be required to be of equal precision. For example, in a clinical trial for a new drug it is not uncommon to include two controls, a placebo and an existing active drug. For regulatory purposes, it often is necessary to demonstrate the magnitude of the activity of the new drug, and therefore the comparison with the placebo is the more important. It is not always necessary to demonstrate to the regulatory agency that the new drug is more effective than the existing drug. But for the purposes of the pharmaceutical company's marketing efforts, in fact, the second comparison is likely to be the more important. This latter comparison would generally be two-sided. Such considerations should be taken into account before determining how to optimally allocate the available experimental resources to different competing test treatments and the controls.

A final brief note concerning nomenclature. We suggest that the word "control" should be used rather than "standard" because the latter sometimes refers to a *known* benchmark value; this is the case, e.g., in the physical sciences (although, not always in the biological sciences). Clearly, if the comparisons are made with a known benchmark then the device of blocking cannot be used.

We again express our gratitude to the authors for this state-of-the-art survey and to the editor for giving us an opportunity to comment on it.

ACKNOWLEDGMENTS

This research was partially supported by the United States Army Research Office through the Mathematical Sciences Institute of Cornell University. We are indebted to Professor Charles Dunnett for his insightful comments and suggestions on an earlier draft of this article.

ADDITIONAL REFERENCES

BECHHOFER, R. E. (1969). Optimal allocation of observations when comparing several treatments with a control. In *Multivariate Analysis II* (P. R. Krishnaiah, ed.) 463–473. Academic, New York.

BECHHOFER, R. E. and NOCTURNE, D. J. M. (1972). Optimal allocation of observations when comparing several treatments with a control. II: 2-sided comparisons. *Technometrics* **14** 423–436.

BECHHOFER, R. E. and TAMHANE, A. C. (1983a). Design of experiments for comparing treatments with a control: Tables of optimal allocations of observations. *Technometrics* **25** 87–95.

BECHHOFER, R. E. and TAMHANE, A. C. (1983b). Incomplete block designs for comparing treatments with a control (II): Optimal designs for one-sided comparisons when $p = 2(1)6$, $k = 2$ and $p = 3$, $k = 3$. *Sankhyā Ser. B* **45** 193–224.

BECHHOFER, R. E. and TAMHANE, A. C. (1985). Tables of admissible and optimal balanced treatment incomplete block (BTIB) designs for comparing treatments with a control. *Selected Tables Math. Statist.* **8** 41–139. Amer. Math. Soc., Providence, R. I.

DUNNETT, C. W. (1955). A multiple comparison procedure for comparing several treatments with a control. *J. Amer. Statist. Assoc.* **50** 1096–1121.

GUPTA, S. S. and SOBEL, M. (1958). On selecting a subset which contains all populations better than a standard. *Ann. Math. Statist.* **29** 235–244.

HOCHBERG, Y. and TAMHANE, A. C. (1987). *Multiple Comparison Procedures*. Wiley, New York.

RASHED, D. H. (1984). Designs for multiple comparisons of control versus test treatments. Ph.D. dissertation, Dept. Statistics, Virginia Polytechnic Institute and State Univ.

Comment

William I. Notz

Sam Hedayat, Mike Jacroux, and Dibyen Majumdar are to be congratulated on this very thorough survey of optimal designs for comparing test treatments with a control. This paper is an excellent starting point for anyone wishing to do research in this area and it is a nice reference for those of us actively engaged in such research. Unfortunately, any such survey begins to go out of date the moment it is completed as research goes ever forward. The authors can do nothing about that, however.

William I. Notz is Associate Professor, Department of Statistics, The Ohio State University, 1958 Neil Avenue, Columbus, Ohio 43210.

Let me begin my comments by describing the history of my own involvement in this area of research. If nothing else, this will at least add a little historical color.

