A SAS macro for constrained randomization of group-randomized designs

M. Ashraf Chaudhary*, Lawrence H. Moulton

Department of International Health, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, MD 21205, USA

Keywords:
SAS®
Constrained randomization
Stratified group-randomized trial

Group-randomized study designs are useful when individually-randomized designs either are not possible, or will not be able to estimate the parameters of interest. Group-randomized trials often have small number of experimental units or groups and strong geographically-induced between-unit correlation, thereby increasing the chance of obtaining a “bad” randomization outcome. It has been suggested to highly constrain the design through restriction to those allocations that meet specified criteria based on certain covariates available at the baseline. We describe a SAS® macro that allocates treatment conditions in a two-arm stratified group-randomized design that ensures balance on relevant covariates. The application of the macro is illustrated using two examples of group-randomized designs.

1. Introduction

The reasons to randomize the units in a cluster-randomized trial include: to assure both the investigator and the general scientific community that there was impartiality of treatment assignment; to avoid hidden biases in treatment assignment; to improve the chance of having a good distribution of relevant characteristics across the treatment arms; to provide a convenient means of implementing treatment assignment; for ease of masking; and to provide a basis for statistical inference [1,2]. The approach implemented in our macro program can help attain the goals implicit in the first three of these reasons, which are important to the success of an experiment, although not the last reason, which is not relevant under model-based inference. The overall goal is not just to raise the probability of a “good” treatment assignment, but to make sure all are satisfied with the resultant allocation.

There are many papers in the statistical literature on group-randomized trials regarding whether to match or not, and how to do the analysis. There is comparatively little information, however, about strategies other than basic pair matching or stratification for achieving a balanced randomization of units to study arms. This may be because the general principles of blocking and stratification are well known. Attention to those principles, however, may be insufficient when dealing with small numbers of groups of highly variable humans who interact in complicated networks.

In individually-randomized designs, larger sample sizes ensure balance on key variables between the trial arms. Community-randomized trials typically have small numbers of randomization units with perhaps only 4–20 groups to be randomized. Even with stratification schemes, such as in pair-matched studies, one can be the victim of “bad luck,” obtaining a treatment allocation that is substantially unbalanced with respect to one or more baseline covariates. In community-randomized trials, we often have many relevant covariates, i.e. the ones that may be related to the study outcome including but not limited to geographic distribution, availability of

---

* The code for the SAS macro CCRA_V1.0 is available at http://www.createbiostats.org/prod01.html or by writing directly to: Mohammad A. Chaudhary.
* Corresponding author.
E-mail address: mchaudha@jhsph.edu (M.A. Chaudhary).
0169-2607/$ – see front matter © 2006 Elsevier Ireland Ltd. All rights reserved.
facilities (such as clinics), mean levels of education and composition, etc. If even one of these characteristics is unbalanced across treatment arms, the study results may be suspect. It is difficult to adjust adequately for group-level covariates, adding to the imperative for balance at the start.

Several authors have explored the role of balancing covariates between treatment groups in the design of cluster-randomized trials and have suggested procedures that can be used to achieve balance using alternative approaches [3–6]. These include highly constraining the design through restriction to those allocations that meet specified criteria. The idea is to constrain so as to ensure marginal balance on relevant covariates—this is similar to the minimization technique for sequentially enrolled trials [6], except that the balancing is done up front [3,4]. When the number of experimental units is sufficiently large and or the number of covariates is small, the standard techniques of blocking and stratification may be sufficient. However, even in such cases it may be wise to simulate and identify the probability of obtaining a “bad” randomization. This was done in a group-randomized trial of pneumococcal vaccine with 38 geographically defined units, with eligible population size as the primary covariate of interest [7]. Many group-randomized trials, however, will have smaller numbers of units, and more covariates for which an imbalance at the start of the trial could prove an embarrassment and complicate inference at the trial’s conclusion.

