The University Of Sheffield.

Methods for adjusting survival estimates in the presence of treatment crossover – a simulation study

Nicholas Latimer, University of Sheffield

Collaborators: Paul Lambert, Keith Abrams, Michael Crowther and James Morden

Please note these are not final results: do not disseminate results without permission from authors

The University Of Sheffield.

University of Leicester

ICR
The Institute of Cancer Research

Contents

- Extensions compared to James' study
- Data generation
- Scenarios
- Performance measures
- Methods
- Results
- Conclusions

Please note: results subject to minor changes – re-running due to minor errors in data generating model

Extensions

- Allow the treatment effect to change over time, based upon a time-dependent covariate
  - Generate a time-dependent covariate that represents the continuous progression of the disease – as this increases, the relative treatment effect falls
  - ➔ This involves relaxing the equal treatment effect assumption
- Allow the treatment crossover decision to be based upon time-dependent covariates, rather than baseline characteristics
- Include 'observational-based' methods
  - IPCW
  - SNM with g-estimation

Data Generation (1)

- Used a two-stage Weibull model to generate underlying survival times and a time-dependent covariate (called 'CEA')
- Longitudinal model for CEA (for *i*th patient at time *t*):

$$cea_i(t) = \beta_{0i} + \beta_1 * \log(t) + \beta_2 * \log(t) * trt_i + \beta_4 badprog_i$$

where $\beta_{0i} \sim N(\beta_0, \sigma_0^2)$

β_{0i} is the random intercept

β_1 is the slope for a patient in the control arm

$\beta_1 + \beta_2$ is the slope for a patient in the treatment arm (all

β_4 is the change in the intercept for a patient with bad prognosis compared to a patient without bad prognosis

- Picked parameter values such that CEA increased over time, more slowly in the experimental group, and was higher in the badprog group

Data Generation (2)

- The survival hazard function was based upon a Weibull (see Bender et al 2005):

$$h(t) = \lambda\gamma t^{\gamma-1}\exp(X\beta)$$

- In our case,

$$X\beta = \delta_1 * trt_i + \delta_2 * badprog_i + \alpha * (cea(t))$$

where δ_1 is the log hazard ratio (the treatment effect)

δ_2 is the impact of a bad prognosis baseline covariate on survival

α is the coefficient of CEA, indicating its effect on survival

- We used this to generate our survival times

➔ So, CEA has an effect on survival

Data Generation (3)

- We then estimated the treatment effect over time (in terms of an acceleration factor) based upon Collett's HR to AF formula:

$$AF = \exp\left(\frac{-(\delta_1 + \alpha\beta_2 * \log(t))}{\gamma}\right)$$

- We used this to 'inflate' survival times of crossover patients, based upon the time-point at which they started receiving the experimental treatment
- NOTE: This equation is wrong

➔ Collett's formula is only applicable when there are proportional hazards, and we do not have proportional hazards due to our time-dependent covariate

