HHS Public Access

Author manuscript

*JAMA Cardiol*. Author manuscript; available in PMC 2019 February 02.

Published in final edited form as:

*JAMA Cardiol*. 2017 November 01; 2(11): 1179–1180. doi:10.1001/jamacardio.2017.2922.

# Restricted Mean Survival Time as a Measure to Interpret Clinical Trial Results

**Dae Hyun Kim, MD, MPH, ScD**[1,2], **Hajime Uno, PhD**[3], and **Lee-Jen Wei, PhD**[4]

[1]Department of Medicine, Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women's Hospital, Boston, MA

[2]Department of Medicine,Division of Gerontology, Beth Israel Deaconess Medical Center, Boston, MA

[3]Department of Medical Oncology, Division of Population Sciences, Dana Farber Cancer Institute, Boston, MA

[4]Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, MA

Clinical trials of cardiovascular therapeutics have been criticized for under-representation of older adults with frailty and multiple chronic conditions. Major strides have been made in recent years to increase representation of this population, as exemplified by clinical trials of transcatheter aortic valve replacement (TAVR).[1,2] While more high-quality evidence is being generated to guide treatment in older adults, the challenge remains to interpret treatment effect in a manner that is informative and intuitive to clinical communities. This is particularly important for frail older adults with limited life expectancy in whom the expected benefit may be too small or remote to justify the risk of treatment-related adverse events. Most cardiovascular trials conventionally report treatment effect in terms of relative risk reduction (e.g., hazard ratio [HR]) and absolute risk reduction (or equivalently, number needed to treat [NNT]) for deaths or cardiovascular events at a specific time point. Recently, restricted mean survival time (RMST) has been proposed as an alternative measure of treatment effect that offers some advantages in design, analysis, and interpretation over the conventional measures.[3–5] In this viewpoint, we explain how different measures of treatment effect are interpreted for evidence-based communication and their caveats using the 5-year follow-up data from the Placement of Aortic Transcatheter Valves (PARTNER) trials as an example.[1,2]

The HR compares the hazard ("instantaneous rate") of a clinical event in treatment group with that in control group. It is usually estimated from a Cox proportional hazards model under the assumption that the hazard in treatment group remains proportional to the hazard in control group at any point during the study period (i.e., proportional hazards assumption). In PARTNER B trial that compared TAVR with medical treatment (Table),[1] a HR of 0.50

**Corresponding author**: Dae Hyun Kim, MD, MPH, ScD, Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital, 1620 Tremont Street Suite 3030, Boston, MA 02120, Phone: 1-617-278-0930; Fax: 1-617-232-8602, dkim12@bwh.harvard.edu.

**Conflict of Interest:** D. H. Kim is consultant to Alosa Health, a non-profit educational organization with no relationship to any drug or device manufacturers. H. Uno and L. J. Wei declare no conflicts of interest.

means that the hazard of death in TAVR patients is always 50% lower than that in patients receiving medical treatment throughout the 5 years. In PARTNER A trial that compared TAVR and surgical aortic valve replacement (SAVR), however, the hazard in the first 30 days was greater in SAVR patients and then became similar to that in TAVR patients. In this situation, the ratio of the two hazards is not constant and interpretation of the 5-year HR is unclear. Moreover, the meaning of HR is not intuitive to clinical communities since HR cannot be interpreted as risk ratio at each time point. When the event rate is very low or the follow-up time is very short, HR is numerically similar to risk ratio. As a ratio measure by itself, it is difficult to appreciate the magnitude of treatment effect based on HR without knowing the hazard in control group. Presenting HR alone tends to influence patients to accept treatment compared with presenting absolute risk difference.[6]

The median survival time difference can enhance intuitive interpretation of treatment effect. In PARTNER B trial, the median survival time for TAVR patients was 19 months longer than that for medically treated patients (Table). While this interpretation is intuitive, the median survival time is insensitive to outliers, such as patients who die shortly after TAVR or those who survive a long time. In addition, the estimation procedure for the median survival time difference can be unstable, as evidenced by wide 95% confidence intervals (CIs). The 95% CI of the median survival time difference from PARTNER A trial indicates that the median survival time of TAVR patients could be 4-month shorter to 13-month longer than that of SAVR patients, which does not exclude a clinically meaningful difference. Moreover, when the cumulative incidence of the event is less than 50%, the median survival time cannot be estimated from the study.

Absolute risk difference is another alternative measure that can be used to improve interpretability of treatment effect.[6] Because it is estimated based on cumulative numbers of events at a given time, it is agnostic to whether the events have occurred earlier or later during the study period. This information may be important in some older patients in whom delaying events by a certain amount of time can be meaningful. The NNT is derived from absolute risk difference (Table). The estimate from the PARTNER B trial suggests that 5 patients need to be treated with TAVR rather than medical treatment to prevent a death in 5 years. Although it may sound intuitive to clinicians, NNT may be difficult for patients to appreciate.[6] When the absolute risk difference is not statistically significant as observed in the PARTNER A trial, NNT cannot be interpreted easily and, therefore, is often omitted from the published report.