While constraining the group-randomized designs one should be cautioned against destroying the validity (independence) of the sample design, which may subject the investigator to criticism [8]. As an example, a completely randomized design is valid if each pair of randomization units has the same probability of being allocated the same treatment. Constraining the design may result in departure from uniformity of these probabilities. Other criteria exist for other designs, but the main idea is one of whether there is independence of treatment assignment between units. Simulations have indicated, however, that this may be a problem only in highly unusual situations: significant departures from nominal Type I error may occur when there is severe correlation among units that also happen to be highly jointly affected by constraints [4]. Although departure from strict validity might be expected to affect randomization-based inference [5], this may matter little in practice.

Based on a recently developed covariate-based constrained randomization algorithm for group-randomized trials [4], we developed a SAS® macro CCRA_V1.0 that implements the proposed algorithm. SAS® is one of the most popular statistical analysis software packages among statisticians and epidemiologists involved in the design and analysis of clinical trials, providing a logical base to develop a group-randomization routine. The payoffs of developing such a macro in SAS® include: already working in a SAS® environment; no licensing of special software; no export of data; no repetition; and no replication. The macro may be used on a routine basis or used across multiple projects.

This paper describes the macro and illustrates its application to examples of two-arm parallel stratified as well as unstratified group-randomized designs.

Section 2 briefly describes the covariate-based constrained randomization algorithm. Section 3 describes the SAS® macro in general terms with input and output datasets. Section 4 is devoted to the application of the macro to two artificial examples of unstratified and stratified group-randomized study designs, respectively. Section 5 summarizes the results.

### 2. Methodology: covariate-based constrained randomization

The objective is to achieve a balance or near balance between the treatment arms of a randomized controlled trial in terms of the available baseline covariates or the functions of the covariates that could be related to the primary outcome(s) of the trial. The individual level continuous as well as the categorical covariates would be summarized for each group in the trial. The summary measures may be means for both continuous and categorical variables or counts for the binary variables. For each covariate, a simple caliper-type criterion may be specified: one could specify that the group means for each trial arm be within quarter of the standard deviation of each other, or within >10%, say. For any criterion used, one would need to specify the means for each covariate for each treatment arm allowing acceptable covariate differences between the arms, first within the strata and then for allocation. For a two-arm stratified group-randomized design, all possible allocations or randomization outcomes are generated within each stratum. For example, if a stratum comprises five groups and three to be allocated to the treatment arm and two to the control arm, the all possible allocations for that stratum would be $C_5^3 = 10$. Within each stratum, the allocations that meet the restrictions are short-listed. Using these short-listed (good) within strata allocations, all possible overall randomization schemes are generated by selecting one good allocation from each stratum. At this stage another set of criterion is established (similar to the one set for within strata) to generate a list of overall good randomization schemes for the trial, which should not be too short or too excessive. The stratum level and the overall level balancing criteria may need adjustment and the macro run over and over again in order to get a reasonable number of overall good randomizations from which one is to be selected at random. Finally, one overall randomization scheme is selected at random from this list and the numbers of times the pairs of groups appear together in the same arm are enumerated as a check of the degree of validity of the design. The macro described in the following section carries out the designated restrictions and displays key information regarding the effects of the restrictions.

### 3. SAS® macro CCRA_V1.0

A randomization plan often includes placing units into relatively homogeneous strata that will be accounted for at the time of analysis in order to reduce variance of the estimated intervention effect. We first make sure that within each stratum, there is balance on the relevant covariates between the study arms (Steps 1–3). At this point, we could just choose at random one allocation from each stratum, however, that would not guarantee overall balance; thus, we generate all possible overall allocations based on the allowable stratum-level
allocations (Step 4). Specifically, we generate all possible overall allocations by taking one allowable allocation from each stratum and short-list only those that meet the overall balance criteria (Step 5). Only one is needed for allocation purposes, but many more are needed to check degree of validity (Step 6). If restrictions do not appear too great, the final allocation is selected at random (Step 7).

Step 1: Generate all possible allocations for each stratum separately by forming combinations of groups in each stratum. The groups in a combination are assigned to a study arm (say arm 1) and the remaining to arm 2.

