Title

Censored patients in Kaplan–Meier plots of cancer drugs: An empirical analysis of data sharing

Authors

Rosen, Kate
Prasad, Vinay
Chen, Emerson Y

Publication Date

2020-12-01

DOI

10.1016/j.ejca.2020.09.031

European Journal of Cancer 141 (2020) 152–161

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Original Research

Censored patients in Kaplan–Meier plots of cancer drugs: An empirical analysis of data sharing

Kate Rosen [a], Vinay Prasad [b], Emerson Y. Chen [c,*]

[a] *School of Medicine, Oregon Health & Science University, Portland, OR, USA*
[b] *Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, USA*
[c] *Division of Hematology Oncology, Oregon Health & Science University, Knight Cancer Institute, Portland, OR, USA*

Received 24 July 2020; received in revised form 8 September 2020; accepted 25 September 2020
Available online 5 November 2020

KEYWORDS
Informative censoring;
Oncology trial;
Survival analysis;
Randomised trial;
Kaplan–Meier curve

Abstract ***Introduction:*** Kaplan–Meier survival analysis, the cornerstone of evaluating efficacy of oncology drugs in randomised controlled trials (RCTs), assumes censored patients are neither healthier nor sicker than those followed. We sought to examine whether censoring patterns differ between the control and experimental arms in one oncology journal that mandates the reporting of the number of patients censored.

Methods: In this retrospective review, proportion of censoring and study design data were gathered from RCTs published in The Lancet Oncology that reported Kaplan–Meier curves between May 2018 and August 2019. Differential censoring rates were analysed at the 1st, 3rd, 6th, and overall time points in each study. Analysis was stratified by curves reporting progression-free survival (PFS) or overall survival (OS) end-points.

Results: Of the 160 articles reviewed, 29 studies with 51 Kaplan–Meier curves were eligible. In both OS (N = 25) and PFS curves (N = 26), the absolute weighted difference in censoring between the control and experimental arms was initially positive, indicating more censoring in the control arm (first time point OS: 0.32%; PFS: 2.00%). The absolute difference then became negative, indicating more censoring in the experimental arm as time progressed (end-of-study OS: −7.54%; PFS: −9.09%).

Conclusion: Differences in censoring between control and experimental arms of cancer RCTs suggest that there could be systematic bias present at various study time points that may influence key results. Further investigation is needed, as possible reasons include study assignment disappointment, inappropriate follow-up length, lack of efficacy, or intolerable

* *Corresponding author*: Assistant Professor of Medicine, Oregon Health & Science University, Knight Cancer Institute, 3181 SW Sam Jackson Park Road, OC14HO, Portland, OR, 97239, USA.Fax: +1 503 494 3257.
E-mail address: cheem@ohsu.edu (E.Y. Chen).

https://doi.org/10.1016/j.ejca.2020.09.031

toxicity, each predominant at specific time points after randomisation.

1. Introduction

The most common primary outcome in cancer randomised controlled trials (RCTs) is a time-to-event end-point, such as overall survival (OS) or progression-free survival (PFS) [1]. The Kaplan–Meier (K-M) survival analysis is frequently used for time-to-event end-points, as the method maximally uses each participant's time-related data. In a K-M analysis, participants contribute to the survival estimate until the event of interest occurs (e.g. death, disease progression, or relapse) or until they are censored (e.g. loss to follow-up or mandatory data lock). The K-M method relies on the assumption of uninformative censoring, meaning there is no systematic difference between the patients censored in the control and experimental arm [2]. In other words, patients who are censored are no more or less likely to experience the event than those followed.

Despite widespread use of the K-M method, little attention has been paid to the proportion of participants censored over time. Most published landmark trials in oncology generally report only the proportion lost to follow-up in the flow diagram and the 'numbers at risk' for the event of interest in their K-M curves. Few, if any, disclose the number of patients censored at each time interval. Therefore, little is known about patterns of censoring, particularly if there are imbalances in the rate of censoring between the control and experimental arms.

In 2018, The Lancet Oncology editors mandated publications presenting K-M plots to include the number of censored patients at each time interval. This provides a unique opportunity to examine censoring in a broad range of trials, thereby eliminating a prior limitation of visually estimating the number of censored patients [3,4]. Here we sought to examine patterns of censoring between control and experimental arms of important published oncology trials throughout study follow-up.

2. Methods

2.1. Overview

This study is a retrospective review of all time-to-event survival data, including proportions of participants being censored, gathered from RCTs published in The *Lancet Oncology* from May 2018 to August 2019. We sought to investigate the relationship of differential censoring between the control and the experimental arms from well-known phase II and III RCTs of oncology drugs. Relevant STROBE guidelines were followed [5]. This study acquired data from only the published literature with no protected health information and did not itself enrol any participants and thus did not meet criteria to be submitted to the local institutional review board.

2.2. Study selection

Given that The *Lancet Oncology* author guidelines began in 2018 requiring manuscripts reporting K-M survival curve data to include both the number of patients at risk and the number of patients censored at each time point [6], we reviewed all RCTs from 1st May 2018 to 31st August 2019. Trials were selected for further review if 1) a RCT with control versus experimental arms was described, 2) at least one K-M survival analysis was used to evaluate the primary end-point, and 3) proportions of censored patients were reported continuously in the published figures.

