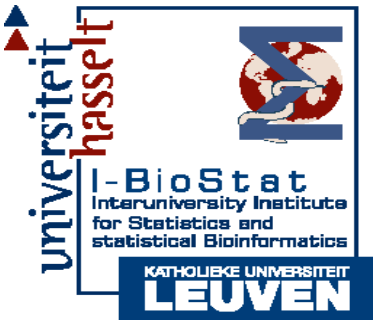
Modelling the effect of risk factors on the time to pregnancy using regression models for survival data

Demeke Kebede Nurgi

promotor :
dr. Goele MASSONNET

Eindverhandeling voorgedragen tot het bekomen van de graad
Master of Statistics Biostatistics

universiteit hasselt



Universiteit Hasselt
Center for Statistics

Modeling the Effect of Risk Factors on the Time to Pregnancy Using Regression Models for Survival Data

By:

Demeke Kebede Nurgi

Supervisor: Dr. Goele Massonnet

Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Biostatistics.

September 2009

Certification

This is to certify that this report was written by Demeke Kebede Nurgi under my thorough supervision.

Dr. Goele Massonnet

……

Signature

Acknowledgements

First of all, I would like to give thanks to Almighty God who created me and who also cares for me.

I would like to express my sincere gratitude to my supervisor Dr. Goele Massonnet for her extensive and continuous guidance and support during the preparation of this report. My most sincere thanks are also expressed to all my professors at Hasselt University for creating such conducive learning environment for me.

Finally, special thanks go to my family members and all my friends for the emotional support and encouragement.

Abstract

Information on the time to pregnancy (TTP) was available for a sample of Danish couples in the age range of 20-35 years who were planning to discontinue contraception to achieve a pregnancy. The information was collected by follow-up over six complete menstrual cycles. The aim was to investigate the effect of risk factors such as smoking status, sperm concentration, alcohol consumption and intake of caffeine on the time to pregnancy. For this purpose, a proportional hazards model, an additive hazards model and a proportional hazards model with time varying effects were fitted. We performed goodness-of-fit for each model which indicates which model is appropriate to model the data.

All the models revealed a similar result with regard to the significance of covariates. Alcohol consumptions in women and sperm concentration were found to be significantly associated with time to pregnancy. The result obtained from all models showed a negative impact of alcohol consumption of women on the monthly probability of conception. This means that the hazard of time to pregnancy decreases if the number of alcoholic drinks increases for the female. The study also revealed a significant positive effect of sperm concentration on the hazard of time to pregnancy. Checking the goodness-of-fit for all models showed there is a problem in the functional form of the covariate sperm concentration (M.ZKON0). A logarithmic transformation of this covariate resulted in a better fit for the additive hazards model.

Key words: Time to pregnancy; Proportional hazards model; Additive hazards model; Martingale.

Table of Contents

List of Tables

List of Figures

1. Introduction

The most crucial event of the reproductive process is the birth of a new offspring. Fertility measures the frequency of this key event. The endpoint for the occurrence of new offspring is typically defined by a live birth. However, it may be defined in a broader sense by a clinical pregnancy or by a conception. Time to pregnancy (TTP) can be defined as the time elapsing from when a couple decides to have a child to clinically recognizable pregnancy. It can be used to compare fertility between groups. However, it is more widely used as a measure of fecundability, a related but slightly different quantity. Fecundability is defined as the probability of achieving conception or a recognized pregnancy per menstrual cycle (El-Shaarawi and Piegorsch, 2002; Curtis et al., 1997).

Alcohol use is associated with altered levels of estrogen and progesterone and irregularities in the menstrual cycles and ovulation (Silva et al., 1999). Similarly a study on experimental animals founds that alcohol is known to decrease steroid hormone concentrations, inhibit ovulation, and interfere with sperm cell transportation through the fallopian tube (Sharma and Chaudhury, 1970). Various studies have addressed the possible association between alcohol intake and fecundity. Most of the studies have found no effect of moderate alcohol intake, whereas a high intake has been associated with reduced fecundability. A large European multicentre study found that women with a preconceptional alcohol intake >14 drinks per week waited longer for a pregnancy than women with no alcohol intake (Olsen et al., 1997). Associations between alcohol consumption and infertility have also been reported by Hakim et al. (1998). Their prospective observational study for 100 menstrual cycles resulted in more than 50% reduction in the probability of conception during a menstrual cycle during which participants consumed alcohol. However, a study on Danish National Birth Cohort showed shorter waiting times for those with a low intake of alcohol compared with non-drinkers. They concluded that, moderate alcohol intake is not strongly associated with subfecundity; smaller amounts of alcohol may have a positive impact on the female reproductive system, perhaps by providing some stress control (Juhl et al., 2001). A study by Curtis et al. (1997) also found no association between fecundability and alcohol use among women and men.

Male reproductive health outcome include sexual function and reproductive hormonal status. However, production of normal sperm has naturally been the central topic of male reproductive health. Sperm concentration is one of the standard evaluations of semen quality. A number of studies have shown a potential decrease in male fertility when a sperm concentration is very low. However, there is no consensus on the value of this limit. From studies in couples attempting to conceive, it has been found that the fertility potential of men decreased when the sperm concentration was under $40*10^6$/ml (Bonde et al., 1998). Similarly a study by Slama et al. (2004) showed the effect of declines in sperm concentration on time to pregnancy.

Cigarette smoke is well established as a reproductive toxin. Research indicates that cigarette smoking is harmful to a woman's ovaries, and the degree of harm is dependent upon the amount and the period of time a woman smokes. Components in cigarette smoke have been shown to interfere with the ability of cells in the ovary to make estrogen and to cause a woman's eggs (oocytes) to be more prone to genetic abnormalities. In general most of the studies support the conclusion of association between female smoking and prolonged time to pregnancy (Baird and Wilcox, 1985; Curtis et al., 1997; Suonio et al., 1990). However, the studies of Baird and Wilcox, (1985) and Suonio et al. (1990) resulted in no association between male smoking habit and time to pregnancy.

The relationship between caffeine intake and fecundity is an issue of controversy. The results of research studies are conflicting and have shown both positive and negative effects of caffeine on fecundity. Hatch and Bracken (1993) found that intake of caffeine from coffee, tea, and caffeinated soft drinks was associated with an increased risk of a delay of conception. On the other hand, Hakim et al. (1998) and Curtis et al. (1997) have not found an association between caffeine intake and time to pregnancy.

This study aims at investigating the effect of risk factors (smoking status of the female, smoking status of the male, sperm concentration, number of drinks for the female and a partner, and intake of caffeine for both female and partner) on the time to pregnancy. For this purpose, a Cox proportional hazards model, an additive hazards model and a Cox proportional hazards model with time varying effects were fitted. We performed goodness-of-fit for each model which indicates which model is appropriate to model the data.

The remainder of this report is organized as follows: Section 1.1 introduces the data set in this study and Section 1.2 provides the basic concepts in survival data analysis. In Section 2 we describe the statistical methods we applied; data exploration, proportional hazards model, proportional hazards model with time varying effects and additive hazards model. Main results of the analyses are presented in Section 3 whereas Section 4 is devoted for discussion.

1.1 Data Description

The data set considered in this study includes a sample of Danish trade union cohabiting members with in the age range of 20-35 years who were planning to discontinue contraception to achieve a pregnancy. Couples without previous reproductive experience who intended to discontinue contraception to become pregnant were eligible for enrolment.

A total of 1113 couples signed up for the study. These couples were contacted regularly to ascertain their intent to discontinue the use of birth control to achieve a pregnancy and were enrolled consecutively as they stopped contraception. Discontinuation of contraception was defined as the date when the last pill was taken, the IUD was removed, or when condom, barrier, or other kinds of birth control were used for the last time. Among couples signed for the study 122 had withdrawn their consent to participate and 153 were ineligible for other reasons, mainly because of previous pregnancies (N = 46) or because they had discontinued contraception before signing up for this study (N = 71). A sample size of 310 couples is available for this study.

The intention was to start follow-up when birth control stopped and continue until a pregnancy was recognized or six complete menstrual cycles were recorded, whichever came first. The observations were censored if the couples dropped out during the study period or when follow-up reached six cycles without pregnancy. The variables in the study are sperm concentration of male in mill/ml(M.ZKON0), smoking status of the female (K.RYG), smoking status of the male (M.RYG), average number of drinks per week for the female (K.ALK), average number of drinks per week for the male (M.ALK), average intake of caffeine in mg per day for the female(K.COF) and for the male(M.COF).

During the period of follow-up, the women daily recorded sexual intercourse and menstrual bleeding, and a semen sample was collected during the menstrual period of each cycle. Participants were encouraged to give semen sample following 3 days of continence, but emphasized that samples obtained outside this time window would still be acceptable. Every month men and women filled in a questionnaire on health, occupational exposures, smoking, alcohol and coffee consumption, intake of medicine, and psychologic stress. The women notified the research team when a pregnancy was diagnosed by a physician or when a commercial pregnancy test was positive (Bonde et al., 1998).

