Implementing Probabilistic Baseline Covariate Adaptive Randomizations for Clinical Trials

Primary Interest Area: Advances in Clinical Trial Methodologies; Probabilistic Baseline Covariate Adaptive Randomizations; Clinical Trial Technological Applications

Authors: Eva R. Miller, Ph.D, Associate Director, Biostatistics and Jim Murphy, Vice President

Affiliation: Interactive Clinical Technologies, Inc. (ICTI), 1040 Stony Hill Road, Suite 200, Yardley, PA 19067

Synopsis: The average expenditure related to development of a drug candidate has doubled over the last decade. Each day that informed decisions about a drug candidate are delayed carries with it opportunity costs, tangible costs, and patient risk. Adaptive trial design is a clinical methodology that is receiving significant attention from pharmaceutical companies and regulatory agencies because it can increase flexibility and therefore efficiency within the clinical trial process. In an adaptive trial, flexibility is derived from decision making rules that are defined in the protocol. The most controversial of these rules is the allocation rule. The allocation rule is the foundation for probabilistic baseline covariate adaptive randomization (pBAR). Probabilistic baseline covariate adaptive randomization can be used to enhance the statistical design of a clinical trial by balancing treatment groups for known covariates. This is especially critical in trials with high strata to patient ratios, where well designed and implemented interactive clinical technology can be used to enable the implementation of increasingly complex pBAR designs.

Adaptive Trial Design Overview and Benefits:

The challenge of developing new medical treatments while minimizing patient safety 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 and industry thought leaders to reexamine traditional clinical trial techniques and find ways to increase the efficiency and safety of the clinical trial process.

One way to address this challenge that is receiving significant attention from pharmaceutical companies, regulatory agencies, and industry thought leaders is adaptive trial design. Adaptive trial design refers to a clinical trial methodology which allows for data that is accumulated during one stage of a trial to be used to influence the way subsequent stages of the trial are conducted. As a result, clinicians applying adaptive trial techniques may be empowered with greater flexibility than is possible with more traditional clinical approaches. This flexibility does not compromise the scientific method because adaptive modifications are made in response to decision making rules that are clearly defined in the protocol. According to Dr. Vladimir

Draglin[1], a leading advocate of carefully designed adaptive trials, the major decision making rules that drive adaptive modifications include:

Sampling Rule: how many subjects will be sampled in subsequent stages
Stopping Rule: when to drop an arm or stop the trial (for efficacy, harm or futility)
Decision Rule: the final decision and interim decisions pertaining to design change (ex. combining phases IIb and III)
Allocation Rule: how subjects will be allocated to the available arms

Adaptive trials can involve any one of the rules listed above, or a combination of these rules. Examples of the adaptive trial modifications that can result from these decision making rules include modifying the sample size, dropping a treatment arm, stopping a study early for success or failure, combining phases (Figure 1), and / or adaptive randomization. Among the rules that influence adaptive modifications, the most controversial and actively debated rule is the allocation rule.

The Role of the Allocation Rule in Covariate Adaptive Randomization

The allocation rule is used in every randomized clinical trial to drive the randomization strategy (the way subjects will be allocated to different treatment arms). In a non adaptive trial, the allocation rule can dictate a noncomplex randomization approach like traditional stratified randomization. However, in an adaptive trial, the allocation can also be used to drive more complex randomization techniques like adaptive randomization. Like simple stratified randomization, the goal of adaptive randomization is to maximize the balance among treatment groups within the analysis subpopulations (strata). However, unlike simple stratified randomization, adaptive randomization uses technology to incorporate balance data accumulated during the trial into each new randomization decision. As a result, a well designed adaptive randomization can help increase the balancing for many factors simultaneously.

