

On the Usefulness of Outcome-Adaptive Randomization

TO THE EDITOR: In a recent article published in *Journal of Clinical Oncology*, Korn and Freidlin¹ compared outcome-adaptive randomization with designs using 1:1 and 2:1 fixed-ratio randomization (2:1 means that the probability of randomization to the experimental arm is 2/3 and to the standard arm is 1/3). The authors found no benefit in using outcome-adaptive randomization compared with fixed-ratio randomization and recommended the latter. This finding is important and interesting because clinical trial designs would be greatly simplified by adopting the straightforward fixed-ratio randomization approach rather than the logistically more involved outcome-adaptive randomization method. The results of Korn and Freidlin rely heavily on simulation studies that focus on a specific outcome-adaptive randomization method² and a limited number of simulation settings. Consequently, one may question the generality of their conclusions. From a theoretic point of view, we provide additional justification and insights into comparisons between outcome-adaptive and fixed-ratio randomization.

Assuming a similar setup as that used by Korn and Freidlin,¹ we considered a comparison between a standard treatment and an experimental treatment with an immediately ascertainable binary end point. We tested the difference in the response rates between the two treatments on the basis of the one-sided normal approximation test with a type I error rate of $\alpha = .1$ and a type II error rate of $\beta = .1$, and

we evaluated the performance of the adaptive and fixed randomization designs on the basis of the expected number of nonresponders (denoted as n_{AR} and n_{FR} for adaptive and fixed randomization, respectively) and the probability that a patient would be a responder.

It is known that the optimal outcome-adaptive design that minimizes the expected number of nonresponders should target the allocation ratio of $\sqrt{p_e} : \sqrt{p_s}$ between the experimental arm and the standard arm,³ where p_e and p_s are the response rates of the experimental and standard treatments, respectively. On the basis of this result, we obtained the minimal expected number of nonresponders for outcome-adaptive randomization designs using the following equation:

$$\min(n_{AR}) = \frac{(z_{1-\alpha} - z_\beta)^2 (\sqrt{p_e q_e} + \sqrt{p_s q_s})^2}{(p_e - p_s)^2}, \quad p_e \neq p_s \quad (1)$$

where $q_s = 1 - p_s$, $q_e = 1 - p_e$, and $z_{1-\alpha}$ and z_β are the $(1 - \alpha)$ th and β th quantiles of the standard normal distribution. Note that if $p_s = p_e$, the optimal outcome-adaptive design becomes equal randomization. It then follows that, compared with the $K:1$ fixed-ratio randomization wherein patients are randomly assigned to the experimental arm with a $K/(1 + K)$ probability, and to the standard arm with a $1/(1 + K)$ probability, the maximum percentage of reduction in the expected number of nonresponders for outcome-adaptive randomization designs is determined by the following equation.

$$\max\left(\frac{n_{FR} - n_{AR}}{n_{FR}}\right) = 1 - \frac{K(\sqrt{p_e q_e} + \sqrt{p_s q_s})^2}{(p_e q_e + K p_s q_s)(K q_e + q_s)} \quad (2)$$

