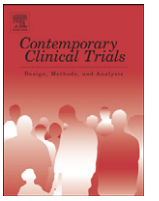
Contemporary Clinical Trials 30 (2009) 171–177

Contents lists available at ScienceDirect

Contemporary Clinical Trials

journal homepage: www.elsevier.com/locate/conclintrial



Adaptive trial design: A general methodology for censored time to event data ☆

Antje Jahn-Eimermacher *, Katharina Ingel

Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI), University of Mainz, 55101 Mainz, Germany

ARTICLE INFO

Article history:
Received 10 July 2008
Accepted 8 December 2008

Keywords:
Clinical trial
Adaptive design
Survival trial
Time to event
Cox regression
Inverse Normal method

ABSTRACT

Adaptive designs allow a clinical trial design to be changed according to interim findings without inflating type I error. The Inverse Normal method can be considered as an adaptive generalization of classical group sequential designs. The use of the Inverse Normal method for censored survival data was demonstrated only for the logrank statistic. However, the logrank statistic is inefficient in the presence of nuisance covariates affecting survival. We demonstrate, how the Inverse Normal method can be applied to Cox regression analysis. The required independence between test statistics of the different stages of the trial can be obtained by two different approaches. One is using the independent increment structure of the score process. The other uses right censoring and left truncating to divide individuals follow-up into per-stage data. Simulation studies show, that performance of the adaptive design does not depend on the method used for obtaining independence. Either way, an adaptive Cox regression analysis is more efficient than an adaptive logrank analysis if nuisance covariates affect survival.

1. Introduction

In many clinical trials the comparison of time to an event between treatment groups is of primary interest. The power to detect a specific treatment effect depends on the number of events, which is the information in survival trials. The sample size necessary to get the required number of events depends on event rates, duration of follow-up and – due to staggered entry – on recruitment rate. At start of a trial there may be some uncertainty about planning characteristics such as treatment effect, recruitment or variability between subjects. It is thus appealing to use information from the current study to adjust the sample size and/or duration of the trial to ensure adequate power. This is of particular concern in survival trials, which are long-term in most cases, so that design modifications may become necessary during the course of the trial.

Adaptive designs have been investigated over the past two decades [1–5]. Other than classical group sequential designs, these methods do not only allow for early stopping, but also for data dependent changes of a trial design without inflating the type I error. All these methods are based on a prespecified combination of test statistics calculated for the different stages of the trial. Control of type I error rate relies on the independence of these stage-wise test statistics. Independence is easily obtained, if subsets of subjects analyzed per stage are disjoint. Thus, all of the methods are directly applicable to instantaneously observed outcomes.

However, in survival trials with censored time to event data, subjects that are at risk over the duration of more than one stage contribute information to each of these stages. As a consequence, test statistics may be correlated and the adaptive designs are not directly applicable for survival response.

In recent years some authors considered how to use adaptive techniques for survival designs. Schäfer und Müller [6] proposed a version of the conditional error function approach [3] for the logrank test to compare two survival distributions. They used the fact, that increments of the logrank test statistic are asymptotically stochastically independent as was shown

☆ This research was supported by a grant of the Deutsche Forschungsgemeinschaft (DFG) Grant JA 1821/1-1.

* Corresponding author. Tel.: +49 6131 172433; fax: +49 6131 172968.
E-mail address: jahn@imbei.uni-mainz.de (A. Jahn-Eimermacher).

1551-7144/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.cct.2008.12.002

for the classical group sequential test designs. The idea of using the independent increment structure of the logrank statistic was also followed by Wassmer [7], who generalized the Inverse Normal method to a comparison of censored survival outcomes. Shen and Cai [8] showed how the variance spending approach of Fisher [4] and Shen and Fisher [9] can be used to adaptively analyse survival data with linear rank tests.

However, a limitation in all those methods is, that they do not consider risk factors other than the treatment available at baseline. Not adjusting for risk factors that have substantial effect on the course of disease can be inefficient and parameter estimates may be biased [10–12].

In the present paper, we propose an adaptive method for survival trials with censored time to event data and staggered entry allowing for covariate adjustment in Cox regression models. The Inverse Normal method is applied to independent test statistics derived from Cox regression analyses and calculated for the different stages of a trial. Two approaches are proposed to obtain independent test statistics. The first one makes use of the independent increment structure of the score process, which was demonstrated in the setting of group sequential designs [13,14]. The second one divides data into different stage data by left truncation and right censoring, each comprising only those data observed during the considered stage of the trial. Test statistics calculated for these per-stage data were shown to be independent. This idea goes back to the work of Keiding et al.[15–17]. They demonstrated how to use data from patients that were under risk, but had no event before interim analysis, to confirm an unexpected interim finding in a confirmatory manner. Bauer and Köhne [1] already discussed the affinity between Keidings idea and adaptive designs, which was picked up in this paper.