The goal of achieving balance for many factors simultaneously is becoming more and more relevant as we begin the transition from the world of blockbuster drugs for mega indications to the world of personalized medicine. Personalized medicine is based on the concept of defining unique but meaningful characteristics of either the patient or the disease form that influence a patient's response to a specific treatment. In order to realize the promise associated with personalized medicine, clinical trials must be designed to demonstrate safety and efficacy within small subgroup populations. Since we must randomize what we plan to analyze (R.A. Fisher[2]), we can expect to see more studies where the number of strata is large relative to the size of the trial (Buyse and McEntegart[3]). Trials of this nature take on properties that are similar to clinical studies involving orphaned drugs and rare conditions, where the available patient population is small relative to the number of strata.

One variety of adaptive randomization that has demonstrated success in achieving balance among studies with high strata to patient ratios is probabilistic baseline covariate adaptive randomization (pBAR). Probabilistic baseline covariate adaptive randomization has gained momentum by enabling better balance to be achieved than is possible using stratified randomization or traditional minimization techniques[4-7] such as those set forth by Taves[4] or Pocock and Simon[5] in the 1970s. Like stratified randomization pBAR is used to balance treatment arms for known covariates, but pBAR has the advantage of being able to handle more factors and does not have the practical problems presented by incomplete blocks often occurring with the use of stratified randomization especially if a trial is stopped.

Probabilistic baseline covariate adaptive randomization (pBAR) is also different from the less contemporary minimization techniques introduced in the 1970's in that some element of chance is applied to every randomization decision, not just for tie-breaking decisions. This is an important distinction because the randomization methodology in pBAR is then probabilistic, rather than deterministic. Also, we are seeing recent examples of complex applications of Pocock and Simon layering multiple replications of a model within a single clinical trial. For example, in an oncology study with five factors and two to three levels per factor, we first consider type of ancillary care administered (as a stratification) before implementing the adaptive randomization algorithm. The technology thus empowers the use of many factors or combinations of factors simultaneously in these allocation rules.

The Debate Surrounding Covariate Adaptive Randomization

An active debate is currently taking place within the statistical community regarding the appropriate application and analysis of various subject allocation techniques. Each subject allocation technique variation applies a different randomization approach to maximize balance among treatment groups within the analysis population. Four randomization approaches at the center of this debate are simple stratified randomization, the optimal design technique, response adaptive randomization and probabilistic baseline covariate adaptive randomization.

Achieving balance among treatment arms can be difficult, especially in studies with more than one strata. As has been clearly documented in the literature, simple stratified randomization is of limited utility if many strata result in too few subjects per stratum (say 2,3,4,or 5). With pBAR, consideration is given to the hierarchy of balancing decisions or the weights of prognostic factors. Hierarchical ordering of prognostic factors or weights of factors may be derived from previous clinical trials or preliminary data. Much consideration goes into determining the appropriate random adaptive algorithm. Atkinson[8] aptly stated: "too small a response to imbalance could lead to nearly random experimentation with a lack of balance in the completed experiment."

Senn[9-12] has argued vigorously against minimization techniques which he does not consider to be superior to simple or stratified randomization. Senn describes minimization as "not based on the theory of experimental design and works on an ad-hoc algorithm that adds together apples and pears". Senn considers Atkinson's optimal design technique[8,13] to be superior to minimization because it is "based upon sound and logical statistical principles[9]".

With Atkinson's[8, 13] optimal design technique, D_A allocates the next subject to treatment group according to the current effect on the diagonal element of the function $(X'X)^{-1}$ of the design matrix that provides the multiplier for the variance of treatment estimate. Atkinson's optimal design technique has an attractive feature in that continuous variables do not need to be categorized. While this method is deterministic as originally developed, it would be easy to introduce a random element into this minimization design. Atkinson's method is rarely implemented because of its complexity and the difficulty of explaining it to non-statisticians. Again to quote Senn "admittedly. The gains from employing this algorithm compared to minimization are modest, but so are the gains of minimization compared to randomization."[9]

Day, Grouin, and Lewis[14] who contributed to the Committee for Proprietary Medicinal Products (CPMP) position on minimization consider that "the scientific community is not of one mind regarding the use of covariate-adaptive randomization procedures." Day, Grouin, and Lewis point out that the direct linkage between randomization and analytical method is a critical underpinning of sound clinical research design and that analytical approaches for minimization are still an open question.