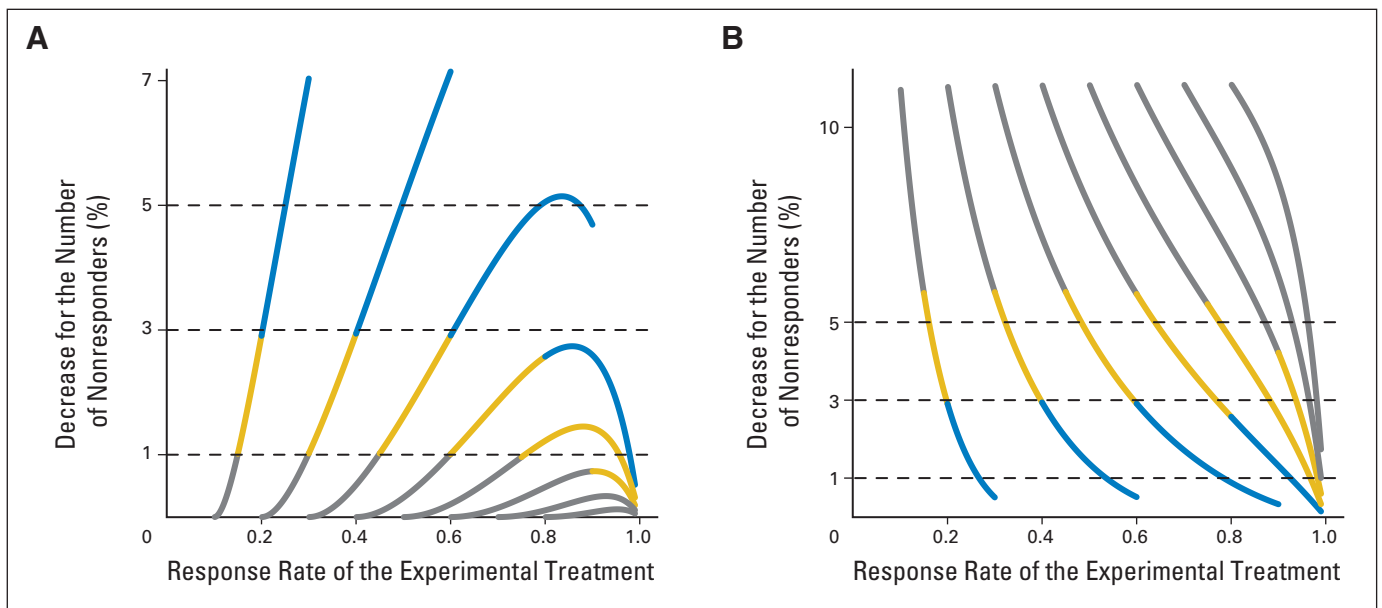


Fig 1. The maximum percentage of reduction in the number of nonresponders for outcome-adaptive randomization with respect to that for (A) 1:1 fixed-ratio randomization and (B) 2:1 fixed-ratio randomization. With respect to the curves from left to right, the response rates of the standard treatment are 0.1, 0.2, ..., 0.8, respectively, and the gray, yellow, and blue segments of each curve indicate that the response rate of the experimental treatment is 0% to 50%, 50% to 100%, and 100% to 200% higher than that of the standard treatment, respectively.

Unlike simulation studies in which the results may depend on the specific methods under consideration and the values of the simulation parameters, the above upper bound is general and holds for all adaptive randomization designs. That is, it is the best that any outcome-adaptive randomization can achieve. Interestingly, this bound does not depend on the prespecified type I and type II error rates. Figure 1A shows the maximum percentage of reduction in the expected number of nonresponders using outcome-adaptive randomization with respect to 1:1 fixed-ratio randomization under various treatment response rates. For clarity of presentation and also for practical relevance, given a specific response rate of the standard treatment, we depict the results only for the case in which the improvement of the response rate by the experimental treatment (compared with the standard treatment) is less than 200%.

We can see that the maximum possible gain (ie, reduction in the expected number of nonresponders) using adaptive randomization increases with the response rate of the experimental treatment. However, in general, such a reduction in the number of nonresponders is rather limited. When the response rate of the experimental treatment is higher than that of the standard treatment by up to 50% (the gray segments of the curves in Fig 1A), the maximum gain using adaptive randomization is typically less than 1%, and even when the response rate of the experimental treatment doubles that of the standard treatment (the yellow segments of the curves in Fig 1A), such a gain is often less than 3%. Figure 1A also suggests that the adaptive randomization method² considered by Korn and Freidlin¹ is not optimal; therefore, their comparisons may slightly favor the fixed-ratio randomization approach. For instance, the adaptive randomization method considered by Korn and Freidlin is inferior to the 1:1 fixed-ratio randomiza-

tion with a larger number of nonresponders (93.5 or 92.9 v 92.4) when the response rates of the standard and experimental treatments are 0.2 and 0.4 (Table 2 in the article by Korn and Freidlin). In contrast, our results show that a well-calibrated optimal adaptive randomization can outperform equal randomization with a smaller number of nonresponders, although such an improvement is typically small (approximately 3%).

