Dose optimization during drug development: whether and when to optimize

Edward L. Korn (ID), PhD,[1,*] Jeffrey A. Moscow, MD,[2] Boris Freidlin (ID), PhD[1]

[1]Biometric Research Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, MD, USA
[2]Cancer Therapy Evaluation Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, MD, USA

*Correspondence to: Edward L. Korn, PhD, Biometric Research Program, Division of Cancer Treatment and Diagnosis, National Cancer Institute, 9609 Medical Center Dr, Rm 5W112, MSC 9735, Bethesda, MD 20892-9735, USA (e-mail: korne@ctep.nci.nih.gov).

Abstract

The goal of dose optimization during drug development is to identify a dose that preserves clinical benefit with optimal tolerability. Traditionally, the maximum tolerated dose in a small phase I dose escalation study is used in the phase II trial assessing clinical activity of the agent. Although it is possible that this dose level could be altered in the phase II trial if an unexpected level of toxicity is seen, no formal dose optimization has routinely been incorporated into later stages of drug development. Recently it has been suggested that formal dose optimization (involving randomly assigning patients between 2 or more dose levels) be routinely performed early in drug development, even before it is known that the experimental therapy has any clinical activity at any dose level. We consider the relative merits of performing dose optimization earlier vs later in the drug development process and demonstrate that a considerable number of patients may be exposed to ineffective therapies unless dose optimization is delayed until after clinical activity or benefit of the new agent has been established. We conclude that patient and public health interests may be better served by conducting dose optimization after (or during) phase III evaluation, with some exceptions when dose optimization should be performed after activity shown in phase II evaluation.

Traditionally, determining the dose level (and schedule) of a new agent or new combination of agents was based on a small phase I dose-escalation trial identifying the highest dose level with an acceptable toxicity profile (recommended phase II dose [RP2D]). The RP2D would then be used to assess activity in a phase II trial with a clinical activity endpoint (eg, objective response rate [ORR] or progression-free survival [PFS]). If the experimental treatment demonstrated sufficient clinical activity, a large, definitive, randomized phase III trial assessing clinical benefit using an efficacy endpoint (eg, overall survival [OS]) would be performed.

It has long been known that for noncytotoxic agents, the maximum tolerated dose may not be necessary to achieve the maximum clinical benefit (1,2). This issue has become more pressing with the advent of targeted therapies (that may work best in combination) and immune therapies (that may work best when given for an extended period). These concerns have led the US Food and Drug Administration (FDA) and others to suggest that randomized dose optimization comparisons between the RP2D and a lower dose level should be performed as part of the initial drug evaluation (3-5). Although the idea of selecting the best dose before commencing definitive efficacy evaluation of new treatments is attractive, whether and at what cost it can be achieved should be carefully examined. In this commentary, we discuss challenges of performing dose optimization early in the drug development process before there is evidence that the treatment at the RP2D is clinically beneficial or even clinically active compared with performing it later. We first discuss strategies for dose

optimization in the early phases of drug development, followed by dose optimization strategies in later stages of drug development, assuming the optimization is needed at all. We end with our recommendations.

Dose optimization in early drug development

Although there are many methodological articles discussing toxicity-driven dose escalation with dose-level activity assessments (6-10), we focus attention here on the situation where one wants to compare in a randomized fashion the clinical activity at the RP2D (where it is assumed the toxicity is acceptable) with the activity at a lower dose level (where it is assumed that a lower dose is always more tolerable). (Randomized comparisons of different drug schedules with the same effective dose level have received less attention because scheduling can usually be guided by the pharmacokinetics of the agent.) Although optimal dose levels could be assessed with the help of biomarkers (11), the comparison of efficacy between the dose levels typically should be based on their clinical activity (eg, ORR or PFS). This is because there are rarely biomarkers available during early drug development that are reproducible and reliable surrogates of clinical benefit or even activity (especially in development of new agents with novel mechanisms of action). Therefore, in most settings, only a randomized comparison of clinical activity provides a

sufficiently reliable way to decide which dose is to be taken for further development (3,12,13).