I first became acquainted with this area of research as a relatively new assistant professor at Purdue. In the Autumn of 1980, Bob Bechhofer came to Purdue as a colloquium speaker. He spoke about results he and Ajit Tamhane had obtained on incomplete block designs for comparing test treatments with a control and which were soon to appear in Bechhofer and Tamhane (1981). One unsolved aspect of the research, which Bob invited those of us in the audience to try and solve, involved constructing finite sets of designs (so-called minimal complete sets of generator designs)

from which all admissible balanced treatment incomplete block designs (BTIBs described by Hedayat, Jacroux and Majumdar) could be generated. Admissibility was defined in terms of the joint confidence coefficient for the elementary treatment-control contrasts, i.e., the contrasts $t_i - t_0$ to use the notation of Hedayat, Jacroux and Majumdar, resulting from a given design. Having exhausted research ideas resulting from my thesis, I was looking for new fields to conquer and so began to work on this admissibility problem. After a few weeks I began to get solutions to the problem. I communicated these to Bob who encouraged me to keep working. Bob also passed my results along to Ajit who noticed several ways to shorten proofs and who provided me with some additional useful results. Ajit and I pooled resources and eventually published our work (Notz and Tamhane, 1983). Türe, then a Ph.D. student of Cheng at Berkeley, was also working on this problem. In 1982 Türe improved upon our methodology and extended our results.

Shortly after completing the article with Ajit, I had a chance to talk with Jack Kiefer, who was at Purdue in the spring of 1981 for the Third Purdue Symposium on Decision Theory and Related Topics, about the problem of finding block designs for treatment-control comparisons. He suggested I try to find designs optimal with respect to ellipsoidal criteria such as D-, A- or E-optimality. These criteria measure, in some sense, the size of the ellipsoidal joint confidence region associated with the space of all linear combinations of the elementary treatment-control contrasts $t_i - t_0$. Notice that this space includes arbitrary contrasts in the test treatments alone as a $v - 1$ dimensional subspace. Jack Kiefer wondered how designs optimal under such criteria might compare with the admissible designs of Bechhofer and Tamhane (1981). It should be kept in mind that in Bechhofer and Tamhane (1981) attention was restricted to the class of BTIB designs and then designs which were admissible in this class were sought. Would BTIB designs remain optimal under the criteria suggested by Kiefer if the class of allowable designs was enlarged to include all block designs?

As a former student of Jack's I knew his suggestions were not to be ignored, so I began working out the structure of such optimal designs. By the end of the summer of 1981 I had an article written which I sent to *The Annals of Statistics*. Unknown to me, that same summer Dibyen Majumdar had independently obtained similar results and had submitted his work to *Journal of Statistical Planning and Inference*. Fortunately a kind editor or referee noticed the two papers and suggested we combine our work into a joint paper for *The Annals of Statistics*. The result was Majumdar and Notz (1983). We found that certain types of BTIBs were optimal so that not all the admissible BTIBs of Bechhofer and Tamhane (1981) were, say, A-optimal but A-optimal BTIBs were admissible. In addition, which designs were optimal (in particular, how many times the control was replicated) depended on the criterion used. Thus for comparing test treatments with a control one does not have the same sort of universal optimality enjoyed by balanced incomplete block designs for estimating all treatment contrasts. A number of researchers, Hedayat, Jacroux and Majumdar being prominent among them, have subsequently greatly extended these early results.

So much for personal history. Some comments on Hedayat, Jacroux and Majumdar's article are now in order. In my opinion, it was the paper by Bechhofer and Tamhane (1981) that sparked the majority of recent research on optimal designs for comparing test treatments with a control. None the less, researchers (myself included) abandoned the approach of Bechhofer and Tamhane fairly quickly and focused attention on more tractable criteria (the influence of Jack Kiefer?) such as A-optimality and MV-optimality. As Hedayat, Jacroux and Majumdar point out, these criteria have nice interpretations in terms of the elementary treatment-control contrasts. Other ellipsoidal criteria do not have this feature, although they also are relatively tractable. When I talk with researchers in multiple comparisons such as Bob Bechhofer, Ajit Tamhane or Jason Hsu, a colleague at Ohio State, I find that they would like to see research return to designs which optimize some features of the rectangular joint confidence regions for the elementary treatment-control contrasts. The first issue I would like to raise, therefore, is what sort of criteria might be useful? Such criteria might be difficult to work with but progress in this direction would make optimal design results more appealing to researchers in the area of multiple comparisons. This issue is not really addressed by Hedayat, Jacroux and Majumdar although any thoughts they might have would be worth hearing.

