

HAZARD RATES AND RESTRICTED MEAN SURVIVAL TIME*

Szilárd Nemes¹

¹Institute of Clinical Sciences, The Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden, nemessz@gmail.com

**Indonesian Journal of Statistics and Its Applications (eISSN:2599-0802)
Vol 3 No 3 (2019), 310 - 319**

Copyright © 2019 Szilárd Nemes. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Restricted Mean Survival Time (RMST) is well-established, but underutilized measure that can be interpreted as the average event-free survival time up to a pre-specified time point. In the last decade RMST received substantial attention and was advocated as an alternative for the Hazard Rate when the proportionality assumption is not met. Currently studies with time-to-event outcomes routinely report survival curves and hazard rates. Research planning assumes extraction of comparative effect measures and variances that facilitates sample size calculations. Here we assessed the possibility of extracting clinically meaningful effect size estimates for RMST based research plans from studies that report survival curves and hazard rates. This assessment was based on simulations using Exponential and Weibull distributions. The simulations suggest that under certain conditions meaningful RMST effect size estimates can be extrapolated from published hazard rates. However, in cases when the proportionality assumption is in doubt (i.e. when RMST have most utility) extraction of meaningful estimates is not feasible.

Keywords: effect size, research design, exponential distribution.

1. Introduction

The Hazard Ratio (HR) is one of the most common statistics in clinical studies. Estimation and interpretation of hazard functions and group wise comparisons assumes constant rates over the follow-up time. This assumption can be relaxed; however, this results in increased complexities in interpretation. One notable exception is the Restricted Mean Survival Time (RMST), that quantifies the expected

* Received Jul 2019; Accepted Aug 2019; Published online on Oct 2019

life expectancy after the intervention up to a time point of interest (Kim et al., 2017; Royston & Parmar, 2013; Uno et al., 2014). When the proportional hazards assumptions of classical survival analysis is violated this approach offers an straightforward and easy to interpret and clinically meaningful estimates (Royston & Parmar, 2011). In clinical studies where there is suspicion of substantial departure from the proportional hazards assumption testing RMST instead of log-rank test is a recommended alternative (Huang & Kuan, 2018). Generally, there is agreement between RMST and HR about the statistical significance of the treatment effect, but regarding treatment-effect measures RMST yielded more conservative estimates than HRs (Trinquart et al., 2016).

One important task for an applied statistician is study design, i.e. making sure that the chose study design and collected data answers in proper fashion the postulated research question. A central aspect of this task is sample size calculations, assuring the collection of enough data for statistical inference. Powering clinical studies with time-to-event outcomes as a long history and a rich literature. However, for the relatively old but newly rediscovered RMST we still lack routines for power calculations. This not so much form a statistical point of view, once we have suitable effect size and variability estimates powering for RMST analysis is a routine task. (Royston & Parmar, 2013). Extraction of effect size and variability estimates, on the other hand is far from being a trivial task. Clinical studies rarely report RMST and routinely resorts presentation of survival curves and/or hazard rates.

In this note we examine the possibility of obtaining feasible effect size estimates for testing RMST from published data. We evaluate if it is possible to obtain robust estimates with simple analytical assumptions.

2. Definitions and Notations

2.1 Restricted Mean Survival Time

The survival function is estimated with the Kaplan-Meier estimator

$$\hat{s}(t) = \prod_{t_i \leq t} \left(1 - \frac{d_i}{Y_i}\right), \quad (1)$$

where Y_i is the number of patients at risk at time t_i and d_i is the number of patients who fail in the interval $t, t + dt$.

The mean survival time is given by

$$E[T] = \int_0^{\infty} S(t) dt \quad (2)$$

however, the mean survival time can be estimated only if don not have any censoring and the life length of every patient is known.

An alternative the mean survival time is the τ -restricted mean survival time. Here, we do not aim to follow-up all patients until their death but to a pre-define length of time.

RMST is estimated as

$$\hat{\mu}(\tau) = \int_0^{\tau} \hat{S}(t) dt \quad (3)$$

where we replace the survival function by its estimate the Kaplan-Meier curve.