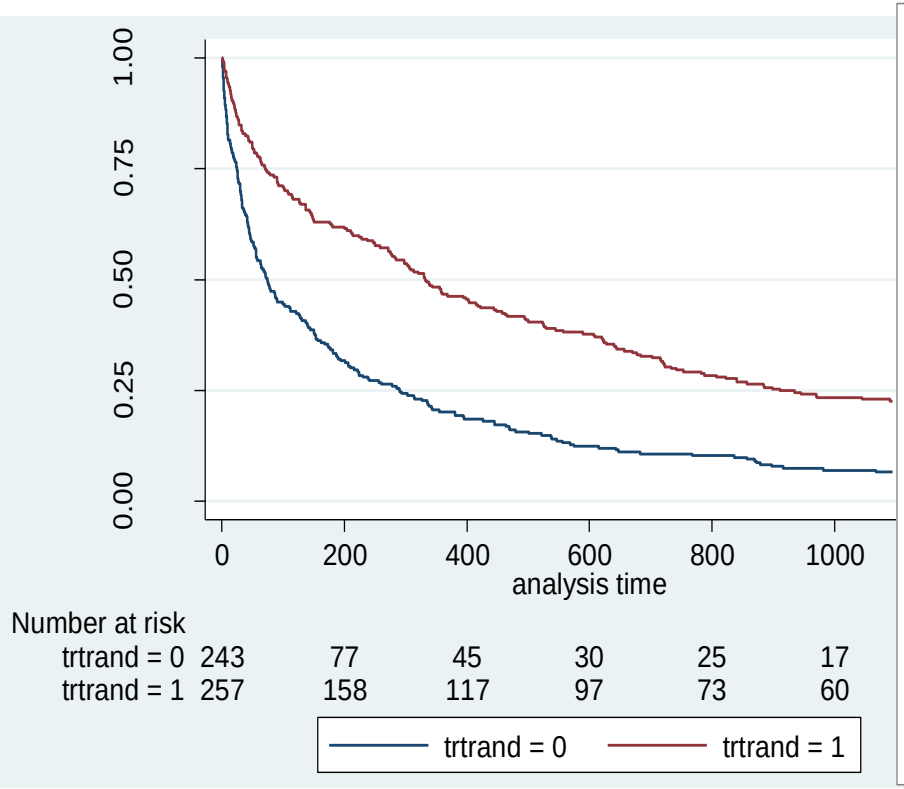
➔ Working on this

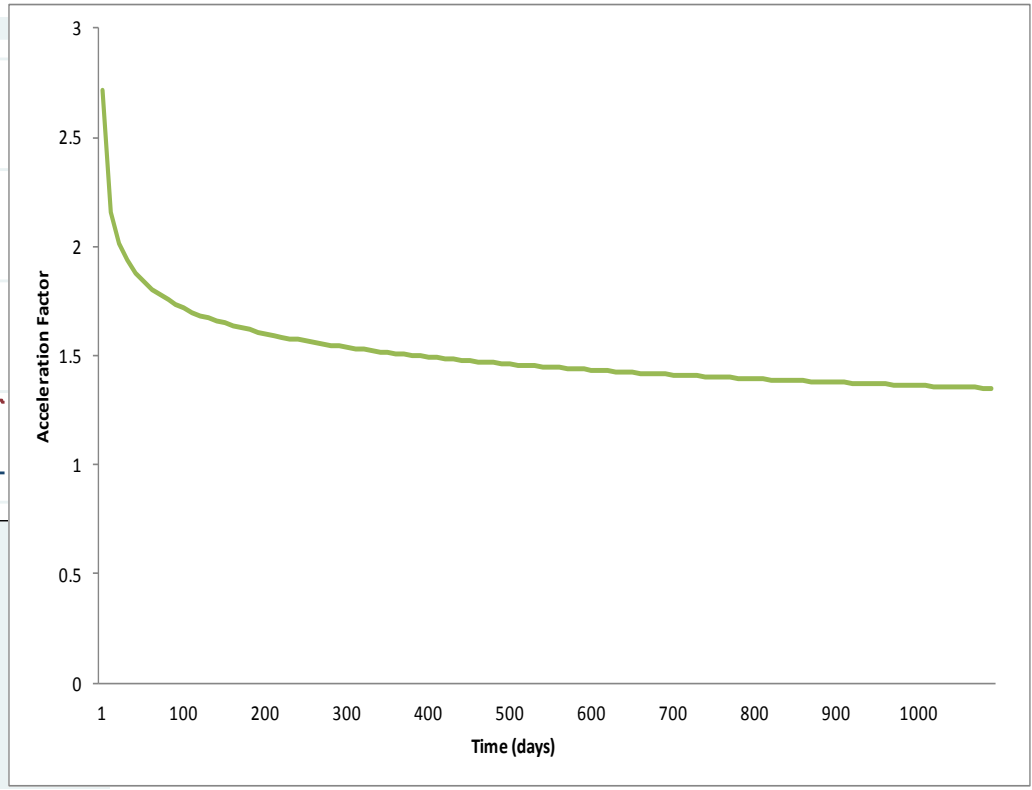
➔ Note – unlikely to change our results as we are still doing what we intended – applying a lower treatment effect to crossover patients

The University Of Sheffield.