My second comment is related to the first and involves the issue of robustness of designs under varying criteria of optimality. As Hedayat, Jacroux and Majumdar point out, for estimating all treatment contrasts balanced incomplete block designs are optimal under a broad class of criteria. Such a general result does not hold for BTIB designs, yet the A-optimal BTIB designs are also MV-optimal and admissible in the Bechhofer and Tamhane (1981) sense. In addition I believe it is not hard to show (for what it is worth) that A-optimal BTIB designs minimize criteria of the form

$$\sum_{i=1}^{v} f(\mathrm{var}(\hat{t}_i - \hat{t}_0))$$

where $\hat{t}_i - \hat{t}_0$ is the best linear unbiased estimate of the elementary treatment-control contrast $t_i - t_0$ and f is a convex, nondecreasing function. It is interesting to speculate, therefore, as to whether or not A-optimal BTIB designs are optimal over a large class of "rectangular" optimality criteria. To answer this, of course, requires coming up with suitable criteria as indicated in my first comment. Similar questions might be raised regarding optimal row-column designs also.

As a third comment, I would like to thank the authors for including material on Bayesian approaches to the design problem. Such approaches seem fairly natural in this setting because often the control treatment is a standard treatment about which we have considerable prior information whereas the test treatments are new and less is known about them. One's prior knowledge about the control should be incorporated into the design and, as one would expect, Bayesian results indicate the effect is to reduce the number of replications of the control. To my knowledge, existing Bayesian results have been obtained by allowing approximate designs and optimal designs are often approximate designs. Although seemingly a hard problem, exact Bayesian design results would be quite interesting. Are the authors aware of any research in this direction?

In summary, Hedayat, Jacroux and Majumdar are to be thanked for a readable and thorough survey article. It is to be hoped that this article will stimulate further research and such research will answer, among other things, the questions I have raised above.

Comment

A. Giovagnoli and I. Verdinelli

This is a very useful survey of many known results on optimal designs of experiments when one of the treatments is a control. It comprises a wide variety of results and it is impossible to comment on each one of them in detail. We shall pick up some general themes.

The first remark is on the choice of the optimality criteria. The title is actually somewhat misleading, because the only optimal designs that are surveyed in it are A- and MV-optimal ones. A-optimality and MV-optimality certainly appear to have very intuitive and appealing statistical interpretations and, according to the authors, are the most widely studied criteria for this type of experimental design. It is a rather disturbing thought, however, that neither of these criteria takes into account the covariances of the estimated treatment-control contrasts.

Besides, other criteria may be relevant in this context. For pilot experiments when the control is taken to be known and the interest lies in testing whether or not the overall effect of the new treatments is appreciable, we may want to contrast the *average* new treatment effect with the old one and minimize $\text{var}(\sum_i \hat{t}_i/v - \hat{t}_0)$, i.e., min var $\sum_i (\hat{t}_i - \hat{t}_0)$, $i = 1, \cdots, v$. This criterion, which can be easily extended to the case of more than just one control, is also mentioned by Majumdar (1986) and it seems appropriate to call it J-optimality because it reduces to minimizing $\text{trace}(JPC_d - P')$, with J the $v \times v$ matrix of all ones. In Giovagnoli and Righi (1985) and Notz (1985), it is shown that certain J-optimal designs are also E-optimal, where E-optimality is defined as minimizing the maximum variance of all the estimated contrasts $\sum_i c_i(t_i - t_0)$ with $\sum_i c_i^2 = 1$, and conversely some sufficient conditions for E-optimality turn out to ensure J-optimality too. Thus although E-optimality does not appear to have a very natural statistical interpretation when there is a control, E-optimal plans may also deserve attention in some cases.

Lastly we would like to stress that in the Bayesian approach, due to the (possibly) different prior assumptions on the test treatments and the control, it is no longer true that designs which are D-optimal for inference on treatment-control contrasts, i.e., which minimize the determinant of the posterior covariance matrix of those contrasts, are always D-optimal for any set of contrasts. Thus in this case it is worthwhile to look at D-optimality too.

J-optimality shares with A- and MV-optimality (and also with E- and D-optimality, and others) the property of being invariant under all relabeling of the test treatments which leave the control unchanged. We believe this invariance under a suitable group to be the key to many results on optimal designs, and in

A. Giovagnoli is Associate Professor, Dipartimento di Scienze Statistiche, Universita degli Studi di Perugia, Via A. Pascoli, 06100 Perugia, Italy. I. Verdinelli is Associate Professor, Dipartimento di Statistica, Prob. e Stat. Appl., Universita di Roma, Piazzale A. Moro 5, 00185 Roma, Italy.