For a two-group comparison (i.e. treatment vs placebo) with $\hat{\mu}_0(\tau)$ and $\hat{\mu}_1(\tau)$ the difference is given by

$$D(\tau) = \mu_0(\tau) - \mu_1(\tau) \quad (4)$$

and estimated as

$$\hat{D}(\tau) = \int_0^\tau \{\hat{S}_0(t) - \hat{S}_1(t)\} dt. \quad (5)$$

Under the null-hypothesis of $\mu_1(\tau) = \mu_0(\tau)$ this difference is normally distributed. Alternatively, one can use the ratio estimator

$$R(\tau) = \frac{\mu_0(\tau)}{\mu_1(\tau)} \quad (6)$$

estimated as

$$\hat{R}(\tau) = \frac{\int_0^\tau \hat{S}_0(t) dt}{\int_0^\tau \hat{S}_1(t) dt} \quad (7)$$

2.2 Proportional Hazards Model

The hazard gives the probability of experiencing the event of interest in an infinitesimal time interval after time t (in a very short interval, i.e. in the next 'moment'). The hazard function is given by

$$h(t)dt = \Pr(t < T < t + d | \min(T, C) \geq t). \quad (8)$$

In most cases we aim to compare 2 (or more) time-dependent functions, the survival function $S(t)$ or the hazard function $h(t)$.

Usually the null hypothesis will be $H_0: S_1(t) = S_0(t)$, where $S_0(t)$ represents the baseline survival (i.e. survival in the control group) and $S_1(t)$ the survival of the treated group. As the survival, and even more the hazard function, can take various forms, the formulation of the alternative hypothesis as $H_A: S_1(t) \neq S_0(t)$ is not practicable. This issue has been debated for a long time within the statistical literature. The solution traces back to the seminal work by Lehmann (Lehmann, 1953) on rank tests, the 'Lehmann alternatives'. The Lehmann alternative assumes that the distribution of the response ($S_1(t)$ or $h_1(t)$ in the present setting) is a specified function of the baseline ($S_0(t)$ or $h_0(t)$).

In the survival analysis setting this leads to the following alternative hypothesis $H_A: S_1(t) = \{S_0(t)\}^\varphi$. Using the relationship between the survival and hazard function this is equivalent to $h_1(t) = \varphi h_0(t)$, i.e. the observed hazard is proportional to the baseline hazard by a factor of φ , yielding the proportional hazards assumptions. This formulation can be extended to include covariate information by assuming that $\varphi = e^{\beta z}$. Here e is the Euler number, equaling approximately 2.72; β is regression coefficients and z is the covariate values. In clinical studies e^β , the hazard rate is often the main statistics of interest.

2.3 The exponential survival models

The exponential distribution is the simplest survival distribution with a constant hazard

rate, $h(t) = h$. The survival function is $S(t) = e^{-ht}$. For the exponential distribution the RMST can be obtained analytically based on equation 2 as

$$\mu(t) = \int_0^\tau \hat{S}(t)dt = \int_0^\tau e^{-ht}dt = \frac{1 - e^{-h\tau}}{h} \tag{9}$$

Finding exact transformation between the parameters of interest, the HR, $D(\tau)$ and $R(\tau)$ is straightforward in the exponential case. Both $D(\tau)$ and $R(\tau)$ are functions of the hazard rate, however the relationship is non-linear. First, we reparametrize $D(\tau)$ and $R(\tau)$ with $h_1 = h_0\varphi$ and equation 4 can be expressed as

$$D(\tau) = \frac{1 - e^{-h_0\varphi\tau}}{h_0\varphi} - \frac{1 - e^{-h_0\tau}}{h_0} \tag{10}$$

The ratio estimator (equation 7) is given by

$$R(\tau) = \frac{(1 - e^{-h_0\varphi\tau})}{\varphi(1 - e^{-h_0\tau})} \tag{11}$$

3. Simulation Study

3.1 Exponential distribution

In this simplistic simulation we aimed to confirm numerically the appropriateness of equations 10 and 11. We simulated 1000 exponentially distributed time to events with in a hypothetical 2 arm study. For the placebo arm we assumed a hazard of 1/365 and a hazard rate between the who arm of 0.5. Additionally, we assumed an exponential censoring with hazard of 1/730. The time points of interest τ was set at 100, 200, 300 and 500 days. We re-iterate the simulation 1000 times. For each simulation estimated $\hat{R}(\tau)$ and $\hat{D}(\tau)$ with equations 7 and 8. This required estimation of h_0, h_1 and φ .