When performing dose optimization before the new treatment has shown clinical activity, the 3 timing options are as follows: 1) as part of the phase I trial using expansion cohorts, 2) as a stand-alone trial, or 3) as part of the phase II trial (Figure 1). (Sometimes phase I histology-specific expansion cohorts are designed as phase II trials, and we consider them as phase II trials in this discussion.) In all cases, the patient populations randomly assigned to the doses will need to be appropriate for evaluating clinical activity (and not the typical population used for a phase I trial). If clinical activity can be isolated in a single-arm study (eg, ORR for a single-agent cytotoxic treatment), then the dose-optimization step can also serve as a phase II trial because the study will also assess clinical activity. In the more common situation where clinical activity will need to be assessed in a randomized trial vs a standard treatment (14), incorporation of the dose selection into the phase II trial will involve a 3-armed randomization (high dose, low dose, and control).

Whatever the timing, the challenging part of identifying the optimal dose between 2 dose levels is the required sample sizes necessary to make a reliable selection. Note that one is not just trying to assess whether there is a positive dose-response relationship across multiple dose levels, an easier task, but rather to protect against selecting the low dose as the optimal dose if it has lost a meaningful part of the high-dose activity. Although dose selection is a multifactorial process involving toxicity and pharmacologic/pharmacodynamic considerations, if measures of clinical activity are to be a major factor, then it is important to understand the relationship between sample size and the reliability of clinical activity–based selection. To explore this, it is instructive to review performance of a selection approach designed to limit the probability (to <10%) of choosing the lower dose when the activity of the lower dose is substantially worse than the higher dose. For example, assume that the high dose with a 40% ORR is considered to have promising clinical activity and a dose that has a 20% ORR would not be expected to have meaningful clinical benefit. Table 1 gives the operational characteristics of such a decision rule for different sample sizes. It demonstrates limited ability to choose the lower dose when it is equally or almost equally active (35%-40% ORR) unless the sample sizes are large (100 patients per arm). For example, with 50 or fewer patients per arm, there could be only a 60% probability to choose the lower dose when it is acceptably active (ORR 35%-40% range).

In phase II settings where the new treatment is being evaluated with a time-to-event endpoint (eg, PFS), selection of the dose level could be based on the hazard ratio of the low dose compared with the high dose. It would be reasonable to assume that if this hazard ratio is in the 1-1.1 range, then the high and low doses are likely to have similar clinical benefit. On the other hand, if this hazard ratio is 1.5 or higher, then the low dose may be unacceptably worse. Similarly to the ORR-driven designs, at least 100 patients per arm are needed to allow a reliable dose selection (details not shown).

Dose optimization in later drug development

The traditional approach for dose-level evaluation compares a lesser dose level with a higher dose level after the higher dose level has shown clinical efficacy (15) (Figure 2, C). This approach has some advantages: Large numbers of patients will not be exposed to a treatment that turns out not to have clinical benefit, the treatment can be made available (although possibly at a dose level higher than necessary) while awaiting the results of the dose-level comparison, and the high-dose vs standard treatment effect seen in the phase III trial can be used to help the design of the follow-up dose-level comparison trial. This is the approach that has been used in the past by the FDA via postmarketing requirements (13). A disadvantage of this approach is that because clinical benefit (rather than clinical activity) is being evaluated in the dose-level comparison, a large number of patients will be required. For example, the overall survival–driven noninferiority PROSELICA trial of cabazitaxel 20 mg/m^2 vs 25 mg/m^2 (the approved dose) for postdocetaxel patients with metastatic castration-resistant prostate cancer enrolled 1200 patients (16).

