Clinical Oncology 30 (2018) e75–e80

Contents lists available at ScienceDirect

Clinical Oncology

journal homepage: www.clinicaloncologyonline.net

Original Article

Cancer Biology and Survival Analysis in Cancer Trials: Restricted Mean Survival Time Analysis versus Hazard Ratios

R.P. A'Hern

London, UK

Received 9 December 2017; received in revised form 16 March 2018; accepted 13 April 2018

Abstract

Hazard ratios are commonly used when comparing survival between two groups and make the assumption that the relative event rates do not change markedly during follow-up, i.e. that event rates are proportional. However, there is currently debate about the use of the proportional hazards assumption to summarise the treatment effect in survival analysis compared with restricted mean survival time (RMST) analysis, particularly in cancer trials. In many situations it is unrealistic to assume that relative event rates in two groups will be proportional throughout follow-up and, hence, RMST analysis, which does not make this assumption, may be preferable. Several benefits of the latter approach have been identified but the biological perspective is not often discussed. Biological features such as the patterns of tumour growth and response can also contribute to assessing the relative merit of these two methods; such biological considerations are the subject of this paper. The observation that the most commonly observed approximation to the underlying distribution of time to event data, the lognormal distribution, does not reliably show proportional hazards in the comparison of two groups, lends weight to a statistical approach that is not based on proportional hazards. The proportional hazards assumption should be viewed more critically when estimating treatment effects. An optimum approach may be to include both proportional hazards analysis and RMST analysis when comparing time to event endpoints.

Key words: Cancer biology; Cox regression; proportional hazards; restricted mean survival time analysis; survival analysis

Introduction

There is currently stimulating discussion about the most appropriate method of survival analysis in randomised trials in oncology, particularly focussing on the common use of the proportional hazards assumption to summarise treatment effects compared with restricted mean survival time (RMST) analysis [1–5]. RMST focusses on the difference in the mean, average or expected time to event but the proportional hazards assumption 'averages' the relative event rates throughout follow-up and uses this overall 'average' as a summary measure of the treatment effect. RMST can be most simply thought of as the area under the survival curve.

The salient features of the debate can be illustrated by considering two specific trials. Figure 1A shows overall survival in the two arms of GOG-111 [6,7], a randomised trial that illustrates changing hazards over follow-up. In GOG-111, 410 women with advanced ovarian cancer and residual masses larger than 1 cm after initial surgery were randomly assigned to receive cisplatin (75 mg/m^2) with either cyclophosphamide (750 mg/m^2) or paclitaxel (135 mg/m^2 over 24 h). The overall hazard ratio was 0.71, but it was not constant, varying from 0.53 in year 1 to 1.0 in year 8. The fitted dashed curves in Figure 1A were calculated assuming that survival in the two arms follows a lognormal distribution, the ubiquity of this survival distribution, which is rarely associated with proportional hazards, is discussed below. Royston and Parmar [6] found the power of RMST analysis was on average 90% compared with 83% for the Logrank test when hazards were non-proportional in the context they considered. In GOG-111 the hazard ratio decays (moves towards unity) as follow-up extends but there are also situations that show the reverse pattern, the initial hazard ratio is close to 1 and the later hazard ratio decreases, showing a treatment effect. Trials of immunotherapy agents provide examples of this phenomenon in which proportional hazards may not hold because sufficient time is required for the initial induction of an immune

Address for correspondence: 81 Hillier Road, Battersea, London SW11 6AX, UK.

E-mail address: roger.ahern@icloud.com.

