



**EDITORIAL**

# Endpoint surrogacy in oncology Phase 3 randomised controlled trials

Endpoint surrogacy is an important concept in oncology trials. Using a surrogate endpoint like progression-free survival as the primary endpoint—instead of overall survival—would lead to a potential faster drug approval and therefore more cancer patients with an earlier opportunity to receive the newly approved drugs.

*British Journal of Cancer* (2020) 123:333–334; <https://doi.org/10.1038/s41416-020-0896-5>

**MAIN**

Randomised controlled trial (RCT) is a highly rigorous study design. In oncology, it determines whether a new cancer drug or a new indication of an existing cancer drug has better clinical benefit compared with the standard of care or placebo, and offers the empirical evidence for a regulation authority like the European Medicines Agency or the US Food and Drug Administration (FDA) to approve the new drug or the new indication.<sup>1</sup>

In RCTs, selecting valid endpoints that capture clinical benefit is very important. For measuring efficacy, overall survival (OS) and progression-free survival (PFS) are the most common endpoints, defined as “the time from randomisation until death from any cause” and “the time from randomisation until objective tumour progression or death”, respectively.<sup>2</sup> In Phase 3 trials, OS is the standard endpoint to establish clinical benefit; however, it requires a longer period of follow-up, often relatively larger number of events, and therefore a higher cost compared with other endpoints like PFS.

To address these limitations of OS, the concept of surrogate endpoint was born. According to the US Federal Food, Drug and Cosmetic Act, the term “surrogate endpoint” means a marker (e.g. laboratory measurement, radiographic image) that “is not itself a direct measurement of clinical benefit, and—(A) is known to predict clinical benefit and could be used to support traditional approval of a drug or biological product; or (B) is reasonably likely to predict clinical benefit and could be used to support the accelerated approval of a drug or biological product”. For example, PFS could be the surrogate endpoint to predict the result of OS in Phase 3 trials.<sup>2</sup> With a shorter period of follow-up, often smaller number of events, and therefore a trial with less cost, using a surrogate endpoint as the primary endpoint—instead of OS—would lead to a potential faster drug approval and therefore more cancer patients with an earlier opportunity to receive the newly approved drugs.<sup>3</sup>

Prior to the use of a surrogate endpoint in clinical trials to establish the efficacy of the new drugs, a question to be answered is: how can we know how accurate the prediction is? Conducting validation studies is an answer. To date, numerous validation studies have been conducted. Recently in the *British Journal of Cancer*, Belin et al.<sup>4</sup> published a complete overview of validation studies on PFS for OS, summarising the validity of PFS as a

surrogate endpoint for OS, and describing the characteristics of the methodologies used in those validation studies. Two main findings have been reported by the authors. First, only half (52%) of all 91 validation studies (involving 24 cancer localisations) concluded on the validity of PFS for OS; of those studies, only half (51%; 24 studies) indicated a good endpoint surrogacy (criteria used in this review: trial-level  $R^2 \geq 0.6$ ). Second, most of those validation studies utilised a meta-analytic approach per recommended, but a remarkable heterogeneity in methods and reporting was noticed. Specifically, the methods were heterogeneous in the evaluation at patient or trial level, and in the trial-level evaluation on the aggregate measures based on one arm (median PFS at a clinically meaningful timepoint), or the treatment effects based on two arms (hazard ratio, difference of median, ratio of median).<sup>4</sup>

These findings have profound implication to design future clinical trials. The first main finding emphasises that endpoint surrogacy of PFS for OS is not universally valid, given that all included validation studies were conducted based on different interventions and different indications (e.g. treatment regimen, treatment line, cancer type, cancer stage). For example, there is currently a lack of strong evidence supporting the validity of PFS for OS in immunotherapies.<sup>4,5</sup> Therefore, we encourage future validation studies that should have focus(es) on specific intervention and specific indication. Through such an approach, one can assess how valid a surrogate endpoint is to predict clinical benefit according to each specific intervention and specific indication. To better accomplish validation studies for specific intervention and specific indication, we need a large pool of clinical trials data in which both OS and the surrogate endpoints were captured. An easy access to the data from completed clinical trials from many pharmaceutical industry or government sponsored trials for research use is the key to such success. We also encourage conducting more validation studies focusing on not only PFS, the most common surrogate endpoint for OS, but also other potential endpoints like metastasis-free survival<sup>6</sup> and milestone survival.<sup>7</sup> Furthermore, these potential surrogate endpoints have focus(es) on specific intervention and specific indication.

Corresponding to the second main finding, we agree with Belin et al.<sup>4</sup> that developing and applying recommendations on

the methodology and reporting of validation analyses is important. Related recommendations on reporting have been well established in the ReSEEM guidelines,<sup>8</sup> which provide useful suggestions on methodology to improve consistency in validation studies. Consensus is still needed with regard to more nuanced aspects of methodology, especially at the policy level. For example, there is no consensus on the extent to the prediction indicating a valid endpoint surrogacy (e.g. trial-level or patient-level  $R^2$ ?  $R^2 \geq 0$ ?). Also, no consensus is made on the validation evaluation that whether and how an adjustment by trial characteristics should be applied, and that whether the aggregate measures or the treatment effects should be used. With respect to which aggregate measures or treatment effects should be used, however, it may be too early to develop a universal recommendation, given the current debate between the use of traditional aggregate measures (e.g. median survival) or treatment effects (e.g. hazard ratio) compared with the aggregate measures or treatment effects based on newer endpoints (e.g. milestone survival<sup>7</sup>) and/or newer statistical methods using the landmark analysis<sup>9</sup> or restricted mean survival time.<sup>10</sup> Therefore, we suggest current validation studies to consider any possible aggregated measures or treatment effects in their evaluation of validity.