We have run two simulations according to two scenarios. The first scenario assumed that we have access to φ and h_0 needs to be extracted from survival curves as $h_0 = -\log(\hat{S}(t))t^{-1}$. Figure 1 presents the results of the simulation. The second scenario assumed that both h_0 and h_1 needed to be extracted from the survival curves.

For both scenarios $\hat{D}(\tau)$ and $\hat{R}(\tau)$ were unbiased and almost identical (Figure 1).

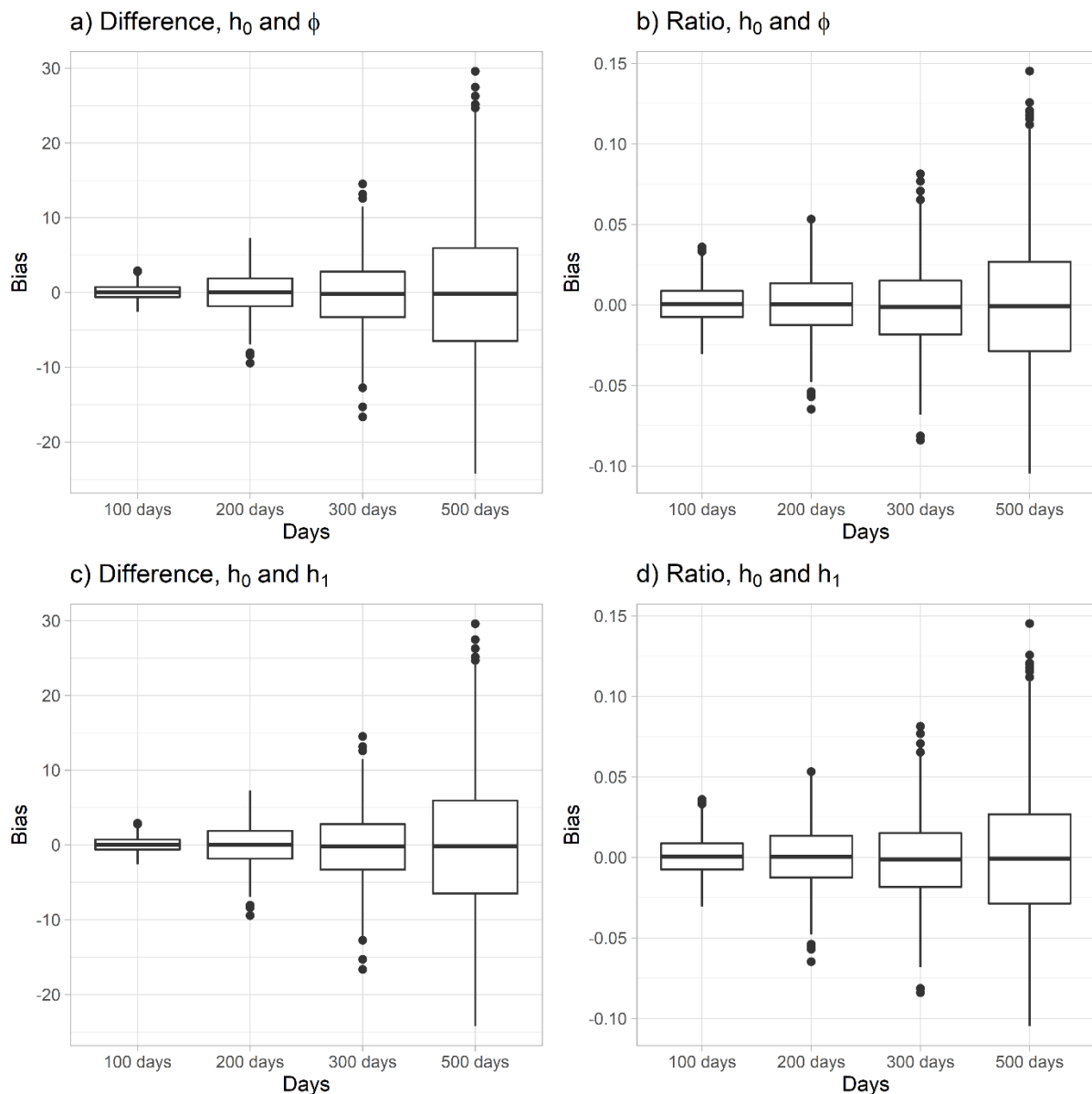


Figure 1: Bias and variability for (a) the difference of two RMST estimates and (b) the ratio of two RMST estimates when the hazard in the placebo group is estimated and the hazard rate is known and (c) the difference of two RMST estimates and (d) the ratio of two RMST estimates when the hazard in the placebo and active arm is estimated.

3.2 Weibull distribution

In this scenario we assessed the feasibility of the outlined algorithm when the assumptions are violated. We simulated survival data with the Weibull distribution. The scale parameter was set to 365, while the shape first was set to 0.5 so the failure rate decreases over time; second to 2 so the failure rate increases with time. As previously

we assumed exponential censoring with hazard of 1/730. The simulation was repeated 1000 times. For each iteration we simulated a data set according to the above described parameters. Fitted Kaplan-Meier curves to the data, extracted the survival probability at 100, 200, 300 and 500 days. Estimation proceeded with the assumption of exponential distribution, which in this case is a clear violation of assumptions. We estimated RMST as $\int_0^\tau S(t) dt = h^{-1}(1 - e^{-h\tau})$. This RMST estimate was then compared to the true RMST estimated with numerical integration of the Weibull survival function with the above listed parameter values.

Figure 2 presents the results of the simulation. As expected a τ dependent negative bias was observed when the failure rate increased with time. Opposite to this, when the failure rate decreases with time the RMST estimated assuming exponential distribution overestimated the true RMST.