1.2 Survival Data Analysis: Basic Concepts

Survival analysis examines and models the time it takes for events to occur. Let T^* denote the survival time which is the time from a well defined time origin to the occurrence of some given event. The following three functions characterize the distribution of T^*.

Let *f* be the probability density function and F the corresponding cumulative distribution function of the (continuous) event time T^*. The basic quantities used to describe failure time data are the survival function

$$S(t) = 1 - F(t) = P(T^* > t)$$
$$= \int_t^\infty f(t)dt \ .$$

A survival function is a nonincreasing function with a value of 1 for t = 0 and 0 at infinity. If T^* is a continuous random variable, then, S(t) is a continuous, strictly decreasing function. A basic quantity, fundamental in survival analysis, is the hazard function. This function is defined as

$$\alpha(t) = \frac{f(t)}{S(t)} = \lim_{h \to 0} \frac{1}{h} P(t \le T^* < t + h \ / T^* \ge t),$$

which may be interpreted as the instantaneous failure rate among those at risk. The survival function may be calculated from the hazard function by

$$S(t) = \exp(-\int_0^t \alpha(s)ds) \ .$$

1.2.1 Censoring

There are certain aspects of survival data, such as censoring that generate great difficulty when trying to analyze the data using traditional statistical models such as multiple linear regressions. A censored observation is defined as an observation with incomplete information.

One possible type of censoring is right censoring, which occurs when only the lower bound for the time of interest is known. When an observation is right censored it means that the information is incomplete because the subject did not have an event during the time that the subject was part of the study. Left censoring is when all that is known is that the individual experienced the event of interest prior to the start of the study, i.e., only an upper bound for the time of interest is available. For interval censoring, the only information available is that the event occurs within some interval of time (Klein and Moeschberger, 1997). The data in this study are right censored; only lower bound for time to pregnancy is known for some individuals. It turns out that handling of censoring and other types of incomplete survival data is surprisingly easy when basing the analysis on models for the intensity function (Martinussen and Scheike, 2006). Next section provides a gentle introduction to counting processes and the martingale theory used to analyze them.

1.2.2 Counting Process

Event time data, where one is interested in the time to a specific event occurs, are conveniently studied by the use of certain stochastic processes. The data itself may be described as a so-called counting process, which is simply a random function of time t, N(t). It is zero at time zero and constant over time except that it jumps at each point in time where an event occurs, the jumps being of size 1.

Assume that n independent subjects are observed over some period of time $[0,\tau]$. For each subject a counting process, $N_i(t)$, that gives the number of events occurring before time t is observed together with possibly additional information in terms of p-dimensional covariates X_i. Models for survival data, or more generally counting process data, are very conveniently formulated through the intensity process $\lambda_i(t)$ which is defined as (Scheike, 2004)

$$\lambda_i(t) = \lim_{h \to 0} \frac{P(N_i(t+h) - N_i(t) \geq 1 / past)}{h} .$$

Define the cumulative intensity by; $\Lambda_i(t) = \int_0^t \lambda_i(s)ds$ such that $M_i(t) = N_i(t) - \Lambda_i(t)$ is a martingale. Let further $N(t) = (N_1(t),...,N_n(t))$ be an *n*-dimensional counting process, $\Lambda(t) = (\Lambda_1(t),...,\Lambda_n(t))$ its compensator, $\lambda(t) = (\lambda_1(t),...,\lambda_n(t))$ an *n*-dimensional intensity process, $M(t) = (M_1(t),...,M_n(t))$ the *n*-dimensional martingale.

2. Statistical Methodology

2.1 Exploratory Data Analysis

In any data analysis it is always a great idea to do some univariate analysis before proceeding to more complicated models. To gain insight into the shape of the survival function for each group we plotted Kaplan-Meier curves for all categorical predictors. We also consider the test of equality across strata for categorical covariates using a log-rank test. A proportional hazards model is also fitted for each continuous covariate one by one to explore the importance of the effect of each covariate on the time to pregnancy in model building. In the following sections statistical analyses employed are discussed.

2.2 Proportional Hazards Model

The most common approach to model covariate effects on survival is the Cox proportional hazards model (Cox, 1972). The Cox model was introduced by Cox (1972) in the context of survival data, and Andersen and Gill (1982) extended it to the counting process framework and gave elegant martingale proofs for the asymptotic properties of the associated estimators (Martinussen and Scheike, 2006).

Suppose we observe n independent and identically distributed (i.i.d.) observations $(T_i, \Delta_i, \mathrm{X_i})$, where $\mathrm{T_i}$ is the right-censored life-time, Δ_i is the indicator telling us whether $\mathrm{T_i}$ is an event time

or censoring time, and X_i is a vector of explanatory variables. The Cox model assumes that the intensity is of the form

$$\lambda(t) = Y(t)\lambda_0(t)\exp(X^T\beta), \tag{2.1}$$

where $X = (X_1, \ldots, X_p)$ is a p-dimensional covariate vector and Y(t) is an at risk indicator which is one if the event is not yet occurred. The parameters of the model are the p-dimensional regression parameter vector β and the non-parametric baseline hazard function $\lambda_0(t)$. If the covariates are time-independent the model assumes that the hazard rates for different values of the covariates are proportional.

Statistical inference in the Cox model is primarily based on maximum partial likelihood. Suppose that there are n independent copies $(N_i(t), Y_i(t), X_i)$, i = 1, . . . , n, observed in some time interval $[0, \tau]$, $\tau < \infty$, and that each $N_i(t)$ has intensity on the Cox form(2.1). The regression parameter β is estimated as the maximizer to Cox's partial likelihood function (Cox, 1972; Martinussen and Scheike, 2006)

$$L(\beta) = \prod_t \prod_i \left(\frac{\exp(X_i^T\beta)}{S_o(t,\beta)} \right)^{\Delta N_i(t)}, \tag{2.2}$$

where $$S_o(t,\beta) = \sum_{i=1}^{n} Y_i(t)\exp(X_i^T\beta) \ .$$

The estimator $\hat{\beta}$ is thus found as the solution to the score equation $U(\beta) = 0$, where

$$U(\beta) = \sum_{i=1}^{n} \int_0^{\tau} \left(X_i - \frac{S_1(t,\beta)}{S_0(t,\beta)} \right) dN_i(t). \tag{2.3}$$

where $S_1(t,\beta)$ is the first order partial derivative of $S_0(t,\beta)$ with respect to β.

If the value of β is fixed, the Nelson-Aalen estimator of the cumulative baseline hazard function $\Lambda_0(t)$ can be estimated as

$$\hat{\Lambda}_0(t,\beta) = \int_0^t \frac{1}{S_0(s,\beta)} dN.(s) \quad ,$$

where $N.(t) = \sum_i N_i(t)$

Given $\hat{\beta}$, as a solution to $U(\beta) = 0$, the Breslow estimate of $\Lambda_0(t)$ is given by

$$\hat{\Lambda}_0(t) = \hat{\Lambda}_0(t,\hat{\beta}) = \int_0^t \frac{1}{S_0(s,\hat{\beta})} dN.(s)$$

The standard asymptotic likelihood inference tests; the Wald, likelihood ratio and score tests, are also valid under the Cox partial likelihood to test hypotheses about β. Let minus the derivative of the score with respect to β be denoted by I(β). We use the notation β_0 to denote the true value of β in the Cox model (2.1). The first test is the usual test based on the asymptotic normality of the (partial) maximum likelihood estimates, referred to as Wald's test. A test of the global hypothesis of H_0: $\beta = \beta_0$ is

$$\chi_w^2 = (\hat{\beta} - \beta_0)^T I(\hat{\beta})(\hat{\beta} - \beta_0)$$

which has a chi-squared distribution with p degrees of freedom if H_0 is true for large samples. The other test statistic, called the likelihood ratio makes use of the log likelihood statistic. It also holds true, using standard arguments from asymptotic theory, that the likelihood ratio statistic

$$\chi_{LR}^2 = -2\log\left(\frac{L(\beta_0)}{L(\hat{\beta})}\right)$$

and the score test statistic

$$\chi_{SC}^2 = U(\beta_0)^T I(\beta_0)^{-1} U(\beta_0)$$

both are asymptotically chi-squared with p degrees of freedom under the null.

2.2.1 Goodness-of-Fit for the Cox Model

As is the case for a linear or generalized linear model, it is desirable to determine whether a fitted Cox regression model adequately describes the data. In a usual linear regression setup, it is quite easy to define a residual for the fitted regression model. However, in regression models presented in this report the definition of the residual is not as clear-cut. A number of residuals have been proposed for the Cox model (Therneau and Grambsch, 2000; Klein and Moeschberger, 1997). Different residuals are useful for examining different aspects of the model.

