Paper 242-26

A SAS® PROGRAM TO PERFORM ADAPTIVE RANDOMIZATION

Carey G. Smoak, Chiron Corporation, Emeryville, CA
Jin-Sying Lin, Chiron Corporation, Emeryville, CA

ABSTRACT

In small to moderate size clinical trials, it is possible for simple randomization (with or without consideration of stratification factors) to result in imbalance of prognostic factors among treatment groups. This imbalance could potentially influence the outcome of the trial. For example, if a greater proportion of "very ill" (as indicated by covariates for severity of illness) subjects are assigned to a particular treatment group, then the outcome of the trial could be influenced by this imbalance. Baseline adaptive randomization improves the chance of having balance among treatment groups with respect to known prognostic factors. This paper describes a SAS program that is a modification of a baseline adaptive randomization procedure proposed by Frane (1998). The procedure is modified to allow for unequal allocation ratios among multiple treatment groups for categorical covariates. The program assigns subjects one at a time to each of the *k* treatment groups ($k \geq 2$), constrained with respect to several pre-specified covariates. Pearson's chi-square test for goodness of fit is utilized to measure treatment imbalance. A new subject is assigned to a treatment group that would achieve the "best" treatment balance. This program was developed with version 6.12 of SAS on a VMS computer.

INTRODUCTION

One potential problem with small to moderate size clinical trials is that simple randomization (with or without taking stratification of prognostic variables into account) may result in imbalance of important covariates among treatment groups. Imbalance of covariates is important because of its potential to influence the interpretation of a trial. For example, the stage of the disease of a subject before the initiation of the study drug could be an important covariate and it would be important to ensure that subjects with more advanced disease are distributed proportionately among the various treatment groups. Suppose that simple randomization was used in a clinical trial to allocate subjects to two treatment groups (one experimental group and one control group). Suppose further that this randomization resulted in imbalance of subjects with more advanced among the two treatment groups, i.e., subjects with more advance disease were not comparable between the two groups studied. Such an imbalance could potentially introduce a bias in the statistical analysis and/or reduce the power of the study.

Baseline adaptive randomization is a procedure which sequentially assigns each new subject to a particular treatment group taking into account the pre-specified baseline covariate values of the new subject as well as all previously randomized subjects. Such adaptive randomization methods can be used to achieve balance in important baseline covariates among treatment groups. The particular type of adaptive randomization procedure presented here is a modification of the procedure proposed by Frane (1998). The procedure is modified to allow for unequal allocation ratios among multiple treatment groups for categorical covariates. For each covariate, treatment imbalance (or deviation from the desired treatment allocation ratio) with the given level of the covariate is evaluated by Pearson's chi-square test statistic for goodness-of-fit. A new subject is assigned to a treatment group that would give us the smallest maximum deviation (this concept is explained in the section titled, "Treatment Allocation") from the desired allocation ratio across the pre-specified covariates. An example clinical trial is used to illustrate this proposed method for performing baseline adaptive randomization. Partial SAS code, based on this example clinical trial, is included.

DESCRIPTION OF THE EXAMPLE CLINICAL TRIAL

The proposed adaptive randomization procedure is illustrated with an example clinical trial. The design of this study is that there are three treatment groups (A, B and C), and subjects are allocated 2:2:1 to treatment groups A, B and C, respectively. The target total number of subjects to be enrolled in this trial is 120. In this example, the first 15 subjects are randomized using simple randomization (e.g., using PROC PLAN). The 16th and subsequent subjects are randomized using the adaptive method. Three baseline covariates (cov1, cov2 and cov3, respectively) are used in the adaptive procedure. In the example, cov1 and cov2 have values of L (for low) and H (for high), while cov3 has three discrete values of 1, 2 or 3.

COVARIATE DISTRIBUTION

Suppose that 25 subjects have already been randomized and the new subject being randomized has the following values: cov1 = H, cov2 = L and cov3 = 2. The distribution of the first 25 subjects for the three covariates is shown in Tables 1-3. Note that only the levels of the covariates corresponding to the levels for the new subject being randomized are shown. The reason for this is that assigning the new subject to a treatment group will only affect the extent of treatment imbalance relating to the actual covariate levels of the new subject being randomized (e.g., the extent of treatment imbalance for the low level of cov1 remains the same regardless which treatment group the new subject, who has high cov1 value, is assigned to).

Table 1. Observed Frequencies for Covariate 1 (H)

	Group A	Group B	Group C	Total
Cov1 (H)	6	8	3	17

Table 2. Observed Frequencies for Covariate 2 (L)

	Group A	Group B	Group C	Total
Cov2 (L)	9	5	4	18

Table 3. Observed Frequencies for Covariate 3 (2)

	Group A	Group B	Group C	Total
Cov3 (2)	2	6	0	8

ADAPTIVE RANDOMIZATION PROCEDURE

To randomize the new subject, the subject is assigned temporarily to each of the three treatment groups and Pearson's chi-square test statistic for goodness-of-fit (testing the desired ratio of 2:2:1) is computed for each of the covariates. Thus a total of nine (three treatment groups times three covariates) statistics are computed. For example, the subject is first assigned to treatment group A, then the corresponding statistics are computed for cov1, cov2 and cov3. The procedure is then repeated for treatment groups B and C, respectively.