Articles were excluded if they met any of the following: multiple experimental arms without a control arm (e.g. two arms testing different doses of the study drug), combination therapies without a control arm (e.g. two experimental arms testing study drug A + B versus study drug A), non-metastatic or non-advanced cancer trials (e.g. adjuvant therapies), or incorrect censoring reporting. Study eligibility was reviewed independently by both KR and EYC, and any discordance was resolved after discussion between both reviewers.

2.3. Data collection

For every eligible study, KR and EYC extracted data related to study title and reference, relevant dates, study design, treatment intervention, primary and secondary end-points, enrolment size, duration of follow-up, and K-M survival analyses. The number of patients at risk and being censored at each time point of the K-M curve of interest were used to calculate the proportion of patients being censored in that study arm (Fig. 1).

2.4. Outcomes

The main outcome, percent of patients censored, was calculated from each reported time point of the survival curve. It is calculated based on the cumulative sum of participants censored by that time point divided by the

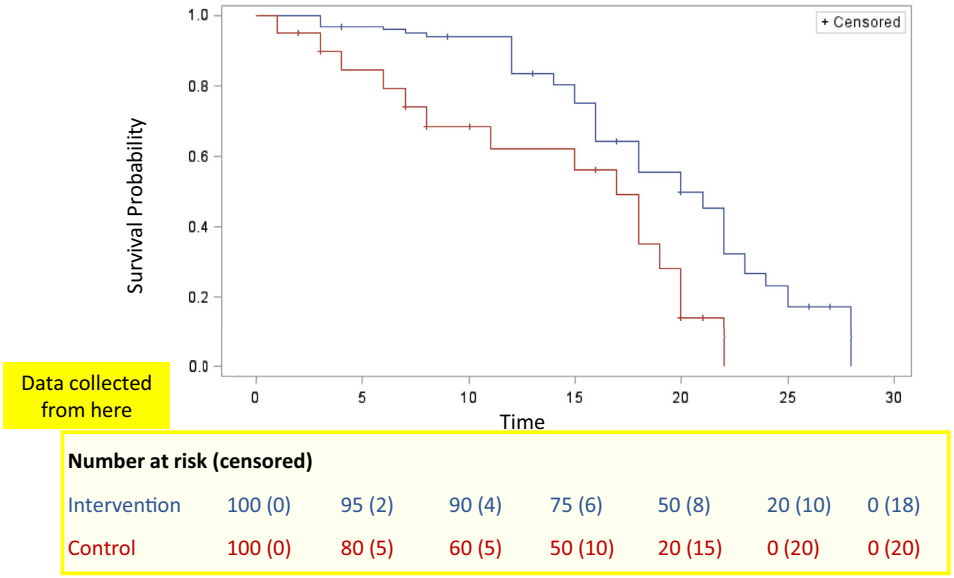

Fig. 1. The example of the calculating rate of censoring using the survival curve.

patients at risk at the start of the study. We specified the first interval, first three intervals, first six intervals, and overall study duration at the time of published analysis as the time points to analyse percent of patients censored over time. For example, using fabricated values in Fig. 1, the percent of patients censored in the first time point of the experimental arm would be (2/100)*100%, third time point would be (6/100)*100%, and overall censoring would be (18/100)*100%. The unweighted absolute % difference in control versus intervention censoring was calculated for each curve at each time point by subtraction. For example, the unweighted absolute difference for Fig. 1 overall censoring by the end of study would be 20%−18% = 2%. A positive value indicates more censoring in the control arm, and a negative value, more censoring in the experimental arm. Weighted averages were then calculated for this difference by adjusting for the relative study enrolment size to all the included studies. We analysed both OS and PFS and graphically presented them with respect to the sample size of each study for every stated time interval.

2.5. Statistical analysis

Difference between the control and experimental arms in all the time points of interest (1st, 3rd, 6th, and overall) was aggregated graphically and statistically tested using the Mann−Whitney test. Comparisons were made using both unweighted (number of K-M analyses) and weighted (by enrolled sample size of each clinical trial) analyses. The numerical difference between control and experimental arms was also compared between the first time point and the end of the study. All statistical testing was conducted using SAS version 9.4 (SAS Institute Inc., Cary NC, USA), otherwise all descriptive calculations and figures were executed using Microsoft Excel and Microsoft PowerPoint.

3. Results

Of 160 studies published in The *Lancet Oncology* between 1st May 2018 to 31st August 2019, 29 were included in our data set [7−36](Fig. 2). Among these 29 unique RCTs representing 14,708 patients with cancer, five (17%) trials reported only OS, and seven (24%) trials reported only PFS and 17 (59%) trials reported both. Three trials reported two OS analysis due to three randomised arms (two experimental, one control). Two of them likewise reported two PFS comparisons.

Among the 29 trials, the most common cancer types studied were haematologic malignancies (7 of 29, 24%) and gastrointestinal cancers (five of 29, 17%, Table 1). Among the 29 trials, 15 (52%) studies were open-label, and 14 (48%) were double-blinded. Five of 29 (17%) trials used supportive care or placebo as the control arm, and 24 of 29 (83%) trials used standard-of-care therapy as the control arm. The median sample size was 484 (range 119−1971), as presented in Table 1.

Among the 51 K-M plots, 25 were OS analyses and 26 were PFS analyses. The median number of time points for which at risk and censored individuals were reported below the K-M plot was 11 (range 5−20) for OS analyses and 10 (5−19) for PFS analyses (Table 1). The length of each time point ranged from 2 to 12 months.