2.2.1.1 Test for proportional hazards assumption

Schoenfeld (1982) proposed residuals for checking the proportional hazards assumption and the residuals are now commonly known as the Schoenfeld residuals. The Schoenfeld residual is the difference between the covariate at the failure time and the expected value of the covariate at this time (Therneau and Grambsch(2000), section 4.6).

We are interested in the test of time invariant effect in the model defined by the extended Cox model with time-varying coefficients

$$\lambda(t) = Y(t)\lambda_0(t)\exp(X^T(t)\beta(t)),$$

in which the relative risk parameters are allowed to depend on time so that the effect of a covariate can change with time. When $\beta(t)$ is not constant, the impact of one or more covariates on the hazard may vary over time. But the restriction $\beta(t) = \beta$ implies proportional hazards, therefore if proportional hazards holds then a plot of $\beta(t)$ versus time will be a horizontal line. When scaled Schoenfeld residuals are defined by the product of the inverse of the estimated variance-covariance matrix of the k^{th} Schoenfeld residual and the k^{th} Schoenfeld residual, Grambsch and Therneau (2000) showed that

$$E(s_{kj}^*) + \hat{\beta}_j \approx \beta_j(t_k),$$

where s_k^* is the scaled Schoenfeld residual and $\hat{\beta}$ is a coefficient from an ordinary fit of the Cox model. This suggests plotting the $s_{kj}^* + \hat{\beta}_j$ versus time as a method to visualize the nature and extent of nonproportional hazards.

The time-varying regression coefficients of the extended Cox model described above can be written as

$$\beta_j(t) = \beta_j + \theta_j g_j(t) \ , \tag{2.4}$$

where $g_j(t)$ is a specified function of time. One typical application of this type of testing is to let $g_j(t)$=log(t) (Cox, 1972) for j = 1, . . .p. The interest is in testing the hypothesis H_0: θ=0 with $\theta = (\theta_1,...,\theta_p)$. When the $g_j's$ are known functions, then the model with coefficients (2.4) is still a Cox model. Therefore, the score test, likelihood ratio test and Wald test can be used to test the hypothesis H_0 (Martinussen and Scheike, 2006; Therneau and Grambsch, 2000).

We denote the score function by $U = (U_1^T, U_2^T)^T$, where the first component is the derivative of partial likelihood with respect to β, and the second component is the derivative with respect to θ. Similarly we denote empirical information matrix written as a block matrix reflecting two parameter vectors by $(I_{kl})_{k,l=1,2}$ with its inverse denoted by $(I^{kl})_{k,l=1,2}$.The score test statistics may thus be written

$$T(G) = U_2^T(\hat{\beta},0) I^{22}(\hat{\beta},0) U_2(\hat{\beta},0) \tag{2.5}$$

where $\hat{\beta}$ denotes the maximum partial likelihood estimator under the null.

The score test statistic is asymptotically chi-square distributed with p degrees of freedom under the null. The GLOBAL test in R, which can be performed using the function cox.zph, approximates $T(G)$.

2.2.1.2 Cox-Snell residuals for assessing the overall fit of the model

Residuals commonly used to assess the overall fit of the Cox model are Cox-Snell residuals (Box-Steffensmeier and Jones, 2004; Klein and Moeschberger, 1997). Suppose that the proportional hazards model (2.1) has been fitted.

The Cox-Snell residuals are defined as

$$r_i = \hat{\Lambda}_0(T_i)\exp(X_i^T\hat{\beta}) \text{ , i = 1, . . . , n,}$$

where $\hat{\Lambda}_0(t_i)$ is an estimator of the cumulative baseline hazard rate and $\hat{\beta}$ is a vector of estimates of regression parameters of the Cox model.

If the model is correct and the $\hat{\beta}$'s are close to the true values of β then, the r_i's should look like a censored sample from a unit exponential distribution.

Indeed, if the model is correct, we have that

$$S(T^*) = 1 - F(T^*) \sim U[0,1] \text{ and } \Lambda(T^*) = -\ln S(T^*) \sim Exp(1),$$

where

$$S(T^*) = \exp(-\Lambda_0(T^*)e^{X^T\beta})$$

Therefore $r_i = \hat{\Lambda}_0(T_i)\exp(X_i^T\hat{\beta}) = -\ln S(T_i) \sim Exp(1)$.

To test whether the Cox-Snell residuals follow approximately a unit exponential distribution, we compute the Nelson-Aalen estimator of the cumulative hazard rate of r_i's $(\hat{\Lambda}_r(r_i))$. Then we check whether the plot of $(\hat{\Lambda}_r(r_i))$ against r_i is a straight line through the origin with a slope of 1.

2.2.1.3 Checking Functional Form of a Covariate

For a given covariate we would like to see an appropriate functional form to explain the influence of the covariate on TTP. Lin et al.(1993) suggested an important class of test statistics based on the cumulative sums of martingale-based residuals.

The martingales under the assumption that the proportional hazards model is true can be written as

$$M_i(t) = N_i(t) - \int_0^t Y_i(s)\exp(X_i^T\beta)d\Lambda_0(s)\,.$$

Using estimates from the Cox model we can estimate $M_i(t)$ as

$$\hat{M}_i(t) = N_i(t) - \int_0^t Y_i(s)\exp(X_i^T\hat{\beta})d\hat{\Lambda}_0(s)\,. \qquad (2.6)$$

The idea is now to look at different functional of these estimated residuals and see if they behave as they should under the model.

Lin et al. (1993) defined a two-dimensional cumulative residual process as

$$M_c(t,z) = \int_0^t K_z^T(s)d\hat{M}(s) \qquad (2.7)$$

where $K_z(t)$ is an nx1 matrix with elements $I(X_{i1} \le z)$ for i = 1, . . . , n, focusing here on the first continuous covariate X_1. In this case martingale residuals are grouped cumulatively with respect to follow-up time and covariate values. To summarize things further one may integrate over the entire time span to get a process only in z:

$$M_c(z) = \int_0^\tau K_z^T(t)d\hat{M}(t)$$

which can be plotted against z. To assess how unusual the observed process is under the model one may plot it along with a few realizations under the model. To further enhance the objectivity

of the graphical technique, one may complement the residual plot with some test statistics which measure the extremity of the observed process. One may therefore compute the supremum of $M_c(t)$. An unusually large value of this supremum would suggest that the functional form for the covariate may be inappropriate.

2.3 Extended Cox Model with time-varying Regression Effects

In medical studies, it is often expected that effects are time-varying. Since traditional models do not allow for a natural description of such effects, there is a need to extend these models. One may extend the Cox model relaxing some of the assumptions. One extension that seems natural is the model where the relative risk parameters are allowed to depend on time so that the effect of a covariate can change with time (Martinussen and Scheike, 2006; O'Quigley, 2008). The model assumes that the intensity has the form

$$\lambda(t) = Y(t)\lambda_0(t)\exp(X^T(t)\beta(t)) \tag{2.8}$$

in which the regression coefficients of the Cox model have been replaced by a vector $\beta(t)$ of time dependent regression functions. The baseline $\lambda_0(t)$ function still gives the intensity for an individual with covariates equal to zero. We will focus on modeling time-varying effects assuming constant covariates over time.

With the aim of presenting inferential procedures, the focus is on the estimation of the cumulative regression coefficients $(B(t) = \int_0^t \beta(s)ds)$. Based on $\hat{B}(t)$ an estimator of $\beta(s)$ may be derived, but when focus is on inferential procedures, the cumulative regression coefficients are preferable to the regression coefficients (see Martinussen and Scheike, 2006, p206-207, for a detailed discussion).

Now we focus on how to carry out inference about the time-varying regression coefficients of the extended Cox model. Two important hypotheses of interest are that a regression coefficient is

non-significant, $H_{01}:\beta_p(t) \equiv 0$, or equivalently $H_{01}:B_p(t) \equiv 0$ and the hypothesis that one of the regression coefficient functions is constant, $H_{02}:\beta_p(t) \equiv \beta_p$. The following test statistics are proposed by Martinussen and Scheike (2006) and Scheike (2004).

To test the hypothesis H_{02} one may use a test statistic depending on $n^{\frac{1}{2}}(\hat{B}(t) - B(t)$ and then approximate its distribution by a resampling process (see Martinussen and Scheike, 2006, section 6.4, for a detailed discussion). The following test statistic can be used

$$T_{2s} = n^{\frac{1}{2}} \sup_{t \in [0,\tau]} \left| \hat{B}_p(t) - \frac{\hat{B}_p(\tau)}{\tau} t \right|. \tag{2.9}$$

Similarly, a test for H_{01} can be based on the following test statistic

$$T_{1s} = \sup_{t \in [0,\tau]} \left| \frac{n^{1/2} \hat{B}_p(t)}{\hat{V}_{pp}^{1/2}(t)} \right| . \tag{2.10}$$

$\hat{V}_{pp}(t)$ is the pth diagonal element of $\hat{V}(t)$,

$$\hat{V}(t) = n^{-1} \sum_{i=1}^{n} \hat{Q}_i^{\otimes 2}(t) \tag{2.11}$$

where for a vector $a, a^{\otimes 2} = aa^T$,

$$\hat{Q}_i(t) = \int_0^t (n^{-1} X^T(s) W(s) X(s))^{-1} X_i(s) d\hat{M}_i(s), \text{ and}$$

W(t) is an nxn diagonal weight matrix.

