

# Postprogression therapy and confounding for the estimated treatment effect on overall survival in phase III oncology trials

Alexander D Sherry ,<sup>1</sup> Pavlos Msaouel,<sup>2,3</sup> Timothy A Lin,<sup>4</sup> Joseph Abi Jaoude,<sup>1,5</sup> Ramez Kouzy,<sup>1</sup> Esther J Beck ,<sup>1</sup> Avital M Miller,<sup>1</sup> Adina H Passy,<sup>1</sup> Gabrielle S Kupferman,<sup>1</sup> Eugene J Koay,<sup>6</sup> Clifton David Fuller,<sup>1</sup> Charles R Thomas,<sup>7</sup> Zachary R McCaw,<sup>8,9</sup> Ethan B Ludmir<sup>6,10</sup>

**To cite:** Sherry AD, Msaouel P, Lin TA, *et al.* Postprogression therapy and confounding for the estimated treatment effect on overall survival in phase III oncology trials. *BMJ Oncology* 2024;**3**:e000322. doi:10.1136/bmjonc-2024-000322

Received 08 January 2024  
Accepted 03 April 2024

## ABSTRACT

**Objective** Estimations of the treatment effect on overall survival (OS) may be influenced by post-progression therapies (PPTs). It is unclear how often OS analyses account for PPT effects. The purpose of this cross-sectional analysis was to determine the prevalence of OS analyses accounting for PPT effects in phase III oncology trials.

**Methods and analysis** We screened two-arm, superiority design, phase III, randomised, oncology trials reporting OS from ClinicalTrials.gov. The primary outcome was the frequency of OS analyses adjusting for PPT confounding. Logistic regressions computed ORs for the association between trial-level covariates and the outcome.

**Results** A total of 334 phase III trials enrolling 265 310 patients were included, with publications between 2004 and 2020. PPTs were reported in 47% of trials (157 of 334), and an analysis accounting for PPTs was performed in only 12% of trials (N=41). PPT adjustments were often prespecified (N=23, 56%), and appeared to be more likely in cross-over studies (OR 5.04, 95% CI 2.42 to 10.38) and studies with discordant surrogate-OS findings (OR 2.26, 95% CI 1.16 to 4.38). In key subgroup analyses, PPT analyses were infrequent, including 8% of trials among those studying locoregional/first-line therapy and 11% of trials among those powered for OS.

**Conclusions** Although time on PPTs is an important component of OS, PPTs are rarely considered in OS analyses, which may introduce confounding on estimates of the treatment effect on OS. PPTs and methods to account for their effects on OS estimates should be considered at the time of trial design and reporting.

## INTRODUCTION

Overall survival (OS) is defined in phase III trials as the time from randomisation until death or censoring. Compared with endpoints that are used as surrogates for OS, such as progression-free survival (PFS), OS is often considered more meaningful to patients.<sup>1</sup> Surrogate endpoints such as PFS can be subject to a number of biases, including

### WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Following progression on a clinical trial, patients often receive non-protocol-specified treatments at the discretion of their oncologist, which are associated with observational selection biases. Differential efficacy and differential receipt of postprogression therapies between arms may introduce confounding between arms in estimating treatment effects on overall survival.

### WHAT THIS STUDY ADDS

⇒ This large-scale cross-sectional study of phase III randomised controlled trials found that postprogression therapies were reported in a minority of clinical trials. Attempts to adjust for confounding on overall survival estimates from postprogression therapies were rare.

### HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ There may be confounding in the overall survival estimates of phase III oncology trials. Regulatory agencies and oncologists should consider confounding effects in the treatment estimates of overall survival from postprogression therapies when interpreting current clinical trials. Future trials should consider prespecification of methods to account for postprogression therapy confounding at the time of trial design.

informative censoring and measurement biases; moreover, several groups have also found that surrogate endpoints often have only limited correlations with OS.<sup>2–6</sup> For these and other reasons, OS is typically recommended as the endpoint for use in regulatory approvals in phase III trials.<sup>7–10</sup>

Unlike progression-based surrogate endpoints, however, estimates of the treatment effect on OS may be influenced by treatments received after progression, aka