Step 2: Compute means of covariates for each allocation in each arm and combine the data for the two arms by stratum and allocation ids.

Step 3: Make a pass through these allocations and select those that meet the specified criteria (balance, or near balance on set of covariates). The criteria should not be so tight that we are left few allocations in one or more strata or one or more strata are eliminated altogether. Alternatively, the criteria may not be so relaxed that they do not give a good screening of the allocations. So we may accept and go to Step 4 or relax or tighten the criteria as appropriate or change the stratification going back to Step 1.

Step 4: Generate all possible overall allocations by picking one ‘good’ allocation from each stratum for each overall allocation.

Step 5: Retain only those overall allocations that meet the overall level balancing criteria.

Step 6: As a check on validity, count the number of times a group appears with another group in the same study arm.

Step 7: Selects one overall allocation at random from all the short-listed overall allocations.

The macro requires an input SAS® dataset containing the stratum and group ids, the number of groups to be randomized to the study arm, say arm 1 in each stratum, and the group-level covariate data. The input SAS® dataset should be read into a temporary SAS® dataset d and may have the following variables:

- \( s \)  stratum ID
- \( group \)  group or cluster ID
- \( r \)  number to be randomized to study arm 1 in each stratum
- \( x1, x2, x3, \ldots \)  covariates

For the unstratified design the stratum ID variable \( s \) would have the same value repeated across the dataset. The temporary input SAS® dataset \( d \), the variables \( s, group \), and \( r \) must be named as such, however the covariates may have any valid SAS® names. All variables in the dataset must be numeric.

The macro invoked by the following statement requires five macro input parameters.

\[
\%\text{macro CCR(nvars,unames,idl,odl,seed)};
\]

- \( nvars \)  names of covariates (must equal the number of covariates)
- \( idl \)  a string of numeric values equal to the number of covariates separated by ‘*’ representing the maximum allowable differences between treatment arms for each covariate for allocations within strata. This constitutes the initial balancing criteria. For an allocation to be allowable, all the covariates must satisfy the criteria individually for that allocation
- \( odl \)  a string of values equal to the number of covariates separated by ‘*’ representing the maximum allowable differences between treatment arms for each covariate for overall allocations. This constitutes the overall balancing criteria. For an overall allocation to be allowable, all the covariates must satisfy the balancing criteria individually for that overall allocation
- \( seed \)  random number seed to select one final overall allocation from all the acceptable overall allocations

### 4. Implementation examples

#### 4.1. Example 1: unstratified group-randomized design

In planning an HIV vaccine study in four villages with baseline HIV prevalence rates (%): (1, 3, 9, 11), the unrestricted randomization would be unbiased as well as valid but with 1/3 chance of all vaccine study arm villages having highest, or lowest prevalence. On the other hand one may prefer randomly selecting one of the two allocations (C and D) that have exact balance or one of the four allocations (B, C, D, E) with exact or near balance.

<table>
<thead>
<tr>
<th>Allocation</th>
<th>Vaccine</th>
<th>Control</th>
<th>Mean difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1, 3</td>
<td>9, 11</td>
<td>8</td>
</tr>
<tr>
<td>B</td>
<td>1, 9</td>
<td>3, 11</td>
<td>2</td>
</tr>
<tr>
<td>C</td>
<td>1, 11</td>
<td>3, 9</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>3, 9</td>
<td>1, 11</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>3, 11</td>
<td>1, 9</td>
<td>2</td>
</tr>
<tr>
<td>F</td>
<td>9, 11</td>
<td>1, 3</td>
<td>8</td>
</tr>
</tbody>
</table>

Actually, constraining to C and D is not valid either, because for example, 1 and 9 are never on the same arm, and 1 and 11 are always linked together.

And what if the geography of the situation were such that a river cuts through the communities with 1 and 3 on one side and 3 and 11 on the other side of the river? We would be thinking we had effectively two allocation units.

