Comment

Robert E. Bechhofer and Ajit C. Tamhane

We had little imagined that BTIB designs for comparing test treatments with a control treatment would generate such a wide research interest among the design theorist community when we first proposed this new class of designs in Bechhofer and Tamhane (1981). Naturally we are very pleased and gratified to note the tremendous progress that has been made in the last seven years in the study of these designs and their extensions, with particular emphasis on the optimality question. Hedayat, Jacroux and Majumdar, who have been at the forefront of this research, are to be congratulated for providing a fine survey of the results. We thank the Editor for giving us an opportunity to discuss this survey.

Hedayat, Jacroux and Majumdar focus their attention on A- and MV-optimal designs. Both of these optimality criteria refer to minimizing suitable functions of the variances of the $\hat{t}_i - \hat{t}_0$, but do not take their *correlations* into account. (We follow the same notation as in their article.) Thus the optimal designs derived would seem to be appropriate when the results of the experiment are to be reported in terms of the above *point* estimates accompanied by their estimated standard errors or in terms of *separate* confidence interval estimates of the $t_i - t_0$, $i = 1, \cdots, v$. However, in many applications a *simultaneous* confidence region (or a set of simultaneous confidence intervals) is more appropriate than separate confidence intervals. The following is an example of such an application. It is frequently desired to select one or more of the test treatments for eventual use. The primary selection criterion is the parameter $t_i - t_0$ (test treatments with large values being preferred, say) but there also are secondary criteria such as costs. The precise rules for selection of the test treatments cannot be stated in advance. For example, depending on the experimental results and other side considerations, the two apparently "best" test treatments (in terms of the $\hat{t}_i - \hat{t}_0$ values) may be selected or even the third apparently "best" test treatment may be selected. A set of *simultaneous* confidence intervals guarantees a specified confidence coefficient regardless of which test treatments are selected and for which the corresponding confidence interval estimates are reported. For other examples of applications where simultaneous inference (multiple comparisons) procedures are called for, see Hochberg and Tamhane (1987, Chapter 1).

Under normal theory, operating characteristics of simultaneous inference procedures are generally functions of not only the variances of the $\hat{t}_i - \hat{t}_0$ but also their correlations. It is true that, for example, A-optimality is equivalent to minimizing the sum of the lengths of the axes of the *simultaneous* confidence ellipsoid (assuming the usual normal theory model) for the given contrasts of interest. But curiously, D-optimality, which corresponds to minimizing the volume of the confidence ellipsoid, and which does take into account the variances as well as the correlations, is not a useful optimality criterion for the present problem, as the authors rightly point out.

We believe that the use of these standard optimality criteria due to Kiefer (1958) is questionable in experiments involving multiple comparisons of test treatments with a control because they do not address the typical inferential goals in such experiments. The authors state that "To begin with we need to postulate a model $\cdots$" In the same vein, it is also true that, to decide on a design (optimal or efficient), we need to postulate the types of *inferences* that will be made based on the data collected from the experiment. The authors make a brief reference to this point when they state that "$\cdots$ our primary goal is to determine which among the test treatments might be better than the control $\cdots$" However, we do not think that this goal necessarily translates into "$\cdots$ to estimate the magnitude of each $t_i - t_0$ with as much precision as possible" without reference to how the resulting estimates will be used to determine the apparently better test treatments. In fact, as we explain below, two types of inferential goals are appropriate in these experiments, and both involve taking into account the variances of the $\hat{t}_i - \hat{t}_0$ as well as their correlations.

Often, in exploratory stages of an investigation there are a large number of new candidate test treatments, and the goal is to screen out those that are inferior to the control treatment. For this goal the subset selection formulation of Gupta and Sobel (1958) would appear to be suitable. The test treatments that are selected in this initial experiment can

Robert E. Bechhofer is Professor, School of Operations Research and Industrial Engineering, College of Engineering, Cornell University, Ithaca, New York 14853. Ajit C. Tamhane is Professor, Department of Statistics, Northwestern University, 2006 Sheridan Road, Evanston, Illinois 60201 and is also a faculty member in the Department of Industrial Engineering and Management Sciences, Technological Institute, Northwestern University.