Additional Considerations for Single-Arm Trials to Support Accelerated Approval of Oncology Drugs

FEINAN LU, TAO WANG, YING LU, AND JIE CHEN*

We appreciate the comments by Cao and Pan [1] who highlighted the major points discussed in our paper, "Considerations for Single-Arm Trials to Support Accelerated Approval of Oncology Drugs" [2], and who also expressed concerns about potential biases and other related issues in single-arm trials (SATs). Specifically, Cao and Pan pointed out that, unlike randomized controlled trials (RCTs), SATs are inherently associated with methodological limitations—such as the lack of a control arm—which can complicate the interpretation of treatment effects. They briefly discussed considerations for analytical methods (e.g., propensity score matching), additional endpoints (e.g., patient-reported outcomes [PROs], quality of life [QoL]), subgroup analyses, and control of false positives. They also addressed the use of external controls, individual patient data (IPD), issues related to statistical power and small sample sizes, and various potential sources of bias (e.g., selection bias, information bias) in SATs. Ultimately, they argued that SATs, as a component of a broader evidentiary framework, should incorporate historical controls, real-world evidence (RWE), and confirmatory post-marketing studies [1].

While we agree with most of the points raised by Cao and Pan [1], we would like to emphasize that SATs may only be appropriate in specific clinical contexts—such as rare and/or life-threatening cancers with no efficacious treatment options—where randomized controlled trials (RCTs) are either (1) unethical or infeasible, as outlined as "necessary conditions" for SATs in our paper [2], or (2) unnecessary in cases where outcomes under control conditions are well understood (analogous to testing parachutes [3]), such as the cell therapy approved for synovial sarcoma [4]. Relevant regulatory guidance documents [5, 6, 7] also support the appropriate use of SATs in the context of accelerated approval (AA). The desirable conditions are presented in Sections 3 and 4 of Lu et al. [2], which—together with the necessary conditions in Section 2— constitute a comprehensive framework of prerequisites for the use of SATs to support AA. To address the concerns and critical points raised by Cao and Pan [1], we provide a brief discussion of additional considerations for using SATs in the regulatory approval of oncology drugs.

As discussed in several sections of our paper [2] and in the comments by Cao and Pan [1], a major concern with SATs is the lack of an internal comparator arm, which can introduce biases in comparative effect estimates. The reflection paper by the European Medicines Agency (EMA) [7] summarizes various sources of bias and corresponding mitigation strategies that can be applied during the design, conduct, analysis, and reporting of an SAT. It also acknowledges that these strategies may not fully eliminate bias and that demonstrating unbiased effect estimates may be impossible. To assess potential biases and their magnitude, one may consider conducting sensitivity analyses to quantify these biases [8, 9] and to explore the robustness of study conclusions to various assumptions and sources of bias [10, 11].

An SAT generally relies on an implicit or explicit *external control* (EC) to estimate the therapeutic effects of an anticancer drug. The FDA draft guidance on externally controlled trials (ECTs) [12] states that "if the natural history of a disease is well-defined and the disease is known not to improve in the absence of an intervention or with available therapies, historical information can potentially serve as the control group." An Implicit EC refers to a study-level summary (e.g., an aggregate response rate derived from previous trials or RWE studies) or a literature reported threshold value (e.g., a quantity derived from a meta-analysis). In contrast, an explicit EC involves pre-defined IPD from other trials or from real-world data (RWD) sources such as disease registries and electronic medical records [13, 14]. From a design perspective, regulatory agencies [6, 12] recommend pre-specification of the following key elements in an SAT protocol when using an EC: suitable data sources, baseline eligibility criteria, appropriate exposure definitions and observation windows, clinically meaningful endpoints, analytic methods, and strategies to minimize the effects of missing data and various biases. The estimand framework outlined in ICH E9(R1) [15] should be followed to precisely define the estimand that aligns with the clinical question, as part of efforts to reduce potential biases. Particular attention should be given to potential discrepancies in the frequency and pattern of intercurrent events (ICEs) between the treatment and EC arms when defining the estimand for an ECT (including SATs); see also Chen et al. [16] for general considerations on estimands in RWE studies. Strategies

*Corresponding author: orcid.org/0000-0002-8219-2370.

for handling ICEs should be pre-specified in the protocol or SAP to ensure that the estimated estimand appropriately addresses the clinical question defined in the protocol. For a causal estimand to be identifiable in ECTs, a fit-for-purpose assessment of the RWD should evaluate the validity of the underlying causal assumptions—consistency, positivity, and exchangeability—in addition to standard data quality metrics such as relevance, reliability, and fitness for research use [16, 17]. From an analysis perspective, the FDA draft guidance [12] emphasizes that the analytic method should be capable of "identifying and managing sources of confounding and bias, including a strategy to account for differences in baseline factors and confounding variables between trial arms." Nevertheless, quantitative bias analysis is recommended to assess the sensitivity of study conclusions to various sources of bias when using an EC in SATs [9, 18].