Day, Grouin, and Lewis site research presented at the Society for Clinical Trials and International Society for Clinical Biostatistics on how blinding may be compromised with minimization[15-17]. Their most cogent concerns, however, is both the need for minimization and the appropriate implementation: citing that studies have been compromised by programming algorithms incorrectly, the choice of factors to be included in the minimization algorithm being poorly thought out, or telephone systems and/or web-based systems have been proven unreliable.

The argument for probabilistic minimization techniques is posed cogently by Buyse and McEntegart[3]. Some of their points include:

- A view that the CPMP claims that dynamic methods such as minimization "remains highly controversial" is unfair because it ignores recent methodological literature that encourages wider use of minimization[18,,19]

- Since the method of minimization considers individual factor levels separately from each other, a larger number of factors can be considered than in a stratified randomization design.

- For studies employing minimization techniques the appropriate analytical approach is still an open question: Whether conventional asymptotic tests can be used for analysis or "rerandomization" tests reflecting the order in which patients with different baseline characteristics entered the trial must be employed.

- It has been questioned whether minimization can actually create imbalances with respect to unknown prognostic factors, but this concern has been refuted as Aickin[20] showed conclusively that dynamic randomization was at least as good as simple randomization.

- "An article in a leading medical journal describes minimization as the platinum standard for clinical trials without drawing any adverse response[21].

There is clear divergence of opinion within the statistical community regarding the most appropriate methods of randomization for multi strata clinical trials. Improved balance among treatment arms can be achieved through use of probabilistic baseline covariate adaptive randomization. However, proper implementation of each of probabilistic baseline covariate adaptive randomization requires methodical planning and preparation. While this cost benefit may make pBAR impractical for some studies, significant value can be demonstrated by maximizing the balance in studies with small or fragmented patient populations. Although the debate will continue for some time, we hope that progress continues to be made in defining common ground. The future of personalized medicine literally rests in the balance.

The Development and Implementation of the Probabilistic Baseline Covariate Random Adaptive Algorithm

Interactive Clinical Technologies Incorporated (ICTI) has been implementing adaptive randomization designs in clinical trials since 1999 and has supported more than 75 adaptive randomization studies of varying complexity. Our experience has demonstrated that in order to ensure that the intended randomization objectives are met, each probabilistic baseline covariate random adaptive algorithm requires an extensive series of planning and verification steps.

After the primary efficacy endpoint and analytical model are determined, the appropriate baseline covariate factors to include in the random adaptive algorithm must be established. As a starting point, information about prognostic factors can be derived from clinical judgment, 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 the choice of proportional assignments among treatment groups (See Rosenberger and Lachin[22] for a thorough textual discussion of algorithm development. See also McEntegart[19] delineated some practical considerations for algorithm development such as how best to categorize a continuous variable, when weighting may be appropriate, and choice of the random element.)

The contributions of the prognostic factors, their hierarchical structure and/or their weighting and how these impact balance need to be thoroughly tested. Currently, there is no set of rules for the statistician to use to develop the single definitive random adaptive algorithm or decision making rule; therefore, a goodly amount of thought and work goes into this process for each clinical trial protocol. An adaptive randomization algorithm which follows either a classical or Bayesian theoretical framework is clearly delineated through simulation testing. It is generally agreed that the algorithm selected must demonstrate a high level of convergence for the balancing criteria.

