

MULTILEVEL RANDOMIZATION

Marina Komaroff, Noven Pharmaceuticals, New York, NY

Lois Lynn, Noven Pharmaceuticals, New York, NY

ABSTRACT

Randomization in clinical trials is essential for the success and validity of a study. PROC PLAN is an important SAS® procedure that generates randomization schedules for variety of experimental designs. This procedure was developed for the major types of randomization like simple, block and stratified randomization where the latter controls and balances the influence of covariates. In addition to SAS® documentation, multiple papers were written to explain how to adapt and enhance the procedure with DATA steps and/or PROC FORMAT.

Clinical research in transdermal medicine introduces the situation where a multilevel randomization is required for levels like treatment, location (arm, thigh, back, etc.) and side (left, right, upper, center, etc.) of a patch application while retaining balance at each level and combination of levels. Schedules get especially complicated for cross-over studies where location and side of patch application needs to be rotated by period and balanced as well. To the authors' knowledge, there are no published papers to accommodate these requirements.

This paper introduces a novel concept of multilevel randomization, provides SAS code utilizing PROC PLAN, and a few examples with increasing complexity to generate balanced multilevel randomization schedules. The authors are convinced that this paper will be useful to SAS-friendly researchers conducting similar studies that require multilevel randomization.

INTRODUCTION

The master of statistics R.A. Fisher introduced the concept of randomization along with the blinding technique as tools to minimize bias in selection and allocation of treatments to the experimental unit.^[1] It became the pharmaceutical industry standard to randomize clinical trial interventions. The randomization techniques for major experimental designs were described by Fleiss in a textbook for statisticians.^[2] The SAS ® Institute has developed PROC PLAN to construct the randomization schedules for these major study designs. The SAS/STAT support documentation and multiple papers from conferences demonstrate that the procedure can be adapted for more sophisticated experimental designs like William's Latin Square (WLS).^[3] Programs within this paper support the fundamental concept of randomization where treatment arms are assigned to experimental units (patients) and introduces the novel concept of multilevel assignments.

The multilevel randomization concept is defined for scientific experiments where the experimental units (patients) are randomly allocated to a level-1 parameter which is usually the different treatments under study and simultaneously allocated to a level-2 parameter which is to be randomly allocated to the level-1, and simultaneously to the level-3 parameter that is to be randomly allocated to the level-1 and level-2 parameter, and for each additional level of randomization. For example, a transdermal patch (treatment) must be assigned to a patient, with an anatomical location on the body (e.g., hip, knee) that should be randomly allocated to the patient and to the patch as well, and the same for the side of the body to place the patch (e.g., left, right, and center, upper, lower). These levels may be independent of each other, yet, follow the same specific rules of randomization within one schedule. There is no paper to the authors' knowledge that demonstrates how utilize PROC PLAN for the multilevel randomization situation.

DESCRIPTION

Table 1 summarizes a few examples for multilevel randomization provided in this paper. The illustrative randomization examples are for a phase I, three-way crossover study design with 18 subjects.

EXAMPLE № 1 illustrates how to perform a randomization for level-1, level-2 and level-3 separately and individually using MACRO WLS3x3 where William's Latin Square design was implemented to randomize 18 subjects. The macro can be used for different sample sizes and can be easily adapted for different study designs.

EXAMPLE № 2 demonstrates how to generate a two level randomization schedule, particularly combining level-1 and level-3.

EXAMPLE № 3 demonstrates a three level randomization schedule (level-1, level-2, and level-3).

Example	Level	Variable Name	Format
№1	1	Treatment	A, B, or C
	2	Anatomical Location	Arm, Hip, Knee
	3	Body Side	L(ef), R(ight)
№2	1 & 3	Treatment & Body Side	
№3	1& 2& 3	Treatment & Location & Body Side	

Table 1: Multilevel Randomization

The algorithm and a program for the William's Latin Square Design is described elsewhere.^[3] The authors developed MACRO WLS3x3 that improves and simplifies the task. The macro accommodates an odd and an even number of treatments which is important for the WLS algorithm along with the other requirements: (1) that each formulation occurs only once for each subject; (2) that each treatment formulation occurs the same number of times in each period; and (3) the number of subjects who receive formulation i in some period followed by formulation j in the next period is the same for all $i \neq j$.

