

Evaluating the Precision of TMLE in Vaccine Efficacy Estimation: Two Case Studies

Giulia Zigon and Joshua Havumaki



NX-GBL-RSA-PPT-260014

June 2026

Disclaimers

The following slides concern the application and comparison of statistical models on GSK proprietary data. The new results generated are useful solely for evaluating the applied methodologies.

Prescribing information for GSK's respiratory syncytial virus (RSV) vaccine (recombinant, adjuvanted) ▼ – also known as adjuvanted RSVPreF3 OA is available by scanning the QR code below:



Adverse events should be reported. Reporting forms and information can be found at <https://yellowcard.mhra.gov.uk/> or search for MHRA Yellowcard in the Google Play or Apple Store. Adverse events should also be reported to GlaxoSmithKline on +44 (0) 800 221 441 or UKSafety@gsk.com

Objective

To explore whether Targeted Maximum Likelihood Estimation **TMLE** can be used to improve **the precision** of vaccine efficacy estimates in phase 2 and phase 3 vaccine clinical trial settings

Methods overview

Estimating covariate-adjusted vaccine efficacy

Treatment effect definition

- Cumulative incidence:

$$CI(t) = 1 - S(t)$$

- Vaccine efficacy (VE):

$$VE(t) = 1 - \frac{CI_{vaccine}(t)}{CI_{control}(t)}$$

Modeling strategy

- Super Learner (ensemble ML)
- TMLE (Targeted Maximum Likelihood Estimation)
→ Doubly robust estimation of causal effects with covariate adjustment

Implementations

- TMLE + Super Learner (survival)
- TMLE + Super Learner (Cox-based)
- Parametric TMLE (no ML)
- Benchmark methods
 - Cox model (G-computation)
 - Weibull models (G-computations)
 - Kaplan–Meier + IPW

Case study 1: Efficacy study of GSK's respiratory syncytial virus (RSV) vaccine in adults aged 60 years and above.

Giulia Zigon and Andrea Callegaro



Vaccine Efficacy

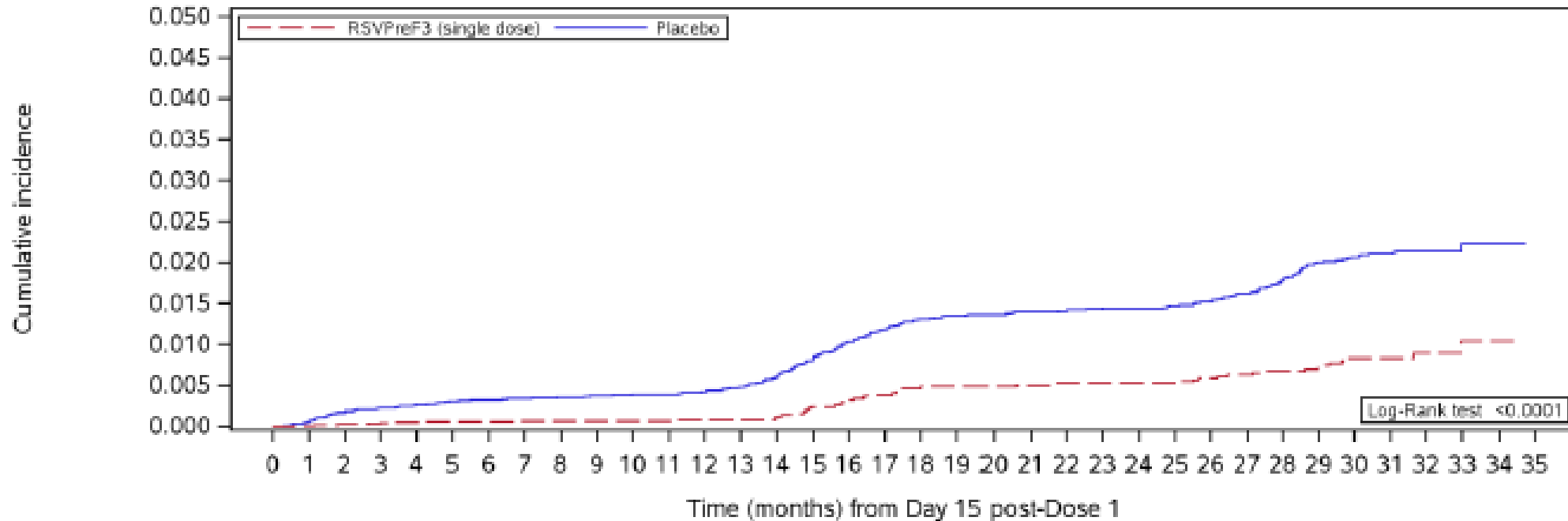
Vaccine efficacy against first occurrence of RT-PCR-confirmed RSV LRTD up to End of Season 1-2 and 3 in Northern Hemisphere, using Poisson method - Single dose - mES