2.3.1 Goodness-of-Fit for the Extended Cox Model

We use martingale residual techniques to validate the fit of the model. Consider the martingale

$$M_i(t) = N_i(t) - \int_0^t Y_i(s)\exp(X_i^T(s)\beta(s))ds\,.$$

We wish to see if the estimates thereof have a behavior consistent with model. The estimate in the vector form is given by

$$\hat{M}(t) = N(t) - \int_0^t \hat{\lambda}(s)ds$$

Based on the estimated martingales, one can construct the cumulative residual process (Martinussen and Scheike, 2006 section 6.9). To assess how unusual the observed process is under the model one may plot it along with a few realizations under the model. To further enhance the graphical technique one can compute the supremum of the cumulative residual process and approximate the quantiles of its limit distribution, under the model, by resampling. An unusually large value of this supremum would suggest that the functional form for the covariate may be inappropriate.

2.4 Additive Hazards Models

An alternative to the proportional hazard model of Cox and the non-parametric extension considered in the previous sections is Aalen's additive hazard model (Aalen, 1989; McKeague, 1988) where the intensity depends linearly on covariates. The proportional hazards model, discussed in the previous sections assumes that the effects of the covariates act multiplicatively on the unknown baseline hazard function. However, the additive hazards model is based on the assumption that the covariates act in additive manner on an unknown baseline hazards rate.

Under the additive Aalen model, we assume that the intensity for the counting process N(t) conditionally on p-dimensional covariate X(t) is of the form

$$\lambda(t) = Y(t)X^T(t)\beta(t) \tag{2.14}$$

where $\beta(t)$ is a p-dimensional regression coefficient and $Y(t)$ at risk indicator.

With the aim of presenting inferential procedures, focus is on estimation of the cumulative regression coefficients since these accommodate inferential procedures. Least-squares estimators are applied using the theory of counting processes.

Let for n independent replicates of ($N_i(t)$, $Y_i(t)$, $X_i(t)$), i = 1, . . . ,n, each $N_i(t)$ has intensity on the form (2.14). For n-dimensional counting process N(t) and cumulative intensity $\Lambda(t) = \int_0^t \lambda(s)ds$, define the n-dimensional martingale as $M(t) = N(t) - \Lambda(t)$ .We thus have that

$$\begin{aligned} dN(t) &= \lambda(t)dt + dM(t) \\ &= X(t)\beta(t)dt + dM(t) \end{aligned} \tag{2.15}$$

Since the increments of the martingale are uncorrelated and have zero mean, this equation suggests that the increments of $\beta(t)dt$, which we write as dB(t), can be estimated by simple multiple linear regression techniques by defining the generalized inverse of X(t) as the p*n matrix

$$X^-(t) = (X(t)W(t)X(t))^- X^T(t)W(t)$$

where W(t) is a nxn diagonal weight matrix.

Therefore from equation (2.15) the estimator can be written as

$$d\hat{B}(t) = X^-(t)dN(t) \quad \Rightarrow \hat{B}(t) = \int_0^t X^-(s)dN(s) \tag{2.16}$$

When X(t) has full rank for all t (asymptotically at least), then $\hat{B}(t)$ is essentially an unbiased estimator of $B(t)$.

We know that the cumulative regression coefficient $B(t)$ is integral of regression function $\beta(t)$. A crude estimate of $\beta(t)$ is given by the slope of our estimate $\hat{B}(t)$ but a better estimate can be

obtained by using a kernel-smoothing technique (Martinussen and Scheike,2006; Klein and Moeschberger, 1997), which we do not pursue here.

For the non-parametric model (2.14) one may wish to test the hypothesis, $H_{01}: \beta_p(t) \equiv 0,$ or equivalently $H_{01}: B_p(t) \equiv 0,$ that a time-varying component is significant or not. The other important hypothesis one wish to test is to decide whether a regression coefficient function is time varying or time-constant. This can be formulated as $H_{02}: \beta_p(t) \equiv \gamma, or$ equivalently $H_{02}: B_p(t) \equiv \gamma t$.

To test if a time-varying component is non-significant (H_{01}) the following test statistic was used.

$$T_{1s} = \sup_{t\in[0,\tau]} \left| \frac{n^{\frac{1}{2}} \hat{B}_p(t)}{\hat{V}^{\frac{1}{2}}{}_{pp}(t)} \right| \tag{2.17}$$

where $\hat{V}_{pp}(t)$ is the p th diagonal element of $\hat{V}(t)$ defined in (2.11)

The second hypothesis H_{02} can be tested by the following simple test statistics

$$T_{2s} = n^{\frac{1}{2}} \sup_{t\in[0,\tau]} \left| \hat{B}_p(t) - \hat{B}_p(\tau)\frac{t}{\tau} \right| \quad . \tag{2.18}$$

The idea is that $\hat{B}_p(\tau)/\tau$ is an estimate of the underlying constant under the null hypothesis. The above test statistics have an asymptotic distribution that can be derived based on the asymptotic distribution of $n^{\frac{1}{2}}(\hat{B}(t) - B(t))$ (Martinussen and Scheike, 2006). The quantiles of this distribution are, however, difficult to obtain, and must be simulated.

2.4.1 Goodness-of-Fit test for Additive Hazards Model

In this section we present graphical and numerical methods for checking the adequacy of the Aalen additive model. The procedures are derived from cumulative sums of martingale-based residuals over follow-up time and/or covariate values.

The underlying martingale residuals are

$$M_i(t) = N_i(t) - \int_0^t Y_i(s) X_i^{T}(s) dB(s) \ .$$

The estimate is given by

$$\hat{M}(t) = N(t) - \int_0^t X(s) d\hat{B}(s) = N(t) - \int_0^t X(s) X^{-}(s) dN(s) = \int_0^t G(s) dN(s),$$

where $G(t) = I - X(t)X^{-}(t)$.

Based on the estimated martingale residuals, a cumulative residual process suggested by Lin et al. (1993) in the context of the Cox model was adopted (see Martinussen and Scheike, 2006, Section 5.7 and Scheike(2004), for a detailed discussion).

One can plot this cumulative residual process against covariate values along with a few realizations under the model to assess how unusual the observed process is under the model. Further, one can compute the supremum of the cumulative residual process and approximate the quantiles of its limit distribution, under the model, by resampling.

3. Results

3.1 Data Exploration

Regarding the proportion of pregnant women at each menstrual cycle, an overview is given in Table 1. It is observed that at the end of the study 222(71.61%) women are pregnant. The proportion of pregnant women decreases across the menstrual cycles with cycle one having the highest proportion and cycle 6 with the lowest proportion.

Table 1: Summary of number of pregnancy (percentage) across menstrual cycles.

Event	**Menstrual Cycles**						**Total**
	1	2	3	4	5	6	
Pregnancy	63(20.32)	58(18.71)	36(11.61)	37(11.94)	15(4.84)	13(4.19)	222(71.61)
Censoring	3(0.97)	1(0.32)	2(0.65)	8(2.58)	0(0)	74(23.87)	88(28.39)

Among a total of 310 females 26.77% are smokers and 28.06% of the male partners are smokers. Mean weekly alcohol intake was 4.06 drinks with a maximum value of 39 and minimum of 0 among women and 9.24 drinks among their male partners with value ranging from 0 to 84 drinks per week. M.ZKON0 (sperm concentration) is one of the covariates of interest with value ranging from 0(mill/ml) to 350mill/ml and a mean of 66.57mill/ml(SD=58.29).

In survival analysis it is highly recommended to look at the Kaplan-Meier curves for the categorical predictors. This will provide insight into the shape of the survival function for each group and give an idea of whether or not the groups are proportional (i.e. the survival functions are approximately parallel).

Figure 1 shows the survival curves for smokers and non-smokers for both genders, and allows a visual comparison of the curves. From the left panel plot we can see that the two curves are not parallel but separate except from the very beginning to approximately 65 days. There is no particular pattern for the right panel curve but the survival curves appear to be close together.