Some have suggested that a practical disadvantage of comparison with a lower dose after the higher dose has shown to be beneficial is that patients may have less interest in participating in such a trial (13,17), although we note that both treatment arms are receiving the (new) treatment that has shown clinical benefit at the higher dose. Another challenge is that pharmaceutical companies may be reluctant to perform such trials (unless required by regulatory authorities). An alternative strategy is to perform the dose-level evaluation as part of the phase III trial (Figure 2, B), using a 3-armed trial with the 2 dose levels and the standard therapy arm. This approach will also lessen the total sample sizes involved because the high-dose arm will not need to be repeated as when a sequential approach (phase III followed by dose comparison testing) is used. An example is given by the E1609 trial, which randomly assigned 1670 resected high-risk melanoma patients to ipilimumab 10 mg/kg, ipilimumab 3 mg/kg, or high-dose interferon (the control treatment) (18). As is typical in these trial designs, the statistical analysis formally compared each of the experimental arms with the control arm, with the choice of recommended dose based on toxicity and informal comparison of the treatment arms if both high- and low-dose arms are statistically better than the control arm. Another advantage of this approach over sequential assessment is that the 2 different dose arms and the standard treatment arm are all available for a simultaneous randomized comparison (3). Two shortcomings of this approach are that it will take longer to determine clinical benefit of the high dose when the experimental treatment works (because of the additional low-dose treatment arm) and that twice as many patients will be treated with an ineffective treatment than if the sequential approach had been used when the experimental treatment is not clinically beneficial (even at the high dose). The latter shortcoming can be mitigated by appropriate interim monitoring (19,20).

To lower the required sample size, one might consider using an intermediate endpoint for the dose level comparisons, for example, the endpoint that was used to show the treatment had clinical activity in phase II or a biomarker possibly developed during the earlier stages of drug development (21). This could be done before or after the phase III trial (Figure 2, A and C) or as part of a 3-armed phase II/III trial design using the intermediate endpoint for the phase II analysis to select one of the dose-level arms to continue to the phase III (Figure 2, B). The phase II/III option is especially attractive when large numbers of patients (~200) are needed to compare the dose levels even with using the intermediate endpoint. There is the hazard in using an intermediate endpoint that it may either miss a clinically relevant difference or identify a difference in dose levels that is not clinically
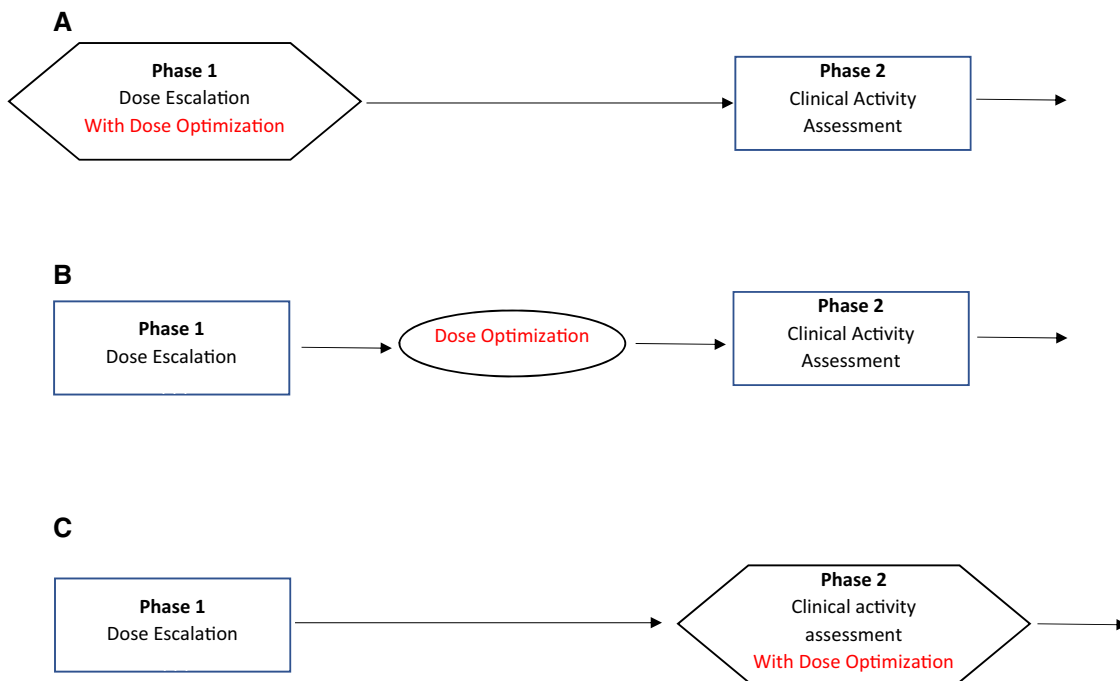
Figure 1. Options for dose optimization in early drug development (before clinical activity is demonstrated).