We found the unweighted proportion of censored patients between control and experimental arms in the OS analyses to be no different statistically in the beginning (2.1% versus 2.1%, p = 0.41) or at the end

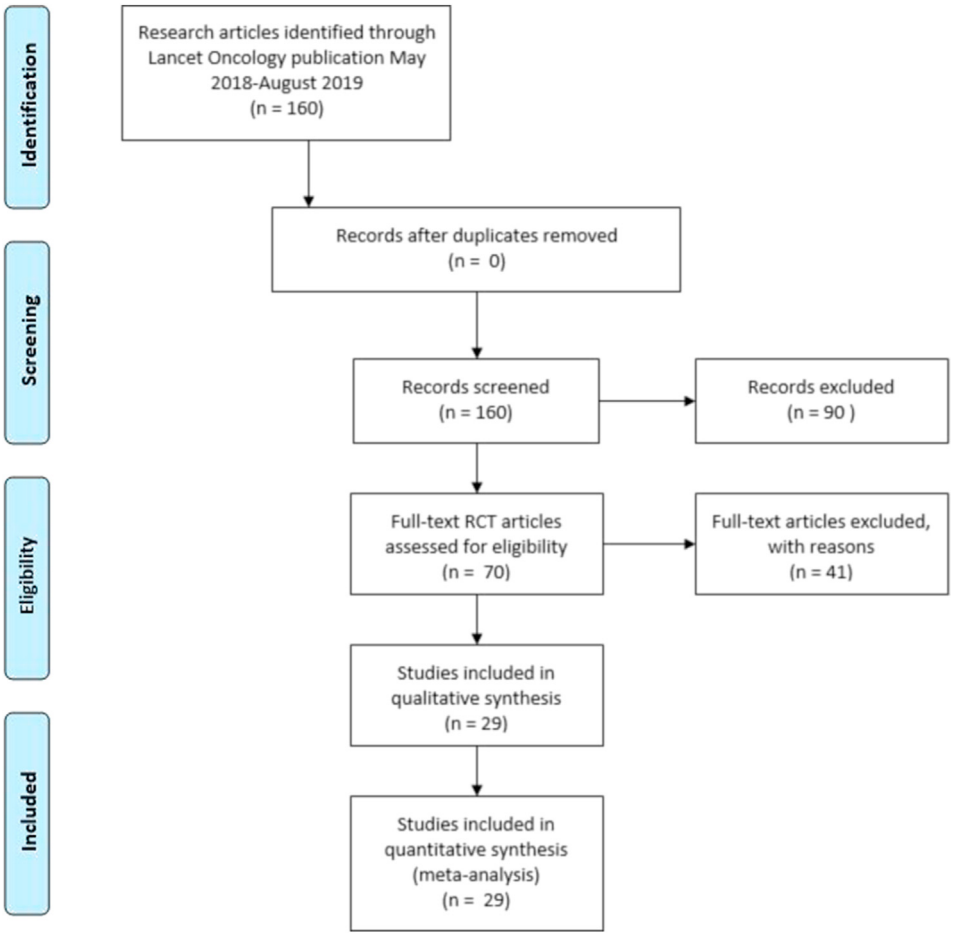

Fig. 2. The consort diagram for included randomised trials.

(33.5% versus 45.3%, p = 0.23, Supplement, eTable 1). Similar results were also observed with PFS despite numerical trends of showing slightly more censoring in the control arm at the first time point before progressing to more censoring in the experimental arm at the end of study (Supplement, eTable 2).

However, when analysing rates of censoring across all studies' time points, the median unweighted absolute difference of censoring in the control versus interventional arm in OS analyses is 0% (range −9.1–14.3%) at the first time point, but then shifted toward the experimental group at −3.7% (range: −29.7–8.8%, p = 0.03) by the conclusion of the studies. Similarly, for PFS, the median unweighted absolute difference in censoring between intervention and control changed from favouring the control group at the first time point (0.9% [range −6.3 to 20.6%]) to favouring the experimental group at the end (−5.4% (range −42.6–27.5%), p < 0.01).

When trials were weighted by the number of enrolled participants, we found the absolute difference between the two arms for OS again show greater censoring in the control arm in the first time point (0.32%) but shifts to intervention arm in the third time point (−0.42%) and continues to widen in the sixth time point (−1.47%) and at the end of the study (−7.54%, Fig. 3). When plotting all studies, the degree of censoring did not correlate with the sample size of the study. Outliers for OS included censoring in the control arm to as high as 14% in the first time point and censoring in the experimental arm to as high as −19% in the third time point.

Similarly for PFS, first and third time points again showed greater censoring in the control arm (2.00% and 1.31% respectively), then shifted to the intervention arm at sixth time point (−3.09%) and at the end of the study (−9.09%, Fig. 4). Again, the degree of censoring did not correlate with the sample size. Extreme outliers were also observed in several studies, including 21% in the first time point, 45% in the third, −17% in the sixth, and −43% in the last time point.

4. Discussion

4.1. Key results

We find a significant trend in censoring for both OS and PFS in cancer clinical trials. Specifically, early censoring appears to occur preferentially in control arms,

Table 1
Basic characteristics of the selected 29 studies with 51 Kaplan–Meier curves.