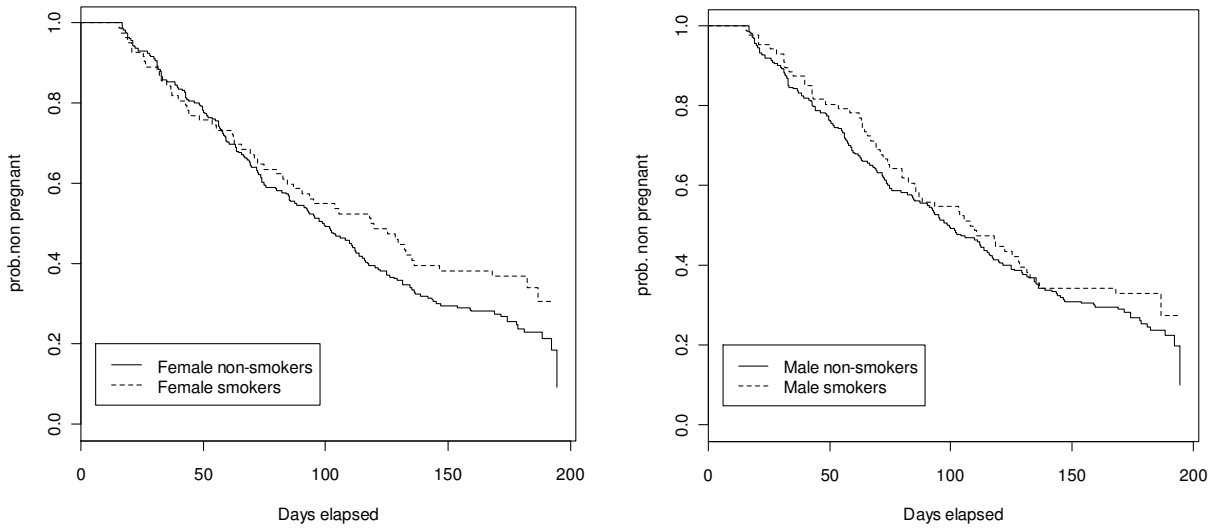


Figure 1: Kaplan-Meier curve for female smoking status (left panel) and male smoking status (right panel)

We also consider the test of equality across strata for categorical covariates using log-rank test. For the covariate K.RYG (smoking status of female) the log-rank test of equality for the group is not significant with a p-value of 0.133. Moreover, the log-rank test of equality across strata for the covariate M.RYG has a p-value of 0.321.

For the continuous variables we used a univariate proportional hazards model (i.e this model is fitted to each covariates one by one) to explore whether or not to include the predictor in the final model. Successive fitting of proportional hazards model to these covariates resulted in p-values 0.141, 0.151, 0.027, 0.468 and 0.0006 for M.COF, K.COF, K.ALK, M.ALK and M.ZKON0 successively in that order. We considered a result of univariate analysis to build our model. Hosmer and Lemeshow (1998) recommend to consider covariates significant at 0.25 level of significance in univariate analysis. We started by including covariates with p-value ≤ 0.25 and all pairwise interactions in the model. Then a back-ward elimination procedure and knowledge of importance of the covariate in affecting TTP were considered to get a final model.

3.2 Proportional Hazards Model

The initial fitting of the Cox proportional hazards model to data in R yielded the following results in Table 2. The likelihood-ratio test for testing the null hypothesis that all $\beta's$ are zero (overall effect of the covariates) is statistically significant (p-value = 0.0005). The effect of individual covariates is tested by Wald-test (z^2). The Wald-test showed that the variable sperm concentration (M.ZKON0) has a significant effect on time to pregnancy. The effect of alcohol consumption (K.ALK) for females is also borderline significant at significance level of 0.05.

Table 2: Parameter estimates and standard error of proportional hazards model

Variable	**Estimate**	**exp(Estimate)**	**se(Estimate)**	**z**	**p-value**
K.ALK	-0.0276	0.973	0.0144	-1.91	0.05600
K.RYG	-0.2274	0.797	0.1582	-1.44	0.15000
M.ZKON0	0.0038	1.004	0.0011	3.50	0.00046
M.COF	0.0003	1.000	0.0002	1.32	0.19000

Likelihood ratio test=20.1 on 4 df, p=0.0005 n= 310

The exponentiated coefficients (also called relative risks) are interpretable as multiplicative effects on the intensity. For example, holding the other covariates constant, an increase in the number of drinks (K.ALK) by one unit decreases the monthly probability of conception by a factor of exp(-0.0276) = 0.973 on average. Similarly, as sperm concentration (M.ZKON0) increases by one unit, and all other variables are held constant, the rate of conception increases by 0.4%**.** The negative coefficient for estimate of K.ALK means that the alcohol consumption is supposed to reduce the intensity of TTP.

3.2.1 Goodness-of-Fit for the Cox Model

3.2.1.1 Test for proportional hazard assumption

As mentioned, a test for the proportional-hazards assumption (PH) was obtained from cox.zph, which computes a test for each covariate along with a global test for the model as a whole. The result is presented in the following table.

Table 3: Test for proportional hazards assumption

	rho	**chisq**	**p-value**
K.ALK	0.0826	1.95	0.163
K.RYG	-0.0992	2.16	0.142
M.ZKON0	0.0784	1.29	0.257
M.COF	-0.0625	1.19	0.275
GLOBAL	NA	6.85	0.144

The GLOBAL test suggests that there is no departure from the standard Cox form with p-value = 0.144. Moreover the individual covariate tests support the proportional hazards assumption is fulfilled for all the covariates in the model.

Figure 2, shows the plot of scaled Schoenfeld residuals versus time for all covariates in the model. The sold line represents a scatter plot smooth to the data points with broken lines as a confidence interval.

A systematic departure from a horizontal line indicates non-proportional hazards. These plots also support the above test that the PH assumption is satisfied for all the covariates in the model.

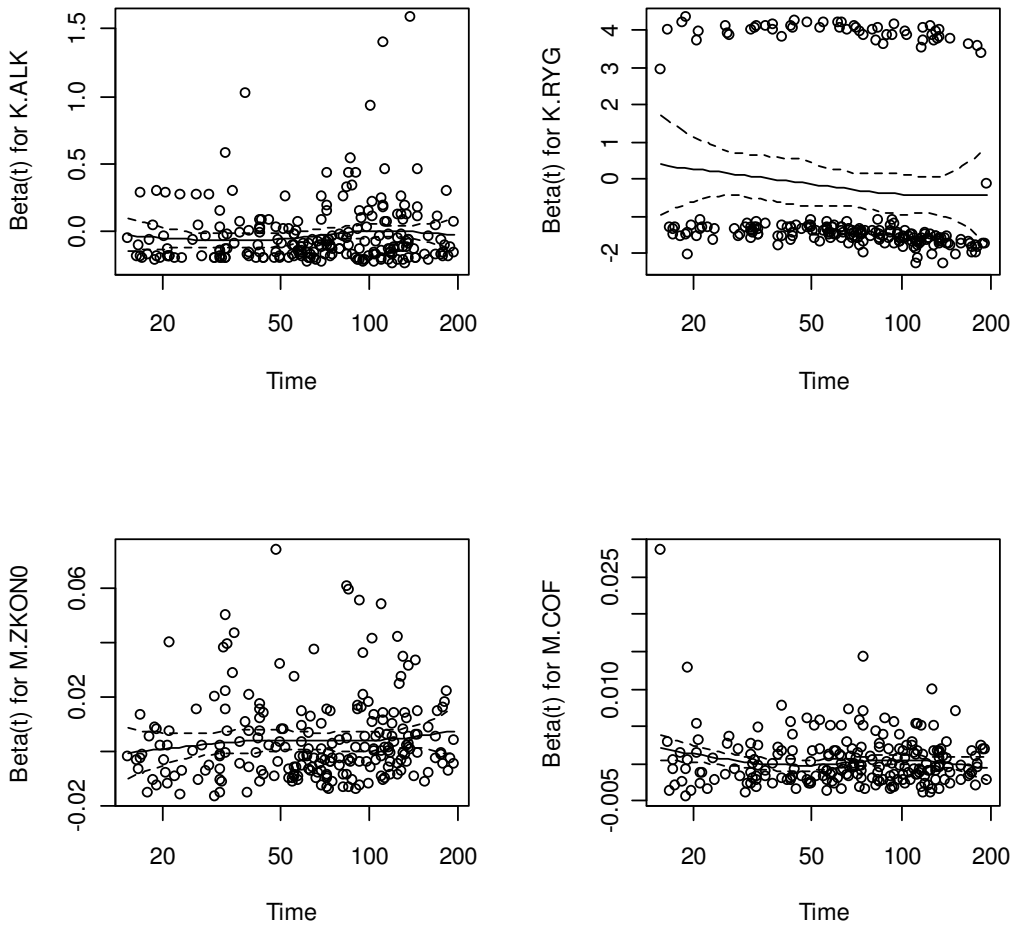


Figure 2: Plots of scaled Schoenfeld residuals against time for each covariate in a model

3.2.1.2 Cox-Snell residuals for assessing the overall fit of the model

We see from the Cox-Snell residual plot (Figure 3) that there are no large departures from a straight line. This shows that the model gives a reasonable fit to the data.