Table 1. Probabilities[a] of selecting a lower dose level in a trial design that has a 0.10 probability of selecting the lower dose when its response rate is lower by 20%[b]

	Difference in response rates between low-dose and high-dose arms		
Sample size per arm	$p_H = 40\%$ $p_L = 20\%$ $p_L - p_H = -20\%$	$p_H = 40\%$ $p_L = 35\%$ $p_L - p_H = -5\%$	$p_H = 40\%$ $p_L = 40\%$ $p_L - p_H = 0\%$
20	.10	.35	.46
30	.10	.50	.65
50	.10	.60	.77
100	.10	.83	.95

[a] Probabilities are estimated from 10^6 simulated trials.
[b] The lower dose level is selected if the 1-sided lower 90% confidence limit for the difference in response rates is greater than −20%.
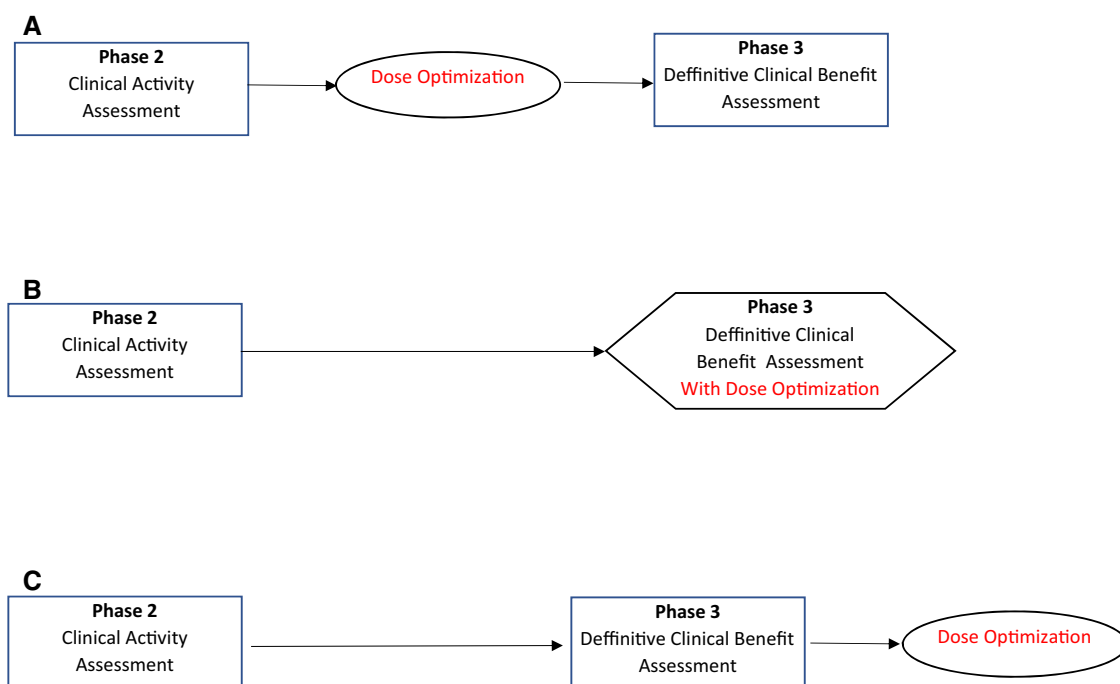

Figure 2. Options for dose optimization in later drug development (after clinical activity is demonstrated).

meaningful. For example, in the PROSELICA trial, the higher dose had a statistically significantly higher Prostate-specific antigen (PSA) response rate than the lower dose (16), which would have led to the incorrect recommendation if this endpoint had been used to select the dose.

An example using a biomarker intermediate endpoint is given by the noninferiority phase II trial (22) of abiraterone at a low dose (250 mg with a low-fat meal) vs standard dose (1000 mg fasting) in castration-resistant prostate cancer that used the change in PSA (12 weeks vs baseline) as the endpoint. The interpretation of the results of the trial was controversial (23-29); we focus here on some design issues. First, although availability of a validated biomarker is more likely later in drug development, variability of the biomarker and its relationship to clinical benefit may still not be fully understood, making its use for dose optimization problematic. Secondly, the use of a continuous biomarker instead of a binary or time-to-event endpoint should, in theory, offer some efficiency gains in term of a smaller required sample size. For example, the abiraterone trial randomly assigned only 72 patients. However, the PSA-change noninferiority margin for this trial was very wide, suggesting a more definitive trial design would have required a larger sample size. Finally, biomarker endpoint trials are frequently identified as phase II, suggesting a follow-up phase III trial with a definitive endpoint. Therefore, a better alternative strategy may be to go directly to the phase III trial with a clinical-benefit endpoint, with interim monitoring allowing for early stopping for inefficacy.