Here is the listing of the input SAS® dataset for this example (Fig. 1).

\[
\begin{array}{rrrrrr}
\text{Obs} & s & group & r & x1 & x2 \\
1 & 1 & 111 & 2 & 1 & 1 \\
2 & 1 & 222 & 2 & 11 & 1 \\
3 & 1 & 333 & 2 & 9 & 3 \\
4 & 1 & 444 & 2 & 3 & 3 \\
\end{array}
\]

Fig. 1 – CCRA_V1.0: input SAS dataset.
validity of the design. Groups 111 and 333 (prevalence 1 and 9) always are in the same arm, as are groups 222 and 444 and the other combinations never appear together. Clearly, this design is not technically valid, although it has enough attractive features that may be acceptable. The final over all randomization (rno 2) is displayed in Fig. 4.

Thus, if arm 1 was the intervention (vaccine) arm, units 222 and 444 (prevalences 3 and 11) would be in that arm, and 111 and 333 would be in the control arm.

In this case, one might relax the prevalence criterion to “mean diff is <3 percentage points” and add “one intervention to the west of the river, one to the east”—this would mean drawing either B (1, 9 get Vac) or E (3, 11 get Vac). If geography were irrelevant we could just exclude A and F, ending up with a more valid design, and less open to accusations of rigging the randomization.

4.2. Example 2: stratified group-randomized design

Now we consider a stratified cluster (clinic) randomized trial currently at the planning stage to determine if the ‘opt out’ strategy of provider-initiated routine HIV counseling at the start of TB treatment increases uptake of counseling in the Eastern Cape Province of South Africa [9]. By including HIV counseling as part of the ‘package’ of a routine medical evaluation administered to all adult TB patients, the TB patients would no longer be asked whether they wanted HIV counseling. Patients could refuse counseling, just as a patient can refuse any aspect of medical evaluations, but by providing HIV counseling universally to all TB patients, it is believed that fewer patients will refuse—will ‘opt out’. A total of 20 clinics broken down into urban, rural and suburban strata are to be randomized to either intervention (opt out) or control (already counseling). Patients could refuse counseling, just as a patient can refuse any aspect of medical evaluations, but by providing HIV counseling universally to all TB patients, it is believed that fewer patients will refuse—will ‘opt out’. A total of 20 clinics broken down into urban, rural and suburban strata are to be randomized to either intervention (opt out) or control (already counseling). Patients could refuse counseling, just as a patient can refuse any aspect of medical evaluations, but by providing HIV counseling universally to all TB patients, it is believed that fewer patients will refuse—will ‘opt out’. A total of 20 clinics broken down into urban, rural and suburban strata are to be randomized to either intervention (opt out) or control (already counseling). Patients could refuse counseling, just as a patient can refuse any aspect of medical evaluations, but by providing HIV counseling universally to all TB patients, it is believed that fewer patients will refuse—will ‘opt out’.

Now we consider a stratified cluster (clinic) randomized trial currently at the planning stage to determine if the ‘opt out’ strategy of provider-initiated routine HIV counseling at the start of TB treatment increases uptake of counseling in the Eastern Cape Province of South Africa [9]. By including HIV counseling as part of the ‘package’ of a routine medical evaluation administered to all adult TB patients, the TB patients would no longer be asked whether they wanted HIV counseling. Patients could refuse counseling, just as a patient can refuse any aspect of medical evaluations, but by providing HIV counseling universally to all TB patients, it is believed that fewer patients will refuse—will ‘opt out’. A total of 20 clinics broken down into urban, rural and suburban strata are to be randomized to either intervention (opt out) or control (already counseling). Patients could refuse counseling, just as a patient can refuse any aspect of medical evaluations, but by providing HIV counseling universally to all TB patients, it is believed that fewer patients will refuse—will ‘opt out’.

Fig. 4 – CCRA_V1.0: finally selected overall randomization.