Once the design of the randomization strategy has been established, ICTI works directly with the client to develop system requirements for the random adaptive algorithm. ICTI uses simulation testing during the development phase of the random adaptive algorithm system module to:

- Address the robustness of the probabilistic baseline covariate random adaptive algorithm to conditions not under the researcher's control, especially the order in which the subjects enroll[5]. As Pocock and Simon stated, the random adaptive algorithm must accommodate this form of variation that is not under the investigator's control. "Re-sequencing" -- a technique for reordering seeded data 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 for strata or subgroups. This may be accomplished by using multiple randomization schedules against the same simulation data set. Some authors refer to this as "re-randomization" but we do not use that term since it is also used in studies involving actual patient reassignments to treatment regime.

- Validate the implementation within the IXRS (phone and internet based interactive clinical technology).

Randomization reports are designed to facilitate the decision-making process by showing resultant balances overall and for subgroups or strata. Also, individual listings of subjects randomized (in simulations and later in the live study) show the stratification factors of the subjects, the randomization counts upon which the treatment assignment was made, the random number, the probabilities, and the logical decision making rules (in footnotes).

Case Study Using the Adaptive Biased Urn Randomization in Small Strata When Blinding is Impossible

Schoeten[23] adapted Wei's biased urn randomization methodology especially for situations in which blinding is inappropriate or sample sizes are very small. Urn randomization was introduced by Wei[24-26], a Bayesian statistical theorist. In urn randomization the assignment of probabilities is adapted to the degree of imbalance in relation to the number of patients already entered into the trial. In Schoeten's method extra balls can be added to the urns at the beginning of the study and are added after each randomization. With this methodology, it is not possible for an investigator to guess the next treatment assignment, even with the knowledge of the previous assignments.

In our real-world example clinical-trial we employed Schoeten's methodology for a rare medical condition within a pediatric population. The clinical trial had three treatment groups (labeled A, B, and C) representing different dose levels of the active study drug (see Figure 2). Given the rare medical condition, the total anticipated sample size was small (under 50 subjects). The design called for one continuous factor (Factor 1) with two levels to be considered. The design is a particularly interesting application because it required attenuation of a second continuous factor (Factor 2 with three levels) for clinical considerations: only treatments A and B were administered to patients in the highest level of Factor 2, and only treatments B and C were administered to patients in the lowest level of Factor 2. Balancing was not required for Factor 2; that factor was just used to accomplish the desired clinical dose regulated attenuation. Balance within this randomization design was achieved overall and for Factor 1.

As shown in the example above, probabilistic baseline covariate adaptive randomization is especially easy to tailor to specific protocols.

Conclusions

Baseline covariate adaptive randomization approaches are flexible, easy to customize for specific clinical trial statistical designs and easy to implement using IXR Systems (phone and internet based interactive clinical technology). They are especially valuable in clinical trials of small and moderate sample size and studies in which subgroup analyses are planned. It is very important to do thorough simulation testing on a random adaptive algorithm before implementation because there is no one definitive algorithm for each clinical trial protocol. A probabilistic baseline covariate adaptive randomization 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.

In view of the lively statistical debate about the appropriateness of minimization techniques for subject allocation, it is highly recommended to incorporate the subject allocation procedure into the clinical trial protocol (including the actual probabilistic baseline covariate random adaptive algorithm within an appendix) along with the planned analytical approach and to submit these for regulatory review and approval before undertaking the clinical trial. While planning and implementing probabilistic baseline covariate adaptive randomization does require significant effort, the benefit that is achieved by maximizing the safety and efficacy information generated from each patient results in a high return on investment.

References

1 V Dragalin 'Classification of Adaptive Designs', presentation IBC 2nd Annual Adaptive Designs Conference, Princeton, NJ November 7, 2005.

2 RA Fisher. *Statistical Methods, Experimental Design and Scientific Inference*, J.H. Bennet, ed. (Oxford University, Oxford) 1956

3 Buyse, M. and D. McEntegart (2004). "Achieving balance in clinical trials." Applied Clinical Trials **13**(5): 36-40.

4 DR Taves. 'Minimization: A new method of assigning patients to treatment and control groups', *Clinical Pharmacology and Therapeutics* 15(3), pp.443-453, 1974.

