Targeted Maximum Likelihood Estimation for Restricted Mean Survival Time in time-to-event data

A case study using a previous non-randomised PAS study

Speaker: Michael Seath

Contents

1. Motivation
2. Introduction to TMLE
3. TMLE to estimate Restricted Mean Survival Time (RMST) in a time-to-event setting
4. Case study: a non-randomised Post Authorisation Safety (PAS) study

The focus of this presentation is on application. For a theoretical underpinning to this estimation method, see Diaz et al (2019) https://doi.org/10.1007/s10985-018-9428-5.

Application is with the accompanying R package survtmlerct (https://github.com/idiazst/survtmlerct). This is **not on CRAN**.

Motivation

Motivation

- Observational studies are non-randomised.
  - Creates potential confounders -> biased estimates.
- Time-to-event data is even more challenging:
  - Different censoring mechanisms, not just different baseline characteristics.
  - Hazard ratio has complicated causal interpretation.

GSK

Motivation

Applying TMLE to estimate Restricted Mean Survival Time in time-to-event data:

- accounts for confounders without over-dependence on parametric modelling assumptions
- accounts for confounding from censoring and baseline characteristics
- estimates a quantity with a more straight forward causal interpretation than the hazard ratio.

Causal inference: an overview

- **Objective**: Identify a target causal estimand and estimate it.
  - Causal estimand: A function of potential outcomes.
    - Potential outcomes are variables describing what *would* happen under hypothetical scenarios (counterfactuals)
- **Example target estimand**: Average Treatment Effect (ATE)

$$ATE = E[Y(1) - Y(0)]$$

$$ATE = E_{\boldsymbol{W}}[E[Y|A = 1, \boldsymbol{W}] - E[Y|A = 0, \boldsymbol{W}]]$$

  - *Y*: outcome; Y(1) & Y(0) are potential outcomes when one is treated & untreated.
  - *A*: treatment, ***W***: confounders
- **Key causal assumptions**:
  - Causal consistency: observed outcome under treatment a equals corresponding potential outcome: Y(a) = Y|A=a
  - No unmeasured confounders: we can **collect enough confounders W** to condition on to treat the data as if randomized
  - Positivity/overlap: every subject has a non-zero probability of receiving either treatment
- **Adjustment for confounding** is crucial in observational studies

Introduction to TMLE

GSK

Common methods for confounding adjustment

G-computation

Relies on correct specification of the outcome model E(Y|T, X)

$\hat{E}[Y|T=1, \boldsymbol{X}]$ & $\hat{E}[Y|T=0, \boldsymbol{X}]$
Generate predicted outcome under both exposure levels

$\widehat{ATE}$ = mean difference in **predicted outcomes** between Treated and Untreated

