

Keep your balance

As drugs for orphan diseases have limited target populations, clinical programmes tend to be small in terms of patient numbers and trials must be designed to account for this. **Dr Eva Miller** and **Jim Murphy** discuss the role of probabilistic baseline adaptive randomisation in ensuring that patients are comparable for all selected covariates

The challenge of developing new medical treatments while minimising patient risk is nothing new to the pharmaceutical industry. However, due to steeply rising drug development costs and escalating patient safety concerns, there is increasing pressure on pharmaceutical companies to re-examine traditional clinical trial techniques and find ways to increase the efficiency and safety of the clinical trial process. Where this becomes particularly important is in studies involving orphan drugs, medical devices, surgical procedures, and terminally ill patients. In these clinical situations, small available patient populations make maintaining balance among treatments especially demanding.

The concept of balance hinges on the need to ensure that patients on different treatment arms are well-matched for all factors, other than treatment, that might affect outcome. These covariates depend on the disease, but may include age, gender, concomitant medications and co-existing conditions.

In a large trial, the natural tendency is to converge towards achieving balance simply by chance as more patients are enrolled. With smaller trials, or larger trials with many sub-groups, a greater degree of rigour is required.

Adaptive trial design is one way to address this challenge, and this is receiving significant attention from both pharmaceutical companies and regulatory agencies. It is a conceptual clinical trial methodology that allows for modifications to take place after the trial has started, without compromising the scientific method. Examples of adaptive trial modifications include combining Phases II and III (see Figure 1), dropping a treatment arm, modifying the sample size, simply stopping a study early for success or failure, or – the focus of this article – probabilistic baseline adaptive randomisation. For all forms of adaptive design, the importance is in maintaining scientific validity – achieved by clear delineation in the protocol of the decision-making rules governing the adaptation.

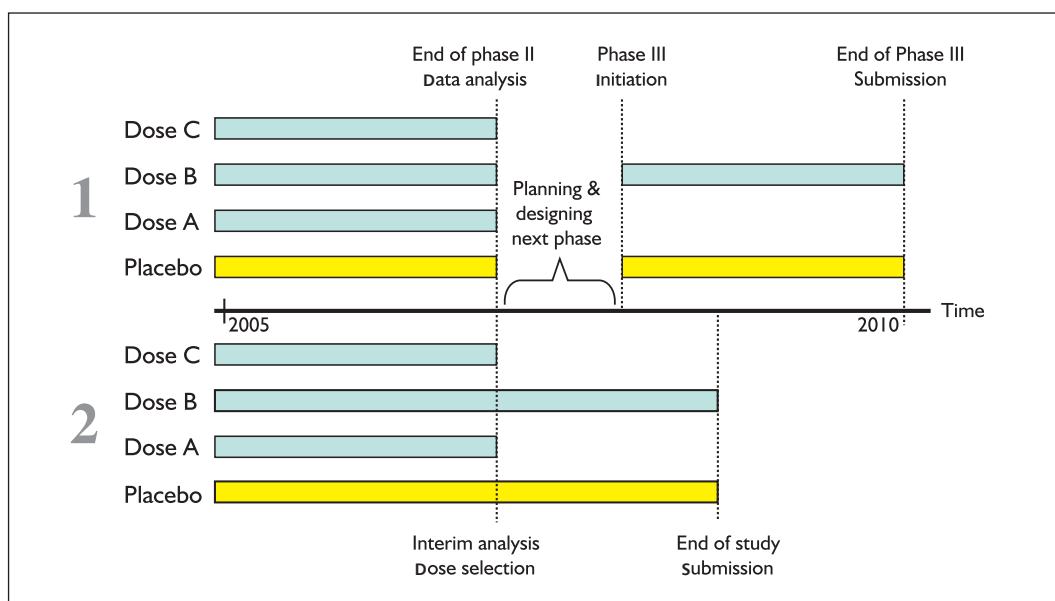


Figure 1: Comparison of traditional (1) and adaptive trial (2) designs for Phases II and III. An adaptive trial design combining the two phases saves time and, as patients from Phase II may participate in Phase III as well, provides additional long-term data.

Randomising to avoid imbalance

Probabilistic baseline adaptive randomisation (pBAR) ensures better balance than is possible using simple randomisation, stratified randomisation or traditional minimisation techniques. It is different from the minimisation techniques introduced in the 1970s in that some element of chance is applied to every randomisation decision, not just to tie-breaking decisions. This is an important distinction – the methodology is then probabilistic rather than deterministic, which means that the investigator cannot know for certain which treatment group the next patient will be assigned to, even if the study is unblinded.