Data Generation (4)

- We then selected parameter values in order that 'realistic' datasets were created:





Data Generation (5)

We made several assumptions about the 'crossover mechanism':

1. Crossover could only occur after disease progression (disease progression was approximately half of OS, calculated for each patient using a beta(5,5) distribution)
2. Crossover could only occur at 3 'consultations' following disease progression
   - These were set at 21 day intervals
   - Probability of crossover highest at initial consultation, then falls in second and third
3. Crossover probability depended on time-dependent covariates:
   - CEA value at progression (high value reduced chance of crossover)
   - Time to disease progression (high value increased chance of crossover)
   - This was altered in scenarios to test a simpler mechanism where probability only depended on CEA

➔ Given all this, CEA was a time-dependent confounder

Scenarios

Variable	Value	Alternative
Sample size	500	✖
Number of prognosis groups (prog)	2	✖
Probability of good prognosis	0.5	✖
Probability of poor prognosis	0.5	✖
Maximum follow-up time	3 years (1095 days)	✖
Multiplication of OS survival time due to bad prognosis group	Log hazard ratio = 0.5	✖
Survival time distribution	Alter parameters to test two levels of disease severity	✓
Initially assumed treatment effect	Alter to test two levels of treatment effect	✓
Time-dependence of treatment effect	Treatment effect received depends upon CEA at time of crossover. However set α to zero in some scenarios. Also include additional treatment effect decrement in crossover patients in some scenarios to approximate a continued reduction in treatment effect over time in these patients	✓
Probability of switching treatment over time	Test two levels of treatment crossover proportions	✓
Prognosis of crossover patients	Test three crossover mechanisms in which different groups become more likely to cross over	✓

Scenarios

Variable	Value	Alternative
Sample size	500	✗
Number of prognosis groups (prog)	2	✗
Probability of good prognosis	0.5	✗
Probability of poor prognosis	0.5	✗
Maximum follow-up time	3 years (1095 days)	✗
Multiplication of OS survival time due to bad prognosis group	Log hazard ratio = 0.5	✗
Survival time distribution	Alter parameters to test two levels of disease severity	✓ (2)
Initially assumed treatment effect	Alter to test two levels of treatment effect	✓
Time-dependence of treatment effect	Treatment effect received depends upon CEA at time of crossover. However set α to zero in some scenarios. Also include additional treatment effect decrement in crossover patients in some scenarios to approximate a continued reduction in treatment effect over time in these patients	✓
Probability of switching treatment over time	Test two levels of treatment crossover proportions	✓
Prognosis of crossover patients	Test three crossover mechanisms in which different groups become more likely to cross over	✓

Scenarios

Variable	Value	Alternative
Sample size	500	✗
Number of prognosis groups (prog)	2	✗
Probability of good prognosis	0.5	✗
Probability of poor prognosis	0.5	✗
Maximum follow-up time	3 years (1095 days)	✗
Multiplication of OS survival time due to bad prognosis group	Log hazard ratio = 0.5	✗
Survival time distribution	Alter parameters to test two levels of disease severity	✓ (2)
Initially assumed treatment effect	Alter to test two levels of treatment effect	✓ (4)
Time-dependence of treatment effect	Treatment effect received depends upon CEA at time of crossover. However set α to zero in some scenarios. Also include additional treatment effect decrement in crossover patients in some scenarios to approximate a continued reduction in treatment effect over time in these patients	✓
Probability of switching treatment over time	Test two levels of treatment crossover proportions	✓
Prognosis of crossover patients	Test three crossover mechanisms in which different groups become more likely to cross over	✓

Scenarios

Variable	Value	Alternative
Sample size	500	✗
Number of prognosis groups (prog)	2	✗
Probability of good prognosis	0.5	✗
Probability of poor prognosis	0.5	✗
Maximum follow-up time	3 years (1095 days)	✗
Multiplication of OS survival time due to bad prognosis group	Log hazard ratio = 0.5	✗
Survival time distribution	Alter parameters to test two levels of disease severity	✓ (2)
Initially assumed treatment effect	Alter to test two levels of treatment effect	✓ (4)
Time-dependence of treatment effect	Treatment effect received depends upon CEA at time of crossover. However set α to zero in some scenarios. Also include additional treatment effect decrement in crossover patients in some scenarios to approximate a continued reduction in treatment effect over time in these patients	✓ (8) (12)
Probability of switching treatment over time	Test two levels of treatment crossover proportions	✓
Prognosis of crossover patients	Test three crossover mechanisms in which different groups become more likely to cross over	✓