TREATMENT ASSIGNMENT

Once the test statistics have been computed, a table of all the test statistics is constructed. For example, Table 4 shows the calculated test statistics for our example of randomizing the 26th subject in this trial. For each treatment, the maximum test statistic (i.e., the largest

test statistic across the three covariates) is then obtained. The new subject being randomized is then assigned to the group with the smallest maximum test statistic. In the example in Table 4, this new subject is assigned to treatment group C because it has the smallest maximum test statistic. If there had been a tie among treatment groups for the smallest maximum test statistic, then the subject would have been randomly assigned (with equal probability) to one of the treatment groups for which there was a tie.

Table 4. Calculated Test Statistics

	Group A	Group B	Group C
Cov1 (H)	0.194	0.750	0.333
Cov2 (N)	1.658	0.605	1.526
Cov3 (1)	3.500	5.722	2.667
Maximum Statistic	3.500	5.722	2.667

The reason for assigning subjects to the treatment group with the smallest maximum test statistic is that, larger test statistics indicate more deviation from the desired allocation ratio (or more treatment imbalance). Therefore, the subject is assigned to the treatment group that gives the "least" treatment imbalance.

SAS CODE:

TEMPORARY ASSIGNMENT TO EACH TREATMENT GROUP

```
**** This part of the program is in a loop.  The loop is      ****;
**** repeated until the new subject has been temporarily ****;
**** assigned to each of the treatment groups.  In this     ****;
**** example the loop is repeated 3 times (once each for ****;
**** groups A, B and C).                                     ****;

**** data temp contains the covariates for all previously  ****;
**** randomized subjects and for the new subject being ****;
**** randomized.                                       ****;

**** &tgnum is treatment group number. ****;

data temp;
      set laball;
      if "&tgnum" = "1" then tgroup = 'A';
      else if "&tgnum" = "2" then tgroup = 'B';
      else if "&tgnum" = "3" then tgroup = 'C';
run;
```

SUMMARIZE COVARIATES AMONG TREATMENT GROUPS

```
proc freq data=temp;
      tables cov1 * tgroup / out=cov1_gp sparse noprint;
run;

**** The proc freq is repeated for each covariate. ****;
```

CALCULATE TEST STATISTIC FOR EACH COVARIATE

```
**** L = Low ****;

%if &cov1 = L %then %do;

      proc transpose data=cov1_gp out=n_cov1;
            var count;
            **** select subjects with low levels of cov1 ****;
            where cov1_gp='L';
      run;

      data gfit1_&tgnum(keep=chisq_&tgnum group);
            length group $10;
            set n_cov1;

            group='COV1_L';

            **** Total (Cov1=L) ****;
            total=col1+col2+col3;

            **** Calculate expected cell values. ****;

            **** &wt1-3 are the weights (2:2:1) for the ****;
            **** treatment allocation.                  ****;

            **** &wt123 is the sum of &wt1 through &wt3. ****;

            exp1=total*(&wt1/&wt123);
            exp2=total*(&wt2/&wt123);
            exp3=total*(&wt3/&wt123);

            **** Calculate the cell chi-square test statistic ****;

            cellchi1=((col1-exp1)**2)/exp1;
            cellchi2=((col2-exp2)**2)/exp2;
            cellchi3=((col3-exp3)**2)/exp3;

            **** Calculate chi-square test statistic ****;

            chisq_&tgnum=cellchi1+cellchi2+cellchi3;

      run;

%end;

**** The calculation of the chi-square test statistic is ****;
**** repeated for each covariate.                        ****;

**** The loop continues until the new subject has been     ****;
**** temporarily assigned to each of the treatment groups ****;
**** and chi-square test statistics have been calculated  ****;
**** for all the covariates.                              ****;
```

CONCLUSION

For small to moderate size clinical trials, the baseline adaptive randomization method could be useful in providing a means to better achieve treatment balance with respect to several important prognostic factors. A SAS program has been developed to perform the adaptive randomization procedure proposed. The SAS program assigns a subject to a treatment group that is associated with the least treatment imbalance.

REFERENCES

Frane, J.W. (1998), " A Method of Biased Coin Randomization, Its Implementation, and Its Validation," *Drug Information Journal*, 32, 423-432.

CONTACT INFORMATION

Carey G. Smoak
Chiron Corporation
4560 Horton Street, MS U315
Emeryville, CA 94608
(510) 923-2304 - work phone
(510) 923-3297 - fax
Carey_Smoak@chiron.com