VOLUME 34 · NUMBER 28 · OCTOBER 1, 2016

JOURNAL OF CLINICAL ONCOLOGY

STATISTICS IN BRIEF

Restricted Mean Survival Time: An Obligatory End Point for Time-to-Event Analysis in Cancer Trials?

Roger P. A'Hern

Medical Statistician (retired), London, United Kingdom.

Published online ahead of print at www.jco.org on August 8, 2016.

Editor's note: Statistics in Brief articles are short communications regarding statistical methods or issues. They are designed to alert and educate the readership about a method or issue that may be unfamiliar to or underused by the clinical research community. Each article will serve as a short primer and may refer the reader to additional sources for detailed information regarding both background and application.

Author's disclosures of potential conflicts of interest are found in the article online at www.jco.org. Author contributions are found at the end of this article.

Corresponding author: Roger P. A'Hern, 81 Hillier Rd, Battersea, London SW11 6AX, United Kingdom; e-mail: roger.ahern@icloud.com.

0732-183X/16/3428w-3474w/$20.00

DOI: 10.1200/JCO.2016.67.8045

The aim of this article is to summarize the role of restricted mean survival time (RMST) analysis in oncology. Everyone is familiar with the use of median survival, or more generally with median time to event (where the event could be progression or treatment failure), to summarize long-term patient outcome. However, many will be less comfortable with the concept of mean, average, or expected survival, even though averages are widely used in other contexts. However, the mean survival time is useful and an important analytic tool.

It is common to assume that medians are additive,[1] but an interesting anomaly arises if outcome is considered only in terms of median time to event. Table 1 shows a hypothetical example of a trial in advanced cancer, in which the median times to progression (ie, progression-free survival [PFS]) in the control and experimental arms are 12 months and 18 months, respectively, and the median survival post progression (SPP) is 24 months in both arms. Many would surmise that the median overall survival (OS) in the control arm would be 36 months, but if both PFS and SPP have an exponential distribution, the median OS would be 42.5 months.

This anomaly does not arise when considering means because for any time-to-event distribution, the sum of the means is always equal to the mean of the sums. In addition, if survival analysis is based on mean time to event, then conditional on SPP being the same in both arms, pure logic would dictate that whatever the duration of SPP, the true difference in mean PFS should be similar to the true difference in mean OS. RMST analysis, which is based on the mean, therefore overcomes the proportional hazards (PH; typically using Cox regression) handicap of a decaying summary measure (the hazard ratio) as the SPP gets larger (noted by Broglio and Berry[2]).

There are therefore good grounds for using RMST difference as a summary statistic for the comparison of outcome, meta-analysis,[3] and assessment of surrogacy. This paradigm has a practical basis; published studies typically present the difference in medians rather than the difference in means, but these two statistics will be highly correlated. The extension in median OS was noted to be similar to the extension in median PFS in a review of 55 phase III studies in nine cancer types.[4] This has also been separately noted in studies of advanced colorectal cancer[5,6] and advanced breast cancer.[6,7] However, one study[8] found that the median OS difference was approximately 40% less than the median PFS difference in non–small-cell lung cancer.

The use of RMST does not overcome any of the problems posed by limited patient follow-up. RMST is a time-dependent measure and is typically calculated over a defined period that has adequate follow-up; hence the description restricted mean survival time. Note that because PFS events happen sooner than deaths, the true additivity of the relationship, mean PFS + mean SPP = mean OS, may be obscured. For example, if an extreme restriction of follow-up (up to 72 months) is considered in the above example, then the difference in mean PFS is 7.3 months and the difference in mean OS is 4.6 months, despite the fact that they are both equal to 9 months in unrestricted analysis.

A number of authors[9-12] have advocated the use of RMST as an end point for the comparison of survival in randomized trials, with both the difference in RMST and the ratio of RMST having been considered as comparative measures. This article focuses on the use of RMST difference for comparative analysis. An important justification for the use of RMST difference is that it is appropriate for any time-to-event relationship between trial arms, whereas the PH model represents only the PH subclass of models.[9] PH models can be extended to incorporate time-dependent effects, but this can lead to loss of interpretability and increase in trial size.