Now we consider a stratified cluster (clinic) randomized trial currently at the planning stage to determine if the ‘opt out’ strategy of provider-initiated routine HIV counseling at the start of TB treatment increases uptake of counseling in the Eastern Cape Province of South Africa [9]. By including HIV counseling as part of the ‘package’ of a routine medical evaluation administered to all adult TB patients, the TB patients would no longer be asked whether they wanted HIV counseling. Patients could refuse counseling, just as a patient can refuse any aspect of medical evaluations, but by providing HIV counseling universally to all TB patients, it is believed that fewer patients will refuse—will ‘opt out’. A total of 20 clinics broken down into urban, rural and suburban strata are to be randomized to either intervention (opt out) or control (already counseling). Patients could refuse counseling, just as a patient can refuse any aspect of medical evaluations, but by providing HIV counseling universally to all TB patients, it is believed that fewer patients will refuse—will ‘opt out’.

Next, we display the number of times the pairs of groups appear together in the same arm as a check of the degree of
means are then computed for each covariate and for each allocation in each arm dataset separately and merged together to compute the differences in the means between the two arms for each covariate. These differences are compared against the values of the within stratum-level balancing criteria. The allocations which satisfy these initial criteria are then displayed. In this example 42 allocations out of a total of 110 qualify with 30 in stratum 1, 4 in 2, and 8 in 3. Based on these acceptable allocations in each stratum, overall allocations are generated by selecting one allocation from each stratum. Out of a total of $3 \times 30 \times 4 \times 8 = 2880$ possible overall allocations, only 16 allocations qualified the overall allocation level criteria. A listing of these 16 allocations displays the stratum ids and within-stratum allocation ids $rno$. Since there are three strata the number of observations in this display is $3 \times 16 = 48$. We now compute the number of times a unit appears with another unit in the same arm as a check for the validity of allocations. The frequencies are shown for all possible $C_2^2 = 190$ pair-wise combinations of groups to be randomized. Clearly the maximum count can be 16 and minimum 0 as a pair may appear within an arm in all of the 16 qualifying overall allocations or may not appear at all. Lastly one allocation, to be used to implement the randomization for the trial, is selected at random from these 16 qualifying overall allocations and displayed. Most of the listings for this example are too long to be shown here. The final overall randomization produced by the macro is shown in Fig. 6.

![Fig. 5 – CCRA_V1.0: input SAS dataset.](image)

![Fig. 6 – CCRA_V1.0: finally selected overall randomization.](image)

## 5. Discussion

The usual experimental techniques of blocking and stratification would reduce the variance only to a certain limit for group-randomized designs where only a handful of groups or communities are to be randomized. Substantial gains in efficiency may be realized in such situations by exercising covariate-based constrained randomization when the covariates are potentially related to the outcome. Use of constrained randomization may at least guard against ruining a large scale multimillion dollar study for hidden potential imbalance.

The two examples illustrate the application of the macro to group-randomized designs with and without stratification. Although in both the examples we have a balance in terms of the units allocated to the two study arms, the macro handles imbalance and has the flexibility of assigning any number of units to a study arm in a stratum. The macro may also be used for pair-matched studies, in which case within-stratum balance is no longer relevant.

One has to be particularly careful in setting up the balancing criteria and closely examine the output to verify that the criteria are not too tight or too relaxed. We have phrased the criteria in terms of the differences between covariate means. In some situations it may be preferable to constrain the ratio of means, in which case the macro may easily be modified. An important limitation of the macro is that it is currently applicable to two-arm designs. However, the randomization of groups to more than two study arms may be achieved with repeated application of the macro.

When venturing into the territory beyond the conventional stratified designs (randomized blocks, pair-matched, and so on), one runs the risk of producing a design that technically is biased or not valid. In the design of experiments nomenclature, a design is biased if, across the randomization units, there is any difference in probability of assignment to a given treatment. This problem rarely arises in practice. More problematic is the validity of a design: a completely randomized design is valid if each pair of randomization units has the same probability of being allocated the same treatment [8].

The SAS® macro discussed is simple to use; however, a certain level of experience in working with SAS® is required.

### Acknowledgements

This research was supported in part by NICHD grant R01-HD38209 and by the Bill and Melinda Gates Foundation as part of the Consortium to Respond Effectively to the AIDS-TB Epidemic (CREATE) project.

### References