MACRO TO GENERATE RANDOMIZATION SCHEDULE (WILLIAM'S LATIN SQUARE)

```

/*%WLS3x3: Parameters for William's Latin Square for 3x3 blocks
(level=1, * Level is an assigned number, also assigned to the output
Levelc=period, * Levelc is the text for each Level of randomization
Trt =%str("A" "B" "C"), * Trt are the treatments for this particular level
Trtn =3, * Trtn is the number of treatments in this particular level
For an even number of trtn the result is a multiple of the provided even number.
For an odd number of trtn the result becomes 2x times the odd number provided.
Block =3) * Block is the number of blocks to accommodate the desired sample size
*/

%MACRO WLS3x3(level=num, levelc=period, trt=mytrt, trtn=3, block =);
PROC PLAN seed=&seed1;
  FACTORS block=&block ordered
    ORDER=&trtn random
    &levelc=&trtn ordered;
  TREATMENTS trt=&trtn cyclic;
  OUTPUT out=firsts&level trt cvals=(&trt) random;
run;
PROC SORT DATA=firsts&level; by block order ; run;
PROC TRANSPOSE DATA=firsts&level out=outss&level prefix=&levelc;
  by block order;
  var trt;
  id &levelc;
run;

*** MIRROR and interlace ***;
DATA outss&level._1(keep=block order &levelc.10 &levelc.15 &levelc.20)
  outss&level._2(keep=block order &levelc.25 &levelc.30 &levelc.35);
  set outss&level;
  format &levelc.10 &levelc.20 &levelc.30 &levelc.15 &levelc.25 &levelc.35 $10.;
  &levelc.10=&levelc.1; &levelc.35=&levelc.10;
  &levelc.20=&levelc.2; &levelc.25=&levelc.20;
  &levelc.30=&levelc.3; &levelc.15=&levelc.30;
x=MOD(&trtn,3);
call symput("x",x);
run;

DATA f1(rename=(&levelc.10=&levelc.1 &levelc.15=&levelc.2 &levelc.20=&levelc.3 ));
  set outss&level._1; run;
DATA f2(rename=(&levelc.25=&levelc.1 &levelc.30=&levelc.2 &levelc.35=&levelc.3 ));
  set outss&level._2; order= order + &trtn ; run;

PROC SORT DATA =f1; by block order; run;
PROC SORT DATA =f2; by block order; run;
%if %eval(&x) = 0 %then %do;

```

```

DATA f12_&level(drop=sequence sequencec)
  seq(keep=sequence sequencec &levelc.1 &levelc.2 &levelc.3);
  set f1 f2;
  by block order;
subject=_n_;
output f12_&level;
if block=1 then do;
  sequence=order;
  sequencec=trim(left(&levelc.1))||"-"||trim(left(&levelc.2))||"-"||
    trim(left(&levelc.3));
  output seq; end;
run;
%end;
%else %do;
DATA f12_&level(drop=sequence sequencec)
  seq(keep=sequence sequencec &levelc.1 &levelc.2 &levelc.3);
  set f1;
  by block order;
subject=_n_; output f12_&level;
if block=1 then do;
  sequence=order;
  sequencec=trim(left(&levelc.1))||"-"||trim(left(&levelc.3));
  output seq; end;
run;
%end;
* add seq number and name ***;
PROC SORT DATA=f12_&level; by &levelc.1 &levelc.2 &levelc.3;
PROC SORT DATA=seq; by &levelc.1 &levelc.2 &levelc.3; run;
DATA f12_final&level;
  merge f12_&level(in=a)
        seq(in=b);
  by &levelc.1 &levelc.2 &levelc.3;
run;
PROC SORT DATA=f12_final&level;
  by block order sequence;
run;
%mend;

```

EXAMPLE № 1

This example demonstrates randomization for level-1, level-2, and level-3 separately and independently of each other.