Regarding *endpoints selection*, most SATs supporting AA use surrogate endpoints to measure immediate or intermediate anticancer activity rather than longer-term, clinically meaningful time-to-event endpoints (e.g., overall survival), as the latter may not be adequately characterized in SATs [7, 19, 20]. PROs and QoL measures may also be incorporated in SATs to capture patient-centric effects—particularly treatment benefits beyond tumor response (e.g., symptom relief) compared to baseline—which can enhance real-world relevance and support regulatory approval [21]. However, caution should be exercised, as PROs and QoL outcomes may be over- or underestimated due to missing data in long-term follow-up SATs [22, 23].

Confirmatory *subgroup analysis* (SA) in SATs is often approached with caution due to inherent limitations, such as the lack of an internal comparator arm, small sample sizes, and susceptibility to bias, which can render SA results unreliable. In general, SA in SATs is conducted exploratorily, without pre-specified statistical power, to provide supporting evidence on the consistency of efficacy across subgroups. In some cases, a pre-specified SA based on molecularly defined biomarkers is conducted in SATs, with appropriate multiplicity adjustments, to evaluate whether a biomarker-modifying effect exists [7, 24, 25].

As for *statistical power and sample size*, regulatory guidance documents on RWE studies (including SATs with IPD ECs) recommend that a statistical analysis plan (SAP) be developed in advance and submitted to the relevant regulatory agency prior to study initiation [7, 12, 26]. The SAP should include, at a minimum, clearly defined analyses for primary and secondary estimands, statistical power, sample size, and methods for controlling the probability of erroneous conclusions. In particular, the sample size of an SAT should be sufficiently large to provide a reliable answer to the clinical question, taking into account the planned analysis and the criteria for trial success [12, 26]. For SATs using a fixed threshold control, classical statistical methods for binary outcomes and duration of response can be used to determine the sample size required to detect a clinically meaningful and statistically significant treatment effect compared

to the fixed threshold value [27]. For SATs using IPD as ECs, in addition to conventional sample size determination methods for detecting meaningful differences with desired power [28, 29, 30, 31, 32], additional statistical considerations include: (1) Identification and evaluation of fit-for-purpose RWD sources—Research-oriented RWD (e.g., disease registries, prospective cohorts) is generally preferred over transactional RWD (e.g., claims data). The choice of RWD source impacts the effective sample size—the number of patients eligible to serve as ECs. (2) Type I error inflation—Type I error may be inflated if the response rate in the EC drifts toward the extremes (0 or 1) from a hypothetical value. (3) Analytical methods for estimating treatment effects—Different methods (e.g., Bayesian dynamic borrowing, PS matching, regression) rely on different assumptions and may yield varying results, affecting power and required sample size [33, 34]. (4) Separation of treatment effects from bias—The sample size must be sufficiently large to distinguish true treatment effects from potential bias in the analysis. (5) Simulation studies—Simulations are useful for exploring the interactions among key design factors such as power, sample size, type I error, analytical methods, bias, and the causal gap. Sample size re-estimation during the course of the study may be implemented upon regulatory agreement [35]; see also Chen et al. [36] for further discussion on trial design and analysis using external data.

Finally, we would like to express our gratitude to Cao and Pan for their thoughtful comments and to the Journal Editor for the opportunity to further elaborate on additional considerations in using SATs to support the AA of oncology drugs.

FUNDING

REFERENCES

[1] CAO, S. and PAN, J. (2025). Comments on "Considerations for Single-Arm Trials in Oncology Drug Accelerated Approval". *The New England Journal of Statistics in Data Science.*

[2] LU, F., WANG, T., LU, Y. and CHEN, J. (2025). Considerations for Single-Arm Trials to Support Accelerated Approval of Oncology Drugs. *The New England Journal of Statistics in Data Science* **3**(1) 16–27.

[3] SMITH, G. C. and PELL, J. P. (2003). Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. *BMJ* **327**(7429) 1459–1461.

