Platform trials are complex innovative designs in which patients are simultaneously randomized to several experimental treatments and to a common comparator group.1 Moreover, as the study progresses, new treatments can be added dynamically to the ongoing trial, and nonefficacious treatments can be discontinued during planned interim analyses. By comparing several treatments against a shared, common control group and offering an open-ended infrastructure allowing to quickly add and remove treatments to the study, platform trials have the potential to accelerate clinical development and bring efficacious treatments to patients faster. These designs are thus increasingly being used in different therapeutic areas, especially in oncology and, more recently, against COVID-19.2

It is often claimed that one of the greatest advantages of platform trials is their potential to use a single infrastructure (eg, common centers, investigators, independent data monitoring committee, protocol, recruitment procedures, database, and so forth), thus offering certain administrative and logistic advantages that further translate into a greater economic efficiency. In their article, Park et al3 compared the cost and time of setting up, conducting, and analyzing a platform trial with other development strategies, including a series of independent randomized clinical trials. The authors demonstrated that, via simulations informed by the famous Systemic Therapy in Advancing or Metastatic Prostate Cancer: Evaluation of Drug Efficacy (STAMPEDE) trial, the cost and time of planning and conducting a platform trial is larger than that of a single trial but smaller than other sequential development options including several trials. Even though in the absence of publicly available data on the cost and setup time of clinical trials, the authors had to resort to a survey of experts to inform their analysis, the article3 illustrates that, when the intention is to generate efficacy data on several treatments, it is more efficient in the long run to do so via a platform trial. Otherwise, if, for example, the read-out of a specific treatment is likely to condition the investigation of the other experimental treatments, then a more staggered approach might be preferrable.

While the article from Park et al3 deals with economic aspects pertaining to the design, setup, and analysis of the trial, supplementary savings in cost and time can arise from the structure of the platform trial itself. First, the opportunity to receive 1 of several experimental treatments with a greater probability than receiving the comparator can be attractive for both investigators and participants and can boost recruitment. Second, the sharing of a single control group leads to a decrease in the overall sample size compared with a series of independent studies and, therefore, translates into cost savings. Other innovative approaches could be used to further reduce the sample size of the trial (or improve its power) and, therefore, its cost, such as the use of nonconcurrent controls (ie, participants in the control group who were enrolled in the platform trial before the addition of a new treatment of interest).4 Doing so, however, requires strong assumptions—usually unknown or unverifiable at the time of writing the protocol—about the existence and impact of time trends and can potentially lead to an increased risk of false-positive claims of efficacy.4 The use of nonconcurrent controls is, therefore, more advisable where there is no regulatory mandate to control the risk of false positives, such as in early-phase trials and academic trials, where the drug developer has sole responsibility for weighing the pros and cons of implementing such an approach.1,2,4 Finally, if head-to-head comparisons between several experimental treatments are incorporated in the trial design, platform trials could also indirectly generate a larger economic reward later in the drug development process. Indeed, at the reimbursement stage, health technology agencies prefer evidence to assess clinical and/or cost-effectiveness data from direct
head-to-head comparisons, which reduce uncertainty, rather than having to compare treatments using indirect comparisons. For sponsors, this comes, however, with the risk of displaying less favorable efficacy data than other experimental treatments in the platform if those are in competition.

Simultaneously comparing several treatments with a common control can, however, bring specific challenges that can have a negative economic impact. First, the treatments in the platform trial can each have specific inclusion and exclusion criteria, and enrolling participants who meet the common denominator of these can be difficult, which can delay recruitment. Moreover, progressing the most promising assets of a phase 2 platform trial to phase 3 can create an issue of self-competition between both trials because patients might prefer being enrolled in a confirmatory trial investigating a treatment that has already shown early signs of efficacy rather than in an exploratory trial. Second, given that all treatment groups are compared with the same comparator group, the totality of evidence ultimately delivered by that study for each of these treatments is tied to the control group and will be extremely dependent on the observed control response. Indeed, if by random chance the observed control response is unexpectedly and erroneously high, declaring statistical significance for each treatment in the platform trials will be less likely, and few to no treatments might be progressed to the next stage of development, thus yielding a potential loss of expected reward if these treatments were in truth efficacious. On the contrary, if the observed control response is unexpectedly and erroneously low, several nonefficacious treatments might mistakenly be treated as statistically significant. Progressing these assets to the next stage of development could lead to failed subsequent confirmatory trials with associated substantial economic loss. Similar statements could be made about futility analyses, where several efficacious treatments could be mistakenly discontinued if the observed control response is erroneously large (or several nonefficacious treatments could mistakenly pass the futility analysis if the observed control response is erroneously low). These issues would not occur with separate trials as unexpected control responses would be less likely to be observed in multiple trials. In addition, although there is a general consensus that no adjustment for multiple testing would be expected in the confirmatory setting if the treatments investigated were different, regulators could ask to adjust the type I error if the treatment groups correspond to different variants (eg, different doses, regimens, or combinations of the same drug), as they would offer multiple chances of claiming efficacy for the same treatment within the same trial. This could therefore require a larger sample size per treatment group to make up for the loss of power, which, in turn, would further increase the costs of the trial.

In conclusion, sharing a common control group in a platform trial can be viewed as a gift and a curse, and choosing whether to implement a platform trial or a more standard development program is complex since it is contingent on numerous factors that are disease and context dependent. To make such a decision, a clear framework needs to be implemented. Decision-making in the pharmaceutical industry is becoming increasingly more quantitative, and in practice, both development approaches would be compared according to a series of standard metrics, including (1) duration of the clinical program; (2) cost and expected reward of the clinical program; (3) flexibility of the clinical program (eg, reporting of the analyses corresponding to the different treatments while maintaining data and trial integrity, ability to add sites, and so forth); (4) probability of success for each treatment, for all treatments, or for at least 1 treatment, assuming all or some of the treatments have a certain efficacy; (5) probability of (multiple) false-positive findings (eg, achieving statistical significance) for each treatment or for at least 1 treatment, assuming all or some of the treatments are comparator-like; and (6) the extent to which a positive readout informs the success of potential subsequent trials and the ability to de-risk the next phase of development. Using these metrics, governance boards would then make their decisions on whether to progress developing a treatment according to a range of diverse considerations, such as portfolio opportunities and investment priorities.
REFERENCES