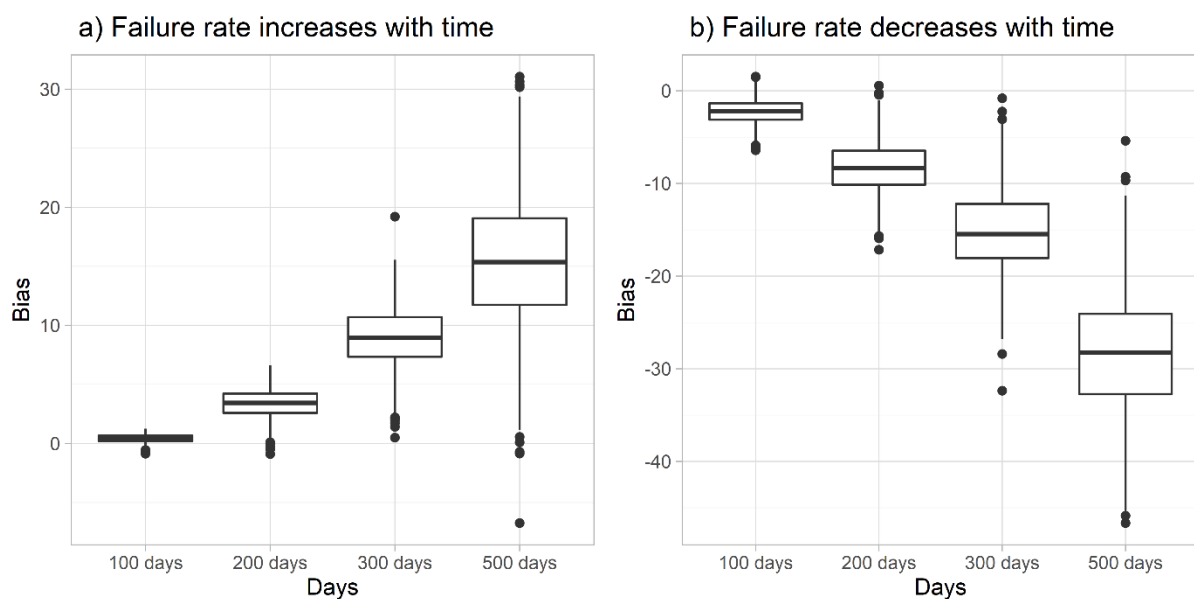


Figure 2: Bias and variability for RMST with wrongly assumed exponential distribution (a) when the true distribution is Weibull with increasing failure rate and (b) when the true distribution is Weibull with decreasing failure rate.

4. Application on real life data

Here we apply the above outlined routines to three studies that were summarized by (Uno et al., 2014). We chose these studies as beside the usual survival curves and hazard rates there were available estimates for both $R(\tau)$ and $D(\tau)$.

The first study by Rajkumar and collaborators (Rajkumar et al., 2010), an Eastern Cooperative Oncology Group study, compared patients treated with low- and high dose of dexamethasone for newly diagnosed multiple myeloma. The authors reported a hazard rate of 0.87 and from the survival curves we could extract that $\hat{S}(40 m) = 0.47$. At 40 months $R(\tau) = 1.06$ and $D(\tau) = 2.2$. Here, it is visible from the survival curves (Figure 3 left down in Rajkumar et al., 2010) that the proportionality assumption is violated. Based on equations 7 and 8 we obtained $\hat{R}(\tau) = 1.04$ and $\hat{D}(\tau) = 1.2$, a clear deviation from the reported values. Likely if estimation would had been restricted to 30

months we would have obtained better estimates.

The second compared single-agent pemetrexed (P) versus the combination of carboplatin and pemetrexed (CP) in first-line therapy for patients with advanced non-small-cell lung cancer (Zukin et al., 2013). The authors reported a hazard rate of 0.95 and from the survival curves we could extract that $\hat{S}(24 \text{ m}) = 0.01$. At 24 months $R(\tau) = 1.49$ and $D(\tau) = 3.9$. Based on equations 7 and 8 we obtained $\hat{R}(\tau) = 1.52$ and $\hat{D}(\tau) = 4.02$.

