

## Efficient Allocation of Patients to Treatment Cells in Clinical Trials With More Than Two Treatment Conditions

**Scott W. Woods, M.D., Diane E. Sholomskas, Ph.D.,  
M. Katherine Shear, M.D., Jack M. Gorman, M.D., David H. Barlow, Ph.D.,  
Andrew W. Goddard, M.D., and Jacob Cohen, Ph.D.**

---

**Objective:** Clinical trials generally allocate patients to equal-sized treatment groups. The authors propose that it may be more efficient to allocate unequal proportions of the total sample size to treatments when more than two treatments are being compared. **Method:** This proposal is illustrated with two examples. One involved a comparison of three treatments and used a dichotomous categorical outcome. The other involved comparison of three treatments and used a continuous measure. **Results:** In both examples, a considerable increase in efficiency was realized by reducing the number of patients assigned to the placebo cell. **Conclusions:** Unequal allocation of patients to treatments should be considered when more than two groups are compared.

**(Am J Psychiatry 1998; 155:1446–1448)**

---

Authorities on clinical trial design generally recommend allocating patients to equal-sized treatment groups (1–4). This recommendation is based on the demonstration, in studies comparing two treatment groups, that equal allocation of patients between groups generally maximizes the power of the statistical comparison to detect a difference as significant. Unequal allocation of patients to treatments has been advocated previously in the comparison of two treatments when the trial must be performed at a fixed cost and one treatment is more expensive than another (5) and when one treatment is more hazardous than another (6).

When more than two treatments are being compared, we propose that it may be more efficient to allocate unequal proportions of the total sample size to treatments, even when costs and hazards of the treat-

ments are equal. Study designs with more than two treatments are not uncommon. For example, by searching MEDLINE for papers indexed as clinical trials and appearing in *The American Journal of Psychiatry* from January 1990 to May 1997, we identified 28 papers that reported randomly assigning patients to three or more groups (references available on request). In none of these were patients clearly randomly assigned prospectively to unequal-sized groups to improve trial efficiency.

In designs with three or more treatments, investigators often are interested ultimately in conducting post hoc pairwise comparisons between relevant pairs of treatments. The expected effect size is always larger for some pairs than for others; therefore, the power in the comparison is higher for some pairs than for others. In such a case, trial efficiency may be increased by assigning a lower proportion of patients to those cells which participate only in higher-power pairwise comparisons. We illustrate this proposal with two examples.

### EXAMPLE 1: COMPARISON OF THREE TREATMENTS USING A DICHOTOMOUS CATEGORICAL OUTCOME

A study compares a new antidepressant medication, A, with a standard treatment, B, and a pill placebo, C. The principal outcome measure is the proportion of patients determined to be treatment responders versus the proportion determined to be treatment nonre-

---

Received Nov. 26, 1997; revision received April 30, 1998; accepted May 12, 1998. From the Connecticut Mental Health Center, Yale University School of Medicine Department of Psychiatry; the University of Pittsburgh; Columbia University, New York; Long Island Jewish Hillside Hospital, Glen Oaks, N.Y.; Boston University; and New York University. Address reprint requests to Dr. Woods, Connecticut Mental Health Center, Yale University School of Medicine Department of Psychiatry, 34 Park St., New Haven, CT 06508; scott.woods@yale.edu (e-mail).

Dr. Cohen passed away in January 1998.  
Supported by U.S. Public Health Service grants MH 45963 and MH 45966.

The authors thank Helena C. Kraemer, Ph.D., for reviewing the manuscript.

sponders. Treatment B is expected from previous work to produce 60% responders and placebo C, 30%. If new treatment A could produce 80% responders, the investigators believe that the difference between treatments A and B would be clinically meaningful; therefore, they desire sufficient power to detect treatment A as significantly different from treatment B if the difference between the response rates is as large as or larger than that between 80% and 60%, with a power of 0.80 and alpha set at 0.05 (two-tailed). The investigators also wish the design to have power to demonstrate that treatment A and B were each superior to placebo, thus documenting that each treatment is effective in the sample and enhancing the validity of the comparison between treatments A and B as a comparison between effective treatments. They consider the validity imparted by the comparisons of treatment B versus placebo C and treatment A versus placebo C to be important and so require power of 0.90, although the tests may be one-tailed because these hypotheses are clearly unidirectional.

