RANDOMIZATION STRATEGIES FOR CLUSTERED CLINICAL SYNDROMES: EFFECTS ON STATISTICAL POWER

Barry D. Lebowitz, Ph.D.,[1] Yuanjia Wang, Ph.D.,[2] Naihua Duan, Ph.D,[2] Charles F. Reynolds III, M.D.[3]
Naomi M. Simon, M.D.,[4] Sidney Zisook, M.D.,[2] M. Katherine Shear, M.D.[2]
[1] University of California, San Diego; [2] Columbia University;[3] University of Pittsburgh; [4] Massachusetts General Hospital

INTRODUCTION

Complicated Grief (CG) is representative of a number of clinical syndromes that occur in clusters. Such clusters may arise from familial genetics (as in schizophrenia or bipolar disorder) or in response to commonly experienced stressful life events (as in survivors of trauma). Multiple survivors of a single death, related to each other in the same social or familial network, may all develop symptoms of CG sufficient to impair functioning. **We address the question of how this clustering influences statistical power in an RCT.**

This situation was encountered during the development of randomization strategies for a 4-site NIMH- and AFSP-supported RCT of CG monotherapies and combined treatments. Sampling of clusters is a commonly used strategy in survey sampling (Fleiss, 1986; Kish, 1995) and randomization of clusters is often used in field trials of organization-level interventions in schools, hospitals, and communities (Cornfield, 1978; Donner and Klar, 1994; Donner, 1998). In these trials, all individuals in a cluster are offered the same intervention. Such clustering strategies are rarely used in therapeutics research.

METHODS

We have launched a 4-cell placebo controlled factorial design RCT comparing the efficacy of pharmacotherapy (citalopram) and targeted psychotherapy (complicated grief therapy, CGT) alone and in combination. We considered several randomization strategies for clusters e.g.: randomly selecting one participant from each of the clusters or sequentially enrolling all members of the cluster (Wingaards-de Meij, 2008). **We chose a "Split-plot" design (Fleiss, 1986) in which each treatment component is randomized at the appropriate level respectively.** More specifically, the split plot design we specified will: (a) **randomize therapy condition (CGT vs. no CGT) at the group level** (couple, family, etc.), so that related participants from the same group are assigned to the same therapy condition (they either all get CGT, or all get no CGT), to avoid possible contamination and other concerns; and (b**) randomize medication condition (medication vs. placebo) at the individual level** as there is less concern about contamination for this treatment component. There is a long and established history of the split-plot design in agriculture and engineering.

RESULTS

T**he split-plot design, compared to individual level randomization, maintains power for the pharmacotherapy aim of the trial but reduces power for the psychotherapy aim as a joint function of the degree of intercorrelation within the cluster and the proportion of clusters in the overall sample.** As shown in the data table and Figure 1, with a sample size of 220 where within cluster correlation is .50, power is reduced from .84 to .82 as the proportion of **clustered participants increases from 0 to .50. With a within-cluster correlation of 0.90, power is reduced further to .74 when 50% of the participants are related.**

ACKNOWLEDGMENTS

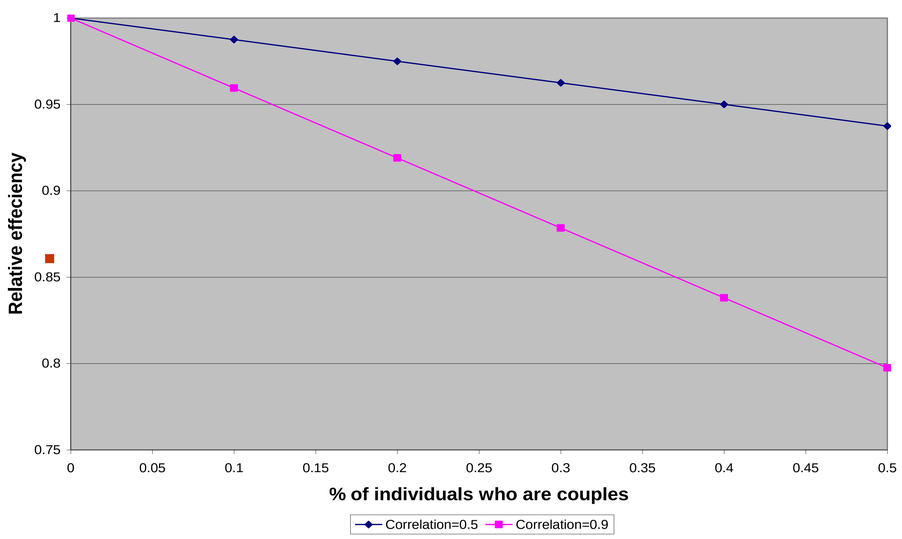
This study is supported by NIMH Grants No. MH060783, MH085308, MH085288, and MH085297 and a grant from the American Foundation for Suicide Prevention.