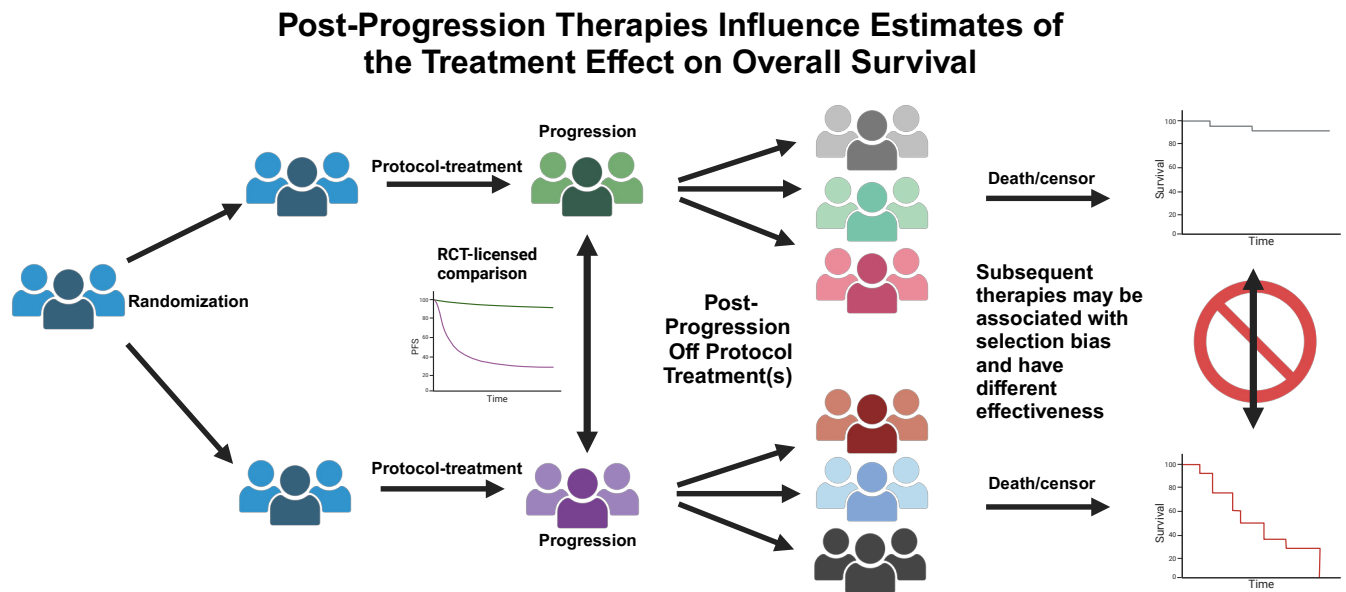


© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

### Correspondence to

Dr Ethan B Ludmir;  
ebludmir@mdanderson.org



**Figure 1** Postprogression therapy effects on estimates of the treatment effect on overall survival. Created using Biorender. com. RCT, randomised controlled trial.

postprogression therapies (PPTs), since disease progression prompting the initiation of PPT is generally considered an event for surrogate endpoints such as PFS.<sup>11–13</sup> In some cases, PPTs have been shown to suggest an OS benefit in randomised trials of clinically inert or harmful drugs (in which patients in the ineffective experimental arm received more effective PPTs, compared with patients in the control arm who crossed over to the ineffective experimental therapy following progression).<sup>14 15</sup> Not accounting for PPTs can also produce the opposite effect and lead to underestimation of the OS benefit, for example, if patients in the less effective arm cross-over to the more effective therapy after progression.<sup>16</sup> Thus, PPTs may have a considerable impact on estimating the treatment effect on OS and subsequent interpretations for clinical practice (figure 1).<sup>17–19</sup> Because the scope of this concern in oncology is poorly understood, the present empirical study was undertaken to evaluate the frequency of phase III trials attempting to account for PPT effects in OS analyses.

## METHODS

A cross-sectional analysis of published, phase III, randomised, controlled, oncology trials was conducted by screening ClinicalTrials.gov in February 2020 without date limitations.<sup>20</sup> Key inclusion criteria were (1) interventional anticancer therapy being tested, (2) a two-arm superiority design and (3) the reporting of matured OS findings as a primary or secondary endpoint (online supplemental figure 1). The maturation of OS findings was defined according to each trial. This study complies with the Strengthening the Reporting of Observational Studies in Epidemiology guidelines.<sup>21</sup> Data were curated from all available publications (manuscript reporting

the primary endpoint as well as any updated analyses, including all supplemental data through February 2024), ClinicalTrials.gov, and the trial protocol (English only), if available, by a single reviewer. For simplicity, surrogate endpoints such as disease-free survival, event-free survival and time to progression were collectively referred to as PFS, which is the most commonly used surrogate survival endpoint.

The primary objective of this study was to evaluate the frequency of trials attempting to statistically account for PPT confounding in an analysis of OS. Any statistical attempts to account for PPTs were considered; several examples include censoring the OS analysis at the time of PPT initiation, inverse probability of censoring weighting, and rank-preserving structural failure time models.<sup>22</sup> Whether any PPT adjustments were prespecified versus post hoc was also recorded; in the absence of a published protocol or statistical analysis plan, it was assumed that analyses not clearly defined as prespecified were post hoc.

Because PPT confounding is present when salvage therapies are available, a sensitivity analysis was performed for trials evaluating the neoadjuvant, definitive, adjuvant, maintenance, first-line metastatic solid tumour or first-line haematological malignancy settings. A second sensitivity analysis was performed including only trials powered for OS with a primary or coprimary endpoint of OS.

Continuous variables were summarised by median and IQR, and categorical variables were summarised by frequency. The trends of PPT adjustments over time were evaluated by ordinary least squares linear regression, in which the slope ( $m$ ) of the regression represented the rate of change. Univariable binary logistic regression models examined associations of trial-level variables with the outcome of any PPT adjustment to calculate ORs and

95% CIs. Multivariable analysis was not performed due to low numbers of the event of interest. All tests were two sided. Missing data were not encountered for this analysis of public data. The significance ( $\alpha$ ) level was set at 0.05. Analyses were performed by using SAS V.9 and plots were created using Prism V.10 (GraphPad, La Jolla, California, USA).