Level 1 – Randomization for Treatment Arms: A, B, and C.

```
%WLS3x3(level=1, levelc=period, trt=%str("A" "B" "C"), trtn=3, block=3);
```

Subject	Block	Order	Sequence	Sequencec	Period1	Period2	Period3
1	1	1	1	B-C-A	B	C	A
2	1	2	2	A-B-C	A	B	C
3	1	3	3	C-A-B	C	A	B
4	1	4	4	A-C-B	A	C	B
5	1	5	5	C-B-A	C	B	A
6	1	6	6	B-A-C	B	A	C
7	2	1	1	B-C-A	B	C	A
8	2	2	3	C-A-B	C	A	B
9	2	3	2	A-B-C	A	B	C
10	2	4	4	A-C-B	A	C	B
11	2	5	6	B-A-C	B	A	C

Subject	Block	Order	Sequence	Sequenc	Period1	Period2	Period3
12	2	6	5	C-B-A	C	B	A
13	3	1	2	A-B-C	A	B	C
14	3	2	3	C-A-B	C	A	B
15	3	3	1	B-C-A	B	C	A
16	3	4	5	C-B-A	C	B	A
17	3	5	6	B-A-C	B	A	C
18	3	6	4	A-C-B	A	C	B

Output 1: Multilevel Randomization: Level 1-Treatment Arm

The number of sequences for treatment arms is 6. The randomization schedule is presented in Output 1 and the unique sequences are presented in Output 2. It can be checked that the randomization follows the rules of WLS design where each pair of treatments is listed an equal number of times.

Sequence	Period1	Period2	Period3	Frequency	Cumulative Frequency
1	B	C	A	3	3
2	A	B	C	3	6
3	C	A	B	3	9
4	A	C	B	3	12
5	C	B	A	3	15
6	B	A	C	3	18

Output 2: The Table of Unique Treatment Sequences

Level 2 – Randomization of Anatomical Location: Arm, Hip, and Knee.

The same macro can be used to randomize location of the body independently of treatment arms.

```
%WLS3x3(level=2, levelc=location, trt=%str("Arm" "Hip" "Knee" ), trtn=3, block=3);
```

Subject	Block	Order	Sequence	Sequenc	Location		
					Period1	Period2	Period3
1	1	1	1	Arm-Knee-Hip	Arm	Knee	Hip
2	1	2	2	Knee-Hip-Arm	Knee	Hip	Arm
3	1	3	3	Hip-Arm-Knee	Hip	Arm	Knee
4	1	4	4	Hip-Knee-Arm	Hip	Knee	Arm
5	1	5	5	Arm-Hip-Knee	Arm	Hip	Knee
6	1	6	6	Knee-Arm-Hip	Knee	Arm	Hip
7	2	1	2	Knee-Hip-Arm	Knee	Hip	Arm
8	2	2	1	Arm-Knee-Hip	Arm	Knee	Hip
9	2	3	3	Hip-Arm-Knee	Hip	Arm	Knee
10	2	4	5	Arm-Hip-Knee	Arm	Hip	Knee
11	2	5	4	Hip-Knee-Arm	Hip	Knee	Arm
12	2	6	6	Knee-Arm-Hip	Knee	Arm	Hip
13	3	1	1	Arm-Knee-Hip	Arm	Knee	Hip
14	3	2	3	Hip-Arm-Knee	Hip	Arm	Knee
15	3	3	2	Knee-Hip-Arm	Knee	Hip	Arm

Subject	Block	Order	Sequence	Sequenc	Location		
					Period1	Period2	Period3
16	3	4	4	Hip-Knee-Arm	Hip	Knee	Arm
17	3	5	6	Knee-Arm-Hip	Knee	Arm	Hip
18	3	6	5	Arm-Hip-Knee	Arm	Hip	Knee

Output 3: Multilevel Randomization: Level 2-Anatomical Location of the Patch

The randomization schedule is presented in Output 3. The number of sequences for treatment arms is 6 and presented in Output 4.

Sequence	Location			Frequency	Cumulative Frequency
	Period1	Period2	Period3		
1	Arm	Knee	Hip	3	3
2	Knee	Hip	Arm	3	6
3	Hip	Arm	Knee	3	9
4	Hip	Knee	Arm	3	12
5	Arm	Hip	Knee	3	15
6	Knee	Arm	Hip	3	18

Output 4: The Table of Unique Location Sequences by Period**Level 3 – Randomization of Body Side for Patch Application: Left or Right.**