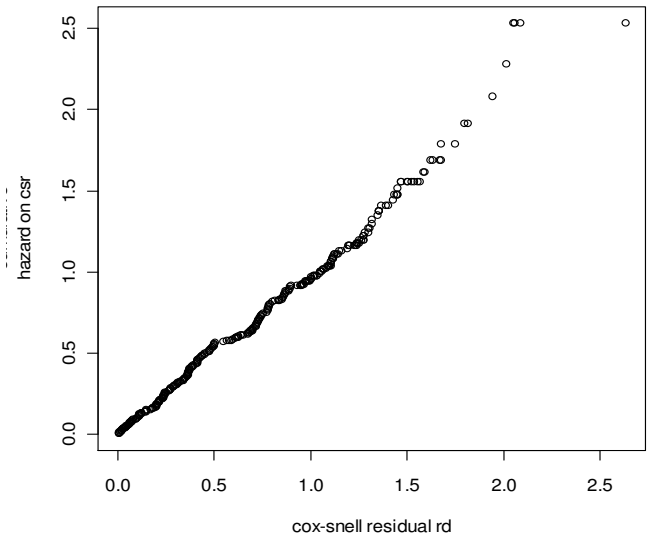


Figure 3: Cox-Snell residuals to assess model fit

3.2.1.3 Checking a functional form of a covariate

Let us examine cumulative residual plots to reveal information about misspecification of the functional form of the continuous covariates in the model. Figure 4 shows the observed cumulative residuals versus continuous covariates with 50 random simulations under the null, which is summarized into p-values by supremum test as 0.208, 0.012 and 0.148 for covariates K.ALK, M.ZKON0 and M.COF respectively. The output of the supremum test for testing if the cumulative residuals behave as they should under the model and figure 4 clearly show that there are problems with the fit of the model, indicated by a p-value of 0.012 for M.ZKON0. This reveals that M.ZKON0 should not be included in the model on its original scale. When M.ZKON0 enters the model on logarithmic scale log(1+ M.ZKON0), the supremum test yields a p-value of 0.03 which indicates there is no improvement of the fit.

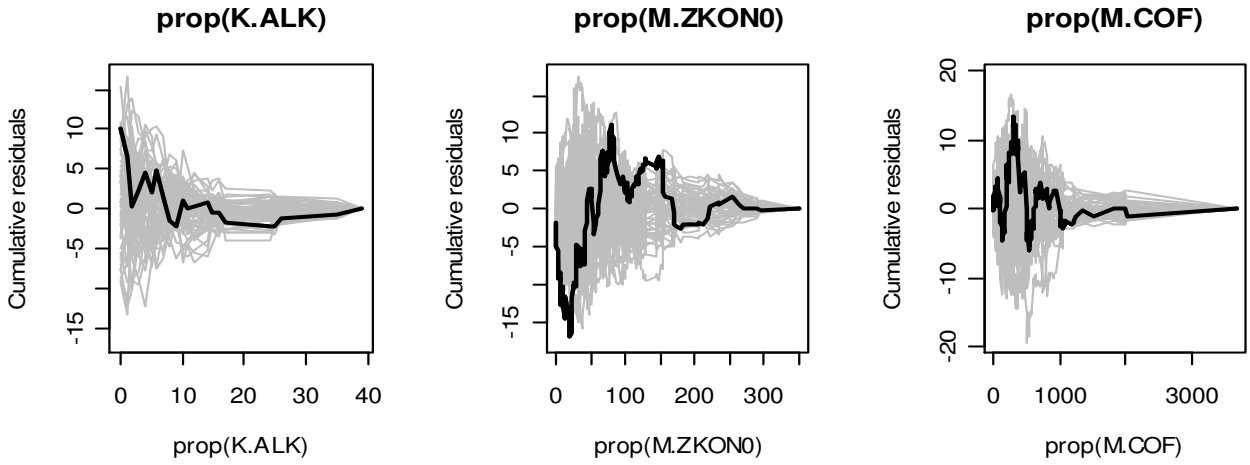


Figure 4: Observed cumulative residuals versus continuous covariates with 50 random realizations under the model

3.3 Extended Cox Model

The tests for non-significant effects and for time invariant effects for the extended Cox model are summarized in Table 4. From the table we see a significant effect of the estimated cumulative regression functions of K.ALK and M.ZKON0. The tests for time invariant effects show that all the covariates effects are not time varying. In figure 5, the estimated cumulative regression function is plotted against time. This gives a description of how the covariate influences the TTP over time.

Table 4: Test for non-significant and time invariant effects of estimated cumulative parameters

	Test for non-significant effects		**Test for time invariant effects**	
Variable	Test statistic	p-value	Test statistic	p-value
K.ALK	4.90	0.000	3.1500	0.412
M.RYG	2.26	0.268	24.3000	0.722
M.ZKON0	4.80	0.000	0.7040	0.260
M.COF	1.50	0.696	0.0492	0.387

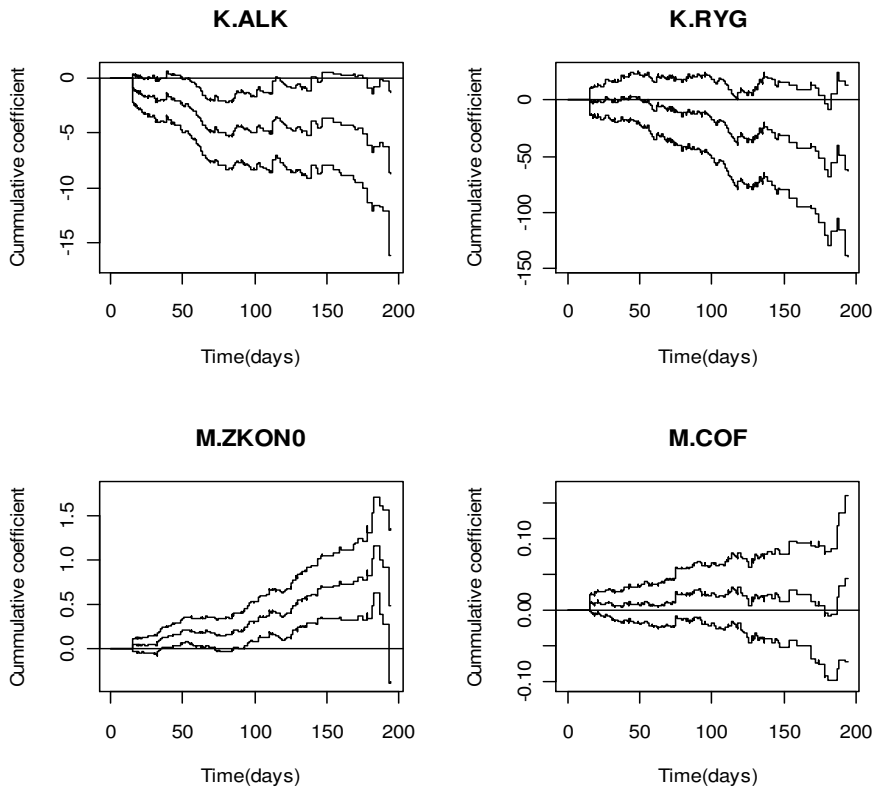


Figure 5: Estimated cumulative regression functions with 95% pointwise confidence intervals based on extended Cox model

Figure 5 depicts the estimated cumulative regression coefficients with 95% pointwise confidence interval. If a regression coefficient is constant over time, the plot of the estimated cumulative regression coefficient versus time should look like a straight line. Figure 5 suggests that the effect of all covariates is constant with time.

3.3.1 Goodness-of- Fit for Extended Proportional Hazards Model

Now, we use cumulative residual plots to check how continuous covariates should be included in the model. Figure 6 shows the cumulative residual plots versus the covariates with 50 random simulations under the null. This reveals that the effect of M.ZKON0 is not well described by the model. The supremum test for the functional form of the covariates gives p-values of 0.216,

0.020 and 0.136 for K.ALK, M.ZKON0 and M.COF, respectively, which indicates that M.ZKON0 should not be included in the model on its original scale.

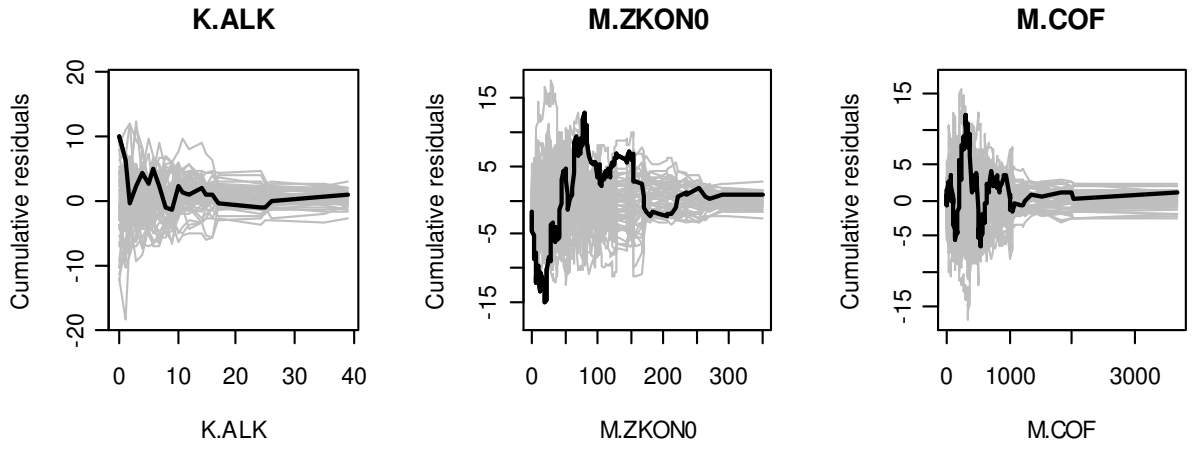


Figure 6: Observed cumulative residuals versus continuous covariates and 50 random realizations under the model.