Probabilistic baseline adaptive randomisation assigns a ‘weight’ to each of the treatment arms as every new patient is enrolled. The weight depends on how many patients are already enrolled in each treatment arm. The weight, or chance of that treatment arm being chosen, is increased if the group is smaller than the others – in other words, the chance of a particular treatment group being chosen is pushed in favour of the assignment that achieves the greatest balance. The important point is that there is still a chance that this will not be the treatment that is assigned however – and the FDA has verbally indicated that there must be at least a 10% chance that this will be the case. This is critical in ensuring the investigator cannot predict which treatment will be assigned next.

There are two methods of pBAR, which differ in the mathematics behind the probability of each treatment being chosen. The first and most commonly used method was developed by Pocock and Simon in 1975.¹ The ‘range’ form of this method involves hypothetically assigning the next subject to each treatment group in turn (see Figure 2). The difference between each pair of treatment groups is then calculated and accordingly, random numbers are assigned to certain treatments. The subject is then designated a

treatment group according to the random number that is allotted to him or her. The greatest likelihood of treatment assignment is to the treatment group or groups with the fewest patients enrolled. Pocock and Simon also set out a dynamic randomisation hierarchical or step-wise approach that uses boundary values to determine which prognostic factor is critical for the balancing decision as each patient is enrolled. This will make one factor over-riding if, by chance, the balance is becoming markedly disturbed.

The second method, originally described by Wei²⁻⁴ and refined by Schoeten,⁵ is known as the ‘urn’ method of randomisation (see Box, page 23). The name comes from the fact that each prognostic factor level is likened to an urn, filled with balls. The idea can be compared to a lottery, where the balls represent the different treatments. When a treatment group is under-represented for a particular factor, more balls are added to represent that treatment group and thus the chances of it being picked, at random, are increased.

The design of an algorithm

At this time, there are no definitive criteria for setting up a pBAR algorithm and the appropriate factors to include in the algorithm must be established specifically for each clinical study. Rosenberger and Lachin offer some excellent statistical explanations on approaches to adaptive randomisation model development.⁶ As a starting point, information about prognostic factors can be derived from clinical judgement, previous studies or historical values reported in clinical literature. After prognostic factors (at baseline) are chosen, decisions need to be made about the nature of the statistical methodological framework, the priority or weight given to each factor within the decision-making rule (even the decision not to use weights actually means that all factors are equally important), and choice of propor-

	T1	T2	T3	[T1-T2]	[T1-T3]	[T2-T3]	Maximum difference
Factor 1	0	0	1	0	1	1	1
Factor 2	3	3	3	0	0	0	0
Difference pre-randomisation							1
Hypothetical values if subject is assigned to treatment arm T1:							
Factor 1	1	0	1	1	0	1	1
Factor 2	4	3	3	1	1	0	1
Difference							2
Hypothetical values if subject is assigned to treatment arm T2:							
Factor 1	0	1	1	1	1	0	1
Factor 2	3	4	3	1	0	1	1
Difference							2
Hypothetical values if subject is assigned to treatment arm T3:							
Factor 1	0	0	2	0	2	2	2
Factor 2	3	3	4	0	1	1	1
Difference							3

Next sequential number from random number list	Treatment
If Diff_{T1} = Diff_{T2} = Diff_{T3}	Assign:
0.0001 ≤ random number ≤ 0.3333	T1
0.3333 < random number < 0.6666	T2
0.6666 < random number < 1.0000	T3
If Diff_{T1} < Diff_{T2} ≤ Diff_{T3}	Assign:
0.0001 < random number ≤ 0.8000	T1
0.8000 < random number < 0.9000	T2
0.9000 < random number < 1.0000	T3
If Diff_{T1} = Diff_{T2} < Diff_{T3}	Assign:
0.0001 ≤ random number ≤ 0.5000	T1
0.5000 < random number < 1.0000	T2

Figure 2: The Pocock and Simon range method of pBAR. The table (left) shows the hypothetical assignment of a new patient, who has both Factor 1 and Factor 2, to each of the treatments. The difference is then calculated and the most favourable treatment assignment noted (in this case, either T1 or T2). The table (right) illustrates the relationship between the calculation of difference and the assignment of random numbers. Thus, in this case, a random number of 0.0001–0.5 would assign the patient to T1, whereas a number >0.5 would assign the patient to T2.


tional assignments among treatment groups.