Endpoint	RSVPreF3				Placebo				VE			
	N	n	T(year)	n/T (per 1000)	N	n	T(year)	n/T (per 1000)	%	LL	UL	P-value
RT-PCR-confirmed RSV LRTD	12466	7	6865.9	1.0	12494	40	6857.3	5.8	82.58	57.89	94.08	<0.0001
	RSVPreF3 (Single dose)				Placebo				VE			
	N	n	T(year)	n/T (per 1000)	N	n	T(year)	n/T (per 1000)	%	LL	UL	P-value
	12469	30	14662.6	2.0	12498	139	17269.0	8.0	67.18	48.19	80.04	<0.0001
	RSVPreF3 (Single dose)				Placebo				VE			
	N	n	T(year)	n/T (per 1000)	N	n	T(year)	n/T (per 1000)	%	LL	UL	P-value
	12468	48	19748.8	2.4	12498	215	27363.6	7.9	62.91	46.74	74.79	<0.0001

Source: Table 14.2.1.1 (VE Analysis 1 Unblinded Report, 13 August 2022 [GSK Study Report 212494, VE1])
 Table 14.2.1.81 (VE Analysis 3 Unblinded Report, 11 August 2023 [GSK Study Report 212494, VE3])
 Table 8.2.1.183 (23JUL2024 17:48 GMT)

Cumulative incidence curves for RSV LRTD reported up to End of Season 3

Cases reported up to End of Season 3



Number at risk

RSVPreF3 (single dose)	12468	12313	12227	12127	11979	11243	7991	5470	4935	4878	4841	4813	4755	4672	4448	3787	1536	55
Placebo	12498	12317	12195	12110	11949	11611	10843	10069	9785	9703	9639	9561	9424	9243	8785	7446	3014	109

Cumulative number of cases

RSVPreF3 (single dose)	0	1	3	5	6	7	7	8	8	8	8	8	10	10	11	18	22	25	29	30	30	31	32	32	32	32	35	37	39	41	46	46	47	48	48	48
Placebo	0	8	20	27	32	37	39	41	43	45	47	47	51	57	70	92	113	127	140	144	146	149	150	153	153	156	160	170	185	204	209	212	214	215	215	215

Alternative Methods for VE estimation

GSK

Analytical methods

Estimation strategy and covariate adjustment

Outcome

Cumulative incidence ratio evaluated at:

12 months

24 months

32 months (maximum follow-up with LRTD case)

LRTD Lower Respiratory Tract Disease

Adjustment strategy

Analyses performed with increasing levels of covariate adjustment:

Unadjusted model

Core adjustment set

→ Age, Sex, Region, Frailty status

Extended adjustment

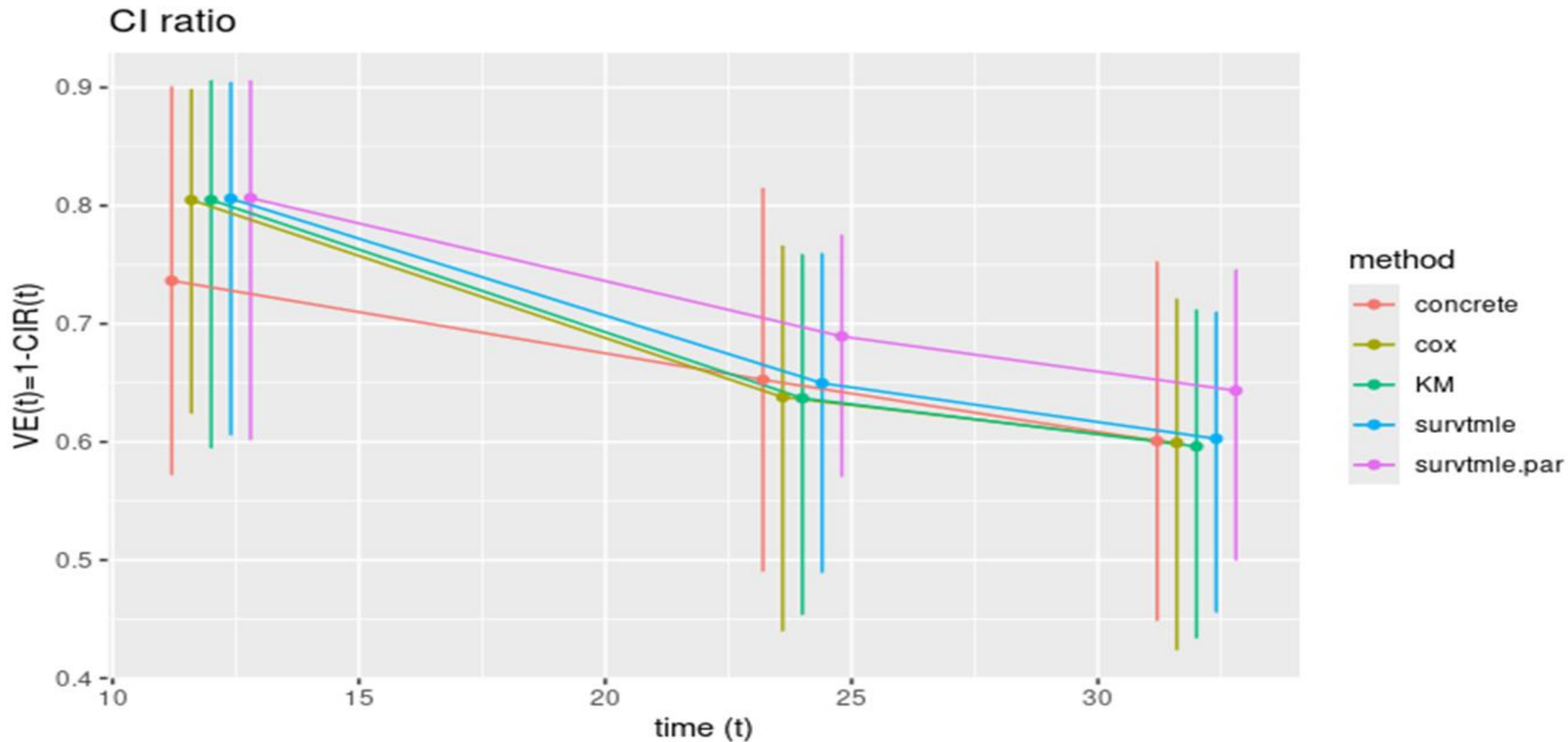
→ + Race

Full adjustment

→ + Ethnicity













Results adjusting for age, sex, region and frailty status





$VE(t)=1-CIR(t)$ timepoints 12, 24 and 32 months



Relative efficiency values for each method and timepoint (vs Kaplan–Meier)

Adjusted CIR – 12, 24, 32 months (Age, sex, region, frailty adjusted)

Method	12 months	24 months	32 months	Overall Pattern
Kaplan–Meier	Reference	Reference	Reference	Baseline (stable)
Cox	 115%	 87%	 83%	Increasing variance over time
Concrete	 71%	 79%	 75%	Consistently higher variance
survtmle	 106%	 118%	 115%	Consistently reduced variance
survtmle.par	 115%	 123%	 67%	Lower early, less stable at 32 months

-  Most methods agree on CIR trends, but not all
-  Concrete and survtmle.par introduce deviations in estimates
-  Variance inflation + estimate instability (Concrete) = weakest option
-  survtmle provides the most robust results (consistent CIR + lower variance)

Conclusions

Main finding

TMLE with Super Learner provides **robust and efficient estimation of vaccine efficacy** compared to traditional approaches

- ✓ TMLE with ML (Super Learner)
- ✓ Comparable point estimates to standard methods
- ✓ Improved precision vs Kaplan–Meier (lower variance)
- ✓ Stable across adjustment strategies
- ✓ No convergence issues observed
- ✓ Robust to time scale (days vs months)

⚠ Limitations

- ! Precision does not always improve with additional covariates
- ! Computational cost (\approx 5–6 hours)