The same macro can be used to randomize sides of the body independently. Having in mind multilevel randomization as next-step, we can notice that the number of body sides (2: Left, Right) is less than the number of treatment arms (3: A, B, C) and anatomical locations (3: Arm, Hip, Knee). We still can use the macro knowing that only two sequences are possible where body sides will be alternated for each of three periods: LRL or RLR. The randomization schedule is presented in Output 5.

```
%WLS3x3 (level=3, levelc=bsite, trt=%str("LRL" "RLR"), trtn=2, block=9);
```

Subject	Block	Order	Sequence	Body Sides
1	1	1	1	LRL
2	1	2	2	RLR
3	2	1	2	RLR
4	2	2	1	LRL
5	3	1	2	RLR
6	3	2	1	LRL
7	4	1	2	RLR
8	4	2	1	LRL
9	5	1	1	LRL
10	5	2	2	RLR
11	6	1	2	RLR
12	6	2	1	LRL
13	7	1	2	RLR
14	7	2	1	LRL
15	8	1	1	LRL
16	8	2	2	RLR

Subject	Block	Order	Sequence	Body Sides
17	9	1	2	RLR
18	9	2	1	LRL

Output 5: Multilevel Randomization: Level 3-Body side (Left or Right)**EXAMPLE № 2**

It is not an easy task to merge schedules for different levels together because the same number of treatments needs to be randomized to the different locations and sides of the body. This example randomizes subjects to the treatment arms where the transdermal patch gets applied randomly to different sides of the body (Left or Right) and keeps balance for the treatments, body sides, and treatments within body sides (e.g. level-3 within level-1).

For this purpose, the first Latin square for level-1 (treatment arms) gets assigned to the first record to the Latin square for level-3 in the first block that was randomly generated in **EXAMPLE № 1**; and second Latin square for level-1 (treatment arms) gets assigned to the second record from block=1 of the Latin Square for level-3. See the code below.

```

** MERGE THE RESULTS OF Level=1 and Level=3 *****;
PROC SORT DATA= f12_final1 out=l1; by block sequence period1 period2 period3; run;
DATA l1_;
  set l1;
  if sequence <4 then seq_level3=1; * first Latin Square;
  else seq_level3=2;                * second Latin Square;
run;

PROC SORT DATA= f12_final3 out=l3(where=(block=1) rename=(sequence=seq_level3));
  by block sequence; run;

PROC SORT DATA=l1; by seq_level3; run;
PROC SORT DATA=l3;  by seq_level3; run;

DATA l13;
  merge l1_(in=a)
        l3(in=b keep=seq_level3 bsitel1);
  by seq_level3;
  site1=substr(bsitel1, 1, 1);
  site2=substr(bsitel1, 2, 1);
  site3=substr(bsitel1, 3, 1);
  p1=trim(left(period1))||"/"||trim(left(site1));
  p2=trim(left(period2))||"/"||trim(left(site2));
  p3=trim(left(period3))||"/"||trim(left(site3));
run;

PROC SORT DATA=l13; by subject block order sequence; run;

```

This randomization schedule is illustrated in Output 6.

Subject	Order	Sequence	Treatment/Body Side		
			Period1	Period2	Period3
1	1	1	B/L	C/R	A/L
2	2	2	A/L	B/R	C/L
3	3	3	C/L	A/R	B/L
4	4	4	A/R	C/L	B/R
5	5	5	C/R	B/L	A/R
6	6	6	B/R	A/L	C/R
7	1	1	B/L	C/R	A/L
8	2	3	C/L	A/R	B/L
9	3	2	A/L	B/R	C/L
10	4	4	A/R	C/L	B/R
11	5	6	B/R	A/L	C/R
12	6	5	C/R	B/L	A/R
13	1	2	A/L	B/R	C/L
14	2	3	C/L	A/R	B/L
15	3	1	B/L	C/R	A/L
16	4	5	C/R	B/L	A/R
17	5	6	B/R	A/L	C/R
18	6	4	A/R	C/L	B/R

Output 6: Multilevel Randomization: Level-1 and Level-3 - Treatment Arm / Body side

Output 7 shows that each treatment arm has the same number of Left or Right in each period (e.g., one time A/R and A/L) while retaining the WLS design for treatment arms and body sides as well.