Characteristic	N (%)	
Number of unique studies	29	
Number of Kaplan–Meier survival curves	51	
Cancer type (N = 29)		
Haematologic malignancies	7 (24%)	
Gastrointestinal	5 (17%)	
Lung	4 (14%)	
Gynecologic	3 (10%)	
Melanoma	3 (10%)	
Brain	2 (7%)	
Breast	2 (7%)	
Prostate	2 (7%)	
Sarcoma	1 (3%)	
Study design (N = 29)		
Open-label	15 (52%)	
Double-blinded	14 (48%)	
Randomisation scheme (N = 29)		
1:1	20 (69%)	
2:1	8 (28%)	
2:1:1	1 (3%)	
Control arm (N = 29)		
Standard of care chemotherapy	5 (17%)	
Supportive care or placebo	24 (83%)	
Sample size[a]	484 (119–1971)	
Median follow-up length (months)[a]	19.2 (7.3–82.0)	
End-point (N = 51)	**Overall Survival**	**Progression-Free Survival**
Number of curves	25 (57%)	26 (43%)
Number of time points[a]	11 (5–20)	10 (5–19)
Length of time points (months)[a]	3 (2–12)	3 (2–12)

[a] Median (range).

gradually shifting toward experimental arms by the end of the study. Several points can be raised from these results.

First, early censoring favouring the control arm may be due to unintentional unblinding. Common side-effects of cancer drugs may help patients and physicians guess which arm they have been assigned. They may also dislike the control arm in open-label studies. Disappointment associated with the control arm may cause patients to drop out, or heighten evaluation bias among the treating physicians [37].

Such problem could be avoided by using the standard therapy deemed by each local standard rather than a weak control arm that is widely available at all sites [38]. Use of a double-blinded placebo and blinded independent review of outcomes could also preserve data integrity post-randomisation. Another possibility could be poor efficacy or intolerable side-effects seen early in trials [3,4,39], although it is unclear why this would preferentially occur in the control arm.

In our analysis, one trial, the registration trial for quizartinib, was notable for frequent early censoring of the placebo arm, with over 10% excess censoring in the control arm [11]. Notably, the US Food and Drug Administration noted this imbalance, and it became the subject of an Oncology Drug Advisory Committee discussion. Specifically, the FDA asked if, "substantial differences between study arms in early censoring and randomised not treated (RNT) rather than the treatment effect of quizartinib could have given rise to a small but statistically significant overall survival (OS) advantage" [40]. A simulation exercise suggests that even under the modest assumption that some censored patients do worse than average, the trial will likely lose significance.

Late censoring in the experimental arm could result from a number of causes. The most likely explanation is premature reporting, sometimes noted as right censoring [37]. Particularly for trials with positive results, fewer participants would have experienced death or progression in the experimental arm, and therefore more participants would be censored simply because of delayed trial entry. Another possibility is that unusual adverse events not immediately detected in the study but accumulating gradually, resulting in dose discontinuation, may cause preferential loss to follow-up, or attrition bias [37]. This explanation could be confirmed or refuted if authors provided the reasons for censoring, or if they collected time to treatment failure data as an alternate end-point for sensitivity analysis [4]. Such practical end-point, which was more commonly used in the past, accounts for both toxicities and efficacy and is less vulnerable to measurement bias and interval censoring [37,41].

Outliers exist at all time points and in both directions for both PFS and OS. Without censoring reported at each time point and only as overall numbers in a consort figure, as remains the current standard, censoring discrepancy at various time points might be missed in many oncology trials. Several studies did discuss end-points affected by post-discontinuation therapies or did conduct sensitivity analyses related to censoring, but never discussion of censoring imbalance [18,30]. By studying the outliers in Figs. 3 and 4, we were able to isolate studies that exhibited imbalanced censoring, a subset of which may influence results.

Our analysis lends itself to several concrete suggestions for improved reporting of clinical trial results. Authors could include a supplementary figure that shows the K-M plot of censored individuals over time (This plot would treat censoring as the event), which would allow visual inspection for imbalances in censoring. Study authors could also report the numbers being censored, numbers at risk, and numbers of events at every time point in the landmark K-M plot for complete transparency. Differences over time, particularly loss to follow-up, may require detailed explanation.