According to the method of Cohen for estimation of sample sizes for differences between rates or proportions (7), the required sample size for each group in the A versus B comparison is approximately 80. For the comparison of B versus C, the required sample size for each group, if they were equal sized, is approximately 46. Since treatment B must have 80 patients, we may size the C sample so that the harmonic mean ( $n' = 2n_B n_C / (n_B + n_C)$ ) of sample C and sample B is 46. Solving the formula for  $n_C$  where  $n' = 46$  and  $n_B = 80$  yields a value for the sample size for the placebo group of approximately 32. Thus the B versus C comparison of 80 versus 32 patients will have a power of 0.90. Power for the A versus C pairwise comparison and the power of the omnibus comparison of the three unequal-sized groups based on the chi-square contingency test are both  $>0.995$ . Thus,  $80+80+32=192$  patients may be randomly assigned to groups in this study at a ratio of A:B:C of 5:5:2 rather than the 240 that would have been needed if equal  $80+80+80$  samples were used, reducing the cost of the trial by 20%.

#### EXAMPLE 2: COMPARISON OF THREE TREATMENTS USING A CONTINUOUS MEASURE

A study compares a new antipsychotic medication, A, with a standard treatment, B, and a pill placebo, C. The principal outcome measure is the pre-post endpoint reduction in the total Brief Psychiatric Rating Scale (BPRS) score. Treatment B is expected from previous work to reduce the BPRS score by a mean of 10 points ( $SD=15$ ) and placebo C, a mean of 3 points ( $SD=15$ ). If new treatment A could reduce the BPRS score by a mean of 15 points ( $SD=15$ ), the investigators believe that the difference between treatments A and B would be clinically meaningful. Power and alpha values are set as they were in example 1. According to the method of Cohen for estimation of required sample

sizes for differences between means (7), the required sample size for each group in the comparison of A versus B is approximately 160. Given that B will need 160 subjects, solving the harmonic mean formula (given in example 1) for the comparison of B versus C yields a sample size for the placebo group of 54. The trial has adequate power for all planned comparisons but requires 374 patients rather than 480, reducing the cost of the trial by 22%.

#### DISCUSSION

In both examples, a considerable increase in efficiency was realized by reducing the number of patients assigned to the placebo cell. The resulting unequal allocation scheme retained the desired level of power for all planned comparisons. This was possible because of the larger effect size and consequently greater power of the pairwise comparisons of active treatment versus placebo relative to the comparison between two active medications.

The examples shown were restricted to studies comparing three treatments. The potential for increasing trial efficiency can also be extended to studies of four or more treatments. A trial employing a similar unequal allocation method among five treatments funded by the National Institute of Mental Health (8) is nearing completion. Use of the unequal allocation scheme led to a 20% reduction in the cost of the trial.

A relatively small number of patients in one of the treatment cells could pose a risk to a clinical trial in some circumstances. These circumstances include a smaller effect size than expected, a higher placebo response rate than expected, fewer patients recruited than expected, or greater attrition than expected. Investigators may wish to safeguard against these possibilities by making conservative estimates of outcome or attrition. Another strategy, used in both examples, is to require higher power in these comparisons. These safeguards will necessarily reduce the gains in efficiency achieved. Investigators should also be aware that power will be lower for within-group comparisons in the smallest group.

Because of the smaller numbers in the placebo cells in both examples, any individual patient's chance of receiving placebo was relatively low (approximately one in six in example 1 and one in seven in example 2). This feature may benefit a clinical trial by making it more attractive to patients and referring clinicians and thus enhancing recruitment.

In summary, we have described an efficient method for determining required sample sizes when studies contain three or more groups. The examples illustrate how clinical trials may benefit in the planning stage from statistical power calculations for each important planned comparison. Similarly, at the publication stage, interpretation of trial results would benefit from reporting and discussing statistical power for each of these comparisons.

**REFERENCES**

1. Armitage P, Berry G: *Statistical Methods in Medical Research*, 2nd ed. Oxford, England, Blackwell Scientific, 1987
2. Friedman LM, Furberg CD, DeMets DL: *Fundamentals of Clinical Trials*, 2nd ed. Littleton, Mass, PSG, 1985
3. Fleiss JL: *The Design and Analysis of Clinical Experiments*. New York, John Wiley & Sons, 1986
4. Winer BJ, Brown DR, Michels KM: *Statistical Principles in Experimental Design*, 3rd ed. New York, McGraw-Hill, 1991
5. Nam J-M: Optimum sample sizes for the comparison of the control and the treatment. *Biometrics* 1973; 29:101–108
6. Gail M, Williams R, Byar DP, Brown C: How many controls? *J Chronic Dis* 1976; 29:723–731
7. Cohen J: *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ, Lawrence Erlbaum Associates, 1988
8. Sholomskas DE, Barlow DH, Cohen J, Gorman JM, Moras K, Papp L, Shear MK, Woods SW: Drug/behavior treatment of panic: study design, in 1990 Annual Meeting New Research Program and Abstracts. Washington, DC, American Psychiatric Association, 1990, p 162