A sensitivity analysis approach for the causal hazard ratio in randomized and observational studies

Rachel Axelrod

Tel Aviv University

ISDSA 2024

joint work with Daniel Nevo

The IMvigor211 study

THE LANCET

Volume 391, Issue 10122, 24 February–2 March 2018, Pages 748-757

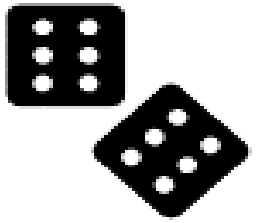
Articles

Atezolizumab versus chemotherapy in patients with platinum-treated locally advanced or metastatic urothelial carcinoma (IMvigor211): a multicentre, open-label, phase 3 randomised controlled trial

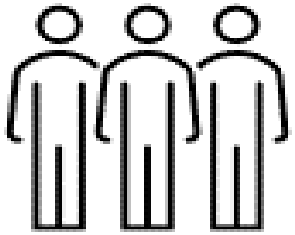
Prof Thomas Powles MD [a], Ignacio Durán MD [b c], Michiel S van der Heijden MD [d], Yohann Loriot MD [e], Nicholas J Vogelzang MD [f], Ugo De Giorgi MD [g], Prof Stéphane Oudard MD [h], Prof Margitta M Retz MD [i], Daniel Castellano MD [j], Prof Aristotelis Bamias PhD [k], Aude Fléchon MD [l], Gwenaëlle Gravis MD [m], Prof Syed Hussain MD [n], Toshimi Takano MD [o], Ning Leng PhD [p], Edward E Kadel III BS [p], Romain Banchereau PhD [p], Priti S Hegde PhD [p], Sanjeev Mariathasan PhD [p], Na Cui PhD [p]...Alain Ravaud MD [q]

The IMvigor211 study

Randomized
Controlled trial



625 patients with
urothelial carcinoma



51% received Atezolizumab (Atezo)
49% received standard chemotherapy
(Chemo)

Standard analysis

- **Goal:** Estimate and compare between the Atezo and Chemo treatment effects on the cancer patients.
- The standard analysis: estimation of the hazard ratio (HR) using the Cox model.

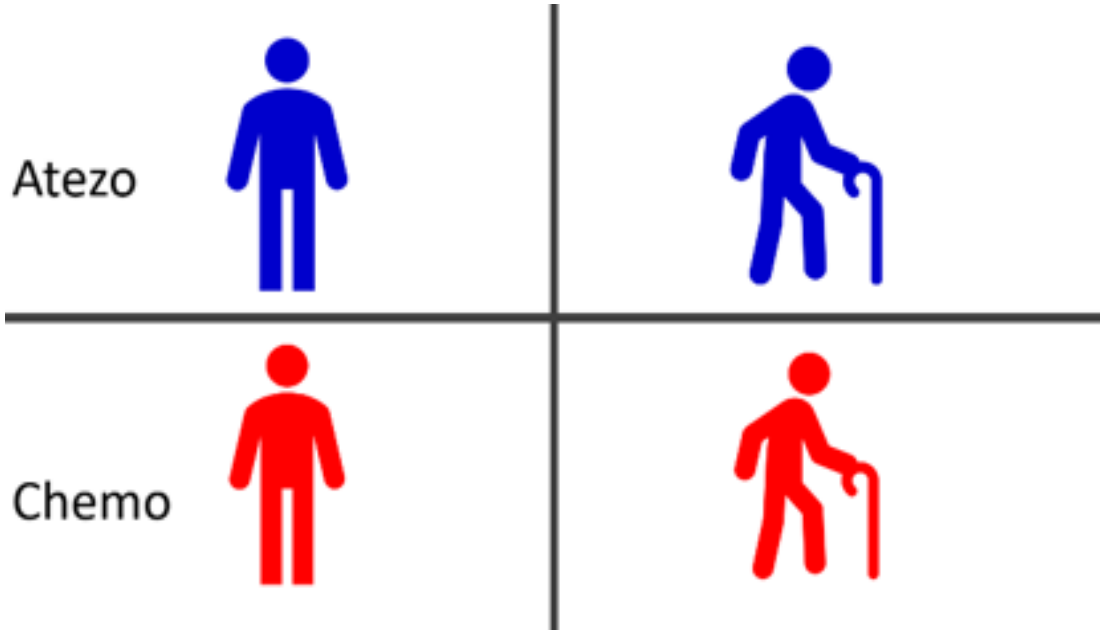
Standard analysis

- **Goal:** Estimate and compare between the Atezo and Chemo treatment effects on the cancer patients.
- The standard analysis: estimation of the hazard ratio (HR) using the Cox model.
- **The HR suffers from an unclear causal interpretation.**