### Patient and public involvement

Patients were not involved as contributors in this study.

## RESULTS

After screening 785 phase III, randomised trials, a total of 334 therapeutic, interventional, two-arm, superiority-design trials enrolling 265 310 patients were included (online supplemental figure 1). Primary endpoint publication dates ranged from 2004 to 2020. The majority of trials evaluated metastatic solid tumours (N=218, 65%) (table 1). OS was the primary or coprimary endpoint of 168 trials (50%). Most trials were open-label (N=206, 62%). Of the 128 double-blind studies, unblinding occurred at the time of progression in 20 trials or at the time of an interim analysis/database lock in 27 trials. Unblinding procedures were not specified in the available trial documentation for 66 trials. In 46 trials, cross-over was allowed, but not required, and in 3 trials, cross-over was required, resulting in a total of 49 cross-over trials (15%). Of these 49 trials, OS was the primary endpoint in 7 trials (with a placebo control arm in 5 of the 7 trials) (online supplemental table 1). In 41 trials, cross-over was prohibited. Rates of cross-over from the control arm to the experimental therapy were reported in 42 trials; in these trials, a median of 55% of eligible patients in the control arm crossed over to the experimental therapy after progression (IQR 37%–76%).

The overall interpretation of PFS and OS was discordant in 107 trials (32%) (eg, superior PFS with no difference in OS). Of trial-level characteristics studied, cross-over rate was most associated with discordance between PFS and OS findings. Of 42 trials reporting cross-over rates, 16 trials had discordant PFS and OS findings. Higher rate of cross-over to the experimental therapy was associated with an increased odds of discordance between PFS and OS (median cross-over rate for discordant trials 74% vs 47% for concordant trials; per each additional percentage point of cross-over, OR 1.03, 95% CI 1.01 to 1.07,  $p=0.03$ ).

Overall, PPTs were reported in any format by 157 trials (47%). However, only 12% of trials (41 of 334 trials) reported an OS analysis adjusted for PPTs. The most typical approach for addressing PPT confounding was simple censoring at the start of PPT ( $n=19$ ). Nine trials used multiple approaches to account for PPT impact on OS (table 2). At least one PPT analysis was preplanned in 56% of trials (N=23), with the remaining trials using only post hoc analyses (N=18 trials, 44%). There was no discernable increase over time in trials attempting to account for PPT ( $m=0.02$ , 95% CI  $-1.29$  to  $1.33$ ,  $p=0.97$ )

(online supplemental figure 2). Trials with cross-over were more likely to account for PPT confounding (33% vs 9%, OR 5.04, 95% CI 2.42 to 10.38,  $p<0.0001$ ). Of trials with PPT analyses, there was no significant association between cross-over trials and prespecification of PPT analyses (50% vs 60%, OR 0.67, 95% CI 0.18 to 2.37,  $p=0.53$ ). Furthermore, trials with a discordance between PFS and OS were more likely to account for PPT confounding (19% vs 9%, OR 2.26, 95% CI 1.16 to 4.38,  $p=0.02$ ). PPT analyses conducted by trials with discordant findings were not more likely to be post hoc compared with trials with concordant findings (40% vs 48%, OR 0.73, 95% CI 0.21 to 2.53,  $p=0.62$ ). There was no significant association between the use of OS as a primary endpoint, use of double-blinding or trial sponsor and PPT adjustment (table 1).

PPTs may be more readily available and/or more effective after the first instance of progression compared with subsequent progressions. Thus, a sensitivity analysis was performed on 187 trials evaluating localised, first-line metastatic and first-line haematological settings (excluding second-line or later metastatic or haematological trials). The characteristics of these 187 trials were similar to the full analysis (online supplemental table 2). Only 8% of trials (N=15) attempted to account for PPT, which was also consistent with the primary analysis. Trials with cross-over were similarly associated with increased odds of PPT analysis (21% vs 5%, OR 5.41, 95% CI 1.81 to 16.55,  $p=0.002$ ), although the association of trials with PFS-OS discordance was reduced (12% vs 6%, OR 2.06, 95% CI 0.66 to 6.07,  $p=0.19$ ).

A second sensitivity analysis was performed evaluating 168 trials powered for OS differences, with OS as the primary or coprimary endpoint (ie, excluding trials where OS was a secondary endpoint). Trial characteristics were largely similar to the overall analysis, except for a decreased representation of cross-over studies (online supplemental table 3). In this subset of trials, only 11% of trials evaluated PPT confounding (N=18), reproducing the overall findings. However, the association of trial covariates and PPT confounding analyses was decreased, including for cross-over trials (21% vs 10%, OR 2.53, 95% CI 0.53 to 9.23,  $p=0.19$ ) and trials with PFS-OS discordance (20% vs 7%, OR 1.13, 95% CI 0.37 to 3.10,  $p=0.82$ ).

## DISCUSSION

In this large-scale analysis of contemporary two-arm superiority-design phase III oncology trials with matured OS endpoints, 88% of trials did not attempt to account for confounding introduced by PPTs. These findings were consistent across sensitivity analyses for trials evaluating definitive/first-line therapy and trials with a primary (including coprimary) endpoint of OS. Taken together, this study suggests that the effects of PPTs on estimates of the treatment effects on OS may be under-evaluated in phase III oncology studies.