In some clinical settings, an agent is determined to have clinical benefit based on a single-arm trial demonstrating a durable ORR dramatically better than seen with standard therapies, avoiding the need for a randomized trial vs a standard therapy. Dose optimization will still require a randomized trial between the doses if the RP2D had tolerability issues, with sample size considerations the same as previously given. For example, sotorasib (960 mg) was given accelerated approval by the FDA for previously treated non-small-cell lung cancer patients with a *KRAS G12C* mutation based on a 36% ORR seen in a 124-patient phase II trial (30,31) that followed 59 non-small-cell lung cancer patients treated in a phase I trial (32); currently, an ongoing 170-patient randomized trial (NCT04933695) is comparing the ORR of 960 mg vs 240 mg doses. It is an open question as to whether ORR is the best primary endpoint to assess clinical benefit for dose optimization in this randomized trial. However, if it is an acceptable endpoint for assessing clinical benefit, one could argue that the optimization could have occurred during the phase II trial (33).

Tolerability component of dose optimization

An additional consideration in the timing of dose optimization is that in phase I studies, the recommended phase II dose is typically determined by high-grade toxicities that occur over a short period of time, which is usually the first 1 or 2 cycles of therapy. There are many possible phase I trial designs (34), and some designs accommodate increased sample sizes to more accurately estimate a dose level with acceptable toxicity. However, these small, early-phase trial designs are not intended to accurately establish a tolerable dose that reflects the rate of dose adjustments or discontinuations in later cycles (for any reason) or lesser-grade side effects experienced over a longer period of time by a larger number of patients. The tolerability of a recommended dose, like its clinical benefit, can only be determined by

the experience of patients treated over time in circumstances similar to routine clinical practice.

In addition, patient populations and drug exposure often vary substantially as drugs and drug combinations develop from early phase to late stage studies. In early-phase studies, patients may be more heavily pretreated and have a greater number of comorbidities than might be found in later-stage studies. Also, because patients in early-phase studies have more advanced disease and often may not be optimally selected for clinical benefit, they may receive fewer cycles of therapy than patients may receive in later-phase studies. In addition, as toxicity experience with novel agent or agent combinations develops, so too will symptom management strategies to ameliorate toxicity symptoms. For all of these reasons, the tolerability assessment in early-phase studies, with heterogenous populations of patients with more advanced disease, can be very different from the assessment in later-phase studies with more homogenous populations and earlier stage disease. These considerations suggest dose optimization may be more meaningfully done in the later stages of drug development.

Recommendations

If most experimental treatments that started in phase I trials were eventually shown to improve patient care, then optimizing the dose level as early as possible would indeed be the best strategy. Unfortunately, this is not the case, with only an estimated 7% of the agents in phase I trials eventually leading to FDA approval (35,36). Figure 3 gives a rough idea of the relative numbers of extra patients that would need to be treated with ultimately unapproved agents for performing dose optimization during early vs late development. As a minor point, these large numbers of patients unnecessarily treated also reduce the pool of patients available to be included in other clinical trials. Furthermore, only a small fraction of approved therapies may require postmarketing dose reexamination because of tolerability issues. The relative disadvantage of early optimization is even worse than shown if a validated biomarker is developed during phase II or phase III, which would allow a smaller optimization trial. Therefore, early dose optimization before one knows that the experimental treatment has any clinical activity and that the tested dose is not tolerable is suboptimal.

