Issues in Type I error rate control in adaptive clinical trials with treatment selection

Martin Posch[1], Willi Maurer[2], and Frank Bretz[2]

[1] Section of Medical Statistics, Medical University of Vienna Spitalgasse 23, A-1090 Vienna, Austria
e-mail: martin.posch@meduniwien.ac.at

[2] Clinical Information Sciences, Novartis Pharma AG, 4002 Basel, Switzerland

Session: "Treatment Selection Designs in Clinical Trials"

Abstract. Interest in adaptive designs for confirmatory clinical trials has increased in the past few years. A particularly appealing application is the use of adaptive designs in combined phase II/III studies with treatment selection at interim. These studies would start comparing several treatments with a control. One (or more) treatment(s) would then be selected after the first stage based on the available information at an interim analysis, including interim data from the ongoing trial, external information and expert knowledge. Recruitment would continue, but now only for the selected treatment(s) and the control, possibly in combination with a sample size reassessment. The final analysis of the selected treatment(s) include(s) the patients from both stages and is performed such that the overall type I error rate is strictly controlled, thus providing confirmatory evidence of efficacy at the final analysis. In this talk we describe two approaches to control the type I error rate in adaptive designs with sample size reassessment and/or treatment selection. The first method adjusts the critical value using a simulation based approach, which incorporates the number of patients at an interim analysis, the true response rates, the treatment selection rule, etc. We discuss the underlying assumptions of simulation based procedures and give several examples where the type I error control is lost if some of the assumptions are violated. The second method is an adaptive Bonferroni-Holm test procedure based on conditional error rates of the individual treatment-control comparisons. We show that this procedure controls the type I error rate, even if a deviation from a pre-planned adaptation rule or the time point of such a decision is necessary. Motivated by a real case study, a simulation study is conducted to compare the two methods with respect to power and type I error rate control, being aware that many further considerations are necessary when designing an adaptive design.

Keywords. Interim Analysis, Flexible Designs, Treatment Selection, Multiple Testing

1 Introduction

Interest in adaptive designs for confirmatory clinical trials has increased in the past few years. Adaptive designs use accumulating data of an ongoing trial to decide on how to modify design aspects without undermining the validity and integrity of the trial. A particularly appealing application is the use of adaptive designs in combined phase II/III studies with treatment selection at interim. Such study would start comparing several treatments with a control. One (or more) treatment(s) would then be selected after the first stage based on the available information at interim, including interim data from the ongoing trial, external information and expert knowledge. Recruitment would continue, but now only for the selected treatment(s) and the control and with possibly reassessed sample sizes. The final analysis of the selected treatment(s) includes the patients from both stages and is performed such that the overall type I error rate is controlled at a pre-specified significance level $\alpha \in (0, 1)$, thus providing confirmatory evidence of efficacy at the final analysis (Bauer and Kieser, 1999). The aim is to control the familywise error rate (FWER) in the strong sense, i.e., the probability to reject erroneously a null hypothesis is bounded by α under any configuration of true and false null hypotheses.

In this talk we investigate two approaches for such designs. One is based on the estimation of the type I error rate based on simulated clinical trials assuming a pre-specified adaptation rule. We demonstrate the limitations and underlying assumptions of such simulation based procedures and give several examples where the type I error control is lost if some of the assumptions are violated. The second approach is an adaptive Bonferroni-Holm test procedure based on conditional error rates of the individual treatment-control comparisons. This procedure controls the type I error rate even if neither the adaptation rule nor the time point of the interim analysis are pre-planned.