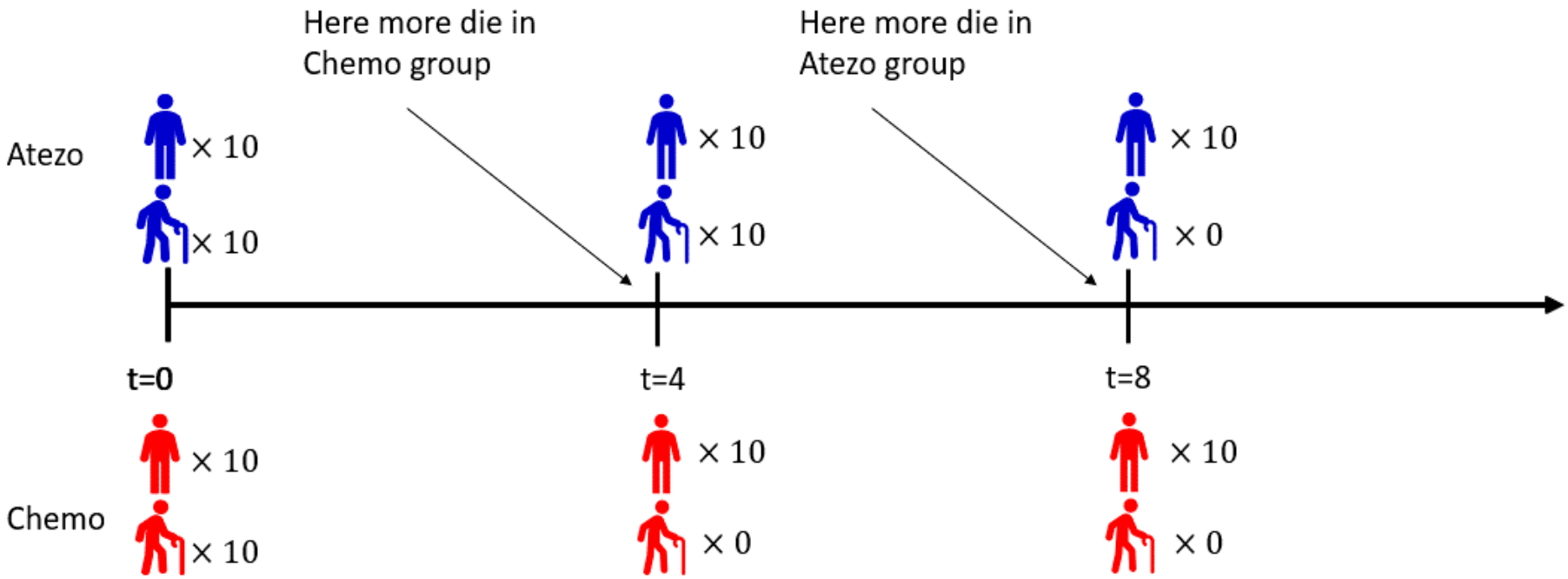
  (Hernán, 2010; Aalen et al., 2015; Martinussen et al., 2020)
- A built-in selection bias in the HR as a parameter.

What is the problem with using HRs?

The treatment groups are heterogeneous!



The selection bias problem of the HR



Even when the treatment is randomized, as time progresses, if the treatment affects survival, the two treatment groups may become less balanced!

Premise

- Although alternatives to the HR were suggested, it remains the most popularly reported measure even in studies targeting causal effects implicitly or explicitly. (Lang and Altman, 2015; Assel et al., 2019)

- Having an alternative HR-like measure that admits a causal interpretation will be useful in practice.

Notations and hazard ratios

- X - a binary treatment. $X \in \{A, C\}$.
- $T^{X=x}$ - the potential survival time had a patient been assigned to treatment group $A = a$, and there was no censoring.
- $\lambda^{X=x}(t)$ - the hazard function of $T^{X=x}$.
- $\lambda(t|X = x)$ - the hazard function of $T|X = x$. Cumulative hazard function $\Lambda(t|X = x) = \int_0^t \lambda(u|X = x)du$