We specifically recommend the following. If there is a question as whether the experimental treatment will be tolerable in the phase III clinical setting or if there is a reliable activity biomarker, then the dose optimization should occur after the treatment has shown activity in phase II, which can provide both activity and tolerability evidence required to make an informed decision regarding the need for dose optimization. Otherwise, if it is believed that dose optimization is advisable even though the agent is tolerable, then it should occur after phase III or, if that is challenging, then as part of the phase III trial. (If the agent has been shown to be unequivocally tolerable in the earlier stages of development, then there may be no need for a dose optimization.) Performing the dose optimization before the agent has shown clinical activity in a phase 2 trial will needlessly expose large numbers of patients to ineffective therapies and slow down drug development. In summary, although early dose optimization may appear an attractive strategy for drug development, it is unclear how it benefits public health and patients if this strategy
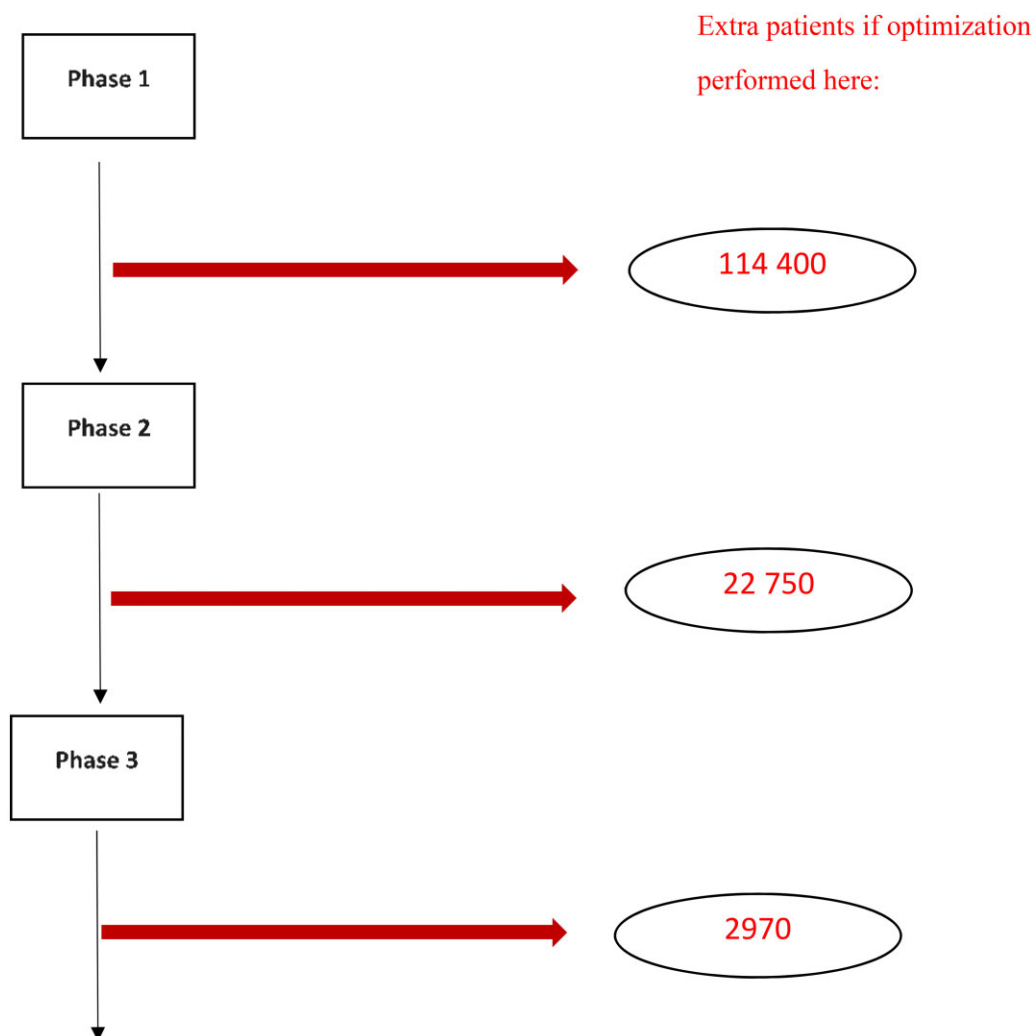
Figure 3. Based on starting 1000 phase I trials, estimated extra numbers of patients treated with an agent that does not receive FDA approval when 200-patient dose optimization studies are performed at different points in the drug development process. Phase success rates are 63.9%, 28.3%, and 45.2%, and eventual FDA approval rates are 6.7%, 10.5%, and 37.0% for phase I, II, and III trials, respectively [see Table 5 published in reference (35).]

increases the chances of selecting an ineffective dose and treating many patients with ineffective therapy.