3.4 Additive Hazards Model

Table 5 shows the summary of the tests for non-significant effects and for time invariant effects for the cumulative regression function using test statistics (2.17) and (2.18) respectively. From the results we see a significant effect of the covariates K.ALK and M.ZKON0. Moreover, the test for time invariant effects shows that all the covariates effects are not time varying.

Table 5: Test for non–significant and time invariant effects of additive hazards model.

	Test for non-significant effects		**Test for time invariant effects**	
Variable	Test statistic	p-value	Test statistic	p-value
K.ALK	3.77	0.002	0.0252	0.321
M.RYG	2.36	0.240	0.1690	0.708
M.ZKON0	3.87	0.001	0.0053	0.192
M.COF	1.47	0.724	0.0004	0.561

In figure 7, the cumulative regression function is plotted against time and gives a description of how the covariate influences the TTP over time. Figure 7 depicts the estimated cumulative regression coefficients with 95% pointwise confidence intervals for each covariate in the model.

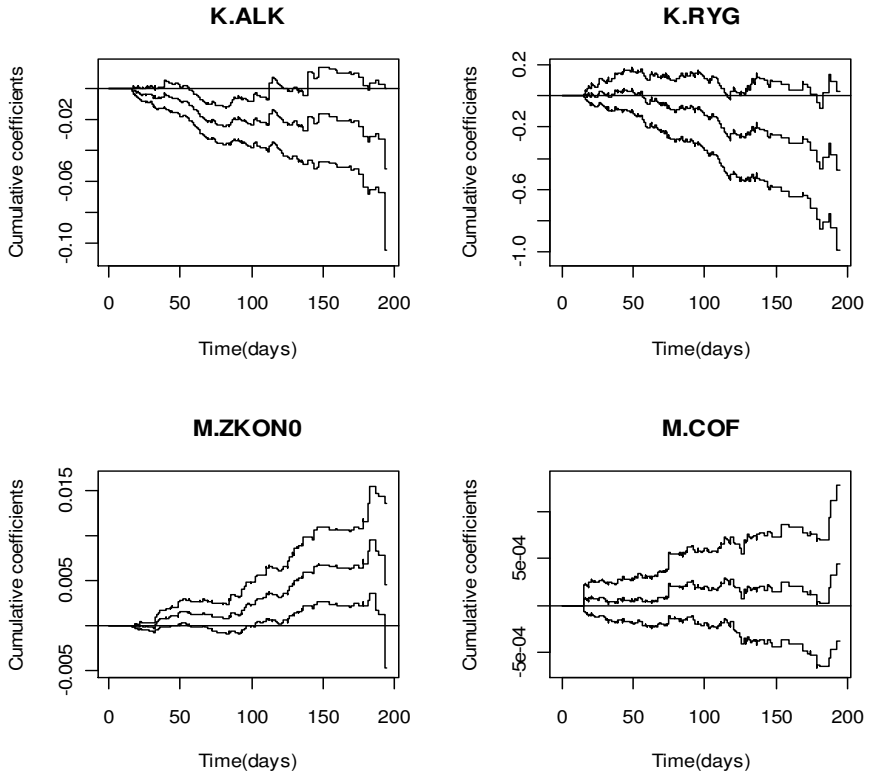


Figure 7: Estimated cumulative regression functions with 95% pointwise confidence intervals based on Aalen's additive model

3.4.1 Goodness-of- Fit for Additive Hazards Model

Now, we use cumulative residual plots to check how continuous covariates should be included in the model. We have three continuous covariates in our model: M.ZKON0 (Sperm concentration), K.ALK (number of alcohol drinks for the female) and M.COF (intake of caffeine for the male). Figure 8 shows the cumulative test processes with 50 random simulations under the null, which is summarized into p-values using a supremum test. The p-values of the supremum test for testing if the cumulative residual processes are consistent with the model are 0.092, 0.042 and 0.112 for K.ALK, M.ZKON0 and M.COF respectively. This shows that there are problems with the fit of the model indicated by a p-value = 0.042 for M.ZKON0. We can also observe that cumulative residual plot of M.ZKON0 in figure 8 shows problem with the fit. This reveals that M.ZKON0 should not be included in the model on its original scale.

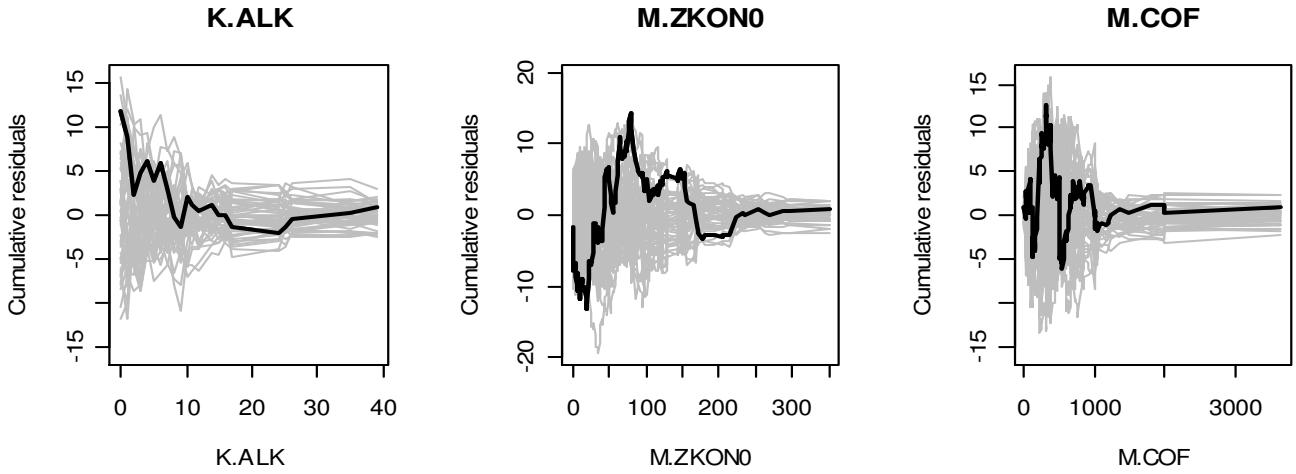


Figure 8: Observed cumulative residuals versus continuous covariates and 50 random realizations under the model.

Figure 9 shows the cumulative residual plots when M.ZKON0 enters the model on a logarithmic scale (log(1+ M.ZKON0)). From Figure 9 we can see that the cumulative residual processes are consistent with the model when compared with Figure 8. Moreover, when log(1+M.ZKON0) enters the model, the p-value of the supremum test jumps from 0.042 to 0.148. This reveals that one should use Log-M.ZKON0 to get an acceptable fit for the additive hazards model.

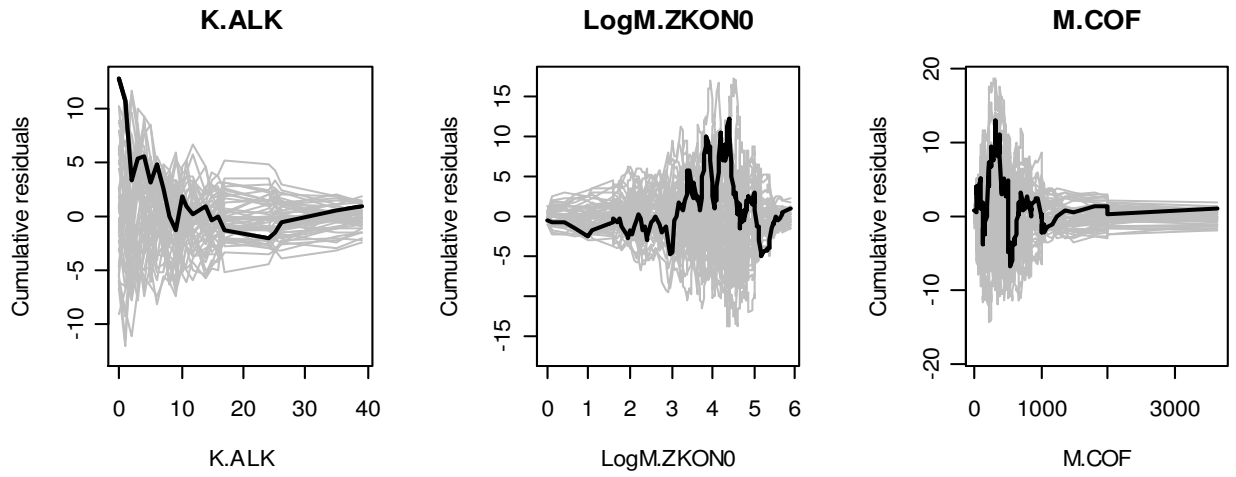


Figure 9: Observed cumulative residuals versus continuous covariates and 50 random realizations under the model.

