Anthony C. Atkinson

Commentary on 'Designs for dose-escalation trials with quantitative responses'

Article (Accepted version) (Refereed)

STATISTICS IN MEDICINE
Statist. Med. 2000; **00**:1–6 *Prepared using* simauth.cls *[Version: 2002/09/18 v1.11]*

Discussion of "Designs for dose-escalation trials with quantitative responses" by R.A. Bailey

A. C. Atkinson*

Department of Statistics, London School of Economics, London WC2A 2AE. E-mail: a.c.atkinson@lse.ac.uk

1. Introduction

Professor Bailey has written an important and interesting paper in which experimental designs for dose escalation are developed, improved and illuminated by her understanding and experience of the properties of block designs (see [BaileyBailey2008]). Given the clarity of her exposition I have few specific comments.

2. Caution in First Application of a Higher Dose

The interpretation of dose escalation studies as block designs assumes that individual cohorts are homogeneous. Even if they are, the Te Genero trial shows the perils that can occur for the first cohort. Recommendation 7.5.3 of [Senn, Amin, Bailey, Bird, Bogacka, Colman, Garrett, Grieve, and Lachmann suggests that simultaneous treatment of all subjects is inappropriate. Should a similar recommendation be made for the new dose in each cohort, not just the first, in case a tolerance threshold is suddenly crossed?

3. Classes of Design: quadratic trend in the response

The general method of development of designs in the paper is to move from principles of design to assessment using tools from optimum design theory. However, the comparisons can only be over the class of designs considered. Are there designs not in this class that have slightly higher values of, perhaps, the A-optimum criterion mentioned in Professor Bailey's §6? If such designs exist, are they only slightly more efficient than the designs found here?

As one example, in the customary approach to block designs the treatment effects τ_i are arbitrary constants. But, with ordered doses there may be a smooth change of response. Is it worthwhile modelling this, for some of the larger designs of Figure 3, as a polynomial trend? The effect on the design may be appreciable; for a single cohort the optimum design for a quadratic model allocates equal numbers of subjects to the minimum, maximum and central dose, or across the two central doses if the number of doses is even.

4. D-optimum Design for Quadratic Trend

I assume that the dose levels are equally spaced levels of a single factor x. Equally spaced intervals in logdose are another possibility although scaling of the placebo then needs special attention. The model for the expected response of a subject in cohort k who gets dose x_i is $\mu_k + \beta_1 x_i + \beta_2 x_i^2$. I take the conditions (i) of Figure 3, corresponding to the last designs in Table I. With the standard design there are ten cohorts and up to 11 dose levels, so that $k = 1, \ldots, 10$; in all there are 12 parameters. Since I am interested in parametric curves, I used D-optimality which minimises the volume of confidence regions for parameter estimates. When, as here, the cohort sizes are fixed the D-optimum design for all parameters is the same as the D_S-optimum design for the parameters of interest β_1 and β_2 ([Atkinson, Donev, and TobiasAtkinson et al.2007, p. 206]).

The algorithm used to find the optimum design is a simple modification of sequential design construction ([Atkinson, Donev, and TobiasAtkinson et al.2007, §11.2]). The resulting design for 200 subjects is in my Table I. The important feature is that the D_S-efficiency of the design in Figure 3 relative to this design is 60.05%. The two designs have some features in common. Both put about half the subjects in any cohort on the highest dose. One important difference is that this new design allocates many more patients to placebo - 82 as opposed to 27; another is the allocation of several subjects to central doses for the later cohorts as is to be expected.

Table I. A D-optimum design for a 10 cohort design when the dose response is quadratic. The relative D_S-efficiency of Design (i) of Figure 3 is 60.05%

	(i) $n = 10, m = 20$										
Dose	0	1	2	3	4	5	6	7	8	9	10
Cohort 1	10	10	0	0	0	0	0	0	0	0	0
Cohort 2	10	0	10	0	0	0	0	0	0	0	0
Cohort 3	10	0	0	10	0	0	0	0	0	0	0
Cohort 4	10	0	0	0	10	0	0	0	0	0	0
Cohort 5	10	0	0	0	0	10	0	0	0	0	0
Cohort 6	10	0	0	0	0	0	10	0	0	0	0
Cohort 7	9	0	0	0	1	1	0	9	0	0	0
Cohort 8	8	0	0	0	4	0	0	0	8	0	0
Cohort 9	5	0	0	3	3	0	0	0	0	9	0
Cohort 10	0	0	0	10	0	0	0	0	0	0	10
	(200 subjects)										

5. Compound Designs

A desirable feature of the designs of Figure 3, perhaps due to the 'Diversity Principle', is that the truncated designs from early stopping are close to the recommended design

for a smaller number of cohorts. This property is unlikely to hold for designs such as that in my table. Other criteria of optimality might give designs with better properties in this respect. It is however not necessary to choose between such criteria; compound criteria can be used in which a weighted product of efficiencies is maximized ([Atkinson, Donev, and TobiasAtkinson et al.2007, cap. 21]). The purpose, as in Professor Bailey's §9, is to find designs that perform well for a variety of criteria.

6. The Halving Principle

The construction of the new designs in Professor Bailey's Sections 3 and 7 by the use of a few straightforward principles is appealing and should encourage their use. However, the design in my table is such that slightly less than half of the subjects in cohort k receive dose k for $k = 7, 8$ and 9. I think this design would be excluded by the 'Halving Principle'. Were such designs explored?

7. Designs for Categorized Responses

The design problem is made more general and simpler to solve by considering quantitative responses, when the designs do not depend on the values of the parameters of the linear model, in my case μ_k, β_1 and β_2. A difficulty for designs for categorical response, such as those considered by [Zhang, Sargent, and MandrekarZhang et al.2006], is that the designs depend on the unknown parameters of the logistic models used to categorise the response variables. However, if the effects being explored are small, the results of [CoxCox1988] indicate that designs for normal theory linear models are likely to be efficient. In any application numerical exploration will be needed to determine the region over which designs for continuous response have adequate efficiency.

REFERENCES

[Atkinson, Donev, and TobiasAtkinson et al.2007]. Atkinson, A. C., A. N. Donev, and R. D. Tobias (2007). *Optimum Experimental Designs, with SAS*. Oxford: Oxford University Press.

[BaileyBailey2008]. Bailey, R. A. (2008). *Design of Comparative Experiments*. Cambridge: Cambridge University Press.

[CoxCox1988]. Cox, D. R. (1988). A note on design when response has an exponential family distribution. *Biometrika 75*, 161–164.

[Senn, Amin, Bailey, Bird, Bogacka, Colman, Garrett, Grieve, and LachmannSenn et al.2007]. Senn, S., D. Amin, R. A. Bailey, S. M. Bird, B. Bogacka, P. Colman, A. Garrett, A. Grieve, and P. Lachmann (2007). Statistical issues in first-in-man studies. *Journal of the Royal Statistical Society, Series A 170*, 517–579.

[Zhang, Sargent, and MandrekarZhang et al.2006]. Zhang, W., D. Sargent, and S. Mandrekar (2006). An adaptive dose-finding design incorporating both toxicity and efficacy. *Statistics in Medicine 25*, 2365–2383.