The third study considered investigate the safety and efficacy of adding bevacizumab to fluorouracil, leucovorin, and oxaliplatin (FOLFOX6) for the adjuvant treatment of patients with stage 2-3 colon cancer (Allegra et al., 2012). The authors reported a hazard rate of 0.95 and from the survival curves we could extract that $\hat{S}(60 \text{ m}) = 0.83$. At 60 months $R(\tau) = 1$ and $D(\tau) = 0.3$. Based on equations 7 and 8 we obtained $\hat{R}(\tau) = 1$ and $\hat{D}(\tau) = 0.24$.

5. Discussion

When planning a clinical trial one question that needs to be considered how well the possible competitors fare and compared with them how much effect we would like to see. With the effect size assumed variance set one can proceed with sample size calculations. In this note we have focused on obtaining effect size estimates from published studies in a hypothetical situation when statistical powering is done for RMST. RMST is a well-established, yet underutilized measure that can be interpreted as the average event-free survival time up to a pre-specified, clinically important time point. (Kim et al., 2017; Uno et al., 2015) and may improve the accuracy of the prognostic estimates that influence clinical decisions and information given to patients (Couchoud et al., 2017). Studies with time-to-event endpoints generally resort to summaries the data with survival curves and hazard rates. Statistical inference is generally based hazard rates. If the proportional hazards assumption is met then the log-rank test is generally a more powerful test, however RMST based approaches can be of value in certain situations (Huang & Kuan, 2018). In roughly, 19 % of oncology clinical trials using time-to-event endpoint there is evidence of non-proportional hazards, and in this case RMST offer a more parsimonious estimates while keeping the conclusions (Rulli et al., 2018). Additionally, RMST is a viable alternative to the HR if causality is of interest (Stensrud et al., 2018). Just as HR (or any other regression estimands) RMST can be adjusted either by using pseudo-observations and generalized estimating equations (Andersen et al., 2004) or by integrating an adjusted Kaplan-Meier estimator with inverse probability weighting (Conner et al., 2019).

