The Potential of Adaptive Designs

Norbert Victor, Meinhard Kieser, Tim Friede
Institut für Medizinische Biometrie und Informatik, Universität Heidelberg
Im Neuenheimer Feld 305
D-69120 Heidelberg, Germany
email: victor@imbi.uni-heidelberg.de

1. Introduction

In the late seventies, designs for clinical trials with interim analyses have been introduced for ethical, economical, and practicability reasons (Pocock, 1977). The flexibility and facilities of such designs have been continuously improved, e.g. by the so-called *error spending function* approach (Lan & DeMets, 1983). The most important step towards more flexibility, however, was the introduction of the adaptive designs by Bauer and Köhne (1994) where control of the Type I error rate (global level α) is achieved by Fisher's combination test. This was the starting point for a multitude of further work that created an imposing tool for design adaptations during the ongoing trial. It became evident that data driven design modifications are possible which go far beyond sample size adaptation and which can be performed without inflation of the nominal Type I error. In this paper we are drawing attention to some of these possibilities of design adaptation.

2. Adaptation of the Sample Size

The issue of sample size adaptation based on interim results of a trial has attracted special attention recently. Two different types of designs have been considered: (i) designs with interim analyses where early stopping of the trial is possible; here, the primary endpoint is compared between the groups requiring an unblinding of the treatment allocation; (ii) the "internal pilot study" design, where after a certain number of patients nuisance parameters are re-estimated and the sample size is adapted based on the estimates of these parameters. In case (ii) the estimation can be done with or without unblinding the treatment allocation of the patients. Friede (2000) has shown that for designs with internal pilot study and normally distributed data "blinded" adaptation of the sample size can be done without an inflation of the Type I error rate. This is an important result because in the case of unblinding during the ongoing trial, the interim analysis must be performed by an Independent Data Monitoring Committee in order to avoid bias. The establishment of such a committee makes the trial more expensive and its organisation cumbersome. An approach which allows to adapt the number of interim analyses and the sample size during the trial (self designing trials) has been proposed by Fisher (1998) and enhanced by Müller & Schäfer (2001).

3. Adaptation of the Hypotheses

Another interesting possibility for the adaptation of the study design is a change of the set of hypotheses after an interim analysis. A reduction of the number of hypotheses to be tested is of interest for studies with multiple endpoints and/or multiple treatment arms. In dose-finding studies such strategies allow, for example, to drop – on the basis of interim results – dosages (and the relating hypotheses) that show unsatisfactory results. The trial is then continued with the remaining dosage groups, and in the final analysis only a reduced number of hypotheses is tested. A hypothesis set concerning multiple endpoints can be reduced analogically. Various kinds of adaptation rules were described, and it was shown that these strategies can lead to a considerable increase of power (Bauer & Kieser, 1999; Kieser, 2000; Friede & Kieser, 2001). Multiple test procedures are available that guarantee control of the experiment-wise error rate in the strong sense for these multiple test problems (Kieser, Bauer & Lehmacher, 1999; Kieser, 2000). Furthermore, even an inclusion of "promising" hypotheses after the interim analysis is possible (Kropf et al., 2000).

4. Change of the Test Statistic

Adaptive designs also allow to choose the test statistic for the final analysis on the basis of the information gained from the interim analysis. Lang et al. (2000) proposed a procedure that uses the "optimal" scores for trend tests estimated from the interim data. A different approach was described by Kieser (2000) where the test statistic with the most favourable power characteristic for the data of the interim analysis is selected for the final evaluation.

5. Theoretical Background

The proof that all these adaptations are possible with control of the Type I error rate makes use of the fact that for the Bauer-Köhne design separate test statistics are calculated from the disjoint samples of the study stages. In fact, the only required property to assure Type I error control under arbitrary (data-dependent) design adaptations is that the distribution of the *p*-value conditionally on the *p*-values from previous stages is stochastically larger or equal than the uniform distribution (Bauer, Brannath & Posch, 2001).

Conclusions

In the last years a number of papers were published dealing with the facilities of a design adaptation in clinical trials. Meanwhile, there is a large tool of methods available that allows a high degree of flexibility without compromising the Type I error rate. Only a few of these possibilities could be touched here, and it could be expected that further methods for design adaptations based on the interim results of a trial will be developed in future. Such designs reduce (in the average) the number of patients and of false negative results and herewith the costs of a trial. Therefore, their routine application in clinical trials should be considered, as recently recommended by the chief statistician of FDA (O' Neill, 2000).

REFERENCES:

Bauer P, Brannath W, Posch M (2001). Flexible two-stage designs: An overview. *Methods of Information in Medicine* **40**, 122-126.

Bauer P, Köhne K (1994). Evaluation of experiments with adaptive interim analyses. *Biometrics* **50**, 1029-1041. Correction in *Biometrics* **52**, 380.

Fisher L (1998). Self-designing clinical trials. *Statistics in Medicine* **17**, 1551-1562

Friede T (2000). Methods for the determination of the sample size in clinical trials with internal pilot study (in German). Doctoral thesis, Medical Faculty, University of Heidelberg.

Friede T, Kieser M (2001). A loss function based approach for dose selection in two-stage dose-response trials. To appear in *Journal of Epidemiology and Biostatistics*.

Kieser M (2000). Methods for decision making in clinical trials with adaptive design (in German). Senior lecturer thesis, Medical Faculty, University of Heidelberg.

Kieser M, Bauer P, Lehmacher W (1999). Inference on multiple endpoints in clinical trials with adaptive interim analyses. *Biometrical Journal* **41**, 261-277.

Kropf S, Hommel G, Schmidt U, Brickwedel J, Jepsen MS (2000). Multiple comparisons of treatments with stable multivariate tests in a two-stage adaptive design, including a test for non-inferiority. *Biometrical Journal* **42**, 951-965.

Lan KKG, DeMets DL (1983). Discrete sequential boundaries for clinical trials. *Biometrika* **70**, 659-663.

Lang T, Auterith A, Bauer P (2000). Trend tests with adaptive scoring. *Biometrical Journal* **42**, 1007-1020

Müller HH, Schäfer H (2001). Adaptive group sequential designs for clinical trials: combining the advantages of adaptive and classical group sequential approaches. To appear in *Biometrics*.

O' Neill R (2000). New approaches in data management and biostatistics. *International Journal of Pharmaceutical Medicine* **14**, 241

Pocock SJ (1977). Group sequential methods in the design and analysis of clinical trials. *Biometrika* **64**, 191-199.

RESUME: Des possibilités d'adaptation du plan d'expérience pendant un essai clinique sont résumés. Il devient évident que ces possibilités – récemment publiés – permettent plus que

l'adaptation de la taille d'échantillon. La contrôle de l'erreur de première espèce est garanti pour toutes ces d'adaptations.