Notations and hazard ratios

- X - a binary treatment. $X \in \{A, C\}$.
- $T^{X=x}$ - the potential survival time had a patient been assigned to treatment group $A = a$, and there was no censoring.
- $\lambda^{X=x}(t)$ - the hazard function of $T^{X=x}$.
- $\lambda(t|X = x)$ - the hazard function of $T|X = x$. Cumulative hazard function $\Lambda(t|X = x) = \int_0^t \lambda(u|X = x)du$
- In an RCT under the standard assumptions $\lambda^{X=x}(t) = \lambda(t|X = x)$, so

$$\frac{\lambda^{X=A}(t)}{\lambda^{X=C}(t)} = \frac{\lambda(t|X = A)}{\lambda(t|X = C)}$$

$$HR^{PO}(t) = HR(t)$$

A causal HR

- In terms of potential outcomes the Standard HR targets

$$HR(t) = \frac{\lim_{dt \to 0} \frac{1}{t} \Pr\left(t \leq T^{X=A} < t + dt \middle| T^{X=A} \geq t\right)}{\lim_{dt \to 0} \frac{1}{t} \Pr\left(t \leq T^{X=C} < t + dt \middle| T^{X=C} \geq t\right)}$$

- A causal HR (Martinussen et al., 2020)

$$HR^C(t) = \frac{\lim_{dt \to 0} \frac{1}{t} \Pr\left(t \leq T^{X=A} < t + dt \middle| T^{X=C} \geq t, T^{X=A} \geq t\right)}{\lim_{dt \to 0} \frac{1}{t} \Pr\left(t \leq T^{X=C} < t + dt \middle| T^{X=C} \geq t, T^{X=A} \geq t\right)}$$

- The $HR^C(t)$ contrasts the hazards under treatment versus no treatment, in the subpopulation of patients who would have survived up to time t regardless of their treatment assignment.

Identifying $HR^C(t)$ using frailty models

- Even in an RCT, $HR^C(t)$ is not identifiable without further assumptions!
- **Intuition:** For each patient, we know for each t whether $T^{X=A} > t$ or $T^{X=C} > t$ (or possibly none if censored).
- So no way to tell for which patients $T^{X=A} \geq t$ **and** $T^{X=C} \geq t$.

Identifying $HR^C(t)$ using frailty models

- Even in an RCT, $HR^C(t)$ is not identifiable without further assumptions!
- **Intuition:** For each patient, we know for each t whether $T^{X=A} > t$ or $T^{X=C} > t$ (or possibly none if censored).
- So no way to tell for which patients $T^{X=A} \geq t$ **and** $T^{X=C} \geq t$.
- **Proposed approach:** a working model such that $T^{X=A}$ and $T^{X=C}$ are correlated through an unmeasured time-fixed frailty variable V.
- Frailty variables are often used to model observed heterogeneity in various survival analysis settings (e.g., clustered data, multi-state models).
- We use frailty variable to model unobserved heterogeneity.

Identification assumptions

- We assume the standard assumptions: SUTVA, randomization or conditional exchangeability, and independent censoring

- We also assume the **working frailty assumptions:**

  1. V is an unmeasured variable from a known parametric family with $E(V) = 1$ and $Var(V) = \theta$
  2. Conditionally on the frailty, the potential survival times are independent:
  $$T^{X=A} \perp\!\!\!\perp T^{X=C} | V$$
  3. The *multiplicity assumption*:
  $$\lambda^{X=x}(t|V = v) = v\psi^{X=x}(t),$$
  where $\lambda^{X=x}(t|V = v)$ is the hazard rate in terms of the potential outcomes given V, and $\psi^{X=x}(t)$ is a function that does not depend on V

Identification of the causal HR: **Randomization**

Proposition

Under **randomization** and the rest of the identification assumptions, the $HR^C(t)$ is identified by

$$\begin{aligned} HR^C(t) &= \frac{\lambda(t|X=A)}{\lambda(t|X=C)} \cdot \varphi\left(\Lambda(t|X=A), \Lambda(t|X=C), \theta\right) \\ &= HR(t) \cdot \varphi\left(\Lambda(t|X=A), \Lambda(t|X=C), \theta\right), \end{aligned}$$

- The function $\varphi(\cdot)$ may take a closed form depending on the specific distribution for V.
- For example, under Gamma frailty:

$$HR^C(t) = \frac{\lambda(t|X=A)}{\lambda(t|X=C)} \cdot \exp\left\{\theta\left[\Lambda(t|X=A) - \Lambda(t|X=C)\right]\right\}$$

Identification of the causal HR: **Observational studies**

- For **observational studies**, it is assumed that for some observed covariates $\boldsymbol{Z}$, we have that $T^{X=x} \perp\!\!\!\perp X | \boldsymbol{Z}$
- Under this assumption, we provide an approximated identification formula for the $HR^C(t)$ as a function of θ
- Identification formula based on Inverse Probability of Treatment Weighting (IPTW). Identification Relies on $\Pr(X = x|\boldsymbol{Z})$
- The approximation should work well when either holds:
  1. The event of interest is rare.
  2. The association between $T^{X=x}$ and the confounders is not too strong.