Nonparametric methods are the gold standard for statistical analysis in situations in which the underlying distribution of the data cannot be confidently established. However, the PH assumption is semiparametric, assuming that there

Table 1. Hypothetical Trial in Which Outcomes for Progression-Free Survival and Survival Post Progression Both Follow an Exponential Distribution

	Median Time to Event (months)				Mean Time to Event (months)		
			OS				OS
Arm	PFS	SPP	Assumed	Actual	PFS	SPP	Assumed and Actual
Control	12	24	36	43	17	35	52
Experimental	18	24	42	51	26	35	61
Difference	6	0	6	8	9	0	9

NOTE. The corresponding distribution of overall survival times is not exponential, but is based on the sum of two exponential distributions. For simplicity, event times have been rounded to whole numbers.
Abbreviations: OS, overall survival; PFS, progression-free survival; SPP, survival post progression.

is a baseline event rate (ie, the hazard rate) by time pattern through follow-up and that groups other than the baseline group have an event rate that is a constant multiple (hazard ratio) of this rate. A benefit of the use of RMST is that it does not make this assumption. Studies indicate that RMST is comparable to PH in detecting differences between arms when hazards are proportional, but better when hazards are not proportional.[10]

If patients experience different event rates in the arms of a trial during follow-up, the relative characteristics of patients at risk in the arms will typically alter with time. Plausibly, this will result in

Table 2. Time Dependency of Treatment Effects on Recurrence in Selected Early Breast Cancer Trialists' Collaborative Group Meta-Analyses[13]

		Years					
		0-1	2-4	5-9	≥ 10		
Type of Treatment	No. of Patients	Recurrence Rate Ratio (SE)					*P*
Chemotherapy meta-analysis[14]							
Taxane plus anthracycline-based regimen *v* same or more (< doubled or approximately doubled) nontaxane cytotoxic chemotherapy	44,251	0.87 (0.03)	0.83 (0.03)	0.89 (0.06)	—		.9
Any anthracycline-based regimen *v* standard CMF (or near-standard CMF)*	14,649	0.83 (0.04)	1.00 (0.05)	0.95 (0.06)	1.2 (0.15)		.001
Any anthracycline-based regimen (eg, standard 4AC)† *v* no chemotherapy	8,575	0.58 (0.05)	0.81 (0.06)	0.89 (0.07)	0.72 (0.11)		< .001
Standard CMF (or near-standard CMF)* *v* no chemotherapy	5,253	0.49 (0.06)	0.78 (0.08)	0.84 (0.09)	0.98 (0.15)		< .001
Tamoxifen *v* nil[15]							
Tamoxifen for approximately 5 years *v* the same management, but no tamoxifen (ER+)	10,645	0.47 (0.05)	0.58 (0.04)	0.68 (0.06)	0.94 (0.09)		< .001
AIs *v* tamoxifen[16]							
Continuous AI *v* continuous tamoxifen (ER+)	9,885	0.64 (0.07)	0.80 (0.07)	0.92 (0.07)	—		.006
Continuous AI *v* tamoxifen then AI (ER+)	12,799	0.74 (0.07)	0.99 (0.08)	0.96 (0.12)	—		.052
Tamoxifen then AI *v* continuous tamoxifen (ER+)	11,798	—	0.56 (0.06)	0.97 (0.07)	0.92 (0.16)		< .001
		Years					
		0	1-2	3-4	5-9	≥ 10	
	No. of Patients	Recurrence Rate Ratio (SE)					*P*
Radiotherapy *v* none after breast-conserving surgery[17]	10,801	0.31 (0.05)	0.53 (0.05)	0.54 (0.07)	0.59 (0.07)	1.17 (0.19)	< .001
		Years					
		0-4	5-9				
	No. of Patients	Recurrence Rate Ratio (SE)					*P*
Radiotherapy *v* none after mastectomy + AD (pN+)[18]	3,131	0.76 (0.05)	0.72 (0.11)				.7
Total	131,787						