Scenarios

Variable	Value	Alternative
Sample size	500	✖
Number of prognosis groups (prog)	2	✖
Probability of good prognosis	0.5	✖
Probability of poor prognosis	0.5	✖
Maximum follow-up time	3 years (1095 days)	✖
Multiplication of OS survival time due to bad prognosis group	Log hazard ratio = 0.5	✖
Survival time distribution	Alter parameters to test two levels of disease severity	✓ (2)
Initially assumed treatment effect	Alter to test two levels of treatment effect	✓ (4)
Time-dependence of treatment effect	Treatment effect received depends upon CEA at time of crossover. However set α to zero in some scenarios. Also include additional treatment effect decrement in crossover patients in some scenarios to approximate a continued reduction in treatment effect over time in these patients	✓ (8) (12)
Probability of switching treatment over time	Test two levels of treatment crossover proportions	✓ (24)
Prognosis of crossover patients	Test three crossover mechanisms in which different groups become more likely to cross over	✓

Scenarios

Variable	Value	Alternative
Sample size	500	×
Number of prognosis groups (prog)	2	×
Probability of good prognosis	0.5	×
Probability of poor prognosis	0.5	×
Maximum follow-up time	3 years (1095 days)	×
Multiplication of OS survival time due to bad prognosis group	Log hazard ratio = 0.5	×
Survival time distribution	Alter parameters to test two levels of disease severity	✓ (2)
Initially assumed treatment effect	Alter to test two levels of treatment effect	✓ (4)
Time-dependence of treatment effect	Treatment effect received depends upon CEA at time of crossover. However set α to zero in some scenarios. Also include additional treatment effect decrement in crossover patients in some scenarios to approximate a continued reduction in treatment effect over time in these patients	✓ (8) (12)
Probability of switching treatment over time	Test two levels of treatment crossover proportions	✓ (24)
Prognosis of crossover patients	Test three crossover mechanisms in which different groups become more likely to cross over	✓ (72)

This combined to 72 scenarios

Scenarios

Variable	Value	Alternative
Sample size	500	✗
Number of prognosis groups (prog)	2	✗
Probability of good prognosis	0.5	✗
Probability of poor prognosis	0.5	✗
Maximum follow-up time	3 years (1095 days)	✗
Multiplication of OS survival time due to bad prognosis group	Log hazard ratio = 0.5	✗
Survival time distribution	Alter parameters to test two levels of disease severity	✓
Initially assumed treatment effect	Alter to test two levels of treatment effect	✓
Time-dependence of treatment effect	Treatment effect received depends upon CEA at time of crossover. However set α to zero in some scenarios. Also include additional treatment effect decrement in crossover patients in some scenarios to approximate a continued reduction in treatment effect over time in these patients	✓
Probability of switching treatment over time	Test two levels of treatment crossover proportions	✓
Prognosis of crossover patients	Test three crossover mechanisms in which different groups become more likely to cross over	✓

Performance measures

Similar to James' study…

- We used bias as our primary performance measure
- Also assessed coverage

However, because our treatment effect is time-dependent there is not a 'true' HR or AF

- Therefore we used restricted mean survival as our true measure. We estimated the truth from our survivor function equations
- This is highly relevant for the context of economic evaluation
- But means that we had to estimate restricted mean survival for each crossover method – not just the adjusted HR or AF

Estimating survival

Three broad approaches assessed (all estimated out to 3 years):

1. **'Survivor function' approach**

   Apply treatment effect to survivor function (or hazard function) estimated for experimental group → calculate AUC

2. **'Extrapolation' approach**

   Extrapolate counterfactual dataset to required time-point (only relevant for RPSFTM/IPE approaches) → calculate AUC

3. **'Shrinkage' approach**

   Use estimated acceleration factor to 'shrink' survival times in crossover patients in order to obtain an adjusted dataset → calculate AUC (only relevent for AF-based approaches)

Methods

Naive methods

- ITT
- Exclude crossover patients (PPexc)
- Censor crossover patients (PPcens)
- Treatment group as a time-dependent covariate (TDCM)
- Treatment crossover as a time-dependent indicator (XOTDCM)