As the adaptive method is based on the Inverse Normal method, the whole framework of group sequential designs can be used to define stopping strategies including stopping for futility. This is in contrast to the approach of Shen and Cheng [18], who generalized the variance spending method to Cox regression analyses.

The paper is organized as follows. In Section 2 we sketch the adaptive Inverse Normal method. The adaptive design for survival data using Inverse Normal is presented in Section 3 with the two approaches obtaining independent test statistics described in Sections 3.1 and 3.2. The performance of the adaptive design is illustrated in Section 4 on simulated data. The paper closes with a discussion in Section 5.

2. Inverse Normal method

Consider a K-stage trial design and let w_k denote the predefined weight of stage k with $w_k>0$ and $\sum_{k=1}^{K} w_k^2 = 1$, $k=1\ldots K$. Let p_k denote the p-value calculated from stage k, $k=1\ldots K$. The test decision after stage i is based on the standardized test statistic

$$Z_i^* := \left(\sum_{k=1}^{i} w_k^2\right)^{-1/2} \sum_{k=1}^{i} w_k \Phi^{-1}(1-p_k). \quad (1)$$

For p_k independent and uniformly distributed under H_0, $\Phi^{-1}(1-p_k)$ are independent and standard normally distributed under H_0. Thus, stopping boundaries from classical group sequential designs for independent, normally distributed increments can be used [2]. As long as the p_k remain independent and uniformly distributed under H_0, information from the current study can be used to change the trial design without inflating type I error. Weights w_k must be used as prespecified and are not allowed to be adapted according to the changed stage sizes. The uniform distribution is usually satisfied only under a point null hypothesis. However, Brannath et al. [19] released this condition to the so-called "p-clud-condition", which is also satisfied for one-sided interval hypotheses. Thus, results derived in Section 3 for point null hypotheses are applicable also for one-sided interval hypotheses. Note, that one-sided p-values have to be used for the Inverse Normal method in order to avoid directional conflicts. Two-sided tests can be derived by performing two one-sided tests.

3. Inverse Normal method for time to event data

Suppose, p-values calculated from analyses at the different stages of the trial were independent and uniformly distributed under H_0. Then, an adaptive design could be derived straightforward applying the Inverse Normal method. The crucial point is to derive independent p-values for survival data as subjects may be at risk over more than one stage of the trial. We propose two different approaches for obtaining independent test statistics. The first one uses the independent increment structure of test statistics. This idea was used by Wassmer [7] for the logrank test statistic, which is sketched in Section 3.1.1 for the sake of completeness although not allowing for covariate adjustment. The approach is generalized in Section 3.1.2 to Cox regression analysis using the independent increment structure of the score test. The second approach divides data into stagewise data, so that test statistics calculated on stagewise data are approximately independent (Section 3.2).

Consider a randomized trial. Let subjects enter the trial staggered and random and let entry times be independent of the survival and censoring times. Censoring is assumed to be independent from survival.

3.1. Independent increment statistics

3.1.1. Logrank test

For sake of illustration, only two treatment groups are considered. With θ being the log-hazard ratio, suppose a null hypothesis $H_0=\{\theta=0\}$ is to be tested against a local alternative $H_1=\{\theta>0\}$. More generally, the null hypothesis may be formulated in terms of equality of the two survivor distributions, as the logrank test does not rely on proportional hazards. Assume, d_k accumulated events are observed at analysis k and let $x_{1k}<x_{2k}<\ldots x_{d_k k}$ be the ordered death times (assuming no ties). Let d_{1ik} be 1, if the death time x_{ik} corresponds to a subject in treatment group 1 and 0, otherwise. Let n_{1ik} and n_{2ik} be the number of subjects at risk in group 1 and 2 at time x_{ik} known at analysis k and $n_{ik}:=n_{1ik}+n_{2ik}$. The logrank test statistic at stage k is defined as

$$LR_k = \frac{\sum_{i=1}^{d_k}\left(d_{1ik}-\frac{n_{1ik}}{n_{ik}}\right)}{\sqrt{\sum_{i=1}^{d_k}\left(\frac{n_{1ik}n_{2ik}}{n_{ik}^2}\right)}}.$$

The logrank test is a powerful nonparametric test, particularly when the alternative to equal hazards is proportional