In this paper we have evaluated if it is possible to extract meaningful effect size estimates that would aid statistical study design for time-to-event outcomes by the means of RMST. We had primarily focused on situation where the statistical has access only to published data, i.e. survival curves and hazard rates. With simplistic assumptions we could show that it is possible to extract meaningful estimates under certain circumstances.

Our starting point was that the studied survival times are exponentially distributed. This is a very common assumption and clinical trials. In this case this assumption was necessary as RMST was calculated based on the survival probability at the restriction point. Likely if we could have used other more flexible distributions the precision of the

estimates would have increased. However, such estimation wasn't possible as parameters of more flexible distributions such as Weibull or Gamma are nonidentifiable in this setting. Estimation possible only if the parameters of the assumed distribution are known at least in one arm of the study (Weir & Trinquart, 2018).

If the distribution of the survival time departs from the exponential distribution (i.e. accelerated failures) and/or there is genuine deviation from the proportional hazards assumptions then this simplistic approach fails. In these cases, extraction of necessary information from published results must resort to a rather diminished pool of research articles that present RMST measures. Alternatively, reconstruction of the individual level data from published Kaplan-Meier curves (Guyot et al., 2012) and estimation of RMST is needed. Though this requires technical expertise that applied researchers might not possess and for a research planning point of view this is a drawback of the method. An additional difficulty that researchers might face is departure from the normal distribution for difference between two arms in restricted mean survival time (Lawrence et al., 2019), that complicates statistical inference.

RMST are useful tools for every statistician, however as statistical science stands today that are more valuable as ad-hoc analyses of collected data based on the 4-step approach (Royston & Parmar, 2011). The applicability of RMST in research planning is further hindered by difficulties in obtaining reliable variance estimates under different grades of censoring. Thus, further research/evaluations are needed to establish routines for research planning.

References

- Allegra, C. J., Yothers, G., O'Connell, M. J., Sharif, S., Petrelli, N. J., Lopa, S. H., & Wolmark, N. (2012). Bevacizumab in Stage II-III Colon Cancer: 5-Year Update of the National Surgical Adjuvant Breast and Bowel Project C-08 Trial. *Journal of Clinical Oncology*, **31**(3): 359-364. doi:10.1200/JCO.2012.44.4711
- Andersen, P. K., Hansen, M. G., & Klein, J. P. (2004). Regression Analysis of Restricted Mean Survival Time Based on Pseudo-Observations. *Lifetime Data Analysis*, **10**(4): 335-350. doi:10.1007/s10985-004-4771-0
- Conner, S. C., Sullivan, L. M., Benjamin, E. J., LaValley, M. P., Galea, S., & Trinquart, L. (2019). Adjusted restricted mean survival times in observational studies. *Statistics in Medicine*, **38**(20): 3832-3860. doi:10.1002/sim.8206
- Couchoud, C., Dantony, E., Elsensohn, M.-H., Rabilloud, M., Ecochard, R., Villar, E., . . . Moranne, O. (2017). Restricted mean survival time over 15 years for patients starting renal replacement therapy. *Nephrology Dialysis Transplantation*, **32**(suppl_2): ii60-ii67. doi:10.1093/ndt/gfw386
- Guyot, P., Ades, A. E., Ouwens, M. J. N. M., & Welton, N. J. (2012). Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Medical Research Methodology*, **12**(1): 9. doi:10.1186/1471-2288-12-9