#### AUTHOR CONTRIBUTIONS

J.Z. wrote the first draft, which was critically revised by M.R.P., X.W., J.L., H.P., R.C.B., G.A.C., W.L. and J.H. All authors approved the final paper.

#### ADDITIONAL INFORMATION

**Ethics approval and consent to participate** Not applicable.

**Consent to publish** Not applicable.

**Data availability** Not applicable.

**Competing interests** The authors declare no competing interests.

**Funding information** J.L. and G.A.C. acknowledge the support for the Siteman Cancer Center from the National Institutes of Health (NIH) Grant P30 CA 091842.

**Note** This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution 4.0 International (CC BY 4.0).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Jianrong Zhang<sup>1</sup>, Meagan R. Pilar<sup>2</sup>, Xiaofei Wang<sup>3</sup>, Jingxia Liu<sup>4</sup>, Herbert Pang<sup>5</sup>, Ross C. Brownson<sup>2,4,6</sup>, Graham A. Colditz<sup>4,6</sup>, Wenhua Liang<sup>7</sup> and Jianxing He<sup>7</sup>

<sup>1</sup>Department of General Practice, Melbourne Medical School; Cancer in Primary Care Research Group, Primary Care Collaborative Cancer Clinical Trials Group (PC4), Centre for Cancer Research, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne; Victorian Comprehensive Cancer Centre, Melbourne, VIC, Australia; <sup>2</sup>Prevention Research Center, Brown School at Washington University in St. Louis, St. Louis, MO, USA; <sup>3</sup>Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, NC, USA; <sup>4</sup>Department of Surgery, Division of Public Health Sciences, Washington University School of Medicine, St. Louis, MO, USA; <sup>5</sup>School of Public Health, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong, China; <sup>6</sup>Alvin J. Siteman Cancer Center, St. Louis, MO, USA and <sup>7</sup>Department of Thoracic Surgery and Oncology, The First Affiliated Hospital of Guangzhou Medical University, State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, Guangzhou Institute of Respiratory Disease, Guangzhou, Guangdong, China  
Correspondence: Jianrong Zhang (jianrong.zhang@unimelb.edu.au)

#### REFERENCES

1. Tannock, I. F., Amir, E., Booth, C. M., Niraula, S., Ocana, A., Seruga, B. et al. Relevance of randomised controlled trials in oncology. *Lancet Oncol.* **17**, e560–e567 (2016).
2. US Food and Drug Administration (FDA). Guidance for industry: clinical trial endpoints for the approval of cancer drugs and biologics. *US Food and Drug Administration*. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-trial-endpoints-approval-cancer-drugs-and-biologics> (2018).
3. Chen, E. Y., Joshi, S. K., Tran, A. & Prasad, V. Estimation of study time reduction using surrogate end points rather than overall survival in oncology clinical trials. *JAMA Intern. Med.* **179**, 642–647 (2019).
4. Belin, L., Tan, A., De Rycke, Y. & Dechartres A. Progression-free survival as a surrogate for overall survival in oncology trials: a methodological systematic review. *Br. J. Cancer*. <https://doi.org/10.1038/s41416-020-0805-y> (2020).
5. Zhang, J., Liang, W., Liang, H., Wang, X. & He, J. Endpoint surrogacy in oncological randomized controlled trials with immunotherapies: a systematic review of trial-level and arm-level meta-analyses. *Ann. Transl. Med.* **7**, 244 (2019).
6. Xie, W., Regan, M. M., Buyse, M., Halabi, S., Kantoff, P. W., Sartor, O. et al. Metastasis-free survival is a strong surrogate of overall survival in localized prostate cancer. *J. Clin. Oncol.* **35**, 3097–3104 (2017).
7. Blumenthal, G. M., Zhang, L., Zhang, H., Kazandjian, D., Khozin, S., Tang, S. et al. Milestone analyses of immune checkpoint inhibitors, targeted therapy, and conventional therapy in metastatic non-small cell lung cancer trials: a meta-analysis. *JAMA Oncol.* **3**, e171029 (2017).
8. Xie, W., Halabi, S., Tierney, J. F., Sydes, M. R., Collette, L., Dignam, J. J. et al. A systematic review and recommendation for reporting of surrogate endpoint evaluation using meta-analyses. *JNCI Cancer Spectr.* **3**, pkz002 (2019).
9. Ascierto, P. A. & Long, G. V. Progression-free survival landmark analysis: a critical endpoint in melanoma clinical trials. *Lancet Oncol.* **17**, 1037–1039 (2016).
10. Ben-Aharon, O., Magnezi, R., Leshno, M. & Goldstein, D. A. Median survival or mean survival: Which measure is the most appropriate for patients, physicians, and policymakers? *Oncologist* **24**, 1469–1478 (2019).