When comparing outcome-adaptive randomization with a 2:1 fixed-ratio randomization, we find the results (Fig 1B) to be remarkably different from those obtained with the 1:1 fixed-ratio randomization. The gain from using outcome-adaptive randomization in terms of reducing the number of nonresponders actually decreases when the response rate of the experimental arm increases (within the range of the response rate that is of interest). This is because when the response rate of the experimental arm increases, the optimal allocation ratio approaches 2:1; therefore, the relative gain of the optimal outcome-adaptive randomization (as compared with the 2:1 fixed ratio randomization) decreases. When the difference in the response rate between the two treatments is small, the gain of the optimal outcome-adaptive randomization can be as large as 10% or more. This finding suggests that in terms of the number of nonresponders, the 2:1 fixed-ratio randomization may not be the best choice under small or moderate effect sizes.

Another criterion used by Korn and Freidlin¹ to evaluate the performance of randomization designs is the probability that a patient will be a responder (denoted as π_{AR} and π_{FR} for adaptive randomization and fixed randomization, respectively). To quantify the largest improvement that adaptive randomization can achieve (with respect to a $K:1$ fixed-ratio randomization design), we obtained the upper

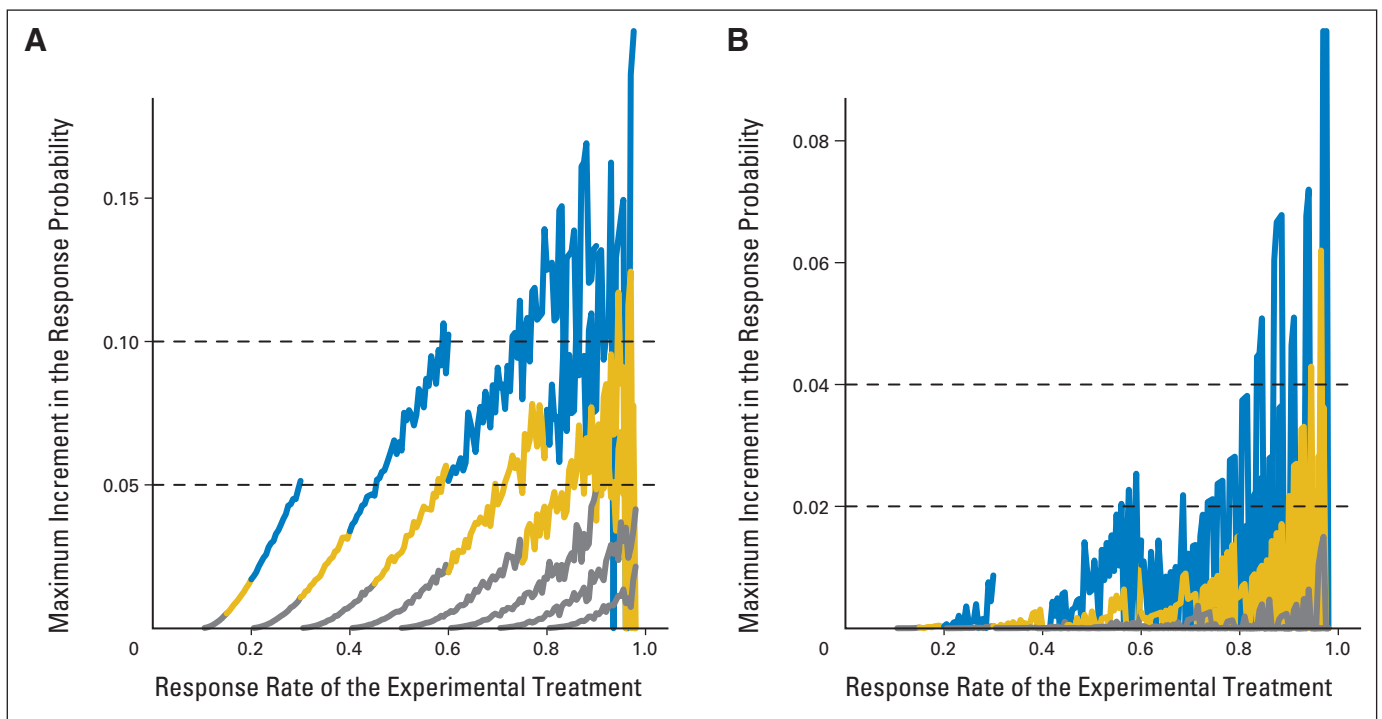


Fig 2. The maximum increment of the probability that a patient will be a responder for adaptive randomization with respect to that for (A) 1:1 fixed-ratio randomization and (B) 2:1 fixed-ratio randomization. With respect to the curves from left to right, the response rates of the standard treatment are 0.1, 0.2, ..., 0.8, respectively, and the gray, yellow, and blue segments of each curve indicate that the response rate of the experimental treatment is 0% to 50%, 50% to 100%, and 100% to 200% higher than that of the standard treatment, respectively.

bound of $\pi_{AR} - \pi_{FR}$ on the basis of a numerical search. To focus on the comparison of the probability that a patient will be a responder, we used a controlled setting in which the adaptive and fixed-ratio randomization designs have the same expected number of nonresponders (in addition to the same type I and type II error rates). Under these conditions, the sample size for adaptive randomization designs must be between $n_{FR}/(1 - p_s)$, when all patients are assigned to the standard treatment, and $n_{FR}/(1 - p_e)$, when all patients are assigned to the experimental treatment, for given values of $p_s < p_e$ and K . We enumerated all possible sample sizes between this range and all possible