

Block Adaptive Randomization

John D. Cook
Department of Biostatistics
P. O. Box 301402, Unit 1409
The University of Texas, M. D. Anderson Cancer Center
Houston, Texas 77230-1402, USA
`cook@mdanderson.org`

January 7, 2011

Abstract

This note proposes a block-adaptive randomization method to limit the length of runs in an outcome-adaptive randomized trial.

1 Motivation

Researchers want to conduct randomized trials, but they are not always happy with the results that randomization produces. Even Ronald Fisher admitted that he would tweak an unfavorable randomization. L. J. Savage reports the following exchange with Fisher [1].

“What would you do,” I had asked, “if, drawing a Latin square at random for an experiment, you happened to draw a [checkerboard]?” Sir Ronald said the thought he would draw again and that, ideally, a theory explicitly excluding regular squares should be developed.

One of the unwelcome possibilities of randomization is a run, a sequence of identical assignments. Runs are not rare in equally randomized trials and they are more common in outcome-adaptive randomized trials because randomization probabilities may exceed $1/2$. Runs are undesirable because they

present a kind of regularity that randomization is intended to prevent. They also cause researchers to question the validity of the procedure producing the randomized assignments.

Basic probability shows that lengthy runs are unlikely, but unlikely events occur when one conducts numerous clinical trials. We have conducted enough adaptively randomized trials at M. D. Anderson Cancer Center to receive a handful of inquiries regarding assignment runs that appear suspiciously long.

2 Method

Equally randomized trials often use block randomization to limit the length of runs. The analogous strategy for adaptive randomization is not immediately clear because the randomization probability is continuously changing in response to new outcomes. We propose the following approach.

Determine a minimum block size m and a maximum block size M . Treatment assignments will be made in blocks of size b where $m \leq b \leq M$. Based on all available patient outcomes, a randomization probability p is computed. See, for example, [2]. The block size b is selected so that p can best be approximated by a fraction with denominator between m and M . For example, suppose a trial has a minimum block size of $m = 4$ and a maximum block size of $M = 8$. If the randomization probability p is 0.68, a block size of $b = 6$ would be used with four assignments to one arm and two to the other.

Each block must contain at least one assignment to each arm. This means that the probability of assigning each arm must be at least $1/M$ and no more than $(M - 1)/M$. The longest possible run would then have length $2M - 2$. This would occur when the inferior arm is the first assignment in one block and the last assignment in the following block.

This method has the benefits of a random block size. When using a fixed block size, a researcher who knows the current randomization probability would know in advance at least the last assignment in a block and possibly other assignments. But this is not possible when the block size is unknown.

The proposed procedure uses a deterministic method for determining block sizes. In theory, a determined researcher could predict the block size by carrying out the necessary calculations. However, the block sizes are not

intuitively predictable. For example, the following sequence of block sizes came out of one simulated trial that had a minimum block size of $m = 2$ and a maximum block size of $M = 8$.

6, 7, 8, 5, 2, 3, 7, 5, 2, 2, 8, 5, . . .

Once a block of b assignments has been computed, the next b patients will be given these assignments. Patient outcomes will have no bearing on randomization probabilities until the assignments in the current block have been handed out. Once a block of assignments is exhausted, all available data will be used to compute the randomization probability p for the next block.

3 Operating characteristics

Preliminary simulations have not shown any differences in the operating characteristics of the proposed method compared to the method implemented in [3]. (In simulations using [3], the minimum randomization probability was selected to be $1/M$ to match the method proposed here.) The selection probabilities for each arm and the number of patients assigned to each arm have been essentially equal, the differences being well within simulation noise.

This is to be expected. Runs are uncommon and so contribute little to the average behavior of the method. While limiting run lengths may not be important for average performance, it can matter to a particular trial.

4 References

- [1] Colin Howson and Peter Urbach. *Scientific Reasoning: The Bayesian Approach* (1989). Open Court Publishing Company.
- [2] J. Kyle Wathen and John D. Cook. *Power and bias in adaptively randomized clinical trials* (2006). Technical Report UTMDABTR-002-06.
- [3] Adaptive Randomization software available on the M. D. Anderson Biostatistics software download site at <http://goo.gl/CIw55>