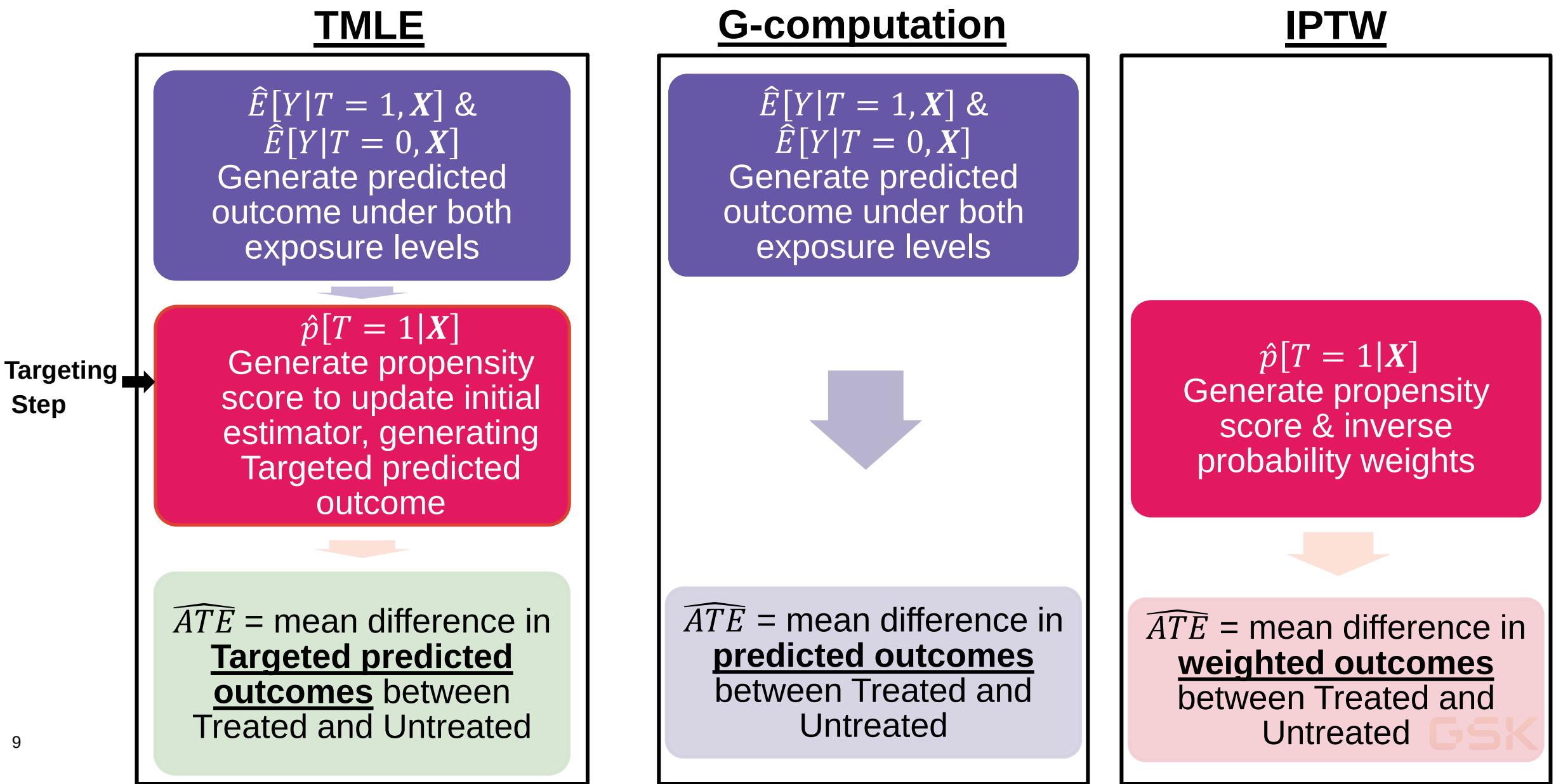
IPTW

Relies on correct specification of the assignment model P(T=1|X)

$\hat{p}[T=1|\boldsymbol{X}]$
Generate propensity score & inverse probability weights

$\widehat{ATE}$ = mean difference in **weighted outcomes** between Treated and Untreated

TMLE: Targeted Maximum Likelihood Estimation



Advantages of TMLE

- **Doubly robust**
  - Consistent estimator (unbiased) if either the propensity or outcome model is correctly specified
  - Efficient estimator (better precision) if both models are correctly specified

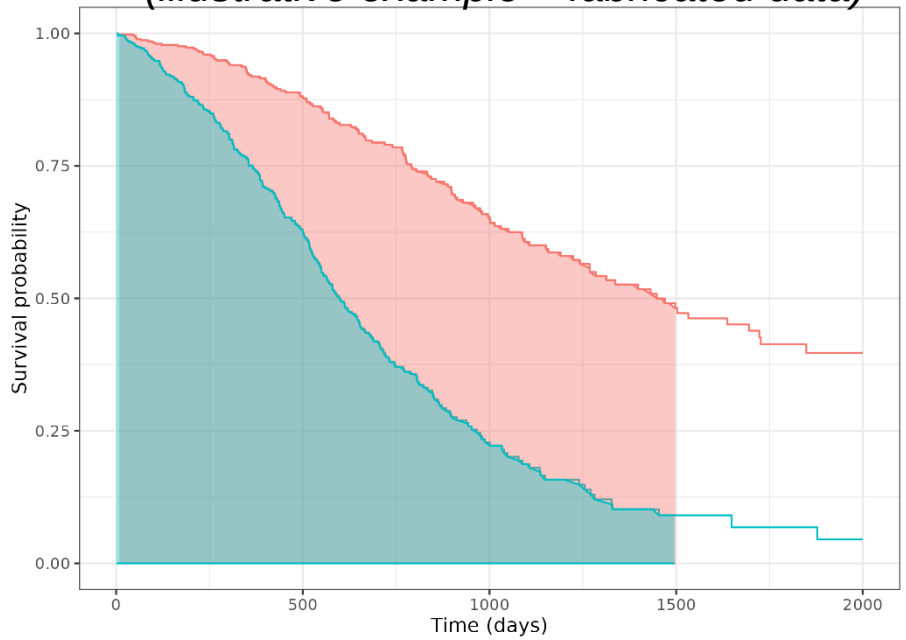
- **Flexible to incorporate machine learning algorithms for both models**
  - No need to specify parametric models
  - Commonly used algorithms: LASSO, random forests, generalized boosted regression etc.
  - Super learner allows combination of many machine learning algorithms

