Jones and colleagues\(^1\) reported final results of the Rapid Administration of Carnitine in Sepsis (RACE) trial, a Bayesian adaptive multiarm trial that evaluated whether levocarnitine, studied at 3 dose levels, reduces the risk of organ failure in patients with septic shock. Using a response-adaptive algorithm, the trial randomized 250 patients between March 5, 2013, through February 5, 2018, to experimental arms with different levocarnitine dose levels and a control arm. The study evaluated treatment efficacy using 2 end points, the variation of the sequential organ failure assessment score at 48 hours from enrollment and overall survival at 28 days from randomization. Based on Bayesian analyses, as planned before the onset of the study, none of the experimental arms showed sufficient evidence of a positive clinical benefit to recommend a confirmatory phase 3 study.

The RACE trial shows that Bayesian modeling combined with rapid and reliable data management procedures allows adaptation of the randomization probabilities using early data, while the study continues the enrollment. Adaptation, in this study, was motivated by the needs of evaluating levocarnitine at different dose levels and of a sequential increase of randomization toward the experimental arm that was more promising and likely to succeed in a subsequent phase 3 study. Adaptive designs to combine exploration of dose levels and final recommendations on whether to proceed with a confirmatory randomized controlled trial are strategic for the efficient development of new treatments.

The use and modeling of multiple end points to adapt randomization, as in the RACE trial, have several motivations. First, early end points (sequential organ failure assessment score at 48 hours) become available shortly after randomization and facilitate adaptation, while other end points that might have a more direct clinical interpretation (survival at 28 days) require more time to be collected. Second, statistical evidence and early estimates of treatment effects can be more pronounced in surrogate or secondary end points. Third, to reduce the risk of bias, final and robust data analyses, at completion of the study, do not need to rely on assumptions or a priori estimates used by the adaptive algorithm to summarize early evidence from multiple end points during randomization.

After an initial period with balanced treatment assignment to 40 patients, the adaptive algorithm of the RACE trial unbalanced randomization toward the most promising arms, with a large proportion (106 [42%]) of the enrolled patients assigned to the highest levocarnitine dose. Importantly, adaptive randomization also assigned an adequate proportion (75 [30%]) of the enrolled patients to the control arm, which, as discussed previously\(^2,3\), is necessary to preserve the probability of a true-positive result.

The RACE trial was designed to select and recommend 1 of the experimental arms for a confirmatory phase 3 study conditionally on a large prediction probability (>90%) of a positive result for a hypothetical subsequent phase 3 study. This metric, as well as the posterior probability of positive treatment effects, appears easier to interpret than \(P\) values that test the null hypothesis of no treatment effects. Importantly, these metrics allow summarization—using Bayesian modeling—of evidence from the control and all experimental arms with different dose levels.

We consider the RACE trial another positive example—as several others in the last 10 years—of Bayesian adaptive design. Emerging new designs such as Bayesian platform trials\(^4,5\) and basket designs\(^1,6\) have great potential to accelerate the development of new drugs,\(^7\) assigning more patients to promising arms relative to traditional balanced randomized designs. While the use of
Bayesian designs becomes more frequent, there also appears to be a trend in the literature toward more complex designs. For example, a design similar to the RACE trial, with an additional component of complexity, would allow the addition, after the onset of the study, of new dose levels to be studied or the use of a continuum of dose levels.

This trend toward adaptive designs (in most cases to evaluate multiple treatments) that become increasingly complex and tailored to endpoints, clinical hypotheses, and organizational needs that are study specific can have positive effects in terms of statistical efficiency, time, and resources to develop new treatments. It also poses the challenge to establish new and appropriate standards on how to communicate, present, and share the study design and the results of adaptive trials. Standards and technologies that are appropriate for relatively simple designs, such as controlled 2-arm balanced designs, are likely to become obsolete and inappropriate for adaptive designs. For example, biostatisticians need to identify good practices and software to facilitate reproducibility of simulation studies that quantify power and other operating characteristics of trial designs. A few summaries of operating characteristics and key concepts followed to design the trial might not be sufficient. Similarly, rigorous standards to show in final analyses absence of potential sources of bias, such as systematic variations of clinical characteristics in the enrolled patients during the randomization period, are necessary for adaptive designs. We are optimistic that the current methodological research, the active role of regulators, and experience from the ongoing Bayesian adaptive trials will align innovations of modern trial designs with appropriate standards in the presentation of designs and results.