**Table 1** Characteristics of trials accounting for postprogression therapy (PPT) bias on overall survival estimates

| Characteristic                                      | PPT analysis  |            |           | OR    | 95% CI        | P value |
|---|---------------|------------|-----------|-------|---------------|---------|
|   | All trials, N | Yes, N (%) | No, N (%) |       |               |         |
| Total trials  | 334           | 41 (12)    | 293 (88)  |       |               |         |
| Disease stage                                       |               |            |           |       |               |         |
| Solid M0  | 65            | 1 (2)      | 64 (98)   | 0.10  | 0.01 to 0.49  | 0.03    |
| Solid M1  | 218           | 29 (13)    | 189 (87)  | Ref   |               |         |
| Haematologic  | 51            | 11 (22)    | 40 (78)   | 1.79  | 0.80 to 3.80  | 0.14    |
| Disease site  |               |            |           |       |               |         |
| Breast  | 50            | 8 (16)     | 42 (84)   | 1.14  | 0.39 to 3.37  | 0.81    |
| Gastrointestinal                                    | 62            | 5 (8)      | 57 (92)   | 0.53  | 0.15 to 1.68  | 0.29    |
| Genitourinary                                       | 43            | 5 (12)     | 38 (88)   | 0.79  | 0.22 to 2.56  | 0.70    |
| Haematologic  | 51            | 11 (22)    | 40 (78)   | 1.65  | 0.61 to 4.64  | 0.33    |
| Thoracic  | 72            | 4 (6)      | 68 (94)   | 0.35  | 0.09 to 1.19  | 0.10    |
| Other*  | 56            | 8 (14)     | 49 (86)   | Ref   |               |         |
| Treatment type                                      |               |            |           |       |               |         |
| Systemic therapy                                    | 322           | 41 (13)    | 281 (87)  |       |               |         |
| Local therapy                                       | 12            | 0 (0)      | 12 (100)  | Ref   |               |         |
| Cooperative group study                             |               |            |           |       |               |         |
| Yes   | 86            | 6 (7)      | 80 (93)   | 0.46  | 0.17 to 1.05  | 0.09    |
| No  | 248           | 35 (14)    | 213 (86)  | Ref   |               |         |
| Industry sponsored                                  |               |            |           |       |               |         |
| Yes   | 284           | 38 (13)    | 246 (87)  | 2.42  | 0.83 to 10.30 | 0.15    |
| No  | 50            | 3 (6)      | 47 (94)   |       |               |         |
| Enrolled patients†                                  | NA            | NA         | NA        | 0.999 | 0.998 to 1.00 | 0.04    |
| Publication date‡                                   | NA            | NA         | NA        | 0.996 | 0.90 to 1.11  | 0.95    |
| Cross-over (allowed or required)                    |               |            |           |       |               |         |
| Yes   | 49            | 16 (33)    | 33 (67)   | 5.04  | 2.42 to 1.38  | <0.0001 |
| No  | 285           | 25 (9)     | 260 (91)  | Ref   |               |         |
| Blinding  |               |            |           |       |               |         |
| Double-blind  | 128           | 15 (12)    | 113 (88)  | 0.92  | 0.46 to 1.79  | 0.81    |
| Open-label  | 206           | 26 (13)    | 180 (87)  | Ref   |               |         |
| Overall survival primary or coprimary endpoint      |               |            |           |       |               |         |
| Yes   | 168           | 18 (11)    | 150 (89)  | 0.75  | 0.38 to 1.44  | 0.38    |
| No  | 166           | 23 (14)    | 143 (86)  | Ref   |               |         |
| Surrogate survival and overall survival discordance |               |            |           |       |               |         |
| Yes   | 107           | 20 (19)    | 87 (81)   | 2.26  | 1.16 to 4.38  | 0.02    |
| No  | 227           | 21 (9)     | 206 (91)  | Ref   |               |         |

Associations are evaluated by unadjusted binary logistic regression.

\*Other disease sites included central nervous system, endocrine, gynaecological, head and neck, paediatric, sarcoma, and skin.

†The median trial enrolment was 585 patients (IQR 362–896).

‡Trends over time are reported in online supplemental figure 2.

§No local therapy trial accounted for PPT, so an odds ratio was not computed.

The implications of this study may impact the interpretations of OS from current phase III oncology trials. In clinical scenarios with salvage PPTs, rather than considering

OS as the time from randomisation to death, OS may be better conceptualised as a composite measure summing multiple landmark times: the time from randomisation



**Table 2** Description of methods used to account for postprogression therapy (PPT) effects on overall survival

| Approaches used to account for PPT                                | Trials* |
|---|---------|
|   | N (%)   |
| Total trials evaluated  | 334     |
| Use of any method to account for PPT                              | 41 (12) |
| Multiple methods to account for PPT                               | 9 (3)   |
| Simple censoring at start of new therapy                          | 19 (5)  |
| Post-progression survival analysis                                | 18 (5)  |
| Rank-preserving structural failure time                           | 7 (2)   |
| Inverse probability of censor weighting                           | 5 (1)   |
| Iterative parameter estimation                                    | 2 (1)   |
| Two-stage method  | 1 (0.3) |
| Marginal structural model   | 1 (0.3) |
| Analysis with data locked before and after cross-over was allowed | 1 (0.3) |