[4] CHAWLA, S. P., PANG, S. S., JAIN, D., JEFFREY, S., CHAWLA, N. S., SONG, P. Y., HALL, F. L. and GORDON, E. M. (2025). Gene and Cell Therapy for Sarcomas: A Review. *Cancers* **17**(7) 1125.

[5] FDA (2023). *Clinical Trial Considerations to Support Accelerated Approval of Oncology Therapeutics—Guidance for Industry.*

[6] NMPA (2023). *Guidance on Single-Arm Trials Supporting Approval of Anticancer Drugs.*

[7] EMA (2024). *Reflection paper on establishing efficacy based on single-arm trials submitted as pivotal evidence in a marketing authorisation–Considerations on evidence from single-arm trials*.

[8] Gray, C., Ralphs, E., Fox, M. P., Lash, T. L., Liu, G., Kou, T. D., Rivera, D. R., Bosco, J., Braun, K. V. N., Grimson, F. et al. (2024). Use of quantitative bias analysis to evaluate single-arm trials with real-world data external controls. *Pharmacoepidemiology and drug safety* **33**(5) 5796.

[9] Gupta, A., Hsu, G., Kent, S., Duffield, S. J., Merinopoulou, E., Lockhart, A., Arora, P., Ray, J., Wilkinson, S. and Scheuer, N. (2025). Quantitative Bias Analysis for Single-Arm Trials With External Control Arms. *JAMA Network Open* **8**(3) 252152–252152.

[10] Ding, P., Fang, Y., Faries, D., Gruber, S., Lee, H., Lee, J. -Y., Mishra-Kalyani, P., Shan, M., van der Laan, M., Yang, S. et al. (2024). Sensitivity Analysis for Unmeasured Confounding in Medical Product Development and Evaluation Using Real World Evidence. *arXiv preprint arXiv:2307.07442*.

[11] Ho, M., Gruber, S., Fang, Y., Faris, D. E., Mishra-Kalyani, P., Benkeser, D. and van der Laan, M. (2024). Examples of applying RWE causal-inference roadmap to clinical studies. *Statistics in Biopharmaceutical Research* **16**(1) 26–39.

[12] FDA (2023). *Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products*.

[13] Hashmi, M., Rassen, J. and Schneeweiss, S. (2021). Single-arm oncology trials and the nature of external controls arms. *Journal of Comparative Effectiveness Research* **10**(12) 1053–1066.

[14] Mishra-Kalyani, P., Kordestani, L. A., Rivera, D., Singh, H., Ibrahim, A., DeClaro, R., Shen, Y., Tang, S., Sridhara, R. and Kluetz, P. (2022). External control arms in oncology: current use and future directions. *Annals of Oncology* **33**(4) 376–383.

[15] ICH (2021). *E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials*.

[16] Chen, J., Scharfstein, D., Wang, H., Yu, B., Song, Y., He, W., Scott, J., Lin, X. and Lee, H. (2024). Estimands in Real-World Evidence Studies. *Statistics in Biopharmaceutical Research* **16**(2) 257–269.

[17] Levenson, M., He, W., Chen, L., Dharmarajan, S., Izem, R., Meng, Z., Pang, H. and Rockhold, F. (2023). Statistical Consideration for Fit-for-Use Real-World Data to Support Regulatory Decision Making in Drug Development. *Statistics in Biopharmaceutical Research* **15**(3) 689–696.

[18] Gruber, S., Lee, H., Phillips, R., Ho, M. and van der Laan, M. (2023). Developing a Targeted Learning-based statistical analysis plan. *Statistics in Biopharmaceutical Research* **15**(3) 468–475.

[19] FDA (2018). *Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics*.

[20] Mittal, A., Kim, M. S., Dunn, S., Wright, K. and Gyawali, B. (2024). Frequently asked questions on surrogate endpoints in oncology—opportunities, pitfalls, and the way forward. *EClinicalMedicine: Part of The Lancet Discovery Science*.

[21] Liu, L., Choi, J., Musoro, J. Z., Sauerbrei, W., Amdal, C. D., Alanya, A., Barbachano, Y., Cappelleri, J. C., Falk, R. S. and Fiero, M. H. (2023). Single-arm studies involving patient-reported outcome data in oncology: a literature review on current practice. *The Lancet Oncology* **24**(5) 197–206.

[22] Di Maio, M. (2023). The value of patient-reported outcomes in single-arm cancer trials. *Cancer Investigation* **41**(5) 491–494.