TMLE to estimate Restricted Mean Survival Time (RMST) in a time-to-event setting

GSK

TMLE: targeting difference in Restricted Mean Survival Time (RMST)

(illustrative example – fabricated data)



Restricted Mean Survival Time:

- Pick a cut-off time τ.
- The target parameter θ (causal estimand) is the difference in mean survival time up to τ:

$$\theta = \int_0^\tau S_1(t)dt - \int_0^\tau S_0(t)dt$$

This is an *estimand:* a function of *potential outcomes*.

$S_1(t)$ is the survival probability at time t under the hypothetical scenario of being assigned treatment 1.

Similar for $S_0(t)$

Interpretation: Participants in cohort 1 survive on average θ days (/weeks/months...) longer than participants in cohort 0 over a period of τ days (/weeks/months...)

Why RMST? Why not keep a hazard ratio?

The hazard ratio is not a causal estimand. The hazard ratio is

$$\lim_{\delta \to 0} \frac{P(T < t + \delta | T > t, A = 1)}{P(T < t + \delta | T > t, A = 0)}$$

definition involves conditioning on post-treatment assignment information

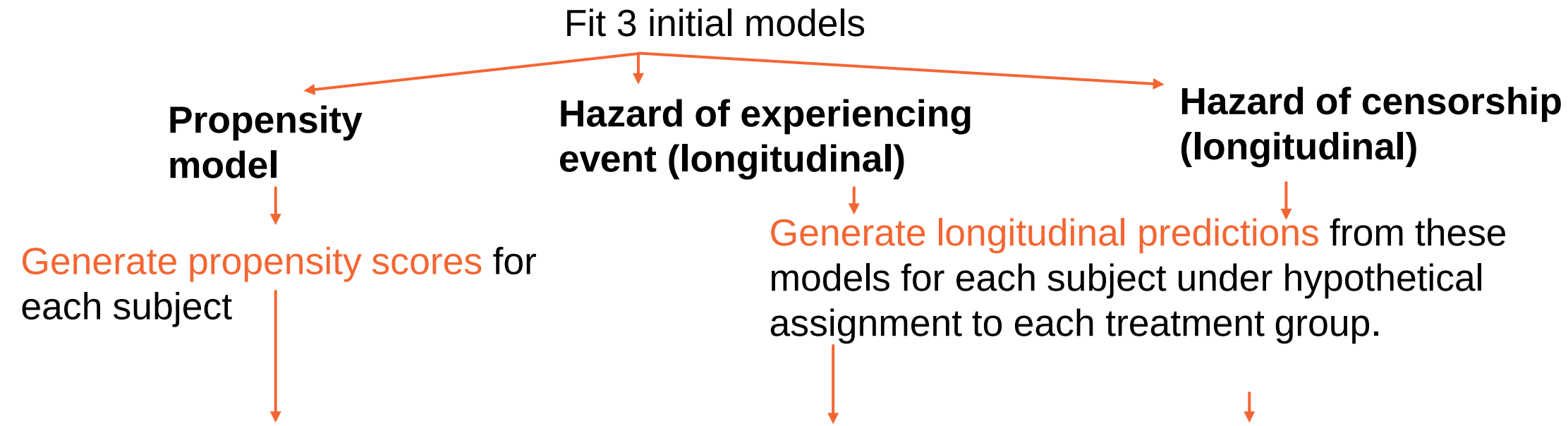
TMLE for time-to-event

Transform time-to-event data to longitudinal

subjid	cohort	time	status
1	Placebo	3	1
2	Active	4	0

subjid	cohort	t	event	censor
1	Placebo	1	0	0
1	Placebo	2	0	0
1	Placebo	3	1	0
2	Active	1	0	0
2	Active	2	0	0
2	Active	3	0	0
2	Active	4	0	1

TMLE: Usage with time-to-event data



By fitting both a propensity model and a hazard of censorship model, estimation adjusts both for

- differences in subject characteristics between the cohorts, AND
- differences in censoring mechanisms between the cohorts