allocations given each of the sample sizes. We then identified the optimal adaptive design that yields the highest increment of the probability that a patient will be a responder, that is, maximum $(\pi_{AR} - \pi_{FR})$, which also has a power of at least $1 - \beta = .9$ with a type I error rate of $\alpha = .1$. Figure 2A shows that, compared with the 1:1 fixed-ratio randomization, although the maximum increment using adaptive randomization can be above 10%, it is mostly below 5% in the common cases in which the response rate of the experimental treatment is 0% to 100% higher than that of the standard treatment (depicted as the gray and yellow segments in the curves).

Because the sample size is discrete (ie, integers), the curves in Figure 2A are not smooth. The lack of smoothness becomes quite obvious toward the end of the curves because at that point in the curve, the required sample size is small (as the difference in the response rates between two treatments becomes large), which makes the discreteness of the sample size more influential. When compared with 2:1 fixed-ratio randomization, the maximum increments of the probability that a patient will be a responder become even smaller and are often less than 2% (see Fig 2B).

In summary, we have obtained the maximum gains that adaptive randomization designs can achieve with respect to fixed-ratio randomization designs. Our theoretic results are generally consistent with the numerical findings of Korn and Freidlin,¹ which suggests that outcome-adaptive randomization might not have substantial advantages over fixed-ratio randomization when the response rate of the experimental treatment is not substantially higher than that of the standard treatment. Because our results do not depend on simulation comparisons with any specific adaptive randomization methods, they generalize and validate the findings of Korn and Freidlin in an important way.

Ying Yuan

The University of Texas MD Anderson Cancer Center, Houston, TX

Guosheng Yin

The University of Hong Kong, Hong Kong

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

REFERENCES

1. Korn EL, Freidlin B: Outcome-adaptive randomization: Is it useful? *J Clin Oncol* 29:771-776, 2011
2. Thall PF, Wathen JK: Practical Bayesian adaptive randomisation in clinical trials. *Eur J Cancer* 43:859-866, 2007
3. Rosenberger WF, Stallard N, Ivanova A, et al: Optimal adaptive designs for binary response trials. *Biometrics* 57:909-913, 2001

DOI: 10.1200/JCO.2010.34.5330; published online ahead of print at www.jco.org on March 21, 2011

Acknowledgment

Supported in part by the National Cancer Institute, Grant No. R01CA154591-01A1 (Y.Y.), and by Grant No. 784010 from the Research Grants Council of Hong Kong (G.Y.).