Estimation of the causal HR

- We cannot estimate $\theta = Var(V)$ from the observed data.
- Instead, we will use θ as a sensitivity parameter. Had we known θ, we could have estimated $HR^C(t)$ from the data.
- From the identification formula, we need to estimate $\lambda(t|X = x)$ and $\Lambda(t|X = x)$.
- We present two estimation methods: Cox-based and kernel-based estimation.
- These methods can be applied for both randomized trials and observational studies.

Estimation of the causal HR

1. **Cox-based estimation:**
   - Assumes that the $HR(t)$ is constant over time and followed by the Cox model
   - Estimates $\lambda(t|X=x)$ and $\Lambda(t|X=x)$ by fitting a standard Cox model

Estimation of the causal HR

1. **Cox-based estimation:**
   - Assumes that the $HR(t)$ is constant over time and followed by the Cox model
   - Estimates $\lambda(t|X=x)$ and $\Lambda(t|X=x)$ by fitting a standard Cox model

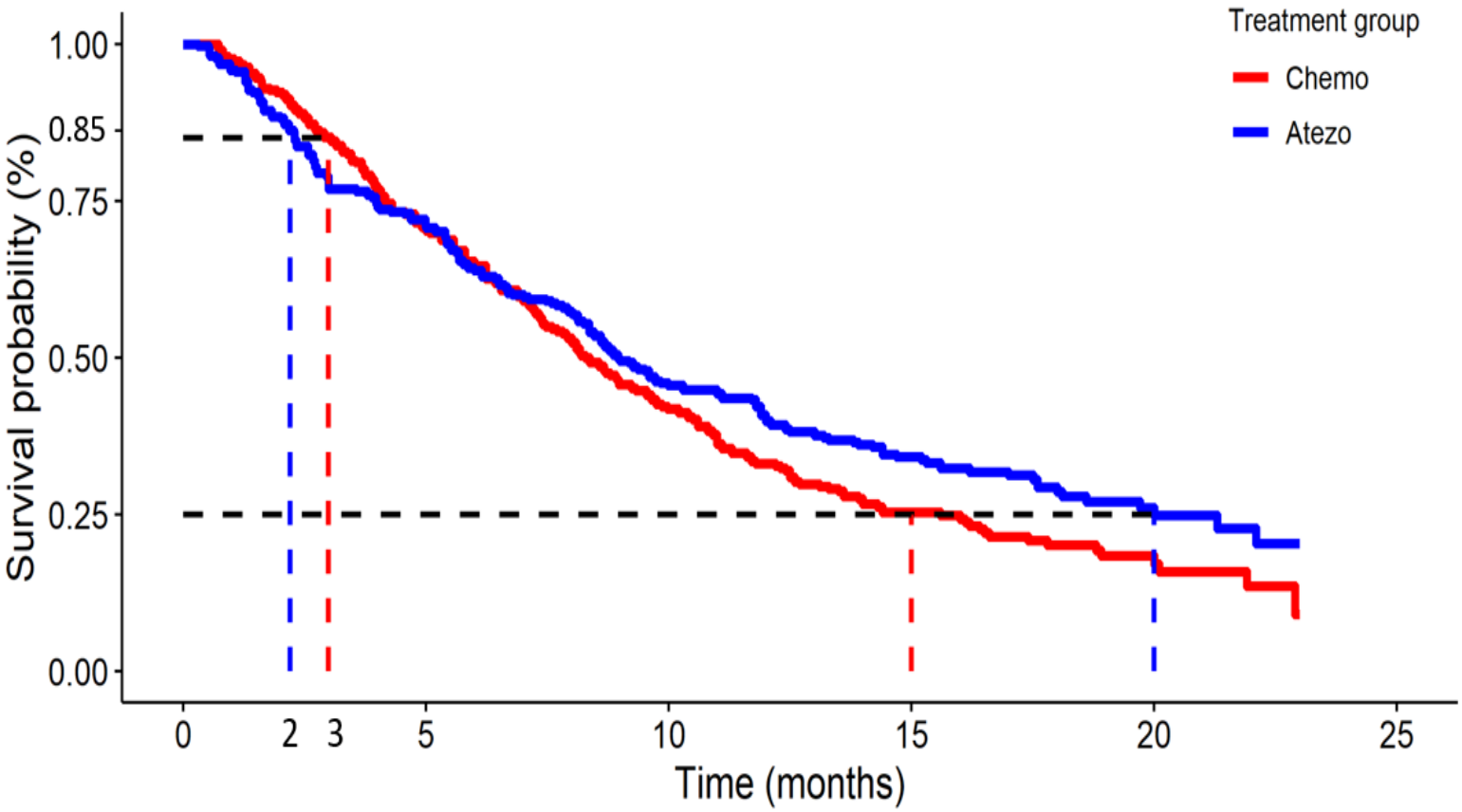
2. **Kernel-based estimation:**
   - No parametric assumptions on $HR(t)$!
   - Estimates $\lambda(t|X=x)$ and $\Lambda(t|X=x)$ at each treatment arm separately using kernel-based estimation under right censoring
   - To improve the estimator performance, we use a time-varying bandwidth, and a boundary kernel.
     (Hess et al., 1999; Müller and Wang, 1990; Muller and Wang, 1994)