NOTE. Figures for trials of bisphosphonates[19] have not been included because the overall reduction in recurrence was nonsignificant. However, event rate ratios for bone recurrence showed a time-dependent pattern: They were 0.75 (years 0-1), 0.83 (years 2-4), and 1.02 (years 5-9).
Abbreviations: A, doxorubicin; AD, axillary dissection; AI, aromatase inhibitor; C, cyclophosphamide; CMF, cyclophosphamide, methotrexate, fluorouracil; ER+, estrogen receptor positive; F, fluorouracil; IV, intravenously; M, methotrexate; pN1, pathologic node positive.
*Standard CMF: Six cycles of C 100 mg/m^2 orally on days 1-14, M 40 mg/m^2 IV on days 1 and 8, and F 500 mg/m^2 IV on days 1 and 8, given once every 4 weeks. Near-standard CMF: six to 12 cycles with same doses as standard CMF and/or C 600 mg/m^2 IV on days 1 and 8, replacing C 100 mg/m^2 orally on days 1-14.
†Standard 4AC: four cycles of A 60 mg/m^2 and C 600 mg/m^2 IV, given once every 3 weeks.

nonproportionality of event rates. It is instructive to consider the incidence of non-PH in a specific, well-studied disease context. For example, Trinquart et al[12] found that 24% of trials violated the proportionality assumption. However, caution should be exercised in accepting this as an estimate of the proportion of treatments likely to display non-PH because trials of ineffective treatments were included in their analysis.

Time dependency of treatment effects on recurrence, in selected Early Breast Cancer Trialists' Collaborative Group meta-analysis publications[13] that showed a treatment effect, were summarized (Table 2). It can be seen that there is evidence of non-PH for all comparisons except for the taxane comparison and the effect of radiotherapy in patients with a mastectomy and axillary dissection. Approximately two of every three patients (84,400 of 131,800) in these trials are likely to have been in a trial which was analyzed assuming PH, but in which the meta-analysis indicates that PH would not be expected to apply. Note that failure to demonstrate a time-dependent effect in a specific trial may reflect lack of power to detect non-PH rather than confirm the existence of PH. A method of analysis such as RMST, which is appropriate for any time-to-event relationship, is therefore valuable.

To plan a trial with RMST as an end point, the sample size needs to be determined for the specific power and significance level that have to be achieved. Royston and Parmar[20] have investigated trial planning issues in detail. They have found that both an RMST test and a PH test can be incorporated in a combined test, with less than a 10% increase in sample size compared with PH alone. These authors have also examined the related issue of the specification of survival functions and follow-up patterns in trial set-up,[21] offering valuable tools for the implementation of RMST as an end point in randomized trials.

In summary, the hazard ratio and RMST difference are complementary techniques that provide alternative methods of summarizing treatment effects. True PH of the treatment effect cannot be assumed in randomized trials in oncology, and RMST analysis, which is more sensitive to non-PH, should always be considered. RMST is relevant to the analysis of studies with time-to event end points, including trials, meta-analyses, and surrogacy assessments. Furthermore, RMST is an important contribution to the analysis of time-dependent data, augmenting classic measures such as Kaplan-Meier estimation, survival estimates at time points, and medians. Trial consortiums can gain confidence in the use of RMST by applying this method to studies that have already been completed and published using PH analysis.

AUTHOR'S DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at www.jco.org.