\*The number of individual approaches used to account for PPT does not add to 41 (total trials accounting for PPT) as 9 trials used multiple methods to account for PPT.

to progression plus the subsequent time(s) on PPT(s) to death (ie, postprogression survival). Importantly, postprogression survival is affected by both the efficacy of PPTs and potential confounding bias, similar to the bias in observational studies (figure 1).<sup>11</sup> Time on PPTs is thus highly relevant to OS in most oncology trials. Time on PPTs may be especially pertinent for trials evaluating locoregional malignancy or trials evaluating first-line therapy, where effective salvage therapies may result in longer postprogression survival times compared with rapidly fatal cancers.<sup>23–26</sup> Not accounting for PPT confounding may result in underestimation or overestimation of OS differences between arms of a randomised controlled trial with conventional comparisons of the upfront randomisation.<sup>14 27</sup> If patients in the control arm cross-over to the more effective experimental therapy following progression, the impact of the experimental therapy on OS may be underestimated due to overestimation of the OS of the control arm.<sup>22</sup> For example, in the PROFILE 1014 trial, the observed OS benefit was weak (HR 0.76, 95% CI 0.5 to 1.05,  $p=0.10$ ), with 84% of patients in the control arm crossing over to crizotinib (the experimental therapy) after progression.<sup>28</sup> However, a prespecified rank-preserving structural failure time model to account for cross-over effects revealed a stronger OS signal in favour of the experimental arm (HR 0.35, 95% bootstrap CI 0.08 to 0.72). Conversely, if patients on the control arm systematically receive less effective PPTs, the control arm may have underestimated OS, resulting in an overestimation of the OS benefits of the experimental arm.<sup>29</sup> For example, the ADAURA trial found that adjuvant osimertinib was associated with better OS compared with placebo after resection of non-small cell lung cancer.<sup>30</sup> However, only 38.5% of eligible patients in the placebo

arm received osimertinib at the time of progression, despite the established efficacy of osimertinib for first-line metastatic disease, which may have biased OS comparisons in favour of the experimental arm.<sup>31 32</sup> Other examples have been reviewed in detail elsewhere, such as the ANNOUNCE trial.<sup>11 33</sup> Due to the biases introduced by uncontrolled PPTs, some contend that cross-over should be protocol driven—either required (when efficacy has been demonstrated in later lines) or prohibited (when efficacy has not been demonstrated in later lines)—rather than simply ‘allowed’ and thus subject to confounding.<sup>31</sup> In this context, it is interesting to note that more trials in our study allowed cross-over rather than required or prohibited cross-over, suggesting there is considerable opportunity for improvement in this aspect of late-phase trial design. A recent study corroborates this notion, finding that PPTs in most phase III trials were substandard for the relevant clinical scenario.<sup>34</sup> Finally, there are important ethical considerations for ensuring trial participants receive effective standard-of-care PPTs, particularly in low-income and middle-income countries with prominent barriers to access to care.<sup>35–38</sup>

While the overall rate of trials attempting to account for PPT confounding was low in this study, several factors were associated with an increased likelihood of PPT analysis. In particular, trials with cross-over were more likely to perform PPT analyses. Because cross-over is a component of trial design, we expected that trials with cross-over would be more likely to perform preplanned, rather than post hoc, PPT analyses. We did not find this correlation among the subset of cross-over trials, although this result may be partly attributable to the small number of cross-over trials in general, which increases the risk of type II error. Trials with a discrepancy between the surrogate survival findings and the OS findings were also more likely to attempt PPT confounding adjustments. This may be partly related to the increased attention given to PPTs in cross-over trials, as we also found that cross-over trials were more likely to be associated with discordant surrogate survival and OS results (which itself is unsurprising in the setting of superior experimental therapy). We also speculated that the correlation between discordant trials and PPT analyses may be related to trialists exploring explanations for their findings after observing discordance; however, trials with discrepant findings did not seem to be enriched for post hoc PPT analyses, although the interpretation of this regression is also limited by type II error risk. Ultimately, however, PPT adjustments should be considered for phase III trials evaluating OS where PPTs are expected at the time of trial design.

While simple censoring was the most common strategy to account for PPT effects, this approach may be particularly prone to selection bias and is generally not recommended.<sup>16 18 22</sup> The key advantages and disadvantages of several selected statistical and trial design approaches are described in table 3. No approach has a universal advantage over the others, and the choice of adjustment should be considered based on each unique clinical trial scenario