Methods

Complex methods

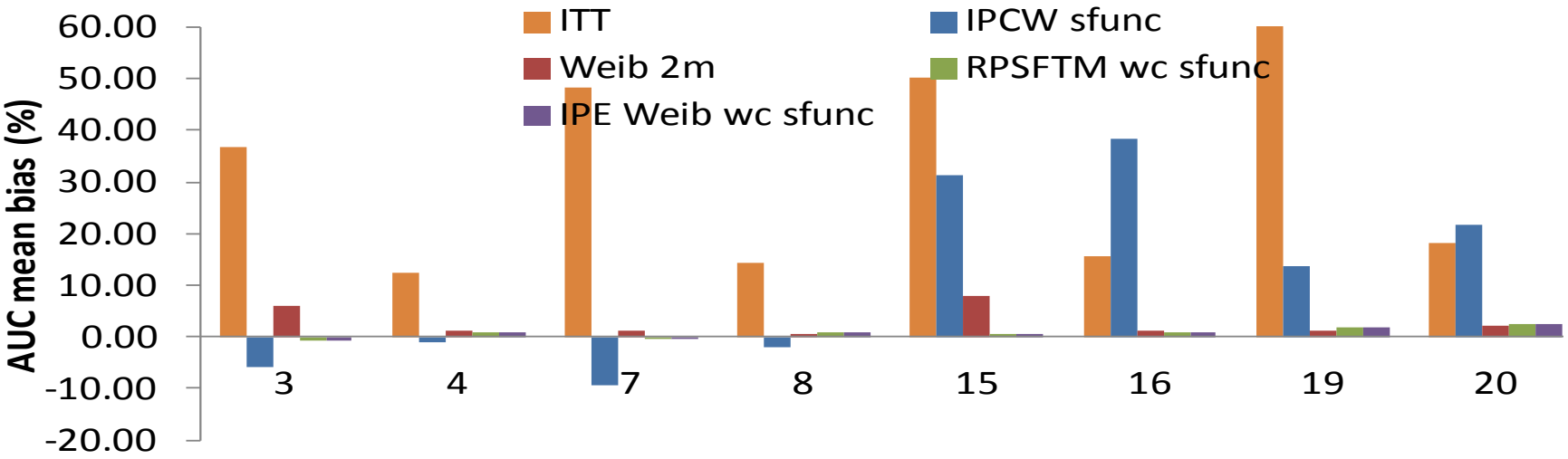
- RPSFTM with log-rank test (with and without covariates)
- IPE algorithm (Weibull and exponential versions, with and without covariates)
- IPCW
- SNM with g-estimation
- Two-stage Weibull method (Weib 2m)

Note, we did not include:

- Walker et al's method due to poor performance in James' study
- Loey and Goetghebeur's method as only for all-or-nothing compliance
- Law and Kaldor's method as fundamentally flawed
- And only included log-rank test version of RPSFTM

Results (1)

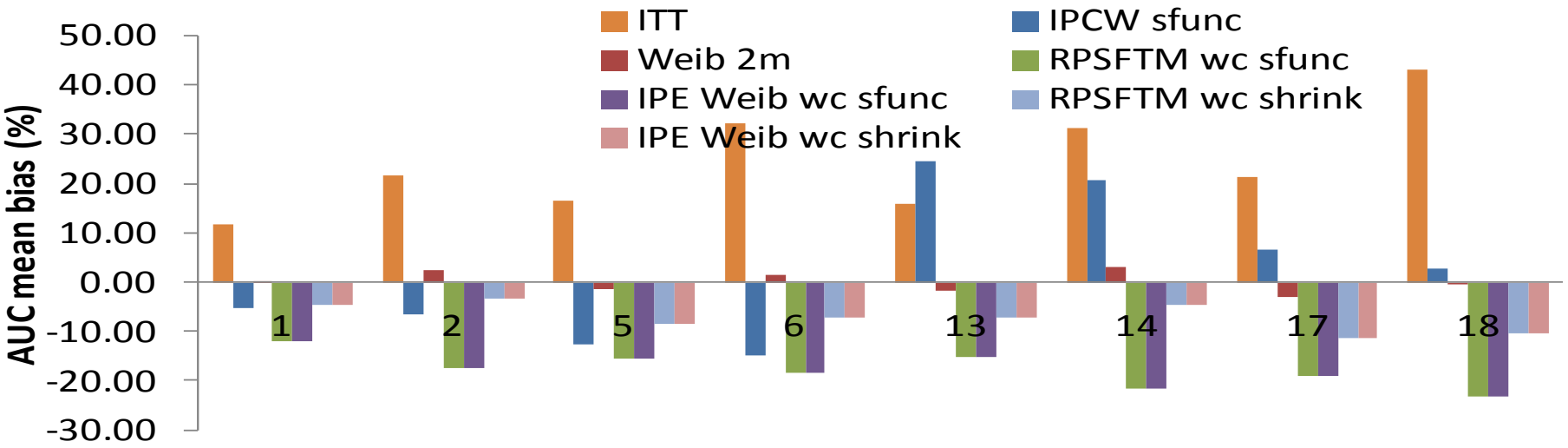
- Randomisation-based methods worked very well when the treatment effect was not time-dependent, eg:



