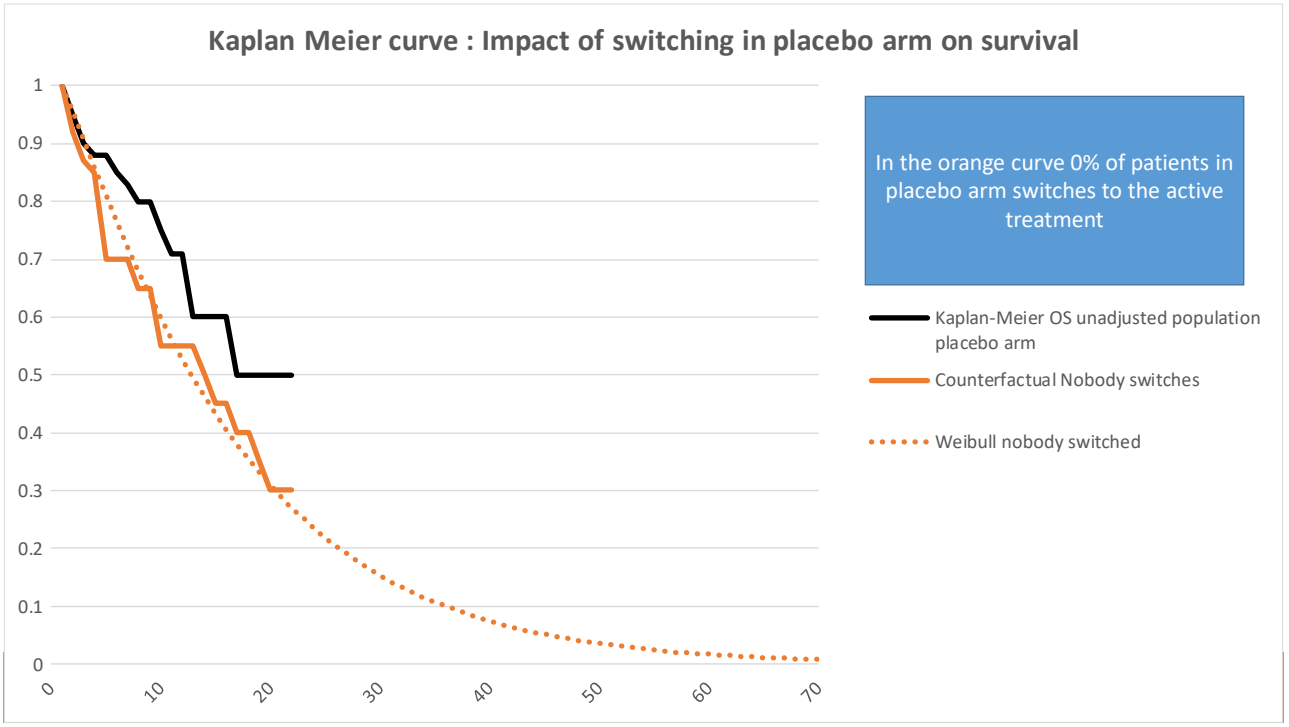
1. Use "RPSFTM / two stage method" t
   from time of switch OR progression for situation that
   1. 0% of the patients switched to novel second line treatment
      - Survival times of patients switching from switch are multiplied with the acceleration factor (RPSFTM)
   2. 100% of the patients switched to novel second line treatment
      - Survival times of patients not switching are from progression multiplied with 1/acceleration factor (RPSFTM)
2. Estimate the weighted mean survival for each decision problem for the placebo arm
   1. Decision problem 1: 70% of patients are assumed to switch and 30% are assumed not to switch in the placebo arm, whereas in the trial 50% switched