Parametric TMLE

Sensitive to model specification

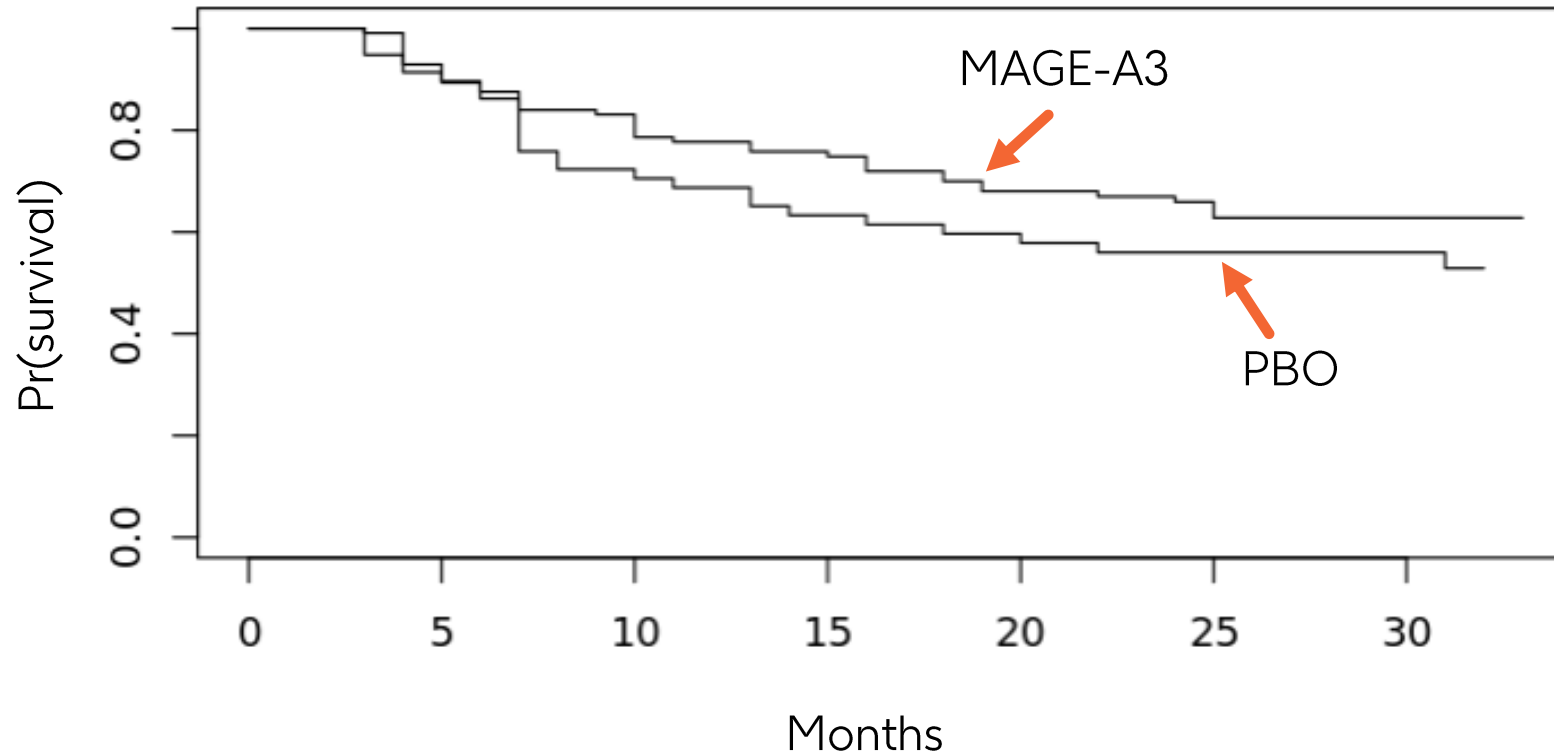
Estimates may deviate from ML-based TMLE

Case study 2: Challenges and Insights from GSK's Historical MAGE-A3 Cancer Vaccine Data

Joshua Havumaki and Andrea Callegaro



Kaplan Meier plot on disease-free survival



Unadjusted HR MAGE-A3 vs. PBO: 0.73 (95% CI: 0.45, 1.2)
















Fig 2B from Vansteenkiste et al. 2013

A large, flowing orange graphic that starts as a thick, rounded shape on the left and tapers into a thin, curved line extending across the top of the page.

