More Balanced Treatment Allocation when Randomization by Center

Ruji Yao[1], Norman Y. Yao[2]
[1]Merck & Co., Inc., NJ, USA
[2]Miller Institute for Basic Research Science, CA, USA

Abstract
For a multi-center clinical trial with randomization by center, the final treatment allocation depends on two random factors: 1) the random code generated prior to and used during the trial and 2) variations in the number of enrolled subjects and their order of randomization. For a design based upon equal assignment over all treatment arms, imbalances are accumulated over all centers. To reduce the noise induced by the first factor, we have introduced a Latin-squares method in [1] to generate a random code for each center. To approach the second issue, we have defined an "optimum" situation in [1], which assumes that the predicted enrollment numbers for each center at the start of a trial are the same as the actual enrolment. Then, we can combine centers with similar predicted size into a Latin-square that achieves treatment balance independent of the order of the randomly generated code. Although the optimum situation rarely arises in practice, it suggests that any information about the final enrollment may help in the overall Latin-squares balancing. Here, we introduce a simple modified Latin-squares method, which assumes that the IVRS system can determine when the final block is being assigned to a given center. We use real world data to demonstrate our approach and to quantify the improvement.

Key Words: combinatorics, randomization by center, Latin-squares

1. Introduction

The question of balancing the final treatment allocation in a large-scale clinical trial is an important practical issue that underlies many pharmaceutical studies [1-7]. Previously, we have introduced a method [1] for treatment allocation, which uses Latin-squares to reduce the imbalance of allocation. Using the idea of an "optimum" situation as presented there, here, we will introduce a modified Latin-squares method to further improve the overall imbalance by reducing the variation of the number of subjects enrolled over all centers. This modification will need some additional information from each center, so we assume that the IVRS system can only use the last block of subjects with the Latin-squares method for treatment assignment. First we discuss the imbalanced treatment allocation from two real world studies and demonstrate how the Latin-squares method can be used to reduce the imbalance. Then, we will show the improvement with this modified method on the same data sets for comparison.

2. Results from two real studies with the issue of imbalance of treatment allocation

- Study 1: A total of 336 subjects were equally randomized in 6 arms with a block size of 6. The observed number of subjects in 6 arms were n=49, 50, 53, 60, 60, 64 and the maximum difference between 2 arms was 15 (64 - 49).
- Study 2: A total of 240 subjects were equally randomized in 3 arms with a block size of 3. The observed number of subjects in 3 arms were n=70, 77, 93 and the maximum difference between 2 arms was 23 (93 - 70).

Below is a summary of simulation results, based on actual time and center of enrollment for each subject. Treatment allocation files were generated with different random seeds.

		Maximum Difference of Each Run			
study	Runs	min	mean	std	max
1	200	4	13.3	4.4	26
2	200	0	9.2	5.3	28

For study 1, the range of maximum is from 4 to 26 and the observed maximum 15 is at the 62% percentile based on 200 simulation runs. For study 2, the range of maximum is from 0 to 28 and the observed maximum 26 is at the 97% percentile. The key issue is that we have no control on such results.

3. Results with Latin-squares Method on data of these two real studies

Let us now consider implementing a Latin-squares based method for these two real studies. Below is the summary of simulation results based on the Latin-squares allocation approach.

For study 1, the maximum difference is decreased by 31% from 26 to 18.
For study 2, the maximum difference is decreased by 50% from 28 to 14.

		Maximum Difference of Each Run			
study	runs	min	mean	std	max
1	200	2	8.23	3.1	18
2	200	0	5.89	3.1	14

4. Remove the variance in actual enrollment across all centers (optimum situation)

Assuming that the predicted numbers of enrollment for all centers are exacted matched with the actual enrollment numbers can provide an advantage. Under this assumption, the variance in the numbers of enrolled subjects over all centers is removed and the results are summarized below:

		Maximum Difference of Each Run			
study	runs	Min	mean	std	max
1	200	0	2.83	0.91	5
2	200	0	1.37	0.93	2

For study 1, the maximum difference has decreased by 81% from 26 to 5.
For study 2, the maximum difference has decreased by 93% from 28 to 2.

This is optimum situation. Although the predicted enrollment may not be the same as the actual enrolled number, it provides an upper bound for the method and also a hints of how to improve the method in a more practical way.

5. Practical approach in reducing the variance in enrollment across centers

When looking at the output from simulations, one notices that only the last block of each center has impact on the cumulated imbalance of treatment allocation. For example, if a center enrolls 11 subjects in a study of block size 4, only the 9^{th} to 11^{th} enrolled subjects need to be combined with subjects from other centers to be benefit from the Latin-squares method. The 1^{st} to 8^{th} subjects cannot be further balanced and hence can only increase effective noise on the system.

With this motivation, we assume that the IVRS system can obtain information from any center to determine when the new block code is the last block for the center in the study. Only the last block of subjects will be needed and included in the Latin-squares method.

For study 1, 11 blocks or 66 subjects excluded.
For study 2, 26 blocks or 78 subjects excluded.