Sequence	Period1	Period2	Period3	Frequency	Cumulative Frequency
1	B/L	C/R	A/L	3	3
2	A/L	B/R	C/L	3	6
3	C/L	A/R	B/L	3	9
4	A/R	C/L	B/R	3	12
5	C/R	B/L	A/R	3	15
6	B/R	A/L	C/R	3	18

Output 7: The Table of Unique Sequences for Level-1 (Treatment Arms) & Level-3 (Body Sides) Randomization**EXAMPLE № 3**

The same idea is applied to randomize into all three levels together. The difference from EXAMPLE № 2 is that we have 3 treatment arms and 3 anatomical locations. In this case, the merge of each unique sequence from level-1&3 (there are three records for each unique sequence) with the first or second Latin square of the level-2 randomization is performed. The output is presented in Output 8; and the unique sequences for three-level randomization are presented in Output 9.

Subject	Sequence	Sequences	Treatment/Location/Body Side		
			Period1	Period2	Period3
1	1	B-C-A	B/Arm/L	C/Knee/R	A/Hip/L
2	2	A-B-C	A/Arm/L	B/Knee/R	C/Hip/L
3	3	C-A-B	C/Arm/L	A/Knee/R	B/Hip/L
4	4	A-C-B	A/Hip/R	C/Knee/L	B/Arm/R
5	5	C-B-A	C/Hip/R	B/Knee/L	A/Arm/R
6	6	B-A-C	B/Hip/R	A/Knee/L	C/Arm/R
7	1	B-C-A	B/Knee/L	C/Hip/R	A/Arm/L
8	3	C-A-B	C/Knee/L	A/Hip/R	B/Arm/L
9	2	A-B-C	A/Knee/L	B/Hip/R	C/Arm/L
10	4	A-C-B	A/Arm/R	C/Hip/L	B/Knee/R
11	6	B-A-C	B/Arm/R	A/Hip/L	C/Knee/R
12	5	C-B-A	C/Arm/R	B/Hip/L	A/Knee/R
13	2	A-B-C	A/Hip/L	B/Arm/R	C/Knee/L
14	3	C-A-B	C/Hip/L	A/Arm/R	B/Knee/L
15	1	B-C-A	B/Hip/L	C/Arm/R	A/Knee/L
16	5	C-B-A	C/Knee/R	B/Arm/L	A/Hip/R
17	6	B-A-C	B/Knee/R	A/Arm/L	C/Hip/R
18	4	A-C-B	A/Knee/R	C/Arm/L	B/Hip/R

Output 8: Multilevel Randomization: Level 1, 2 and 3- Treatment Arm / Location / Body side

Period1	Period2	Period3	Frequency	Cumulative Frequency
A/Arm/L	B/Knee/R	C/Hip/L	1	1
A/Arm/R	C/Hip/L	B/Knee/R	1	2
A/Hip/L	B/Arm/R	C/Knee/L	1	3
A/Hip/R	C/Knee/L	B/Arm/R	1	4
A/Knee/L	B/Hip/R	C/Arm/L	1	5
A/Knee/R	C/Arm/L	B/Hip/R	1	6
B/Arm/L	C/Knee/R	A/Hip/L	1	7
B/Arm/R	A/Hip/L	C/Knee/R	1	8
B/Hip/L	C/Arm/R	A/Knee/L	1	9
B/Hip/R	A/Knee/L	C/Arm/R	1	10
B/Knee/L	C/Hip/R	A/Arm/L	1	11
B/Knee/R	A/Arm/L	C/Hip/R	1	12
C/Arm/L	B/Knee/R	A/Hip/L	1	13
C/Arm/R	A/Hip/L	B/Knee/R	1	14
C/Hip/L	B/Arm/R	A/Knee/L	1	15
C/Hip/R	A/Knee/L	B/Arm/R	1	16

Period1	Period2	Period3	Frequency	Cumulative Frequency
C/Knee/L	B/Hip/R	A/Arm/L	1	17
C/Knee/R	A/Arm/L	B/Hip/R	1	18

Output 9: The Table of Unique Sequence for Three Levels of Randomization (Treatments, Locations and Sides)