- Huang, B., & Kuan, P.-F. (2018). Comparison of the restricted mean survival time with the hazard ratio in superiority trials with a time-to-event end point. *Pharmaceutical Statistics*, **17**(3): 202-213. doi:10.1002/pst.1846
- Kim, D. H., Uno, H., & Wei, L. J. (2017). Restricted Mean Survival Time as a Measure to Interpret Clinical Trial Results. *JAMA Cardiol*, **2**(11): 1179-1180. doi:10.1001/jamacardio.2017.2922
- Lawrence, J., Qiu, J., Bai, S., & Hung, H. M. J. (2019). Difference in Restricted Mean Survival Time: Small Sample Distribution and Asymptotic Relative Efficiency. *Statistics in Biopharmaceutical Research*, **11**(1): 61-66. doi:10.1080/19466315.2018.1527249
- Lehmann, E. L. (1953). The Power of Rank Tests. *Ann. Math. Statist.*, **24**(1): 23-43. doi:10.1214/aoms/1177729080
- Rajkumar, S. V., Jacobus, S., Callander, N. S., Fonseca, R., Vesole, D. H., Williams, M. E., . . . Greipp, P. R. (2010). Lenalidomide plus high-dose dexamethasone versus lenalidomide plus low-dose dexamethasone as initial therapy for newly diagnosed multiple myeloma: an open-label randomised controlled trial. *The Lancet Oncology*, **11**(1): 29-37. doi:[https://doi.org/10.1016/S1470-2045\(09\)70284-0](https://doi.org/10.1016/S1470-2045(09)70284-0)
- Royston, P., & Parmar, M. K. (2013). Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC Medical Research Methodology*, **13**(1): 152. doi:10.1186/1471-2288-13-152
- Royston, P., & Parmar, M. K. B. (2011). The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt. *Statistics in Medicine*, **30**(19): 2409-2421. doi:10.1002/sim.4274
- Rulli, E., Ghilotti, F., Biagioli, E., Porcu, L., Marabese, M., D'Incalci, M., . . . Torri, V. (2018). Assessment of proportional hazard assumption in aggregate data: a systematic review on statistical methodology in clinical trials using time-to-event endpoint. *British Journal of Cancer*, **119**(12): 1456-1463. doi:10.1038/s41416-018-0302-8
- Stensrud, M. J., Aalen, J. M., Aalen, O. O., & Valberg, M. (2018). Limitations of hazard ratios in clinical trials. *European Heart Journal*, **40**(17): 1378-1383. doi:10.1093/eurheartj/ehy770
- Trinquart, L., Jacot, J., Conner, S. C., & Porcher, R. (2016). Comparison of Treatment Effects Measured by the Hazard Ratio and by the Ratio of Restricted Mean Survival Times in Oncology Randomized Controlled Trials. *Journal of Clinical Oncology*, **34**(15): 1813-1819. doi:10.1200/jco.2015.64.2488
- Uno, H., Claggett, B., Tian, L., Inoue, E., Gallo, P., Miyata, T., . . . Wei, L. J. (2014). Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. *Journal of Clinical Oncology*, **32**(22): 2380-2385. doi:10.1200/jco.2014.55.2208

- Uno, H., Wittes, J., Fu, H., Solomon, S. D., Claggett, B., Tian, L., . . . Wei, L.-J. (2015). Alternatives to Hazard Ratios for Comparing the Efficacy or Safety of Therapies in Noninferiority Studies Alternatives to Hazard Ratios. *Annals of Internal Medicine*, **163**(2), 127-134. doi:10.7326/M14-1741 %J Annals of Internal Medicine
- Weir, I. R., & Trinquart, L. (2018). Design of non-inferiority randomized trials using the difference in restricted mean survival times. **15**(5): 499-508. doi:10.1177/1740774518792259
- Zukin, M., Barrios, C. H., Rodrigues Pereira, J., De Albuquerque Ribeiro, R., de Mendonça Beato, C. A., do Nascimento, Y. N., . . . Lilenbaum, R. C. (2013). Randomized Phase III Trial of Single-Agent Pemetrexed Versus Carboplatin and Pemetrexed in Patients With Advanced Non–Small-Cell Lung Cancer and Eastern Cooperative Oncology Group Performance Status of 2. *Journal of Clinical Oncology*, **31**(23): 2849-2853. doi:10.1200/JCO.2012.48.1911