Sensitivity analysis steps

1. Choose the parametric distribution for V and a range of possible θ values
   - We present closed-form formulas for the $HR^C(t)$ under different distribution choices: Gamma, Inverse Guassian and positive stable
   - Instead of θ choose the correspondent Kendall's τ correlation values between $T^{X=A}$ and $T^{X=C}$

     (Oakes, 1989)
2. Estimate all of the identification formula parts: $\lambda(t|X=x)$ and $\Lambda(t|X=x)$
3. Plug in the obtained estimators in the identification formula to obtain an estimator for the function $HR^C(t)$ as a function of τ or θ.

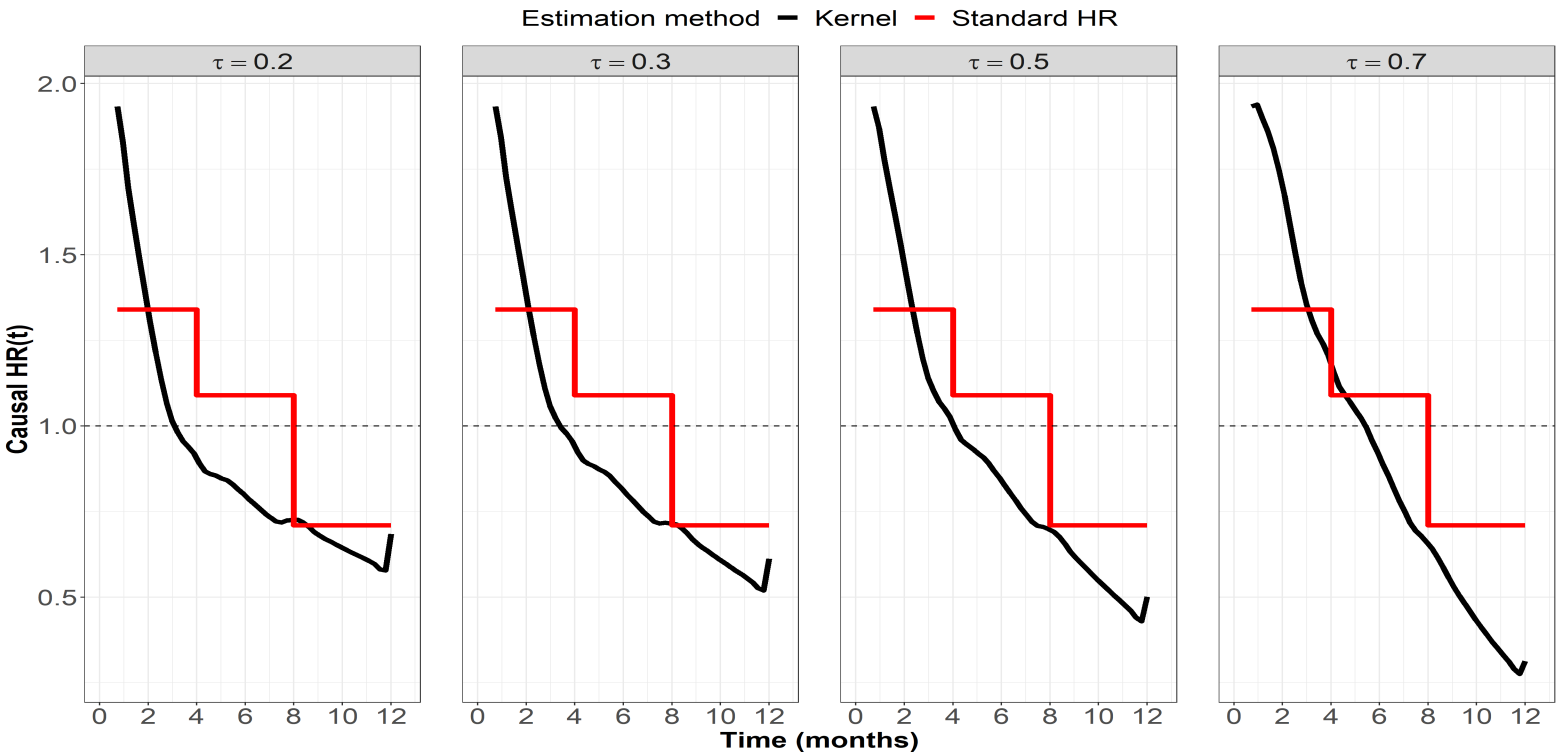
Back to the Motivating example



Survival curves estimated by Kaplan–Meier estimator

Sensitivity analysis results

- Since the proportional hazard assumption is violated, we focus on the kernel-based analysis
- We choose the Gamma distribution and the range of Kendall's $\tau \in (0.1, 0.3, 0.5, 0.7)$. Under the Gamma distribution $\tau = \theta / (\theta + 2)$



Sensitivity analysis results

- Similar qualitative results under inverse-Gaussian frailty and all Kendall's τ.
- $\widehat{HR}^{C}_{\text{kernel}}(t) > 1$ at the start of the study, indicating Atezo treatment is inferior to Chemo in terms of short-term survival.
- After approximately four months, $\widehat{HR}^{C}_{\text{kernel}}(t)$ decreased below one, suggesting Atezo treatment prolongs the survival times of patients would have survived up to four months regardless of their treatment assignment.
- In contrast, Powles et al. (2018) main result was that "Atezolizumab was not associated with significantly longer overall survival than chemotherapy in patients". A piecewise constant HR analysis would have aligned better with our results.

Discussion and take-home messages

What we have talked about today?

- The standard HR is a biased estimator even in randomized trials