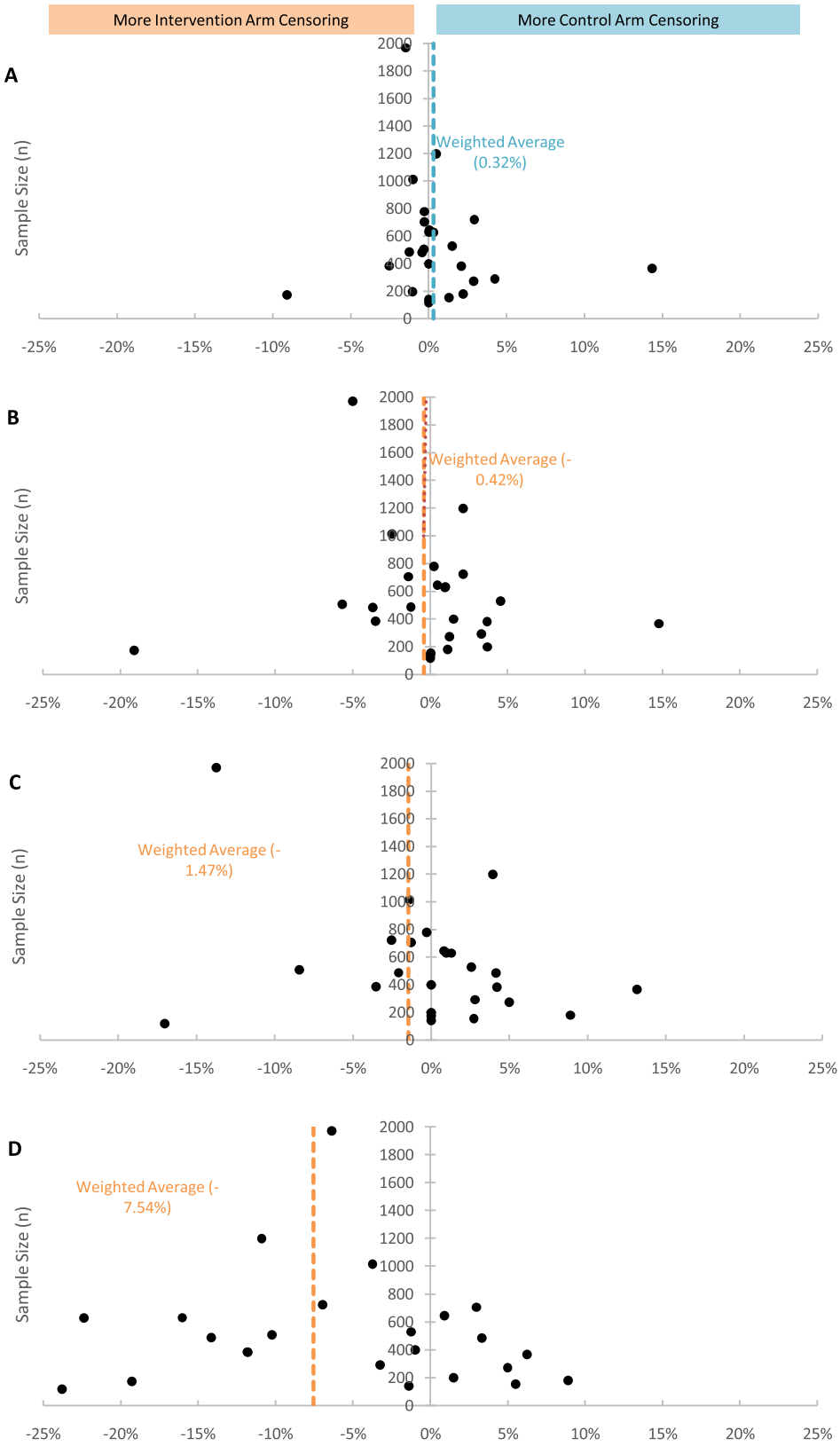

Fig. 3. Weighted absolute difference of % censored patients in overall survival analyses (panels A 1st time interval; B 3rd time interval; C 6th time interval; D overall).