The contributions of the prognostic factors, their hierarchical structure and/or their weighting and how these affect balance need to be thoroughly tested to:

- Address the robustness of the random adaptive algorithm to conditions not under the researcher's control, especially the order in which the subjects enrol.¹ 'Re-sequencing' – a technique for re-ordering patients in a simulation and checking the resultant balances while holding the random adaptive algorithm and random numbers constant – is applied
- Verify the degree to which chance alone will alter the overall balance and the balance of strata or sub-groups. This may be accomplished by using multiple random number schedules against the same simulation dataset. Some authors refer to this as 're-randomisation'
- Validate the implementation within a phone or Internet-based interactive clinical technology by verifying that all boundary values are correct, that definitions stated in the algorithm are appropriately programmed, and that the information from the randomisation module about each patient is being appropriately captured.

Randomisation reports are designed to facilitate the decision-making process by showing resultant balances overall and for sub-groups or strata. The algorithm selected must demonstrate a high level of convergence for the balancing criteria – it must be successful in achieving balance. Moreover, individual listings of subjects randomised (in simulations and later in the live study) must show the stratification factors of the subjects, the randomisation counts on which the treatment assignment was made, the random number, the probabilities, and the logical decision-making rules (in footnotes).

A valuable tool for small populations

Probabilistic baseline adaptive randomisation approaches are flexible, easy to customise for specific clinical trial statistical designs and relatively easy to implement using phone or Internet-based interactive clinical technology. They are especially valuable in clinical trials with small and moderate sample sizes and studies in which sub-group analyses are planned. It is very important to do thorough simulation testing on the algorithm before implementation because there is no single definitive algorithm for each clinical trial protocol. A random adaptive algorithm can be used with confidence when simulation testing results in a high degree of convergence, regardless of the random number schedule, the order of patients enrolling in the study, or the background characteristics of the simulated subjects. While planning and implementing pBAR does require significant effort, the benefit that is achieved by maximising the safety and efficacy information generated from each patient results in a high return on investment. 

Adaptive trial design: A case study using urn pBAR

In this hypothetical example, a trial is designed for a potential treatment for immune thrombocytopenic purpura, a rare autoimmune disease in which the body destroys its own platelets, damaging clotting mechanisms. The design requires consideration of two factors (Factor 1 and Factor 2), which are further split into levels. Factor 1 can be either of two levels (i and ii), whereas Factor 2 can be split into three levels (x, y or z).

There are three treatment arms (A, B and C), representing different dosage levels of the active study drug. Factor 2 is merely used as a clinical attenuation, as it is inappropriate to deliver the medication at certain dosages to those patients with certain levels of Factor 2. Accordingly, patients with level x of Factor 2 are only to be assigned to treatments B and C, and patients with level z of Factor 2 should only be assigned to treatments A and B. Patients with level y of Factor 2 may be assigned to any of the treatments. Since Factor 2 is being used as a statistical control, balancing is not required for Factor 2.

Patient Q is next to enrol. He has a Factor 1 of level i and Factor 2 of level y, so he can be assigned to any of the three treatments. Of the enrolled patients who match Patient Q in terms of Factor 1, nine are assigned to treatment A, six to treatment B and seven to treatment C. On the basis of these figures, the mathematical algorithm assigns the following 'weights' to the treatments: A = 0.25, B = 0.375 and C = 0.375.

Therefore the treatments are assigned within the following random number ranges:

$0.0001 \leq \text{random number} < 0.2500$	A
$0.2500 \leq \text{random number} < 0.6250$	B
$0.6250 \leq \text{random number} < 1.0000$	C

Patient Q is randomly allotted 0.7834. He is assigned to treatment C.

References

- 1 SJ Pocock, R Simon. 'Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial', *Biometrics*, 31, pp103–115, 1975.
- 2 LJ Wei. 'A class of designs for sequential clinical trials', *Journal of the American Statistical Association*, 72, pp 382–386, 1977.
- 3 LJ Wei. 'An application of an urn model to the design of sequential controlled clinical trials', *Journal of the American Statistical Association*, 73, pp559–563, 1978.
- 4 LJ Wei. 'Properties of the urn randomisation in clinical trials', *Controlled Clinical Trials*, 9, pp345–364, 1988.
- 5 HJA Schouten. 'Adaptive biased urn randomization in a small strata when blinding is impossible', *Biometrics*, 51, pp1529–1535, 1995.
- 6 WF Rosenberger, JM Lachin. *Randomization in Clinical Trials: Theory and Practice*, Wiley, 2002.

Eva Miller, PhD

Associate Director, Biostatistics,
Interactive Clinical Technologies Incorporated (ICTI),
Yardley, PA, US

Jim Murphy

Vice-President, Business Development and Marketing,
ICTI, Yardley, PA, US

Tel: +1 267 685 4284
Fax: +1 267 685 4262
E-mail: eva.miller@icti-global.com