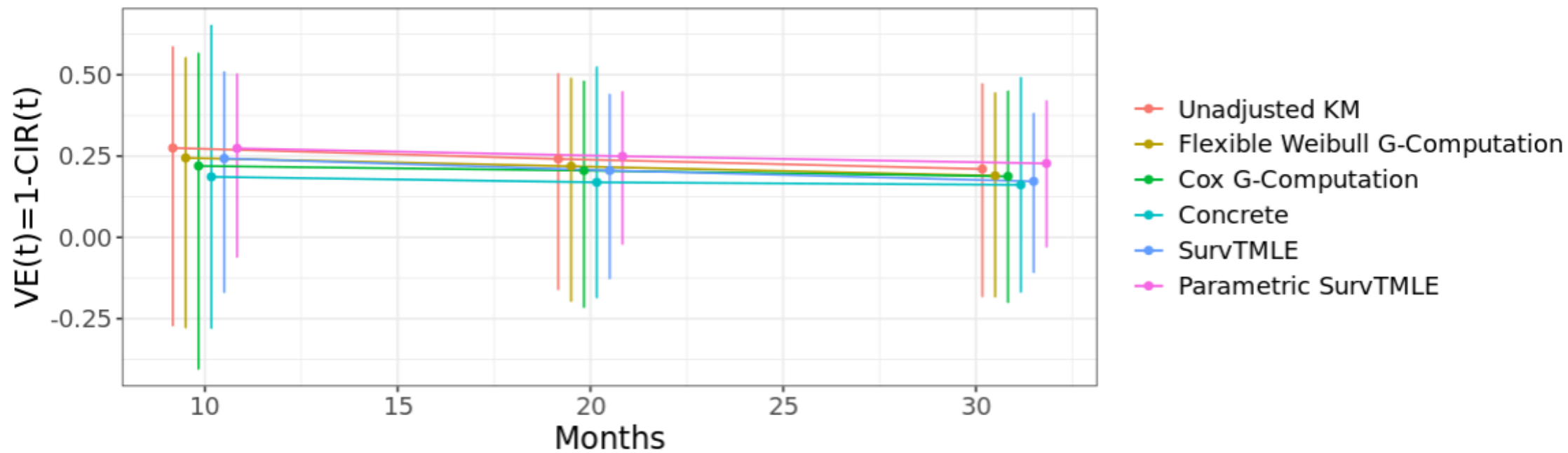
Results

GSK

Gains in efficiency: Relative efficiency values for each method and timepoint
 Variance (log (RR)) of unadjusted KM/variance(log(RR)) of given method

	10 months	20 months	31 months
Flexible Weibull G-Computation	 114%	 99%	 114%
Cox G-Computation	 91%	 100%	 107%
Concrete TMLE	 74%	 86%	 94%
SurvTMLE	 167%	 147%	 191%
Parametric SurvTMLE	 220%	 190%	 197%

Vaccine efficacy Results from all methods



What is happening under the null hypothesis at 31 months?

	Proportion of times null rejected over 5,000 runs
Flexible Weibull G-Computation	2.3%
Cox G-Computation	1.9%
Concrete TMLE	6.3%
SurvTMLE	5.1%
Parametric SurvTMLE	4.7%

Final conclusions

- **Statistical issues:**
 - Gain in power probably due to type I error inflation
 - CVTMLE not yet available for survival outcomes
- **Software issues:**
 - SurvTMLE and Concrete are no longer maintained on CRAN
 - Additional packages explored (e.g., ltmle, TMLE3, MOSS), but **we encountered implementations issues**
 - Machine learning (cross-validation) appears to produce errors in implementation but without clear messages to understand the source
- **Concerns for primary analysis:**
 - Is type I error inflation acceptable?
 - Is it acceptable to use error-prone software that isn't maintained?

References

- Benkeser, D. C., and N. S. Hejazi. "survtmle: Targeted minimum loss-based estimation for survival analysis in R." 2017,
- Cai W, van der Laan MJ (2019+). One-step TMLE for time-to-event outcomes. Working paper.
- Chen, David, et al. "concrete: Targeted Estimation of Survival and Competing Risks in Continuous Time." arXiv preprint arXiv:2310.19197 (2023).
- Coyle, Jeremy. "tmle3: The extensible TMLE framework." *Zenodo* (2021).
- Lendle, Samuel D., et al. "ltmle: an R package implementing targeted minimum loss-based estimation for longitudinal data." *Journal of Statistical Software* 81 (2017): 1-21.
- Van Der Laan, Mark J., and Daniel Rubin. "Targeted maximum likelihood learning." (2006).
- Vansteenkiste, Johan, et al. "Adjuvant MAGE-A3 immunotherapy in resected non–small-cell lung cancer: phase II randomized study results." *Journal of Clinical Oncology* 31.19 (2013): 2396-2403.