It is recommended that the reader perform a full permutation of the subjects to randomly re-assign their numbers for the final randomization schedule that presented in Output 10.

```

PROC PLAN seed=&seed4;
  factors subject=18 random;
  output out=subject;
run;
DATA mysubject;
  set subject;
  mysubject="9"||put(_N_, z2.);
run;
PROC SORT DATA=l123; by subject; run;
PROC SORT DATA=mysubject; by subject; run;
DATA finals;
  merge l123(in=a)
        mysubject(in=b);
  by subject;
  if a or b;
run;
PROC SORT DATA=finals; by mysubject; run;

```

My subject	Treatment		Treatment/Location/Body Side		
	Sequence	Sequenccec	Period1	Period2	Period3
901	6	B-A-C	B/Hip/R	A/Knee/L	C/Arm/R
902	3	C-A-B	C/Arm/L	A/Knee/R	B/Hip/L
903	1	B-C-A	B/Knee/L	C/Hip/R	A/Arm/L
904	6	B-A-C	B/Knee/R	A/Arm/L	C/Hip/R
905	5	C-B-A	C/Hip/R	B/Knee/L	A/Arm/R
906	5	C-B-A	C/Knee/R	B/Arm/L	A/Hip/R
907	4	A-C-B	A/Knee/R	C/Arm/L	B/Hip/R
908	1	B-C-A	B/Hip/L	C/Arm/R	A/Knee/L
909	5	C-B-A	C/Arm/R	B/Hip/L	A/Knee/R
910	4	A-C-B	A/Arm/R	C/Hip/L	B/Knee/R
911	3	C-A-B	C/Knee/L	A/Hip/R	B/Arm/L
912	6	B-A-C	B/Arm/R	A/Hip/L	C/Knee/R
913	2	A-B-C	A/Hip/L	B/Arm/R	C/Knee/L
914	1	B-C-A	B/Arm/L	C/Knee/R	A/Hip/L
915	3	C-A-B	C/Hip/L	A/Arm/R	B/Knee/L
916	2	A-B-C	A/Arm/L	B/Knee/R	C/Hip/L
917	2	A-B-C	A/Knee/L	B/Hip/R	C/Arm/L
918	4	A-C-B	A/Hip/R	C/Knee/L	B/Arm/R

Output 10: Final Randomization Schedule for Three Levels of Randomization (Treatment, Location and Side)

VALIDATION

The **WLS3x3** macro results were reproduced with independent programs using the same seeds. The randomization schedules with different permutations and subsets of levels to accommodate study designs were generated; and all of them passed their quality control checks.

LIMITATIONS/FUTURE PLANS

The **WLS3x3** macro was written for odds and even numbers of formulations of William's Latin Square cross-over design and adapted for this paper for the 3x3 crossover study. The macro would need to be slightly modified to accommodate 4x4 or 5x5 crossover studies.

Another macro is under development. It will perform multiple merges to achieve multilevel randomization in one step for different study designs.

CONCLUSION

This paper is an introduction to multilevel randomization. It provides examples from transdermal medicine research where the multilevel concept is relevant, and illustrates how to create a randomization schedule responding to the design requirements. The macro for William's Latin Square design is presented as the first step, and SAS code for the next steps for multilevel level randomization schedules are presented as examples. They can be used as the basis for more sophisticated study designs and requirements.

ACKNOWLEDGEMENT

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are registered trademarks or trademarks of their respective companies.

REFERENCES

- [1] R.A. Fisher. 1937. **The Design of Experiments** (2 ed.). Edinburgh: Oliver and Boyd.
- [2] Joseph L. Fleiss, Bruce Levin, Myunghee Cho Paik. 2003. **Statistical Methods for Rates and Proportions**, New Jersey: John Wiley & Sons, Inc.
- [3] SUGI 27: Paper 267-27 **Generating Randomization Schedules Using SAS Programming** Chunqin Deng and Julia Graz, PPD, Inc., Research Triangle Park, North Carolina

CONTACT INFORMATION

Marina Komaroff

Associate Director – Biostatistics
Product Development/Clinical Operations
Noven Pharmaceuticals, Inc.
Empire State Building
350 Fifth Avenue, 37th Floor
New York, NY 10118
Tel: 212 299 4202
Email: Mkomaroff@noven.com

Lois Lynn

Manager, Biostatistics
Product Development /Clinical Operations
Noven Pharmaceuticals, Inc.
Empire State Building
350 Fifth Avenue, 37th Floor
New York, NY 10118
Tel: 212 287 0791
Email: LLynn@noven.com