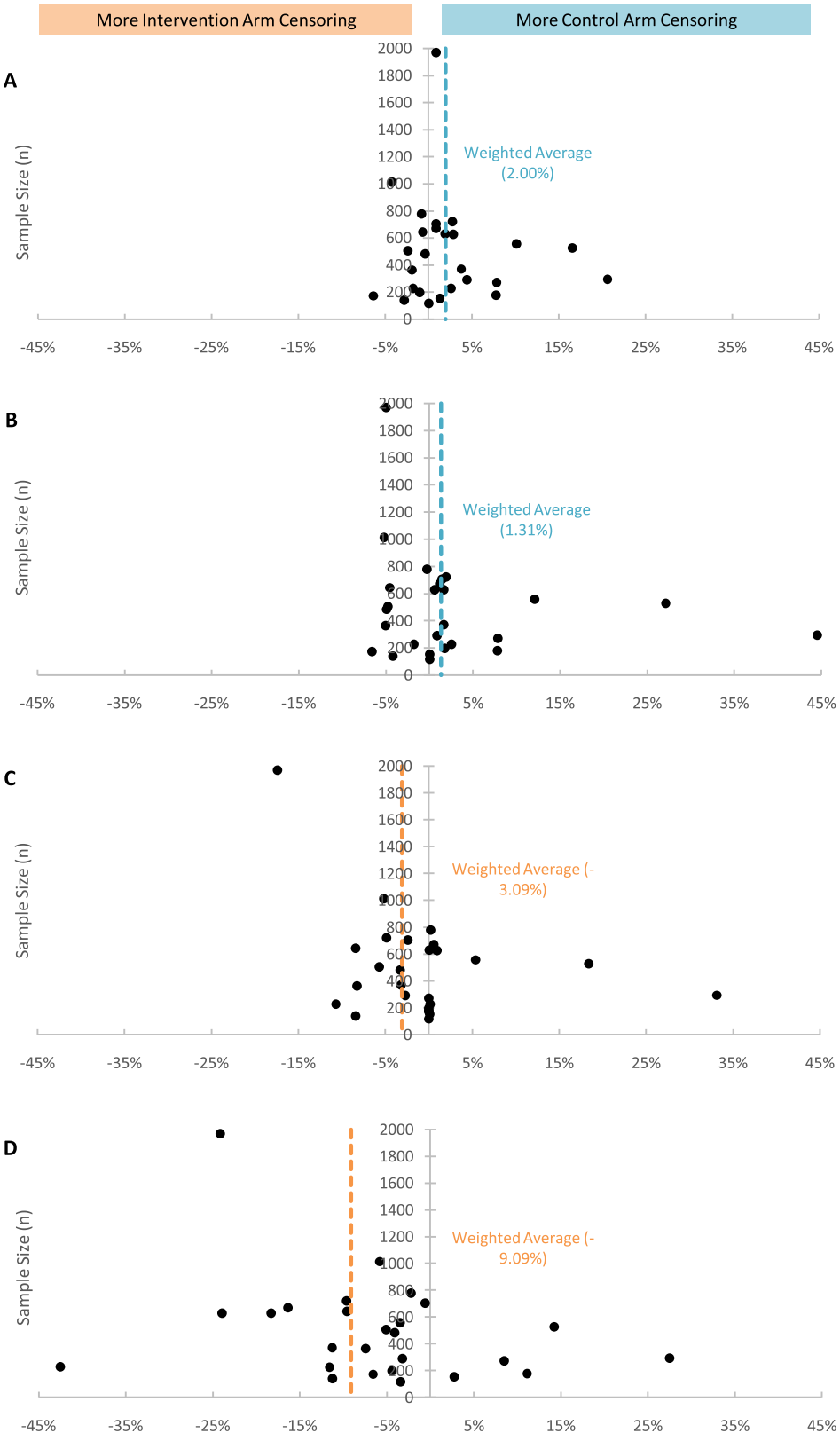

Fig. 4. Weighted absolute difference of % censored patients in progression-free survival analyses (panels A, 1st time interval; B, 3rd time interval; C, 6th time interval; D, overall).

Finally, sensitivity analyses, such as using time to treatment failure or running sample simulations, could be performed. If we assumed censored patients are more or less likely to experience the event of interest (i.e. censoring is informative) do significant results vanish? Data sharing of either de-identified primary data or at least original outputs of statistical analysis programs can further clarify these issues. Further investigation is warranted to explain reasons for differential censoring.

4.2. Limitations

The studies included in our analysis represent only one year of RCTs published in one journal. These effects may not be seen in other populations or fields of study. We hope more clinical trials will report complete survival data sets and examine the proportion of censoring in the data set. Although only a specific subset of RCTs was included here, they nonetheless represent a variety of study designs, thereby making direct comparisons somewhat challenging. In particular, some studies reported time points every 3 months, others every 2 or 6 months in the survival curves. These limitations suggest that one could standardise a specific duration reported for all K-M curves (e.g. every 3 months) to be presented at least in the supplemental data. There was also not enough data to examine subgroups by cancer types, study design, and drug class, which others could do; once more comprehensive data is reported by future trials.

4.3. Conclusion

Because of a new reporting standard implemented at The *Lancet Oncology*, we identified a trend of censoring favouring control arms at the start of RCTs followed by censoring favouring experimental arms late in the study in both OS and PFS analyses (Supplement, eFig. 1). Notably, several trials exhibit striking censoring imbalance, and in one case this was noted by regulatory agencies. Differences in rates of censoring should trigger further exploration as the cause, and sensitivity analyses and bias-resistant end-points could be performed to ensure results remain valid. If other journals were to follow suit, then the K-M assumption of uninformative censoring would be violated less frequently.

Author Contribution

Kate Rosen, Validation, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualisation. Vinay Prasad, Conceptualisation, Methodology, Formal analysis, Writing - review & editing, Supervision, Funding acquisition. Emerson Y. Chen, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Visualisation, Project administration.

Funding

OHSU Knight Cancer Institute.

Conflict of interest statement

Vinay Prasad reports Arnold Ventures funding; Johns Hopkins Press, Medscape; honoraria for grand rounds/lectures from universities, medical centres, non-profits, and professional societies; consulting role in UnitedHealthcare; receiving speaking fees from Evicorefees; other roles in Plenary Session podcast of Patreon backers.

Emerson Chen reports honoraria from Horizon CME.

All other authors have nothing to declare.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejca.2020.09.031.

References

[1] Kay A, Higgins J, Day AG, Meyer RM, Booth CM. Randomized controlled trials in the era of molecular oncology: methodology, biomarkers, and end points. Ann Oncol 2012;23:1646–51.

[2] Templeton AJ, Amir E, Tannock IF. Informative censoring - a neglected cause of bias in oncology trials. Nat Rev Clin Oncol 2020 Jun;17(6):327–8. https://doi.org/10.1038/s41571-020-0368-0. PMID: 32273582.

[3] Prasad V, Bilal U. The role of censoring on progression free survival: oncologist discretion advised. Eur J Canc 2015;51:2269–71.

[4] Templeton AJ, Ace O, Amir E, Vera-Badillo F, Ocana A, Pond GR, et al. Influence of censoring on conclusions of trials for women with metastatic breast cancer. Eur J Canc 2015;51:721–4.

[5] von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007;370:1453–7.

[6] Randomised trials in the Lancet: formatting guidelines. Lancet 2018. https://www.thelancet.com/pb/assets/raw/Lancet/authors/Rctguidelines.pdf.

[7] Barlesi F, Vansteenkiste J, Spigel D, Ishii H, Garassino M, de Marinis F, et al. Avelumab versus docetaxel in patients with platinum-treated advanced non-small-cell lung cancer (JAVELIN Lung 200): an open-label, randomised, phase 3 study. Lancet Oncol 2018;19:1468–79.

[8] Bisogno G, Jenney M, Bergeron C, Gallego Melcon S, Ferrari A, Oberlin O, et al. Addition of dose-intensified doxorubicin to standard chemotherapy for rhabdomyosarcoma (EpSSG RMS 2005): a multicentre, open-label, randomised controlled, phase 3 trial. Lancet Oncol 2018;19:1061–71.