- We presented identification formula for the causal HR as a function of an interpretable sensitivity parameter
- We proposed two estimation methods

What we have not talked about today?

- Simulation results: when each estimation method is better? how our methods preform in the presence of confounders? how the estimation is affected by violation of the frailty assumptions?
- Application of our framework in a real observational study example

$$HR^C(t) = \frac{\lambda(t|X=A)}{\lambda(t|X=C)} \cdot \varphi\left(\Lambda(t|X=A), \Lambda(t|X=C), \theta\right)$$

Thank you!

Paper:

R. Axelrod & D. Nevo. "A sensitivity analysis approach for the causal hazard ratio in randomized and observational studies", *Biometrics (2022)*

R package: `CausalHR`

Github: `https://github.com/xlrod1/CausalHR`

Reproducibility:

Github: `https://github.com/xlrod1/Reproducability_R_code`


axelrod1@mail.tau.ac.il


@RachelAxelrod2

References

Aalen, O. O., Cook, R. J., and Røysland, K. (2015). Does cox analysis of a randomized survival study yield a causal treatment effect? *Lifetime data analysis*, 21(4):579–593.

Assel, M., Sjoberg, D., Elders, A., Wang, X., Huo, D., Botchway, A., Delfino, K., Fan, Y., Zhao, Z., Koyama, T., et al. (2019). Guidelines for reporting of statistics for clinical research in urology. *The Journal of urology*, 201(3):595–604.

Hernán, M. A. (2010). The hazards of hazard ratios. *Epidemiology*, 21(1):13.

Hess, K. R., Serachitopol, D. M., and Brown, B. W. (1999). Hazard function estimators: a simulation study. *Statistics in medicine*, 18(22):3075–3088.

Lang, T. A. and Altman, D. G. (2015). Basic statistical reporting for articles published in biomedical journals: the" statistical analyses and methods in the published literature" or the sampl guidelines. *Int J Nurs Stud*, 52(1):5–9.

Martinussen, T., Vansteelandt, S., and Andersen, P. K. (2020). Subtleties in the interpretation of hazard contrasts. *Lifetime Data Analysis*,