The authors report no conflicts of interest for this work

POWER WITH SPLIT PLOT DESIGN

% of individuals who are couples	Correlation between related subjects	Effective sample size	Power
0	0.5	220	**0.840**
0.1	0.5	217	**0.834**
0.2	0.5	214	**0.830**
0.3	0.5	211	**0.823**
0.4	0.5	209	**0.819**
0.5	0.5	206	**0.816**
0	0.9	220	**0.840**
0.1	0.9	211	**0.823**
0.2	0.9	202	**0.808**
0.3	0.9	193	**0.788**
0.4	0.9	184	**0.770**
0.5	0.9	175	**0.747**

Figure 1. Effects on power (Relative Efficiency) measured by the ratio of the effective sample size and the nominal sample size (n=220) with split-plot design and individual-level randomized design as a function of percentage of individuals who are couples. Power estimates calculated using PS 3.0.12 (Dupont and Plummer, 1990)



DISCUSSION

In the context of clustered eligibility for a 2X2 factorial RCT we considered, but rejected, block randomization of the pharmacotherapy condition. **Block randomization would enhance statistical power and precision by blocking the randomization at the group level**:for example, in a husband-wife cluster, either "husband gets citalopram and wife gets placebo", or "wife gets citalopram and husband gets placebo". This would disallow "both get citalopram" and "both get placebo". By blocking, we would reduce the uncertainty in the comparison between citalopram and placebo because the common elements in the family are cancelled out when comparing husband and wife. This is equivalent to the common practice in animal experiments to block by litter. However, **this option would introduce clinical and ethical complications.** We chose to randomize individuals in the cluster to citalopram vs. placebo without blocking, i.e., the same way "singleton" participants are randomized.

CONCLUSION

For the CGT condition, we may lose power and precision by randomization at the group level, compared to randomization at the individual level, due to intra-group correlation. For example, even though we have a nominal sample size of 2 when we enroll a couple into the study, the effective sample size is lower because the statistical information gained from the second related participants is typically less than the information gained from an unrelated participant, due to the two participants from the same group being correlated. **The effect on power (seen in terms of relative efficiency or the design effect) is determined by the strength of the intra-group correlation.** Assuming the correlation between related subjects is 0.5, when 100% of the samples enrolled are couples, our power is still adequate (79%). Assuming the correlation between related subjects is 0.9, when 90% of the samples enrolled are couples, our power is below 65% which is insufficient. In an extreme scenario, if both members of a couple are identical, i.e. perfectly correlated, there is no new information provided by the added participant.

.

REFERENCES

Fleiss J. The Design and Analysis of Clinical Experiments. New York: John Wiley & Sons, 1986

Kish L.. Survey Sampling. New York: John Wiley & Sons, 1995

Cornfield J. Randomization by group: a formal analysis. Am J Epidemiol. 1978;108:100–102*m*

Donner A , Klar N. Cluster randomization trials in epidemiology: theory and application. Journal of Statistical Planning and Inference. 1994;42:37-56.

Donner A . Some aspects of the design and analysis of cluster randomization trials. Applied Statistics. 1998;47:95-113.

Wijngaards-de Meij L, Stroebe M, Schut H, Stroebe W, van den Bout J, van der Heijden PGM, Dijkstra I. Parents grieving the loss of their child: Interdependence in coping. Br J Clin Psychol 2008;47:31-42

Dupont WD, Plummer WD. Power and sample size calculations: a review and computer program. Control Clin Trials 1990;11:116-128