Funding

No funding was used for this study.

Notes

Role of the funder: Not applicable.

Disclosures: None.

Author contributions: Conceptualization: ELK, JAM, BF; Writing-original draft: ELK, JAM, BF; Writing- review and editing: ELK, JAM, BF.

Data availability

No primary data was used in this manuscript. All figures cited are from published sources.

References

1. Gelmon KA, Eisenhauer EA, Harris AL, Ratain MJ, Workman P. Anticancer agents targeting signaling molecules and cancer cell environment: challenges for drug development? *J Natl Cancer Inst*. 1999;91(15):1281-1287.
2. Korn EL, Arbuck SG, Pluda JM, Simon R, Kaplan RS, Christian MC. Clinical trial designs for cytostatic agents: are new approaches needed? *J Clin Oncol*. 2001;19(1):265-272.
3. Shah M, Rahman A, Theoret MR, Pazdur R. The conundrum of oncology drug dosing: more is less and less is more. *N Engl J Med*. 2021;385(16):1445-1447.
4. U.S. Food and Drug Administration. *Project Optimus: Reforming the Dose Optimization and Dose Selection Paradigm in Oncology*. https://www.fda.gov/about-fda/oncology-center-excellence/project-optimus. Accessed November 9, 2022.
5. Zirkelbach JF, Shah M, Vallejo J, et al. Improving dose-optimization processes used in oncology drug development to minimize toxicity and maximize benefit to patients. *J Clin Oncol*. 2022;40(30):3489-3500.

6.   Thall PF, Cook JD. Dose-finding based on efficacy-toxicity trade offs. *Biometrics*. 2004;60(3):684-693.

7.   Hunsberger S, Rubinstein LV, Dancey J, Korn EL. Dose escalation trial designs based on a molecularly targeted endpoint. *Stat Med*. 2005;24(14):2171-2181.

8.   Zhang W, Sargent DJ, Mandrekar S. An adaptive dose-finding design incorporating both toxicity and efficacy. *Statist Med*. 2006;25(14):2365-2383.

9.   Hoering A, LeBlanc M, Crowley J. Seamless phase I–II trial design for assessing toxicity and efficacy for targeted agents. *Clin Cancer Res*. 2011;17(4):640-646.

10. Yuan Y, Lee JJ, Hilsenbeck SG. Model-assisted designs for early-phase clinical trials: simplicity meets superiority. *J Clin Oncol Precis Oncol*. 2019;3:1-12. doi:10.1200/PO.19.00032.

11. Ratain MJ. Targeted therapies: redefining the primary objective of phase I oncology trials. *Nat Rev Clin Oncol*. 2014;11(9):503-504.

12. Center for Drug Evaluation and Research. Food and Drug Administration. *FDA Expansion Cohorts: Use in First-in-Human Clinical Trials to Expedite Development of Oncology Drugs and Biologics Guidance for Industry*. 2022. https://www.fda.gov/media/115172/download. Accessed November 8, 2022.

13. Minasian L, Rosen O, Auclair D, Rahman A, Pazdur R, Schilsky RL. Optimizing dosing of oncology drugs. *Clin Pharmacol Ther*. 2014;96(5):572-579.

14. Rubinstein LV, Korn EL, Freidlin B, Hunsberger S, Ivy SP, Smith MA. Design issues of randomized phase II trials and a proposal for phase II screening trials. *J Clin Oncol*. 2005;23:199-206.

15. Nie L, Lee KY, Verdun N, De Claro RA, Sridhara R. Dose finding in late-phase drug development. *Ther Innov Regul Sci*. 2017;51(6):738-743.

16. Eisenberger M, Hardy-Bessard A-C, Kim CS, et al. Phase III study comparing a reduced dose of cabazitaxel (20 mg/m2) and the currently approved dose (25 mg/m2) in postdocetaxel patients with metastatic castration-resistant prostate cancer-PROSELICA. *J Clin Oncol*. 2017;35(28):3198-3206.