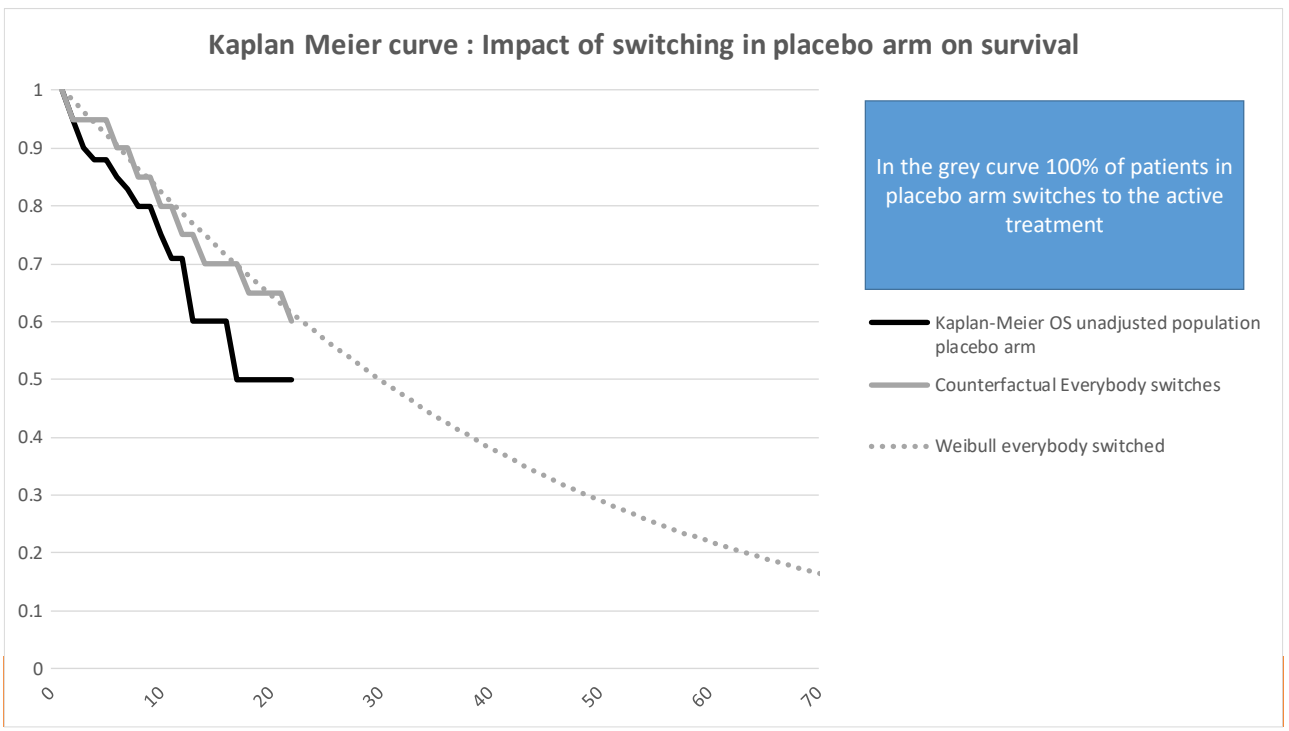
      $S(t)_{70\%\ switch} = 30\%\ S(t)_{0\%\ switch} + 70\%\ S(t)_{100\%\ switch}$
   2. Decision problem 2: 100% of the patients is assumed to switch in the placebo arm
3. Per decision problem compare the new weighted mean survival in placebo arm with that of the active arm.

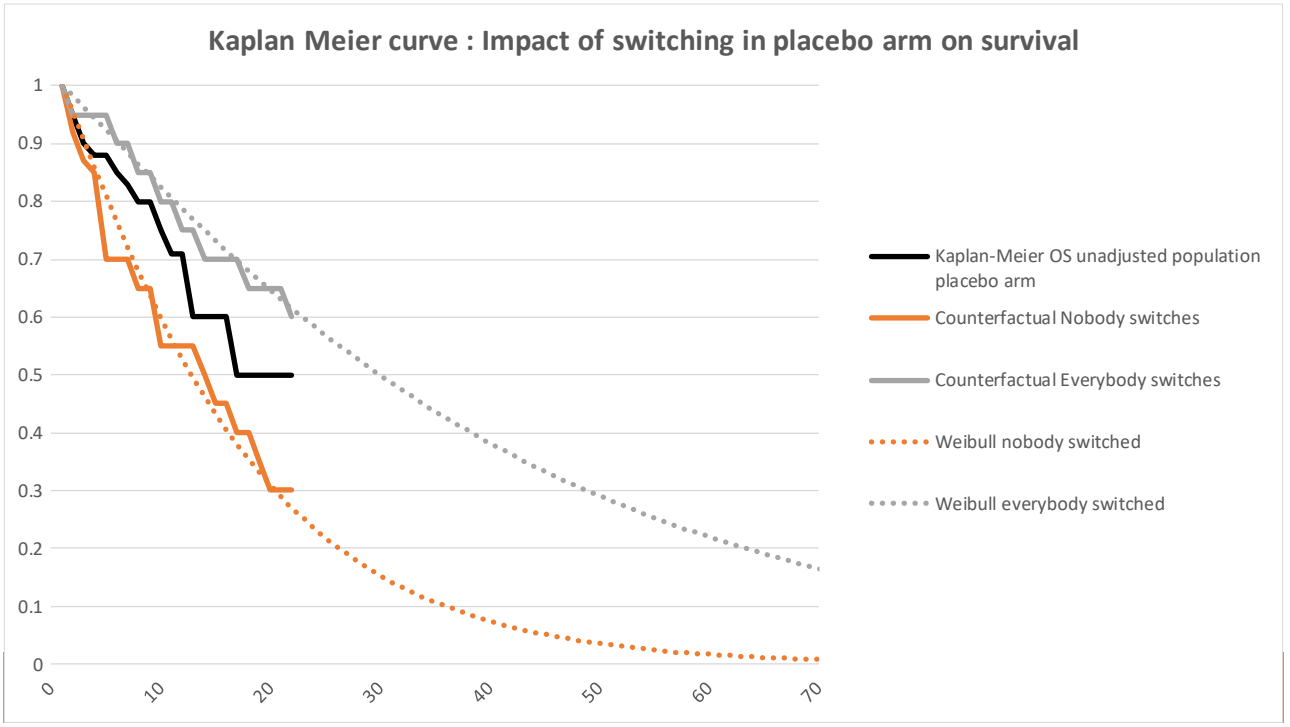
RPSFT (adjust post switch times)

