Additional online material for:

**Summarizing Primary Results in Clinical Trials with a Time-to-event Endpoint: Complementing Different Measures for a Comprehensive Assessment of Treatment Effect**

Bellavia A, Murphy SA

Circulation

**Page 2. Summary table of interpretation and main properties of presented measures**

**Pages 3-5. Additional notes and reference**

**Table 1.** Summary of common measures of treatment effects in randomized clinical trials with time-to-event endpoints. An expanded version of this table, with additional references and links to software material, is available as an Online Appendix.

| Measure | Metric | Perspective | Interpretation | Main properties | Estimation |
|---|---|---|---|---|---|
| **Hazard Ratio (HR)** | Hazard | Overall (average over follow-up time) | Ratio of the hazard rates (instantaneous rate of the events) over time | - Valid interpretation under proportional hazards<br>- Clinical translation is not straightforward if the baseline hazard is not specified<br>- Does not provide any direct information on the cumulative risk of the event<br>- Several issues with causal interpretation | - Cox regression and its extensions<br>- Parametric and flexible parametric regression |
| **Restricted Mean Survival Time Difference (RMSTΔ)** | Time | Overall (over the entire follow-up, or up to a specific time point) | Difference in average event-free survival between 0 and a given time point *t* | - Intuitive interpretation in terms of time differences<br>- Calculated as an average between 0 and a time point *t*. Caution must be taken if the effect varies over time | - Kaplan-Meier<br>- Generalized Linear Models with pseudo-values |
| **Risk Difference (RD, ARD, ARR) / Risk Ratio (or Relative Risk, RR)** | Risk | Time-specific | Difference or ratio in the risk of the event at a given time point *t* | - Requires specific time point specification<br>- Focuses on what you see at a given time point and not on how you get there<br>- Relevant clinical interpretation<br>- Allows deriving number-needed-to-treat | - Kaplan-Meier<br>- log-binomial models<br>- Generalized Linear Models with pseudo-values |
| **Percentile Difference (PD)** | Time | Risk-specific | Difference in time by which a specific proportion of events (risk) is achieved | - Intuitive interpretation in terms of time<br>- Mostly useful when focusing on common events where rates are high (e.g. worsening HF in patients with prior HF hospitalization)<br>- To avoid data extrapolation beyond follow-up, it requires focusing on event probabilities achieved by both groups | - Kaplan-Meier<br>- Quantile regression for censored data |

**Summary of discussed measures with additional notes and references**

**1. Hazard Ratio (HR):** Measure of overall comparison in the rate metric

Interpretation:

Ratio of the hazard rates (instantaneous rate of the events) over time

Advantages and Limitations:

- Valid interpretation under proportional hazards
- Clinical translation is not straightforward if the baseline hazard is not specified
- Does not provide any direct information on the magnitude of the risk of the event [Sutradhar, R. & Austin, P. C. Relative rates not relative risks: addressing a widespread misinterpretation of hazard ratios. Ann. Epidemiol. 28, 54–57 (2018).]
- Several issues with causal interpretation [Hernán MA. The hazards of hazard ratios. Epidemiology (Cambridge, Mass.). 2010 Jan;21(1):13.]

Estimation

- Cox regression model and its extensions
- Parametric and flexible parametric regression models [Royston P, Lambert PC. Flexible parametric survival analysis using Stata: beyond the Cox model. College Station, TX: Stata press; 2011 Sep 10.]

Additional References:

Uno, H. et al. Alternatives to Hazard Ratios for Comparing the Efficacy or Safety of Therapies in Noninferiority Studies. Ann. Intern. Med. 163, 127–134 (2015).

Stensrud, M. J., Aalen, J. M., Aalen, O. O. & Valberg, M. Limitations of hazard ratios in clinical trials. Eur. Heart J. 40, 1378–1383 (2019).