[9] Chekerov R, Hilpert F, Mahner S, El-Balat A, Harter P, De Gregorio N, et al. Sorafenib plus topotecan versus placebo plus topotecan for platinum-resistant ovarian cancer (TRIAS): a multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. Lancet Oncol 2018;19:1247–58.

[10] Clarke N, Wiechno P, Alekseev B, Sala N, Jones R, Kocak I, et al. Olaparib combined with abiraterone in patients with metastatic

castration-resistant prostate cancer: a randomised, double-blind, placebo-controlled, phase 2 trial. Lancet Oncol 2018;19:975–86.
[11] Cortes JE, Khaled S, Martinelli G, Perl AE, Ganguly S, Russell N, et al. Quizartinib versus salvage chemotherapy in relapsed or refractory FLT3-ITD acute myeloid leukaemia (QuANTUM-R): a multicentre, randomised, controlled, open-label, phase 3 trial. Lancet Oncol 2019;20:984–97.
[12] Dummer R, Ascierto PA, Gogas HJ, Arance A, Mandala M, Liszkay G, et al. Overall survival in patients with BRAF-mutant melanoma receiving encorafenib plus binimetinib versus vemurafenib or encorafenib (COLUMBUS): a multicentre, open-label, randomised, phase 3 trial. Lancet Oncol 2018;19:1315–27.
[13] Fizazi K, Tran N, Fein L, Matsubara N, Rodriguez-Antolin A, Alekseev BY, et al. Abiraterone acetate plus prednisone in patients with newly diagnosed high-risk metastatic castration-sensitive prostate cancer (LATITUDE): final overall survival analysis of a randomised, double-blind, phase 3 trial. Lancet Oncol 2019;20:686–700.
[14] Friedlander M, Gebski V, Gibbs E, Davies L, Bloomfield R, Hilpert F, et al. Health-related quality of life and patient-centred outcomes with olaparib maintenance after chemotherapy in patients with platinum-sensitive, relapsed ovarian cancer and a BRCA1/2 mutation (SOLO2/ENGOT Ov-21): a placebo-controlled, phase 3 randomised trial. Lancet Oncol 2018;19:1126–34.
[15] Fuchs CS, Shitara K, Di Bartolomeo M, Lonardi S, Al-Batran SE, Van Cutsem E, et al. Ramucirumab with cisplatin and fluoropyrimidine as first-line therapy in patients with metastatic gastric or junctional adenocarcinoma (RAINFALL): a double-blind, randomised, placebo-controlled, phase 3 trial. Lancet Oncol 2019;20:420–35.
[16] Jackson GH, Davies FE, Pawlyn C, Cairns DA, Striha A, Collett C, et al. Lenalidomide maintenance versus observation for patients with newly diagnosed multiple myeloma (Myeloma XI): a multicentre, open-label, randomised, phase 3 trial. Lancet Oncol 2019;20:57–73.
[17] Jiang Z, Li W, Hu X, Zhang Q, Sun T, Cui S, et al. Tucidinostat plus exemestane for postmenopausal patients with advanced, hormone receptor-positive breast cancer (ACE): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol 2019;20:806–15.
[18] Kim YH, Bagot M, Pinter-Brown L, Rook AH, Porcu P, Horwitz SM, et al. Mogamulizumab versus vorinostat in previously treated cutaneous T-cell lymphoma (MAVORIC): an international, open-label, randomised, controlled phase 3 trial. Lancet Oncol 2018;19:1192–204.
[19] Lombardi G, De Salvo GL, Brandes AA, Eoli M, Ruda R, Faedi M, et al. Regorafenib compared with lomustine in patients with relapsed glioblastoma (REGOMA): a multicentre, open-label, randomised, controlled, phase 2 trial. Lancet Oncol 2019;20:110–9.
[20] Long GV, Dummer R, Hamid O, Gajewski TF, Caglevic C, Dalle S, et al. Epacadostat plus pembrolizumab versus placebo plus pembrolizumab in patients with unresectable or metastatic melanoma (ECHO-301/KEYNOTE-252): a phase 3, randomised, double-blind study. Lancet Oncol 2019;20:1083–97.
[21] Moreau P, Mateos MV, Berenson JR, Weisel K, Lazzaro A, Song K, et al. Once weekly versus twice weekly carfilzomib dosing in patients with relapsed and refractory multiple myeloma (A.R.R.O.W.): interim analysis results of a randomised, phase 3 study. Lancet Oncol 2018;19:953–64.
[22] Moreno C, Greil R, Demirkan F, Tedeschi A, Anz B, Larratt L, et al. Ibrutinib plus obinutuzumab versus chlorambucil plus obinutuzumab in first-line treatment of chronic lymphocytic leukaemia (iLLUMINATE): a multicentre, randomised, open-label, phase 3 trial. Lancet Oncol 2019;20:43–56.
[23] Robak T, Jin J, Pylypenko H, Verhoef G, Siritanaratkul N, Drach J, et al. Frontline bortezomib, rituximab, cyclophosphamide, doxorubicin, and prednisone (VR-CAP) versus rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) in transplantation-ineligible patients with newly diagnosed mantle cell lymphoma: final overall survival results of a randomised, open-label, phase 3 study. Lancet Oncol 2018;19:1449–58.
[24] Saito H, Fukuhara T, Furuya N, Watanabe K, Sugawara S, Iwasawa S, et al. Erlotinib plus bevacizumab versus erlotinib alone in patients with EGFR-positive advanced non-squamous non-small-cell lung cancer (NEJ026): interim analysis of an open-label, randomised, multicentre, phase 3 trial. Lancet Oncol 2019;20:625–35.
[25] Shitara K, Doi T, Dvorkin M, Mansoor W, Arkenau HT, Prokharau A, et al. Trifluridine/tipiracil versus placebo in patients with heavily pretreated metastatic gastric cancer (TAGS): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol 2018;19:1437–48.
[26] Tripathy D, Im SA, Colleoni M, Franke F, Bardia A, Harbeck N, et al. Ribociclib plus endocrine therapy for premenopausal women with hormone-receptor-positive, advanced breast cancer (MONALEESA-7): a randomised phase 3 trial. Lancet Oncol 2018;19:904–15.
[27] van den Bent MJ, Klein M, Smits M, Reijneveld JC, French PJ, Clement P, et al. Bevacizumab and temozolomide in patients with first recurrence of WHO grade II and III glioma, without 1p/19q co-deletion (TAVAREC): a randomised controlled phase 2 EORTC trial. Lancet Oncol 2018;19:1170–9.
[28] Vergote I, Scambia G, O'Malley DM, Van Calster B, Park SY, Del Campo JM, et al. Trebananib or placebo plus carboplatin and paclitaxel as first-line treatment for advanced ovarian cancer (TRINOVA-3/ENGOT-ov2/GOG-3001): a randomised, double-blind, phase 3 trial. Lancet Oncol 2019;20:862–76.
[29] West H, McCleod M, Hussein M, Morabito A, Rittmeyer A, Conter HJ, et al. Atezolizumab in combination with carboplatin plus nab-paclitaxel chemotherapy compared with chemotherapy alone as first-line treatment for metastatic non-squamous non-small-cell lung cancer (IMpower130): a multicentre, randomised, open-label, phase 3 trial. Lancet Oncol 2019;20:924–37.
[30] Zhu AX, Kang YK, Yen CJ, Finn RS, Galle PR, Llovet JM, et al. Ramucirumab after sorafenib in patients with advanced hepatocellular carcinoma and increased alpha-fetoprotein concentrations (REACH-2): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol 2019;20:282–96.
[31] Richardson PG, Oriol A, Beksac M, Liberati AM, Galli M, Schjesvold F, et al. Pomalidomide, bortezomib, and dexamethasone for patients with relapsed or refractory multiple myeloma previously treated with lenalidomide (OPTIMISMM): a randomised, open-label, phase 3 trial. Lancet Oncol 2019;20:781–94.
[32] Tabernero J, Hoff PM, Shen L, Ohtsu A, Shah MA, Cheng K, et al. Pertuzumab plus trastuzumab and chemotherapy for HER2-positive metastatic gastric or gastro-oesophageal junction cancer (JACOB): final analysis of a double-blind, randomised, placebo-controlled phase 3 study. Lancet Oncol 2018;19:1372–84.
[33] Gregorc V, Gaafar RM, Favaretto A, Grossi F, Jassem J, Polychronis A, et al. NGR-hTNF in combination with best investigator choice in previously treated malignant pleural mesothelioma (NGR015): a randomised, double-blind, placebo-controlled phase 3 trial. Lancet Oncol 2018;19:799–811.
[34] Bromberg JEC, Issa S, Bakunina K, Minnema MC, Seute T, Durian M, et al. Rituximab in patients with primary CNS lymphoma (HOVON 105/ALLG NHL 24): a randomised, open-label, phase 3 intergroup study. Lancet Oncol 2019;20:216–28.
[35] Eng C, Kim TW, Bendell J, Argiles G, Tebbutt NC, Di Bartolomeo M, et al. Atezolizumab with or without cobimetinib versus regorafenib in previously treated metastatic colorectal cancer (IMblaze370): a multicentre, open-label, phase 3, randomised, controlled trial. Lancet Oncol 2019;20:849–61.