5 SJ Pocock, R Simon. 'Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial', ***Biometrics*** 31, pp 103-115, 1975.

6 SJ Pocock. *Clinical Trials, a Practical Approach*. Chapter 5, Chichester: Wiley, 1983.

7 T Therneau. 'How many stratification factors are "too many" to use in a randomization plan?' ***Controlled Clinical Trials*** 14: 98-108, 1993.

8 Atkinson, A. C. (1982). "Optimum Biased Coin Designs for Sequential Clinical-Trials with Prognostic Factors." Biometrika **69**(1): 61-67.

9 Senn, S. J. (2004). "Unbalanced claims for balance." Applied Clinical Trials **13**(6): 15-16.

10 Senn, S. J. (1995). "A personal view of some controversies in allocating treatment to patients in clinical trials." Statistics in Medicine **14**(24): 2661-2674.

11 Senn, S. J. (2000). "Consensus and controversy in pharmaceutical statistics (with discussion)." The Statistician **49**: 135-176.

12 Senn, S. J. (2004). "Added Values: Controversies concerning randomization and additivity in clinical trials." Statistics in Medicine **23**(24): 3729-3753.

13 Atkinson, A. C. (1999). "Optimum biased-coin designs for sequential treatment allocation with covariate information." Statistics in Medicine **18**(14): 1741-1752.

14 Day, S., J.-M. Groulin, et al. (2005). "Achieving balance in clinical trials." Applied Clinical Trials **14**(1): 24-26.

15 R. Hills, R. Gray, K.Wheatley, "High Probability of Guessing Next Treatment Allocation with Minimisation by Clinician (abstract)," *Controlled Clinical Trials*, 24, 70S(2003).

16 G. McPherson, M.Campbell, D.Elbourne, "Minimisation: Predictability Versus Balance (abstract), *Controlled Clinical Trials*, 24, 133S (2003).

17 N. Scott and G. McPherson, "Minimisation: Muticentre Clinical Trials (abstract)," *Controlled Clinical Trials*, 24, 175S (2003).

18 N.W. Scott *et al*.,"The Method of Minimization for Allocation to Clinical Trials: A Review," *Controlled Clinical Trials*, 23, 662-674 (2002).

19 McEntegart, D.J. (2003). "The Pursuit of Balance Using Stratified and Dynamic Randomization Techniques: An Overview", Drug Informaion Journal **37**: 293-308.

20 M. Aickin, "Randomization, Balance and the Validity and Efficiency of Design-Adaptive Allocation Methods," *Journal of Statistical Planning and Inference*, 94,97-119 (2001).

21 T.Treasure and K. MacRae, "Minimisation: the Platinum Standard for Trials?," *British Medical Journal,* 317, 362-363 (1998).

22. WF Rosenberger, JM Lachin. Randomization in Clinical Trials: Theory and Practice. Wiley, 2002.

23 HJA Schouten. 'Adaptive biased urn randomization in small strata when blinding is impossible', ***Biometrics,*** 51, pp 1529-1535, 1995.

24 LJ Wei. 'A class of designs for sequential clinical trials', ***Journal of the American Statistical Association*** 72: pp 382-386, 1977.

25 LJ Wei. 'An application of an urn model to the design of sequential controlled clinical trials', ***Journal of the American Statistical Association*** 73: pp 559-563, 1978.

26 LJ Wei. 'Properties of the urn randomization in clinical trials', ***Controlled Clinical Trials*** 9: pp 345-364, 1988.