[23] Gupta, M., Akhtar, O. S., Bahl, B., Mier-Hicks, A., Attwood, K., Catalfamo, K., Gyawali, B. and Torka, P. (2024). Health-related quality of life outcomes reporting associated with FDA approvals in haematology and oncology. *BMJ Oncology* **3**(1) 000369.

[24] Balar, A. V., Castellano, D., O'Donnell, P. H., Grivas, P., Vuky, J., Powles, T., Plimack, E. R., Hahn, N. M., de Wit, R. and Pang, L. (2017). First-line pembrolizumab in cisplatin-ineligible patients with locally advanced and unresectable or metastatic urothelial cancer (KEYNOTE-052): a multicentre, single-arm, phase 2 study. *The Lancet Oncology* **18**(11) 1483–1492.

[25] Oda, Y. and Narukawa, M. (2022). Response rate of anticancer drugs approved by the Food and Drug Administration based on a single-arm trial. *BMC Cancer* **22**(1) 277.

[26] NMPA (2023). *Guidance on the Design and Protocol Contents of Real-World Studies*.

[27] Yao, S., Shang, Q., Ouyang, M., Zhou, H., Yao, Z., Liu, Y. and Luo, S. (2025). Designing Single-Arm Clinical Trials: Principles, Applications, and Methodological Considerations. *Annals of Clinical Epidemiology* 25011.

[28] O'Malley, A. J., Normand, S. Q. L. T. and Kuntz, R. E. (2002). Sample size calculation for a historically controlled clinical trial with adjustment for covariates. *Journal of Biopharmaceutical Statistics* **12**(2) 227–247.

[29] Englert, S. and Kieser, M. (2012). Adaptive designs for single-arm phase II trials in oncology. *Pharmaceutical Statistics* **11**(3) 241–249.

[30] Schmidli, H., Gsteiger, S., Roychoudhury, S., O'Hagan, A., Spiegelhalter, D. and Neuenschwander, B. (2014). Robust meta-analytic-predictive priors in clinical trials with historical control information. *Biometrics* **70**(4) 1023–1032. https://doi.org/10.1111/biom.12242. MR3295763

[31] Viele, K., Berry, S., Neuenschwander, B., Amzal, B., Chen, F., Enas, N., Hobbs, B., Ibrahim, J. G., Kinnersley, N. and Lindborg, S. (2014). Use of historical control data for assessing treatment effects in clinical trials. *Pharmaceutical Statistics* **13**(1) 41–54.

[32] Rahman, R. and Iftakhar Alam, M. (2022). Stopping for efficacy in single-arm phase II clinical trials. *Journal of Applied Statistics* **49**(10) 2447–2466. https://doi.org/10.1080/02664763.2021.1904846. MR4440947

[33] Seeger, J. D., Davis, K. J., Iannacone, M. R., Zhou, W., Dreyer, N., Winterstein, A. G., Santanello, N., Gertz, B. and Berlin, J. A. (2020). Methods for external control groups for single arm trials or long-term uncontrolled extensions to randomized clinical trials. *Pharmacoepidemiology and Drug Safety* **29**(11) 1382–1392.

[34] Rippin, G., Ballarini, N., Sanz, H., Largent, J., Quinten, C. and Pignatti, F. (2022). A review of causal inference for external comparator arm studies. *Drug Safety* **45**(8) 815–837.

[35] Yue, L. Q., Lu, N. and Xu, Y. (2014). Designing premarket observational comparative studies using existing data as controls: challenges and opportunities. *Journal of Biopharmaceutical Statistics* **24**(5) 994–1010. https://doi.org/10.1080/10543406.2014.926367. MR3246540

[36] Chen, J., Ho, M., Lee, K., Song, Y., Fang, Y., Goldstein, B. A., He, W., Irony, T., Jiang, Q. and van der Laan, M. (2023). The Current Landscape in Biostatistics of Real-World Data and Evidence: Clinical Study Design and Analysis. *Statistics in Biopharmaceutical Research* **15**(1) 29–42.

Feinan Lu. Taimei Intelligence, Building 24 F6, 1999 Yishan Road, Shanghai, China.
E-mail address: `feinan.lu@taimei.com`

Tao Wang. Taimei Intelligence, Building 24 F6, 1999 Yishan Road, Shanghai, China.
E-mail address: `tao.wang2@taimei.com`

Ying Lu. Biomedical Data Science, Stanford University, USA.
E-mail address: `ylu1@stanford.edu`

Jie Chen. Taimei Intelligence, Building 24 F6, 1999 Yishan Road, Shanghai, China.
E-mail address: `jiechen0713@gmail.com`