TMLE for time-to-event

Append predictions from trained models

Probability of event/censor under hypothetical of receiving treatment 1/0

subjid	cohort	t	event	censor	Propensity Score	Prob event_1	Prob event_0	Prob censor_1	Prob censor_0
1	Placebo	1	0	0	0.3	0.01	0.01	0.1	0.11
1	Placebo	2	0	0	0.3	0.015	0.015	0.1	0.11
1	Placebo	3	1	0	0.3	0.02	0.015	0.11	0.12
2	Active	1	0	0	0.85	0.001	0.001	0.1	0.11
2	Active	2	0	0	0.85	0.001	0.0015	0.1	0.12
2	Active	3	0	0	0.85	0.0015	0.0022	0.15	0.15
2	Active	4	0	1	0.85	0.002	0.0025	0.15	0.15

TMLE uses predictions to estimate difference in RMST

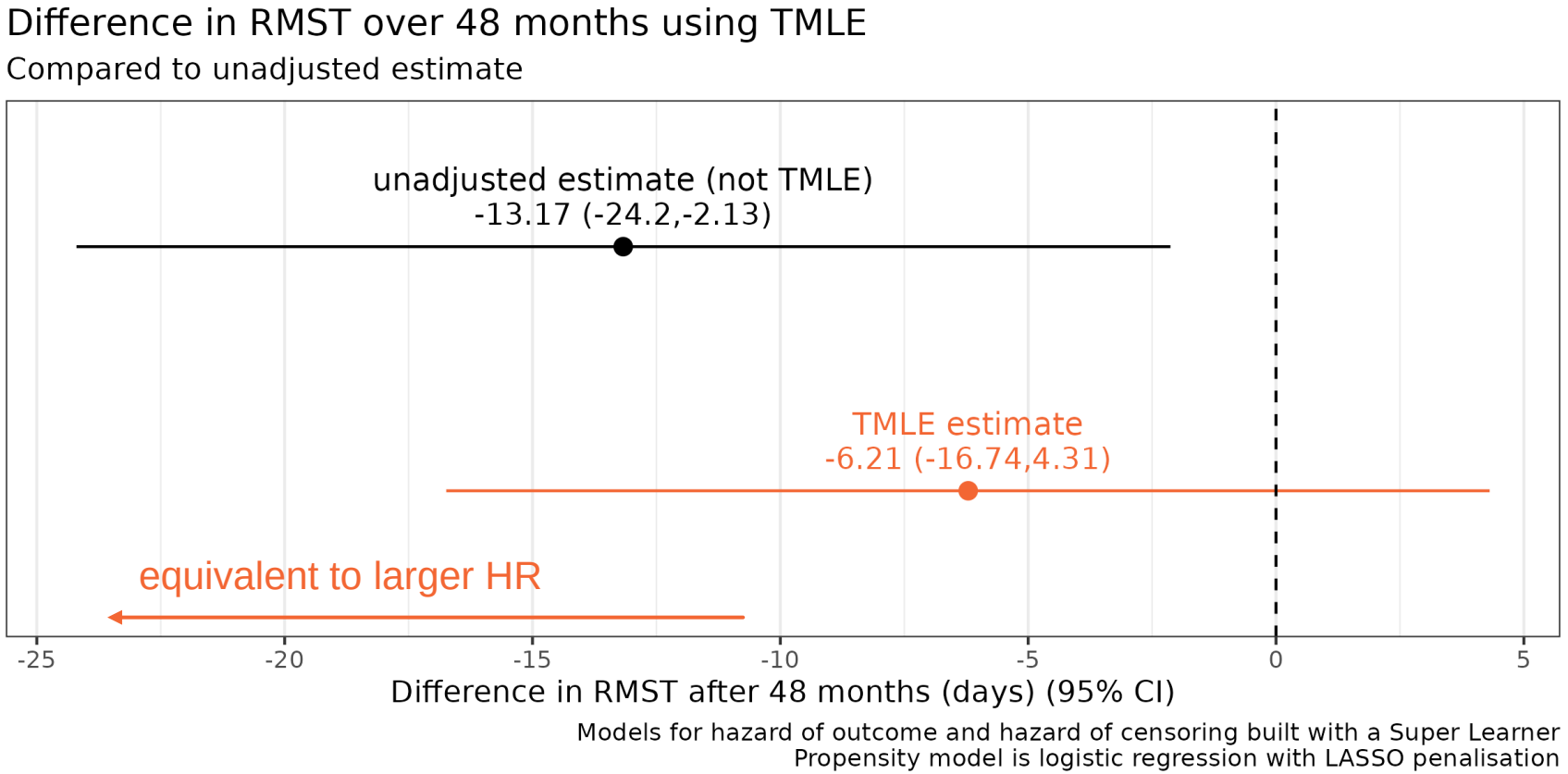
Case study: a non-randomised Post Authorisation Safety (PAS) study