- IPCW method performed poorly when crossover proportion was very high
- SNM method performed poorly
- Naive methods performed poorly
- Two-stage Weibull produced low bias

Results (2)

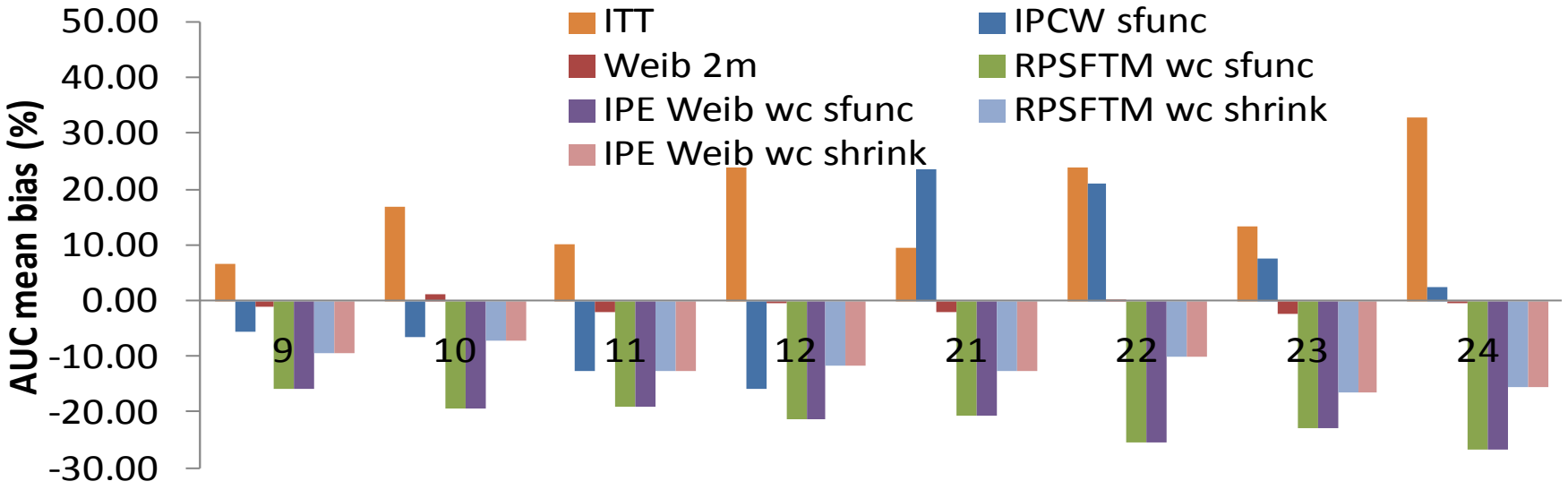
- Randomisation-based methods produced large bias when the treatment effect was time-dependent, eg:



- RPSFTM/IPE 'shrinkage' approach performed better, but this is flawed
- IPCW method performed better than randomisation-based approaches providing crossover proportion was less than 90%, but stilll gave considerable bias

Results (3)

- When there was an additional treatment effect decrement in crossover patients, indicating a particularly strong time-dependent treatment effect, the randomisation-based methods performed even less well:



- IPCW is unaffected by this, and becomes more likely to produce least bias (excluding two-stage Weibull approach)

Conclusions

- RPSFTM / IPE survivor function methods produce very low levels of bias when the treatment effect is not time-dependent
- When the treatment effect (in terms of an AF) is 20-30% lower in crossover patients RPSFTM / IPE survivor function approaches produce high levels of bias (>10%)
  - 'Shrinkage' approaches perform with lower bias but these methods are flawed
- When the treatment effect decrement is >30% IPCW produces less bias than any RPSFTM / IPE variant, providing <90% of at-risk patients crossover
  - But significant bias remains
- Where applicable, two-stage methods are worthy of consideration
- 'Survivor function' approaches generally produce lower bias than 'extrapolation' approaches, due to the loss of information associated with recensoring and the effect of this on the extrapolation