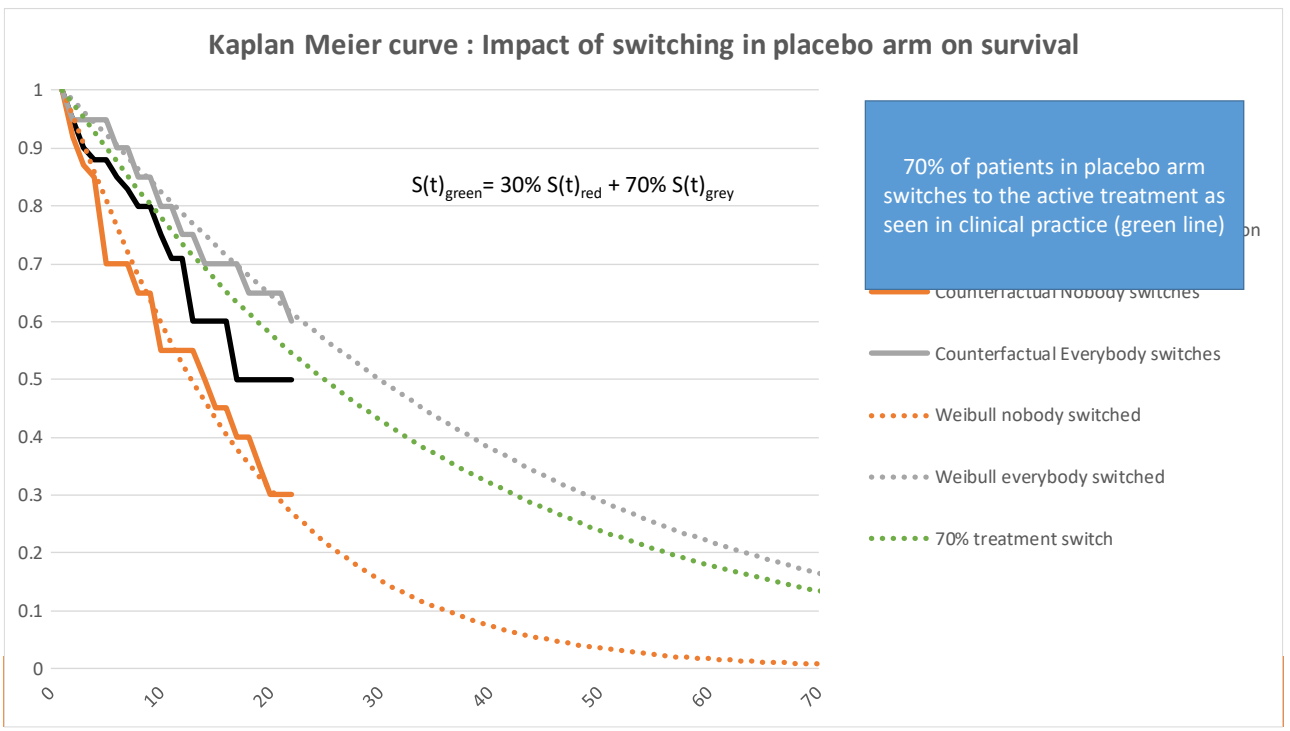


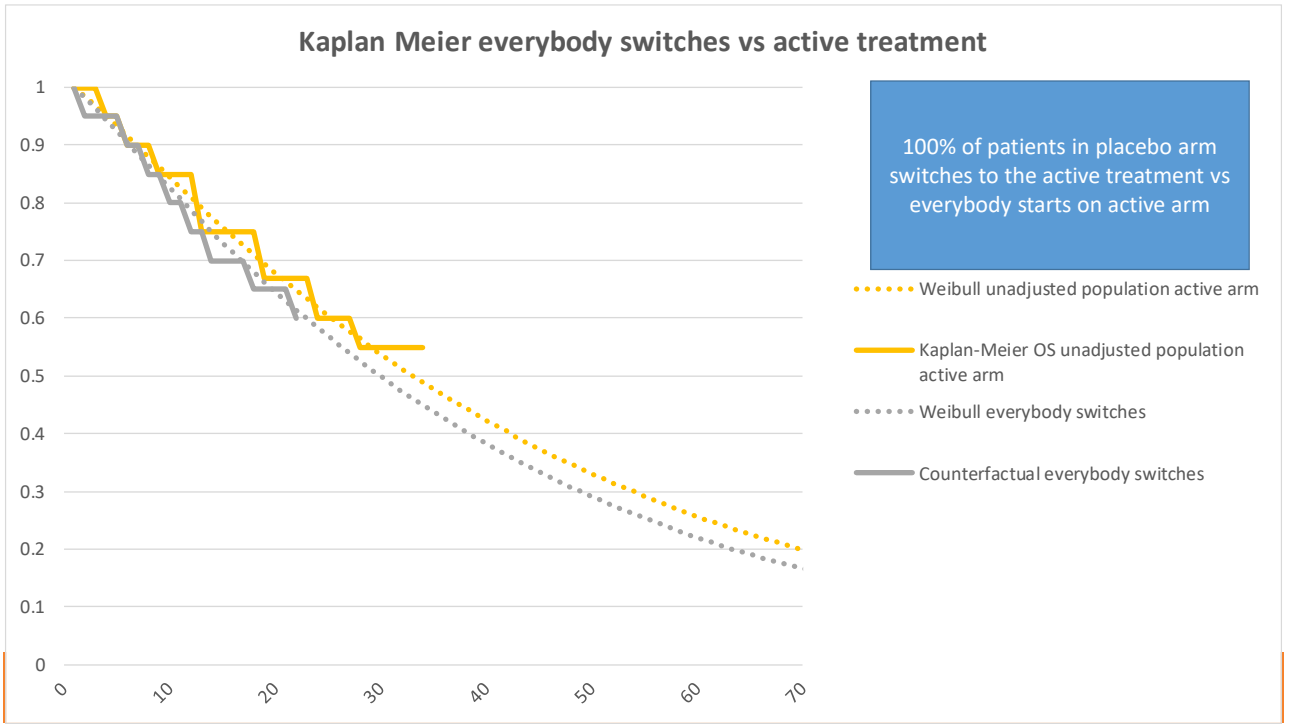



Kaplan Meier curve : Impact of switching in placebo arm on survival
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
10
20
30
40
50
60
70
In the orange curve 0% of patients in placebo arm switches to the active treatment
Kaplan-Meier OS unadjusted population placebo arm
Counterfactual Nobody switches
Weibull nobody switched


Kaplan Meier curve : Impact of switching in placebo arm on survival
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
10
20
30
40
50
60
70
In the grey curve 100% of patients in placebo arm switches to the active treatment
Kaplan-Meier OS unadjusted population placebo arm
Counterfactual Everybody switches
Weibull everybody switched


Kaplan Meier curve : Impact of switching in placebo arm on survival
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
10
20
30
40
50
60
70
Kaplan-Meier OS unadjusted population placebo arm
Counterfactual Nobody switches
Counterfactual Everybody switches
Weibull nobody switched
Weibull everybody switched


Kaplan Meier curve : Impact of switching in placebo arm on survival
$S(t)_{green}$= 30% $S(t)_{red}$ + 70% $S(t)_{grey}$
70% of patients in placebo arm switches to the active treatment as seen in clinical practice (green line)
Counterfactual Everybody switches
Weibull nobody switched
Weibull everybody switched
70% treatment switch



Discussion

- The approach can be used to adjust the survival of the trial to reflect local clinical practice decision problems.
- This can easily be implemented in partition survival model (PSM) framework for health economic modelling purposes
- Alternatives like treatment sequencing models have limitations
  - How to derive efficacy after treatment switch, as in trial often many different treatment sequences are found?
  - If post treatment switch survival is time dependent transitions a more complex PSM model is required.
- All limitations of the cross-over methods apply here
- More difficult situations like switching in active arm can be accounted for