**Table 3** Potential approaches towards addressing postprogression therapy (PPT) confounding of overall survival

| Potential approaches to PPT adjustment                                 | Main advantage  | Main disadvantage  |
|--|---|--|
| Simple censoring at next-line therapy                                  | Simple, easily applied to trials without prespecified PPT.  | Loss of survival information following progression; susceptible to informative censoring.                            |
| Two-stage estimation   | Does not hold common treatment effect assumption.   | Does not adjust for time-dependent confounding following progression.  |
| Inverse probability of censoring weighting                             | Adjusted estimates of the effect of PPT on survival accounting for time-dependent confounding. Adjusts for censoring biases.                | Assumes no unmeasured confounders. More suitable for observational studies.  |
| Iterative parameter estimation/rank-preserving structural failure time | Adjusted estimates of the effect of PPT on survival accounting for time-dependent confounding. Accounts for treatment switching/cross-over. | Assumes that the PPT effect is the same for both treatment arms.   |
| Sequentially multiple randomised assignment trials                     | Least susceptible to confounding. Comparative overall survival estimates are licensed by virtue of rerandomisation                          | Resource-intensive. Requires longer protocol follow-up. Difficult to incorporate new PPTs developed during the trial |

at the time of trial design. Because of the advantages and disadvantages of each statistical approach, sensitivity analyses incorporating multiple methods, with appropriate multiplicity of testing control, may be reasonable. For example, KEYNOTE-024 adjusted for cross-over effects with three models: the simplified two-stage method, rank-preserving structural failure time, and inverse probability of censoring weighting, each of which showed a larger effect size than the unadjusted analysis.<sup>39</sup> BREAK-3 investigators showed that a rank-preserving structural failure time model and an iterative parameter estimation model strengthened the estimated OS treatment effect in a trial where more than half of patients in the control arm crossed over to the experimental therapy at progression.<sup>40</sup> Bayesian approaches have also been suggested and may yield more efficient inferences compared with frequentist approaches.<sup>41–42</sup> Ultimately, however, statistical approaches may not be able to adequately facilitate interpretation of trials with flawed designs and substantial PPT confounding.<sup>11</sup> Sequentially multiple randomised assignment trials (SMART) randomise patients at the original treatment assignment and again at the time of progression to determine PPT allocation.<sup>43–45</sup> This second randomisation to determine PPT allocation at progression removes systematic confounding from PPT selection, thus providing statistical license to compare dynamic treatment regime pairs by virtue of randomisation. Particularly in an era of increasing PPT effectiveness with targeted therapies and immunotherapies, trialists should consider SMART designs to control for systematic PPT confounding.

There are several limitations to consider for this study. Trials were identified from the US-based registry ClinicalTrials.gov, so the findings may not be extrapolatable to global trials or trials conducted prior to the development

of ClinicalTrials.gov in 2000. Data were abstracted by a single reviewer. While all manuscripts related to each trial as well as all online supplemental materials were evaluated for PPT analyses, trials that conducted but did not report these analyses due to presentation bias or other reasons may have led this study to underestimate the extent to which phase III trials have approached PPT analyses. Correlations with trial-level factors and PPT confounding analysis were not adjusted due to a small number of trials with PPT analysis and must be interpreted carefully.

In summary, a large-scale analysis of phase III oncology trials suggests that OS analyses accounting for the effects of PPTs are rare. To clarify the treatment effect estimates on OS, PPT effects should be considered from trial design through analysis, interpretation, and reporting of results. Readers, journal editors, and regulatory agencies should weigh the impacts of PPT confounding in the interpretation of the treatment effect on OS.

#### Author affiliations

<sup>1</sup>Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

<sup>2</sup>Department of Genitourinary Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

<sup>3</sup>Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

<sup>4</sup>Department of Radiation Oncology and Molecular Radiation Sciences, Johns Hopkins Medicine, Baltimore, Maryland, USA

<sup>5</sup>Department of Radiation Oncology, Stanford University, Stanford, California, USA

<sup>6</sup>Department of Gastrointestinal Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

<sup>7</sup>Department of Radiation Oncology and Applied Sciences, Dartmouth College Geisel School of Medicine, Lebanon, New Hampshire, USA

<sup>8</sup>Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

<sup>9</sup>Insitro, South San Francisco, California, USA

<sup>10</sup>Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

X Alexander D Sherry @AlexSherryMD

**Acknowledgements** We acknowledge the late Dr Cullen Taniguchi, former Associate Professor in our department, for sharing his insights through numerous helpful discussions prior to the completion of this study.

**Contributors** ADS, PM, TAL and EBL conceptualised the study. ADS, TAL, JAJ, RK, GSK and EBL curated the data. ADS did the formal analysis. EBL acquired funding. EBL supervised the study and provided study resources. ADS, JAJ and ZRM prepared the visualisations. All authors contributed to the interpretation of the data. ADS wrote the initial draft, and all authors reviewed and revised the manuscript. ADS, as the guarantor, holds full responsibility for the overall content, had full access to all the data, and maintained control over the decision to publish.

**Funding** Supported in part by grant P30 CA016672 from the National Cancer Institute, National Institutes of Health (MD Anderson), the Sabin Family Fellowship Foundation (EBL) and the Fund for Innovation in Cancer Informatics (EBL).