17. Araujo DV, Uchoa B, Soto-Castillo JJ, Furlan LL, Oliva M. When less may be enough: dose selection strategies for immune checkpoint inhibitors focusing on AntiPD-(L)1 agents. *Targ Oncol*. 2022;17(3):253-270.

18. Tarhini AA, Lee SJ, Hodi FS, et al. Phase III study of adjuvant ipilimumab (3 or 10 mg/kg) versus high-dose interferon alfa-2b for resected high-risk melanoma: North American Intergroup E1609. *J Clin Oncol*. 2020;38(6):567-575.

19. Freidlin B, Korn EL. Monitoring for lack of benefit: a critical component of a randomized clinical trial. *J Clin Oncol*. 2009;27(4):629-633.

20. Korn EL, Hunsberger S, Freidlin B, et al. Preliminary data release for randomized clinical trials of noninferiority: a new proposal. *J Clin Oncol*. 2005;23(24):5831-5836.

21. Piccart MJ, Hilbers FS, Bliss JM, et al.; BIG-NABCG Collaboration. Road map to safe and well-designed de-escalation trials of systemic adjuvant therapy for solid tumors. *J Clin Oncol*. 2020;38(34):4120-4129.

22. Szmulewitz RZ, Peer CJ, Ibraheem A, et al. Prospective international randomized phase II study of low-dose abiraterone with food versus standard dose abiraterone in castration-resistant prostate cancer. *J Clin Oncol*. 2018;36(14):1389-1395.

23. Kolesar JM, Liu GX. Low-fat abiraterone food effect is of little consequence. *J Clin Oncol*. 2018;36(14):1385-1386.

24. Tannock IF. Low-fat abiraterone food effect is of great consequence. *J Clin Oncol*. 2018;36(30):3058-3059.

25. Isaacsson Velho P, Eisenberger MA. There is now compelling evidence to further evaluate lower doses of abiraterone acetate in men with metastatic prostate cancer: it should be safer, may be as effective and less expensive. *J Clin Oncol*. 2018;36(30):3059-3060.

26. Tiako Meyo M, Alexandre J, Goldwasser F, et al. Low-dose abiraterone regimen: drug monitoring might be the key. *J Clin Oncol*. 2018;36(30):3061-3062.

27. Woei-A-Jin FJSH, Van Nieuwenhuyse T, van Erp NP, et al. Dose reduction may jeopardize efficacy of abiraterone acetate. *J Clin Oncol*. 2018;36(30):3062-3064.

28. Szmulewitz RZ, Karrison T, Stadler WM, et al. Low-dose abiraterone with food: rebutting an editorial. *J Clin Oncol*. 2018;36(30):3060-3061.

29. Kolesar JM, Liu GX. Reply to I.F. Tannock, P. Isaacsson Velho et al, R.Z. Szmulewitz et al, M. Tiako Meyo et al, and F.J.S.H. Woei-A-Jin et al. *J Clin Oncol*. 2018;36(30):3065-3066.

30. Skoulidis F, Li BT, Dy GK, et al. Sotorasib for lung cancers with KRAS G12C mutation. *N Engl J Med*. 2021;384(25):2371-2381.

31. Nakajima EC, Drezner N, Li X, et al. FDA approval summary: sotorasib for KRAS G12C-mutated metastatic NSCLC. *Clin Cancer Res*. 2022;28(8):1482-1486.

32. Hong DS, Fakih MG, Strickler JH, et al. $KRAS^{G12C}$ inhibition with sotorasib in advanced solid tumors. *N Engl J Med*. 2020;383(13):1207-1217.

33. Ratain MH, Tannock IF, Lichter AS. Dose optimization of sotorasib: is the US Food and Drug Administration sending a message? *J Clin Oncol*. 2021;39(31):3423-3426.

34. Clertant M. Early-phase oncology trials: why so many designs? *J Clin Oncol*. 2022;40(30):3529-3536.

35. Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J. Clinical development success rates for investigational drugs. *Nat Biotechnol*. 2014;32(1):40-51.

36. Zhang SX, Fergusson D, Kimmelman J. Proportion of patients in phase I oncology trials receiving treatments that are ultimately approved. *J Natl Cancer Inst*. 2020;112(9):886-892.