The authors can be contacted at eva.miller@almacgroup.com, and jim.murphy@almacgroup.com

Figure 1 Comparison of Traditional and Adaptive Trial Designs for Phase 2/3

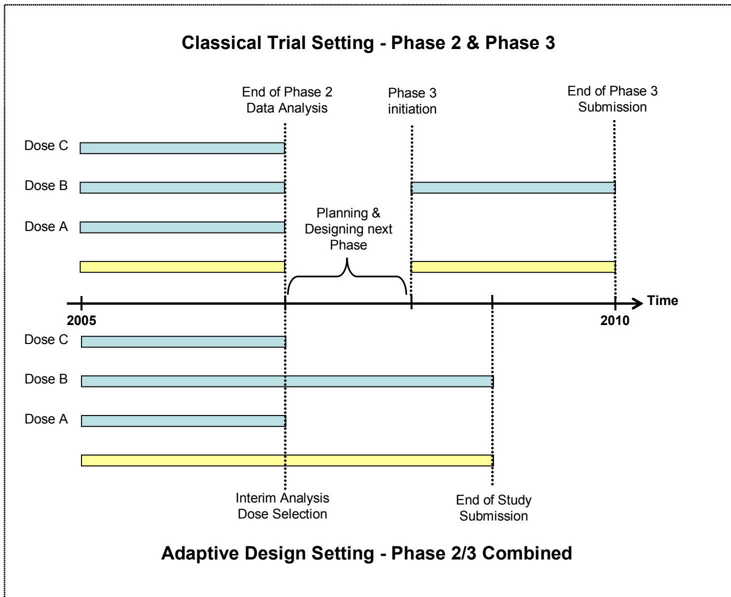

Figure 2 An Example of a Probabilistic Baseline Covariate Adaptive Randomization Design Implementation *(for small sample sizes and when blinding is not possible)*

Adaptive Randomization Using Urn Randomization Methodology

Attenuated Example with Two Factors, Factor 1 and Factor 2
Three Treatment Groups Representing Different Levels of Active Drug — Active Drug

Note to file: from Simulation 2

Subject ID: 111112049

Site Number: 22

Note: If {A+B+C} = 0 then go directly to decision making rules in cells J13..K15.

s= 0

x= 1

	A	B	C
Treatment			
Factor 2: Low level			
Factor 1(i)		1	8
Factor 2: Middle level			
Factor 1(i)	9	6	7
Factor 2: High level			
Factor 1(i)	9	8	

Note: Enter all seven rand count values for each patient being randomized.

	A	B	C
Treatment			
Factor 2: Middle level			
A=	2	2	2
B=	2	2	2
C=	1	1	1
D=	0	0	0
E=	0	0	0
F=	1	1	1
G=	2	2	2
H=	2	2	2
I=	0.375	0.3125	0.3125
J=	0.3333	0.3333	0.3333
K=	-0.0417	0.0208	0.0208
L=	-0.0833	0.0417	0.0417
$P_{(T)}$ =	0.2500	0.3750	0.3750
neg. corrected $P_{(T)}$ =	0.25	0.375	0.375
$P''_{(T)}$ =	0.25	0.375	0.375

Adaptive Randomization Using Urn Randomization Methodology

Attenuated Example with Two Factors, Factor 1 and Factor 2

Three Treatment Groups Representing Different Levels of Active Drug — Active Drug

Subject ID: 111112049 — Factor 1: Level 2

Site Number: 22 — Factor 2: Middle leve

Next sequencial number from rand list:	Random Number:	Treatment Arm Assigned:
520	0.3789	B

			Assign:
If A = B = C = 0			
0.0001	≤ random number < 0.3333	0.3333	A
0.3333	≤ random number < 0.6667	0.6667	B
0.6667	≤ random number < 1.0000	1.0000	C
Factor 2: Low level			
0.0001	≤ random number ≤ 0.5000	0.5000	B
0.5000	≤ random number < 1.0000	1.0000	C
Factor 2: Middle level			
0.0001	≤ random number < 0.2500	0.2500	A
0.2500	≤ random number < 0.6250	0.6250	B
0.6250	≤ random number < 1.0000	1.0000	C
Factor 2: High level			
0.0001	≤ random number < 0.4000	0.4000	A
0.4000	≤ random number < 1.0000	1.0000	B

Figure 3 Interrelationship of Enabling Technologies