**Competing interests** PM reports honoraria for scientific advisory board membership for Mirati Therapeutics, Bristol-Myers Squibb and Exelixis; consulting fees from Axiom Healthcare; non-branded educational programmes supported by Exelixis and Pfizer; leadership or fiduciary roles as a Medical Steering Committee Member for the Kidney Cancer Association and a Kidney Cancer Scientific Advisory Board Member for KCCure; and research funding from Takeda, Bristol-Myers Squibb, Mirati Therapeutics, and Gateway for Cancer Research (all unrelated to this manuscript's content). ZRM reported employment at Insitro (unrelated to this manuscript's content). DF receives unrelated funding and salary support from: NIH NIBIB Research Education Programs for Residents and Clinical Fellows Grant (R25EB025787-01); NIDCR Academic Industrial Partnership Grant (R01DE028290); NCI Parent Research Project Grant (R01CA258827) an NIH/NCI Cancer Center Support Grant (CCSG); and an NSF Division of Civil, Mechanical and Manufacturing Innovation (CMMI) grant (NSF 1933369). DF has received direct industry grant support, honoraria, and travel funding from Elekta AB unrelated to this project and receives direct infrastructure support from the multidisciplinary Radiation Oncology/ Cancer Imaging Program (P30CA016672-44) of the MD Anderson Cancer Center Support (Core) Grant (P30CA016672) and the MD Anderson Programme in Image-guided Cancer Therapy. CRT receives a stipend from the AMA as deputy editor of JAMA Oncology. No other authors report any conflicts of interest.

**Patient and public involvement** Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication** Not applicable.

**Ethics approval** Institutional review board approval was not required due to public availability of the trial-level data.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available on reasonable request. Research data are stored in an institutional repository and will be shared on reasonable request to the corresponding author up to 1 year following publication of the manuscript.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Alexander D Sherry <http://orcid.org/0000-0001-5115-1691>

Esther J Beck <http://orcid.org/0009-0009-6821-8966>

## REFERENCES

- 1 Walia A, Tuia J, Prasad V. Progression-free survival, disease-free survival and other composite end points in oncology: improved reporting is needed. *Nat Rev Clin Oncol* 2023;20:885–95.
- 2 Lin TA, Sherry AD, Ludmir EB. Challenges, complexities, and considerations in the design and interpretation of late-phase oncology trials. *Semin Radiat Oncol* 2023;33:429–37.
- 3 Pasalic D, McGinnis GJ, Fuller CD, et al. Progression-free survival is a suboptimal Predictor for overall survival among metastatic solid tumour clinical trials. *Eur J Cancer* 2020;136:176–85.
- 4 Belin L, Tan A, De Ruyck Y, et al. Progression-free survival as a Surrogate for overall survival in oncology trials: a methodological systematic review. *Br J Cancer* 2020;122:1707–14.
- 5 Prasad V, Kim C, Burotto M, et al. The strength of association between Surrogate end points and survival in oncology: A systematic review of trial-level meta-analyses. *JAMA Intern Med* 2015;175:1389–98.
- 6 Haslam A, Hey SP, Gill J, et al. A systematic review of trial-level meta-analyses measuring the strength of association between Surrogate end-points and overall survival in oncology. *Eur J Cancer* 2019;106:196–211.
- 7 Booth CM, Eisenhauer EA. Progression-free survival: meaningful or simply measurable. *J Clin Oncol* 2012;30:1030–3.
- 8 Walia A, Haslam A, Prasad V. FDA validation of Surrogate endpoints in oncology: 2005–2022. *J Cancer Policy* 2022;34:100364.
- 9 Kemp R, Prasad V. Surrogate endpoints in oncology: when are they acceptable for regulatory and clinical decisions, and are they currently Overused? *BMC Med* 2017;15:134.
- 10 Booth CM, Eisenhauer EA, Gyawali B, et al. Progression-free survival should not be used as a primary end point for registration of anticancer drugs. *JCO* 2023;41:4968–72.
- 11 Msaouel P, Lee J, Thall PF. Interpreting randomized controlled trials. *Cancers (Basel)* 2023;15:4674.
- 12 Korn EL, Freidlin B, Abrams JS. Overall survival as the outcome for randomized clinical trials with effective subsequent therapies. *J Clin Oncol* 2011;29:2439–42.
- 13 Stewart DJ. Before we throw out progression-free survival as a valid end point. *J Clin Oncol* 2012;30:3426–7.
- 14 Valenti V, Jiménez-Fonseca P, Msaouel P, et al. The misleading effect of treatment crossover in randomized trials of therapies with marginal treatment benefit. *Cancer Invest* 2022;40:184–8.
- 15 Fojo T, Amiri-Kordestani L, Bates SE. Potential pitfalls of crossover and thoughts on Iniparib in triple-negative breast cancer. *J Natl Cancer Inst* 2011;103:1738–40.
- 16 Watkins C, Huang X, Latimer N, et al. Adjusting overall survival for treatment switches: commonly used methods and practical application. *Pharm Stat* 2013;12:348–57.
- 17 Haslam A, Gill J, Crain T, et al. The frequency of medical reversals in a cross-sectional analysis of high-impact oncology journals, 2009–2018. *BMC Cancer* 2021;21:889.
- 18 Latimer NR, Abrams KR, Lambert PC, et al. Adjusting survival time estimates to account for treatment switching in randomized controlled trials—an economic evaluation context: methods, limitations, and recommendations. *Med Decis Making* 2014;34:387–402.
- 19 Prasad V, Vandross A. Failing to improve overall survival because post-protocol survival is long: fact, myth, excuse or improper study design? *J Cancer Res Clin Oncol* 2014;140:521–4.
- 20 Sherry AD, Corrigan KL, Kouzy R, et al. Prevalence, trends, and characteristics of trials investigating local therapy in contemporary phase 3 clinical cancer research. *Cancer* 2023;129:3430–8.
- 21 von Elm E, Altman DG, Egger M, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *The Lancet* 2007;370:1453–7.
- 22 Morden JP, Lambert PC, Latimer N, et al. Assessing methods for dealing with treatment switching in randomised controlled trials: a simulation study. *BMC Med Res Methodol* 2011;11:4.
- 23 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* 2010;28:1958–62.
- 24 Petrelli F, Barni S. Correlation of progression-free and post-progression survival with overall survival in advanced colorectal cancer. *Ann Oncol* 2013;24:186–92.
- 25 Meirson T, Goldstein DA, Gyawali B, et al. 25. Meirson T, Goldstein DA, Gyawali B, Tannock IF. review of the monarchE trial suggests no evidence to support use of adjuvant Abemaciclib in women with breast cancer. *The Lancet Oncology* 2023;24:589–93.
- 26 Tannock IF, Goldstein DA, Ofer J, et al. Evaluating trials of adjuvant therapy: is there benefit for people with Resected renal cancer? *J Clin Oncol* 2023;41:2713–7.