https://doi.org/10.1016/j.clon.2018.04.011

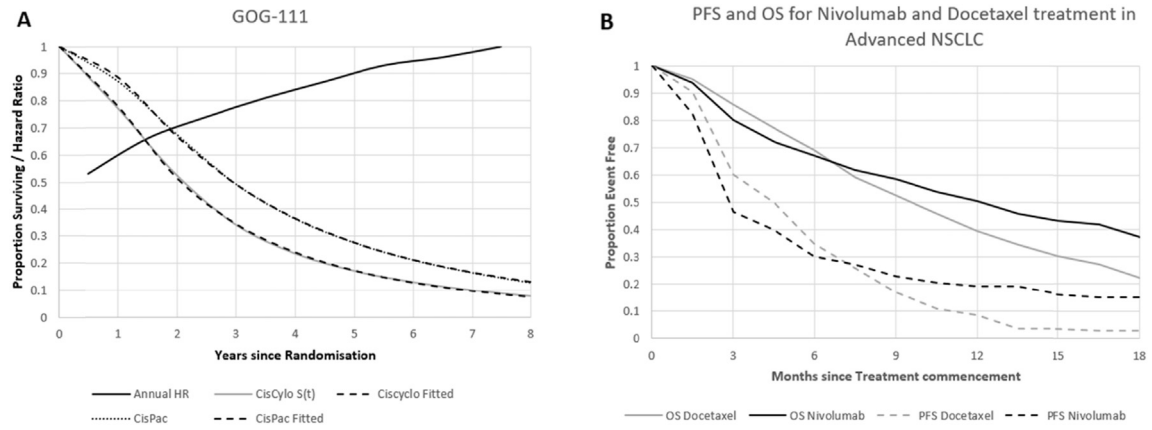

Fig 1. (A) Overall survival curves, together with the annual hazard ratio, for the two arms of GOG-111, a trial comparing two chemotherapy regimens in advanced ovarian cancer. (B) Progression-free and overall survival curves from a trial of nivolumab versus docetaxel in advanced non-small cell lung cancer. CisCyclo, cisplatin plus cyclophosphamide; CisPac, cisplatin plus paclitaxel; Fitted, fitted lognormal survival curve; HR, hazard ratio; OS, overall survival; PFS, progression-free survival.

response that is effective against the tumour. Figure 1B, for example, shows progression-free survival (PFS) and survival curves extracted from a publication of an immunotherapy trial that compared treatment with nivolumab versus docetaxel in advanced non-small cell lung cancer [8] ($n = 582$). Both PFS and overall survival only begin to show a beneficial effect for nivolumab after about 6 months. The hazard ratios were 0.92 (nivolumab: docetaxel, $P =$ ns) and 0.73 ($P = 0.002$) for PFS and overall survival, respectively, but the difference in RMST estimated from the curves was more similar, 0.6 months for PFS and 0.9 months for overall survival (both estimated up to 18 months). Statistical methods are detailed in Supplementary Appendix 1.

Three benefits of calculating mean survival will be briefly described. First, overall survival can be split into two periods when considering advanced cancer: the time to disease progression or PFS and the time from progression to death (survival post-progression; SPP). The simple equality PFS + SPP = overall survival can be best considered statistically in terms of means, i.e. mean PFS + mean SPP = mean overall survival [1]. This would imply that if mean SPP is similar in both arms of a trial, i.e. the biology of the progressive disease is similar in the two arms, the difference in mean PFS in the trial will tend to be similar to the difference in mean overall survival, a pattern consistent with that noted by a several authors in a number of different cancer types [1]. The hypothesis of similar mean SPP in the trial arms, although clearly not guaranteed in any specific trial, forms a valuable hypothesis when extrapolating differences in PFS to realistic differences in overall survival because a novel treatment may extend PFS, but there are no a priori grounds to conclude that disease that is resistant to treatment and leads to progression will lead to different survival in the two trial arms. Saad *et al.* [9] found that the median overall survival is typically three times the median PFS in advanced breast cancer trials. If this observation is applied uncritically to a trial in which the median PFSs in two arms are 7 months versus 9 months then it might be anticipated that the median overall survivals would be 21 months versus 27 months, a 6 month difference. However, the broad pattern across advanced breast cancer suggests the true overall survival difference is more likely to be similar to the difference in median PFS [10] (i.e. 21 months versus 23 months).

A second benefit of using mean survival is that it represents average duration in a specific disease state, e.g. average months progression free, hence it is of direct relevance in calculating quality-adjusted life years (QALYS), an important end point used in cost analysis to establish the cost/benefit balance of a therapy, for example by the National Institute for Health and Care Excellence (NICE) on behalf of the (UK) National Health Service [11]. Finally, in contrast to the hazard ratio, which does not have a straightforward interpretation in terms of duration of survival, RMST analysis gives a simple quantitative measure of the improvement in survival in months or years summarised over the whole period of follow-up over which it is calculated and can give a better measure of treatment effect [5].

It is relevant to ask whether biological considerations can contribute to the debate surrounding these two types of analysis. Robust estimation of treatment effects in small biologically defined subgroups of patients is necessary to accomplish the aims of personalised medicine and proportional hazards analysis may not be the optimum method to estimate such effects [12].

Tumour Growth and Long-term Outcome

The seminal model of tumour growth and treatment response, the log-kill model, was developed by Skipper *et al.* [33] at the Southern Research Institute, Birmingham, Alabama, USA and was founded on the empirical observation that leukaemia L1210 cells in mice grow exponentially up to a fatal size (10^9 cells). This model can be used to demonstrate some important properties of tumour growth and regression in response to treatment, many of which are also