**2. Restricted Mean Survival Time Difference (RMSTΔ):** Measure of overall comparison in the time metric

Interpretation:

Difference in the area under the survival curve (mean survival time) up to a given time point

Advantages and Limitations:

- Intuitive interpretation in terms of time differences
- Calculated as an average between 0 and a time point t. Caution must be taken if the effect varies over time.

- Extrapolation beyond t would require conditioning on the future [https://discourse.datamethods.org/t/restricted-mean-survival-time-and-comparing-treatments-under-non-proportional-hazards/2686]

Estimation:

- Kaplan-Meier (non-parametric, unadjusted)
- Linear models with pseudo-values [Ambrogi F, Iacobelli S, Andersen PK. Analyzing differences between restricted mean survival time curves using pseudo-values. BMC medical research methodology. 2022 Dec;22(1):1-2.]

Additional References:

Kloecker DE, Davies MJ, Khunti K, Zaccardi F. Uses and limitations of the restricted mean survival time: illustrative examples from cardiovascular outcomes and mortality trials in type 2 diabetes. Annals of internal medicine. 2020 Apr 21;172(8):541-52.

Han, L. Breaking Free from the Hazard Ratio: Embracing the Restricted Mean Survival Time in Clinical Trials. NEJM Evid. 2, EVIDe2300142 (2023).

Perego, C. et al. Utility of Restricted Mean Survival Time Analysis for Heart Failure Clinical Trial Evaluation and Interpretation. JACC Heart Fail. 8, 973–983 (2020).

**3. Risk Difference (RD, ARD, ARR) / Risk Ratio (or Relative Risk, RR)**: Time-specific measures in the risk metric

Interpretation:

Difference or ratio in the risk of the event at a given time point

Advantages and Limitations:

- Requires specific time point specification.
- Focuses on what you see at a given time point and not on how you get there
- Relevant clinical interpretation
- Allows deriving number-needed-to-treat - Allows deriving number-needed-to-treat (NNT=1/ARD), which indicates how many patients must be treated to prevent one event occurrence (or the first one, with a composite endpoint) by the time point of interest.

Estimation:

- Kaplan-Meier (non-parametric, unadjusted)
- Log-binomial models [Donoghoe MW, Marschner IC. logbin: an R package for relative risk regression using the log-binomial model. Journal of Statistical Software. 2018 Sep 4;86:1-22.]

- Linear models with pseudo-values [Gabriel EE, Arkema EV, Sachs MC. Direct modeling of relative and absolute risks in register data: mortality risk in sarcoidosis. Annals of Epidemiology. 2022 Feb 1;66:1-4.]

Additional References:

Gerds TA, Scheike TH, Andersen PK. Absolute risk regression for competing risks: interpretation, link functions, and prediction. Statistics in medicine. 2012 Dec 20;31(29):3921-30.

Bellavia A et al. Estimating and Presenting Hazard Ratios and Absolute Risks from a Cox Model with Complex Non-linear Interactions. Conditionally accepted

Austin PC. Absolute risk reductions, relative risks, relative risk reductions, and numbers needed to treat can be obtained from a logistic regression model. Journal of clinical epidemiology. 2010 Jan 1;63(1):2-6.

**4. Percentile Difference (PD):** Risk-specific measure in the metric of time

Interpretation:

Difference in time by which a specific proportion of events is achieved

Advantages and Limitations:

- Intuitive interpretation in terms of time
- Mostly useful when focusing on common events where rates are high and censoring is low (e.g. overall mortality), where a delay in the event time is as relevant as a reduction of risk in informing clinical guidance.
- To avoid data extrapolation beyond follow-up, it requires focusing on event probabilities achieved by both groups

Estimation:

- Kaplan-Meier (non-parametric, unadjusted)
- Quantile regression for censored data [Orsini N, Wolk A, Bottai M. Evaluating percentiles of survival. Epidemiology. 2012 Sep 1;23(5):770-1.]

Additional References:

Bellavia A, Bottai M, Orsini N. Evaluating additive interaction using survival percentiles. Epidemiology (Cambridge, Mass.). 2016 May;27(3):360.