- 27 Isbary G, Staab TR, Amelung VE, *et al*. Effect of crossover in oncology clinical trials on evidence levels in early benefit assessment in Germany. *Value Health* 2018;21:698–706.
- 28 Solomon BJ, Kim D-W, Wu Y-L, *et al*. Final overall survival analysis from a study comparing first-line Crizotinib versus chemotherapy in ALK-Mutation-positive non-small-cell lung cancer. *J Clin Oncol* 2018;36:2251–8.
- 29 Haslam A, Prasad V. When is crossover desirable in cancer drug trials and when is it problematic. *Ann Oncol* 2018;29:1079–81.
- 30 Tsuboi M, Herbst RS, John T, *et al*. Overall survival with Osimertinib in Resected EGFR-Mutated NSCLC. *N Engl J Med* 2023;389:137–47.
- 31 Gyawali B. Problematic Crossovers in cancer drug trials. *Nat Rev Clin Oncol* 2023;20:815–6.
- 32 Tsuboi M, Herbst RS, Wu Y-L. Osimertinib in Resected EGFR-Mutated NSCLC. *N Engl J Med* 2023;389:1342.
- 33 Tap WD, Wagner AJ, Schöffski P, *et al*. Effect of doxorubicin plus Olaratumab vs doxorubicin plus placebo on survival in patients with advanced soft tissue Sarcomas: the ANNOUNCE randomized clinical trial. *JAMA* 2020;323:1266–76.
- 34 Olivier T, Haslam A, Prasad V. Post-progression treatment in cancer randomized trials: a cross-sectional study of trials leading to FDA approval and published trials between 2018 and 2020. *BMC Cancer* 2023;23:448.
- 35 Jalali R, Nogueira-Rodrigues A, Das A, *et al*. Drug development in Low- and middle-income countries: opportunity or exploitation? *Am Soc Clin Oncol Educ Book* 2022;42:1–8.
- 36 Ruff P, Al-Sukhun S, Blanchard C, *et al*. Access to cancer Therapeutics in Low- and middle-income countries. *Am Soc Clin Oncol Educ Book* 2016;35:58–65.
- 37 Pramesh CS, Badwe RA, Bhoo-Pathy N, *et al*. Priorities for cancer research in Low- and middle-income countries: a global perspective. *Nat Med* 2022;28:649–57.
- 38 Rubagumya F, Hopman WM, Gyawali B, *et al*. Participation of lower and upper middle-income countries in clinical trials led by high-income countries. *JAMA Netw Open* 2022;5:e2227252.
- 39 Reck M, Rodríguez-Abreu D, Robinson AG, *et al*. Updated analysis of KEYNOTE-024: Pembrolizumab versus platinum-based chemotherapy for advanced non-small-cell lung cancer with PD-L1 tumor proportion score of 50% or greater. *JCO* 2019;37:537–46.
- 40 Latimer NR, Abrams KR, Amonkar MM, *et al*. Adjusting for the confounding effects of treatment switching—the BREAK-3 trial: Dabrafenib versus Dacarbazine. *Oncologist* 2015;20:798–805.
- 41 Liu F, Li Q. A Bayesian model for joint analysis of multivariate repeated measures and time to event data in crossover trials. *Stat Methods Med Res* 2016;25:2180–92.
- 42 Xu Y, Müller P, Wahed AS, *et al*. Bayesian Nonparametric estimation for dynamic treatment regimes with sequential transition times. *J Am Stat Assoc* 2016;111:921–35.
- 43 Murphy SA. An experimental design for the development of adaptive treatment strategies. *Stat Med* 2005;24:1455–81.
- 44 Wang L, Rotnitzky A, Lin X, *et al*. Evaluation of viable dynamic treatment regimes in a Sequentially randomized trial of advanced prostate cancer. *J Am Stat Assoc* 2012;107:493–508.
- 45 Lorenzoni G, Petracchi E, Scarpi E, *et al*. Use of sequential multiple assignment randomized trials (smarts) in oncology: systematic review of published studies. *Br J Cancer* 2023;128:1177–88.