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.[3–5] It is equivalent to the area under the Kaplan-Meier curve from the beginning of the study through that time point. The RMST difference means gain or loss in the event-free survival time due to treatment vs. control during this period. The PARTNER B trial showed that TAVR patients lived, on average, 13 additional months in 5 years than medically treated patients (Table). In PARTNER A trial, no statistically significant difference was found in the 5-year RMST between TAVR and SAVR. The 95% CI suggests that TAVR patients may live 3 fewer months to 5 more months in 5 years than SAVR patients. As such, summarizing the treatment effect and its 95% CI using the average survival time can be more easily

understood by clinical communities. In addition, RMST difference is valid and interpretable whether or not the proportional hazards assumption is violated. Since RMST difference is based on all patients' exposure times as opposed to HR which depends on the number of events (regardless of exposure times), it is as powerful as log-rank test or HR-based test when the proportional hazards assumption is valid, and can be more powerful when the assumption is violated.[4] This advantage may allow clinical trialists to design non-inferiority studies of cardiovascular treatments with a smaller number of patients.

Effective communication of treatment effect from clinical trials is a prerequisite for shared treatment decision-making. RMST, a robust measure that represents the average event-free survival time in a pre-specified period, may provide useful information on treatment effect that complements conventional measures of relative and absolute risk reductions. Further research is needed to evaluate whether RMST improves patients' understanding and influence treatment decision.

## ACKNOWLEDGMENT

**Funding/Support:** The work was partially supported by R01 HL089778 (NIH/NHLBI), R00 HS022193 (NIH/AHRQ), and R21 AG049385 (NIH/NIA). D. H. Kim is supported by the Paul B. Beeson Clinical Scientist Development Award in Aging (K08AG051187) from the National Institute on Aging, American Federation for Aging Research, The John A. Hartford Foundation, and The Atlantic Philanthropies.

**Role of Funder/Sponsor Statement:** The funder did not have any role in design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

## REFERENCES

1. Kapadia SR, Leon MB, Makkar RR, et al. 5-year outcomes of transcatheter aortic valve replacement compared with standard treatment for patients with inoperable aortic stenosis (PARTNER 1): a randomised controlled trial. Lancet. 2015;385(9986):2485–2491. [PubMed: 25788231]
2. Mack MJ, Leon MB, Smith CR, et al. 5-year outcomes of transcatheter aortic valve replacement or surgical aortic valve replacement for high surgical risk patients with aortic stenosis (PARTNER 1): a randomised controlled trial. Lancet. 2015;385(9986):2477–2484. [PubMed: 25788234]
3. Royston P, Parmar MK. 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 Med Res Methodol. 2013;13:152. [PubMed: 24314264]
4. Uno H, Wittes J, Fu H, et al. Alternatives to Hazard Ratios for Comparing the Efficacy or Safety of Therapies in Noninferiority Studies. Ann Intern Med. 2015;163(2):127–134. [PubMed: 26054047]
5. Uno H, Claggett B, Tian L, et al. Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. J Clin Oncol. 2014;32(22):2380–2385. [PubMed: 24982461]
6. Zipkin DA, Umscheid CA, Keating NL, et al. Evidence-based risk communication: a systematic review. Ann Intern Med. 2014;161(4):270–280. [PubMed: 25133362]
7. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. BMC Med Res Methodol. 2012;12:9. [PubMed: 22297116]

**Table.**

**Measures to Summarize Treatment Effect Using Examples of the Placement of Aortic Transcatheter Valves (PARTNER) Trials***

| Effect Measure | PARTNER A (TAVR vs. SAVR) Estimate (95% CI) | PARTNER B (TAVR vs. MT) Estimate (95% CI) | Caveats |
|---|---|---|---|
| Hazard ratio | 1.01 (0.84 to 1.23) | 0.50 (0.39 to 0.65) | • The hazard is non-intuitive to interpret.<br>• Treatment effect is assumed constant over time; if this is not the case, the interpretation is unclear.<br>• Presenting hazard ratio without absolute event rates can be misleading.<br>• The statistical power depends on the number of events during the follow-up time. |
| Median survival time difference | 4.3 months (−3.9 to 12.5) | 19.3 months (9.4 to 29.2) | • The median survival time is intuitive to interpret.<br>• The median survival time is insensitive to early deaths or long-term survivors.<br>• The estimation procedure is less efficient (i.e., wide 95% Cis). It may be inconsistent with hazard ratio estimates.<br>• When the event rate is low, the median survival time is not observed. |
| Risk difference at time *t (5 years)* | −5.0 % (−12.6 to 2.6) | −22.9 % (−31.1 to −14.6) | • Risk difference is intuitive to interpret.<br>• Risk difference does not reflect the cumulative incidence profiles before time *t*. |
| NNT at time *t (5 years)* | 20.0 (not calculated) | 4.4 (3.2 to 6.8) | • NNT does not reflect the cumulative incidence profiles before time *t*.<br>• It may be intuitive to clinicians, but may not improve patients' understanding.<br>• When the risk difference is not statistically significant, interpretation of NNT is unclear. |
| RMST difference at time *t (5 years)* | 1.0 month (−2.7 to 4.6) | 12.6 months (8.2 to 17.1) | • RMST is intuitively interpreted as the average survival time from baseline to time *t*.<br>• Clinically important time frame to evaluate RMST should be chosen a priori.<br>• In case of low event rate, the CI of RMST can be narrow and provide sufficient evidence for a non-inferiority claim. |

CI, confidence interval; MT, medical treatment; NNT, number needed to treat; RMST, restricted mean survival time; SAVR, surgical aortic valve replacement; TAVR, transcatheter aortic valve replacement.

*Individual patient data were not available to the authors; we reconstructed time-to-event data from the Kaplan-Meier curves in the published papers[1,2] using a well-established algorithm.[7] The hazard ratios (95% confidence intervals) estimated from the reconstructed data were close to the published hazard ratios and 95% confidence intervals.