REFERENCES

1. Saad ED, Katz A, Buyse M: Overall survival and post-progression survival in advanced breast cancer: A review of recent randomized clinical trials. J Clin Oncol 28:1958-1962, 2010
2. Broglio KR, Berry DA: Detecting an overall survival benefit that is derived from progression-free survival. J Natl Cancer Inst 101:1642-1649, 2009
3. Wei Y, Royston P, Tierney JF, et al: Meta-analysis of time-to-event outcomes from randomized trials using restricted mean survival time: Application to individual participant data. Stat Med 34:2881-2898, 2015
4. Redman MW, Goldman BH, LeBlanc M, et al: Modeling the relationship between progression-free survival and overall survival: The phase II/III trial. Clin Cancer Res 19:2646-2656, 2013
5. Tang PA, Bentzen SM, Chen EX, et al: Surrogate end points for median overall survival in metastatic colorectal cancer: Literature-based analysis from 39 randomized controlled trials of first-line chemotherapy. J Clin Oncol 25:4562-4568, 2007
6. Bowater RJ, Bridge LJ, Lilford RJ: The relationship between progression-free and post-progression survival in treating four types of metastatic cancer. Cancer Lett 262:48-53, 2008
7. Sherrill B, Amonkar M, Wu Y, et al: Relationship between effects on time-to-disease progression and overall survival in studies of metastatic breast cancer. Br J Cancer 99:1572-1578, 2008
8. Johnson KR, Ringland C, Stokes BJ, et al: Response rate or time to progression as predictors of survival in trials of metastatic colorectal cancer or non-small-cell lung cancer: A meta-analysis. Lancet Oncol 7:741-746, 2006
9. Royston P, Parmar MKB: The use of restricted mean survival time to estimate the treatment effect in randomized clinical trials when the proportional hazards assumption is in doubt. Stat Med 30: 2409-2421, 2011
10. Royston P, Parmar MKB: 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 13:152, 2013
11. 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 32:2380-2385, 2014
12. Trinquart L, Jacot J, Conner SC, et al: Comparison of treatment effects measured by the hazard ratio and by the ratio of restricted mean survival times in oncology randomized controlled trials. J Clin Oncol 34:1813-1819, 2016
13. University of Oxford: Clinical Trial Service Unit & Epidemiological Studies Unit. https://www.ctsu.ox.ac.uk/research/ebctcg
14. Early Breast Cancer Trialists' Collaborative Group (EBCTCG), Peto R, Davies C, et al: Comparisons between different polychemotherapy regimens for early breast cancer: Meta-analyses of long-term outcome among 100,000 women in 123 randomized trials. Lancet 379:432-444, 2012
15. Early Breast Cancer Trialists' Collaborative Group (EBCTCG), Davies C, Godwin J, et al: Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: Patient-level meta-analysis of randomized trials. Lancet 378:771-784, 2011
16. Early Breast Cancer Trialists' Collaborative Group (EBCTCG), Dowsett M, Forbes JF, et al: Aromatase inhibitors versus tamoxifen in early breast cancer: Patient-level meta-analysis of the randomized trials. Lancet 386:1341-1352, 2015
17. Early Breast Cancer Trialists' Collaborative Group (EBCTCG), Darby S, McGale P, et al: Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer mortality: Meta-analysis of individual patient data for 10,801 women in 17 randomized trials. Lancet 378:1707-1716, 2011
18. EBCTCG (Early Breast Cancer Trialists' Collaborative Group), McGale P, Taylor C, et al: Effect of radiotherapy after mastectomy and axillary surgery on 10-year recurrence and 20-year breast cancer mortality: Meta-analysis of individual patient data for 8135 women in 22 randomized trials. Lancet 383: 2127-2135, 2014
19. Early Breast Cancer Trialists' Collaborative Group (EBCTCG), Coleman R, Powles T, et al: Adjuvant bisphosphonate treatment in early breast cancer: Meta-analyses of individual patient data from randomised trials. Lancet 386:1353-1361, 2015
20. Royston P, Parmar MK: Augmenting the log-rank test in the design of clinical trials in which non-proportional hazards of the treatment effect may be anticipated. BMC Med Res Methodol 16:16, 2016
21. Royston P, Parmar MKB: An approach to trial design and analysis in the era of non-proportional hazards of the treatment effect. Trials 15:314, 2014

AUTHOR'S DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Restricted Mean Survival Time: An Obligatory End Point for Time-to-Event Analysis in Cancer Trials?

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or jco.ascopubs.org/site/ifc.

Roger P. A'Hern
No relationship to disclose