Post Authorisation Safety Study

- Non-randomised – treatment assignment based on actual clinical setting
- Follow up for 2-5 years
- Routine data collection
- Primary objective: non-inferiority of treatment versus comparator for time-to-first-event of a composite endpoint of a number of safety events.
- Low event rates.
- Study originally analysed using Cox PH model, with Inverse Probability of Treatment Weights (IPTW), plus additional covariate adjustment.

TMLE RMST results using PAS study data

Using the propensity model used in the study and our modified propensity model.



TMLE pulls in the estimated treatment difference versus no adjustment

- RMST contextualises the treatment effect in terms of **differences** in survival time
  - with **low event rates** these estimates are **less sensitive** to changes than a hazard ratio

TMLE estimates used:

- Super Learner (LASSO, random forest, MARS, XGBoost) for the hazard models.
- LASSO-penalised logistic regression for the propensity model.

Acknowledgements

- **Stephen Weng –** project oversight
- **Tom Taverner –** aid in navigating the literature for TMLE for survival analysis and providing feedback on methodology.
- **Ali Donald –** support in application to study
- **Lodo Edwards –** running simulations

References

Díaz, I., Colantuoni, E., Hanley, D.F. et al. Improved precision in the analysis of randomized trials with survival outcomes, without assuming proportional hazards. Lifetime Data Anal 25, 439–468 (2019). https://doi.org/10.1007/s10985-018-9428-5

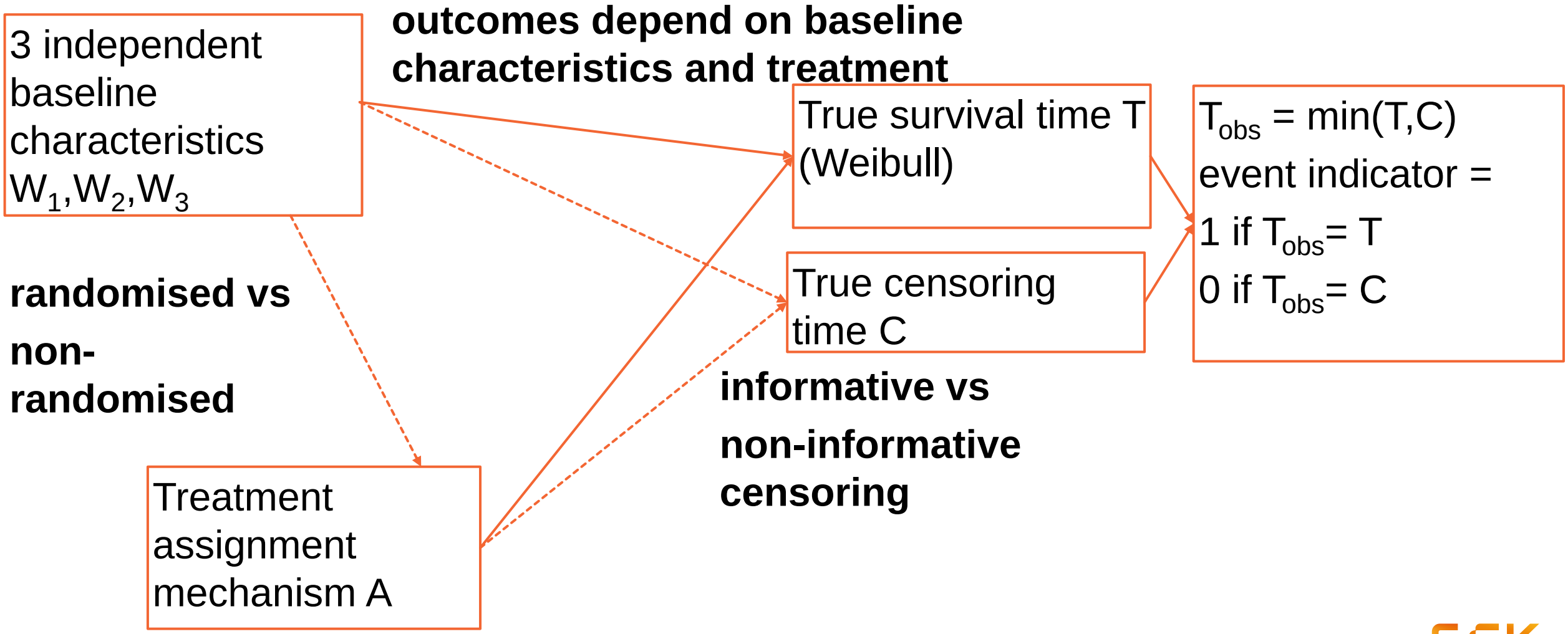
Martinussen, Torben, Causality and the Cox Regression Model (March 1, 2022). Annual Review of Statistics and Its Application, Vol. 9, Issue 1, pp. 249-259, 2022, Available at SSRN: https://ssrn.com/abstract=4065374 or http://dx.doi.org/10.1146/annurev-statistics-040320-114441