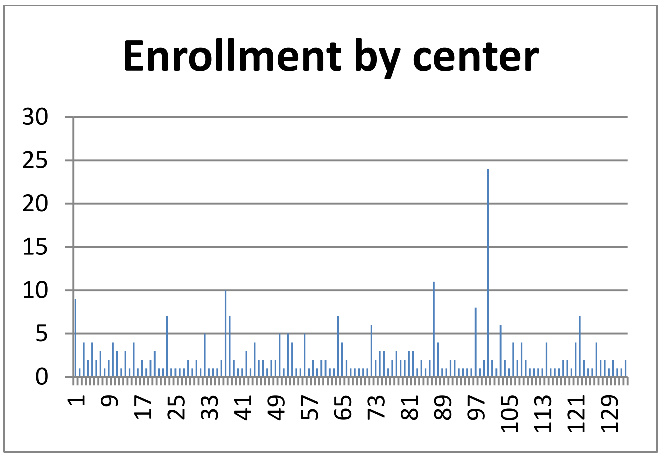
Below is simulation results with the new approach.

		Maximum Difference of Each Run			
study	runs	Min	mean	std	max
1	200	2	7.69	2.80	15
2	200	0	5.17	2.89	13

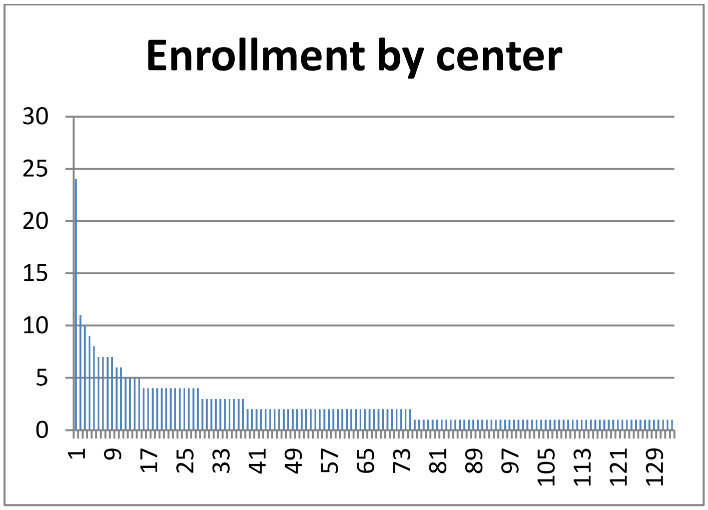
For study 1, the maximum difference is decreased by 42% from 26 to 15.
For study 2, the maximum difference is decreased by 54% from 28 to 13.

6. Graphical display of idea in modified Latin-squares method

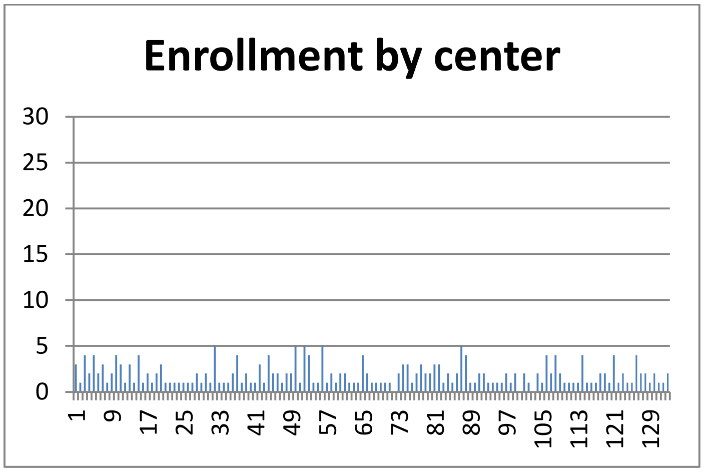
6.1 Plot of number of subjects enrolled (range is from 1 to 24)



6.2 Plot of the number of subjects enrolled (sorted by enrollment numbers. Assuming the numbers are the same as predicted number at the start of the trial. This information is used in the optimum situation to achieve the best possible improvement by a Latin-squares method.)



6.3 Plot of number of subjects enrolled (only the last block of subjects needs to be randomized with the Latin-squares method. All other blocks are already ideal.)



7. Conclusion

As expected, all three Latin-squares based methods outperform the generic centralized randomization approach. In conclusion, we have proposed and analyzed a combinatorics based approach to reducing finite-size imbalances in clinical trials which requires randomization by center.

References

[1] R. Yao and N. Yao, “Balancing Treatment Allocation and Randomization with Combinatorics” Biometrics Section-JSM 2013 Proceedings.

[2] P. Armitage, “The Role of Randomization in Clinical Trials,” Controlled Clinical Trials, 1, 345–352 (1982).

[3] L.A. Kalish and C.B. Begg, “Treatment Allocation Methods in Clinical Trials: A Review,” Statistics in Medicine, 4, 129–144 (1985).

[4] M. Buyse, Applied Cinical Trials, Volume 9, Number 6, pages 32-37, August 2000.

[5] S.J. Pocock and R. Simon, “Sequential Treatment Assignment with Balancing for Prognostic Factors in the Controlled Clinical Trial,” Biometrics, 31, 103–115 (1975).

[6] L.A. Kalish and C.B. Begg, “The Impact of Treatment Allocation Procedures on Nominal Significance Levels and Bias,” Con- trolled Clinical Trials, 8, 121–135 (1987).

[7] D.J. Torgerson and C. Roberts, “Randomisation Methods: Concealment,” British Medical Journal, 319, 375–376 (1999).