2 Simulation Based Approach

To estimate a type I error rate by simulations of clinical trials with a pre-specified adaption rule, it is required that the distribution of the final test statistics (under sample size reassessment) is fully specified under the null hypothesis and does not depend on nuisance parameters. However, in real clinical trials these requirements are often not satisfied and adaptation rules exist, where the distribution of the final test statistics may depend on unknown parameters: e.g.,

1. If nuisance parameters are present, the type I error rate may not be determined by the given null hypothesis but may depend on these nuisance parameters. This is the case, for example, in a parallel group treatment-control comparison, where the adaptation rule depends on absolute response rates (and not only on differences of response rates to the control). Then the type I error rate may depend on the absolute response rates.
2. If the adaptation rule depends on data that is correlated with the primary endpoint, the type I error rate will typically depend on the overall joint distribution. This is of particular importance in clinical trials, because considerations other than the observed results for the primary (efficacy) variable may influence the interim decision and the joint distribution is usually unknown.
3. If several treatments are investigated in a clinical trial, the type I error rate will depend on the (unknown) effect sizes as well as on the configuration of true and false null hypotheses. Since we aim at controlling the FWER in the strong sense, simulations under the assumption that all treatment effects are the same, (including the one from the control group) may not be sufficient.
4. If one deviates from the pre-specified adaptation rule, the type I error rate may not be controlled at level α. This may be the case if, for example, the timing of the interim analyses is triggered by unexpected adverse events and thus deviates from the pre-planned time points.

Based on an actual case study we give an example where the type I error rate under the global null hypotheses (stating that there is no treatment effect in any of the active treatment groups) is smaller than in scenarios where there is a treatment effect in some of the treatment arms. Additionally, we show how deviations from the pre-planned selection rule may lead to an inflation of the type I error rate.

3 The Partial Conditional Error Approach

Bauer and Kieser (1999) proposed a flexible procedure to control the FWER in designs with adaptive treatment selection which is based on the application of the closed testing procedure and combination tests. It does not require to pre-specify the adaption rules a-priori and allows adaptations based on primary and secondary endpoints as well as external data. For treatment control comparisons of normally distributed endpoints König et al. proposed an alternative procedure based on the conditional error rate of the Dunnett test. Similar to the combination test approach it does not require a pre-specification of the adaptation rules.

In this work we propose an adaptive version of the Bonferroni-Holm test. In contrast to the approach based on the Dunnett test it is only based on the (conditional) marginal distribution of the test statistics for the individual treatment control comparisons. If no treatment is dropped (and hence, no adaptation is performed) the classical Bonferroni-Holm test is performed. Only if a treatment is dropped a modified test statistics is applied that is in the spirit of the conditional error approach (Müller and Schäfer, 2004). Applying the conditional error approach directly to the Bonferroni Holm test would require the computation of the conditional error rate given the interim data (i.e. the probability to reject the respective null hypothesis given the interim data and assuming the null hypothesis is true). This would necessitate the evaluation of the joined distribution of all elementary test statistics. However, the adaptive test proposed here is based on the sum of marginal conditional error rates of the elementary tests. The marginal conditional error rate for an elementary hypothesis test is defined as the conditional probability that its test statistics will exceed the Bonferroni adjusted critical value in the final analysis. Thus no information on the joined distribution of test statistics is required to implement the test. Similarly, the sum of conditional error rates of elementary tests was used in [4] to construct a sequential Bonferroni test that asymptotically exhausts the type I error rate in a non-adaptive design setting.

Note that the proposed adaptive test does not rely on binding futility rules. Thus, the type I error is controlled even if one does not apply the pre-specified futility criteria.

For a detailed description of the partial conditional error test see [5].

References

Bauer, P. and Kieser, M. (1999), Combining Different Phases in the Development of Medical Treatments Within a Single Trial, *Statistics in Medicine*, 18, 1833–1848.

König, F., Brannath, W., Bretz, F., and Posch, M. (2008), Adaptive Dunnett Tests for Treatment Selection, *Statistics in Medicine*, 27, 1612–1625.

Müller, H. H. and Schäfer, H. (2004), A General Statistical Principle for Changing a Design Anytime During the Course of a Trial, *Statistics in Medicine*, 23, 2497–2508.

M. Posch and F. Futschik. (2008) A uniform improvement of bonferroni type tests by sequential tests. *JASA*, 103:299–308.

Posch, M., Maurer, W., and Bretz, F. (2008), Type I error rate control in adaptive designs for confirmatory clinical trials with treatment selection at interim, *(submitted)*.