Backup slides

Simulations

Simulation process

Survival data is non-randomised with informative censoring
number of simulations = 10,000



Simulation process - details

Number of simulations = 10,000

n=4000

Independent baseline characteristics

W1 ~ Bern(0.7)

W2 ~ Bern(0.3)

W3 ~ N(0,1)

Treatment assignment

Non-randomised (A depends on W): $\eta_A = 2W_1-3W_2+4W_3+ 6e_0$

Randomised: $\eta_A = e_0$

A ~ Bern(η_A-mean(η_A))

Survival time is Weibull distributed, and depends on A and W

$\eta_T \sim \log(\lambda)-2A+4W_1-3W_2-2W_3+e_1$

T ~ Weibull(k,exp(η_T))

Censoring is informative (depends on A and W) or non-informative (random).
Censoring truncated at t=2000

Informative censoring (depends on A and W): $\eta_C = \log(1000)+A-0.5W_1+0.5W_2+0.1e_2$

Non-informative censoring: $\eta_C = \log(1000)+0.1e_2$

C ~ Exp(1 / exp(η_C)) , truncated at 2000

T_{obs}=min(T,C)

event = 1 if T_{obs}=T , 0 if T_{obs}=C

Random noise variables

$e_i \sim N(0,1)$

Weibull parameters

$\lambda = 5.85 * 10^7$

k = 0.271

Simulation results: headings

- In a randomised, non-informative censoring setting, TMLE increased precision relative to unadjusted KM-based estimation
- For non-randomised, informative censoring, coverage maintained as long as either:
  - the outcome model is correct
  - both the propensity model and the censor model are correct

Simulation results – non-informative censoring

	Data	Method	Bias	MSE	SD point estimate	Coverage (%) of 95% CIs	Mean width of 95% CI
All models correctly specified	Randomised	Unadjusted	0.018	106	10.3	95.7%	41.4
	Randomised	TMLE	-0.01	101	10.1	95.7%	**40.4**
	Non-randomised	Unadjusted	-15.5	349	10.5	70.4%	41.8
	Non-randomised	TMLE	-0.02	150	12.2	94.8%	47.4
Propensity model misspecified*	Non-randomised	TMLE	-0.08	139	11.8	95.2%	46.5
All models misspecified*	Non-randomised	TMLE	10.3	248	11.9	**87.1%**	46.8

* Misspecified = W_2 excluded from the model

GSK

Simulation results – informative censoring and non-randomised

	Method	Bias	MSE	SD point estimate	Coverage (%) of 95% CIs	Mean width of 95% CI
All models correctly specified	Unadjusted	-15.5	339	10.0	70.3	41.2
	TMLE	-0.916	135	11.6	95	45.8
Propensity model misspecified*	TMLE	-0.729	128	11.3	95.2%	45.4
Outcome model misspecified*	TMLE	-0.786	144	12.0	94.9%	46.6
Censor model misspecified*	TMLE	-0.64	138	11.7	95.2%	46.6
Only propensity model correct	TMLE	-0.667	143	12.0	95.3%	46.9
Only outcome model correct	TMLE	-0.705	133	11.5	95.6%	46.0
Only censor model correct	TMLE	9.55	226	11.6	88.3%	45.8

* Misspecified = W_2 excluded from the model

***If the set of Ws that impact censoring mechanism and treatment mechanism are different this may not hold*

gsk.com