[36] Hodi FS, Chiarion-Sileni V, Gonzalez R, Grob JJ, Rutkowski P, Cowey CL, et al. Nivolumab plus ipilimumab or nivolumab alone versus ipilimumab alone in advanced melanoma (CheckMate 067): 4-year outcomes of a multicentre, randomised, phase 3 trial. Lancet Oncol 2018;19:1480–92.

[37] Villaruz LC, Socinski MA. The clinical viewpoint: definitions, limitations of RECIST, practical considerations of measurement. Clin Canc Res 2013;19:2629–36.

[38] Hilal T, Sonbol MB, Prasad V. Analysis of control arm quality in randomized clinical trials leading to anticancer drug approval by the US food and drug administration. JAMA Oncol 2019;5:887–92.

[39] Altman DG. Missing outcomes in randomized trials: addressing the dilemma. Open Med 2009;3:e51–3.

[40] FDA briefing document: oncologic drugs advisory committee (ODAC) meeting May 14, 2019. Administration USFaD; 2019.

[41] Chibaudel B, Bonnetain F, Shi Q, Buyse M, Tournigand C, Sargent DJ, et al. Alternative end points to evaluate a therapeutic strategy in advanced colorectal cancer: evaluation of progression-free survival, duration of disease control, and time to failure of strategy–an Aide et Recherche en Cancerologie Digestive Group Study. J Clin Oncol 2011;29:4199–204.