Table 6 shows the summary of the tests for non-significant effects and for time invariant effects for the cumulative regression function when M.ZKON0 enters the model on logarithmic scale. The test shows a significant effect of the estimated cumulative regression functions of K.ALK and M.ZKON0. The tests for time invariant effects show that all the covariates effects are not time varying. In figure 10, the estimated cumulative regression function ($\hat{B}(t)$) is plotted against time. The slope of these lines is a crude estimate of $\beta(t)$. Positive slopes of the plots occur during periods when increasing covariate values are associated with increases in the intensity function. Negative slopes occur during periods when increasing covariate values are associated with decreases in the intensity function.

Table 6: Test for non–significant and time invariant effects of additive Aalen model when LogM.ZKON0 is used.

	Test for non-significant effects		**Test for time invariant effects**	
Variable	Test statistic	p-value	Test statistic	p-value
K.ALK	3.64	0.005	0.0220	0.348
M.RYG	2.63	0.151	0.2960	0.394
LogM.ZKON0	6.00	0.000	0.1450	0.173
M.COF	1.92	0.417	0.0006	0.328

If a regression coefficient is constant over time, the plot of the estimated cumulative regression coefficient should look like a straight line. It appears from these plots that the effect of all covariates is roughly constant with time. This corresponds with the result in Table 6.

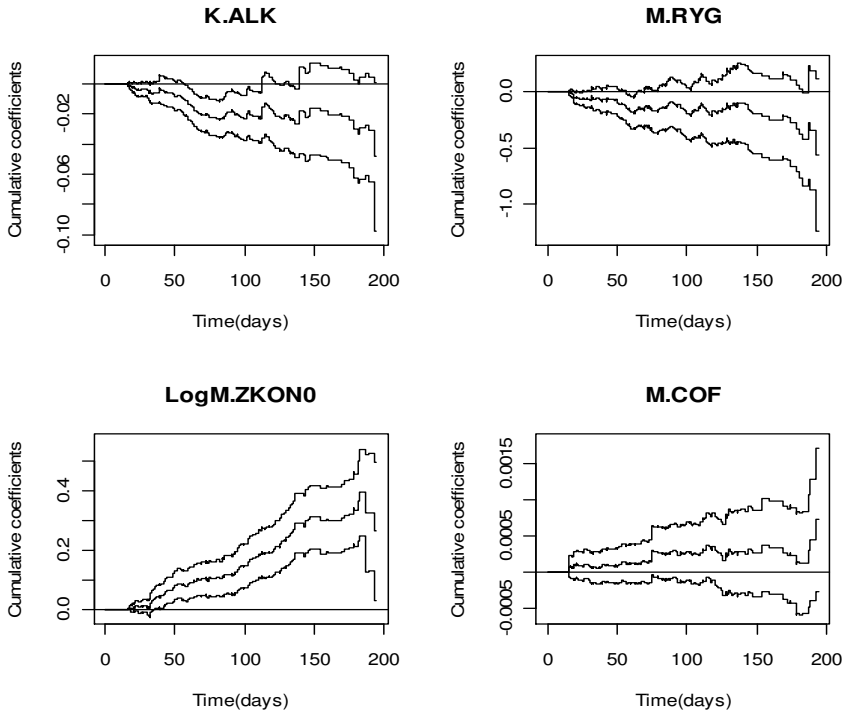


Figure 10: Estimated cumulative regression functions with 95% pointwise confidence intervals based on Aalen's additive model when LogM.ZKON0 is used.

4. Discussion and Conclusion

Studies of time to pregnancy, or waiting time to conception, have proved fruitful in identifying male and female exposures with adverse effects on fertility. We carried out survival data analysis on time to pregnancy from a sample of Danish couples in the age range of 20-35 years who were planning to discontinue contraception to achieve a pregnancy. The aim was to investigate the effect of risk factors (smoking status, sperm concentration, alcohol drinks and intake of caffeine) on the time to pregnancy. For this purpose a Cox regression model, an additive hazards model and a Cox regression model with time varying effects were fitted and goodness of fit was used to compare this models.

The results show that alcohol intake in women was associated with time to pregnancy. The result obtained from all models showed a negative impact of alcohol consumption of women on the monthly probability of conception. These findings are similar to the results of a study by Olsen et al. (1997) and Hakim et al. (1998), However, since the study design and the statistical methods employed are not the same, it is difficult to compare. The study also revealed a significant positive effect of sperm concentration on the hazard of time to pregnancy.

All the models revealed a similar result with regard to the significance of covariates. A test for time invariant effects in the additive model and in the extended Cox model showed the effect is not time-varying for all covariates in the model. This supports the assumption of proportional hazards of Cox proportional hazards model. A test for the functional form of the continuous covariates showed M.ZKON0 should not be included in the model on its original scale for all models. However, when M.ZKON0 enters the model on logarithmic scale, the additive model gave appropriate fit. Therefore, when fitting an additive model, one should use log-M.ZKON0 to get an acceptable fit of the model. But we could not find an appropriate functional form of this covariate for the proportional hazards model. There is no problem with the functional form of other covariates.

Additive and proportional hazard models postulate a different relationship between hazard and covariates and the subject matter rarely indicates which of the models are preferable. Even though these two models all can be used to investigate the effects of risk factors on time to event,

the Cox model has been more popular than the additive risk model. One advantage of the additive model is that time-varying effects are easy to estimate. This model assumes that the covariates act in an additive manner on an unknown baseline hazard rate. The unknown risk coefficients are allowed to be functions of time so that the effect of a covariate may vary over time. Even though there are many advantages in using the additive hazards model, it is not widely used. One reason for this is probably due to the fact that the model only contains nonparametric terms, and that the handling of these terms for inferential purposes is not fully developed.

References

Aalen, O.O. (1989). A linear regression model for the analysis of lifetimes. *Statistics in Medicine* **8**, 907–925.

Aalen, O.O., Borgan, O. and Gjessing, H. K. (2008). *Survival and Event History Analysis*. New York: Springer.

Baird, D.D., Wilcox, A.J. (1985). Cigarette smoking associated with delayed conception. *Journal of the American Medical Association* **253:**2979-2983.

Bonde, J.P., Hjollund, N.H., Jensen, T.K., Ernst, E., Kolstad, H., Henriksen, T.B., Giwercman, A., Skakkebaek, N.E., Andersson, A.M. and Olsen, J. (1998). A Follow-Up Study of Environmental and Biologic Determinants of Fertility among 430 Danish First-Pregnancy Planners: Design and Methods. *Reproductive Toxicology* **12**, 19-27.

Box-Steffensmeier, J. M. and Jones, B.S. (2004). *Event History Modeling.* Cambridge University Press, New York.

Cox, D.R. (1972). Regression Models and Life Tables. *Journal of the Royal Statistical Society, Series* B, **34**:187-220.

Curtis, K.M., Savitz, D.A. and Arbuckle, T.E. (1997).Effects of Cigarette Smoking, Caffeine Consumption, and Alcohol Intake on Fecundability. *American Journal of Epidemiology* **146**:32-41.

El-Shaarawi, A.H. and Piegorsch, W.W. (2002). *Encyclopedia of environmetrics*. Volume 3.

Hakim, R.B., Gray, R.H. and Zacur, H. (1998). Alcohol and caffeine consumption and decreased fertility. *Fertil Steril* **70**, 632-637.

Hatch, E.E. and Bracken, B.M. (1993). Association of delayed conception with caffeine consumption. *American J of Epidemiology* **138**: 1082-1092.

Hosmer, D.W., and Lemeshow, S. (1998). *Applied Survival Analysis, Regression Modeling of Time to event Data, 2nd Edition*. Wiley Series in Probability and Statistics.

Juhl, M., Andersen, A.N., Gronbaek, M**.,** Olsen, J. (2001). Moderate alcohol consumption and waiting time to pregnancy. *Human Reproduction* **16**:2705-2709.

Klein, J.P. and Moeschberger, M.L. (1997). *Survival Analysis: Techniques for Censored and Truncated Data.* Springer-Verlag, New York

Lin, D.Y., Wei, L.J., Z. (1993). Checking the Cox model with cumulative sums of martingale based residuals. *Biometrika* **80**, 557-572

Martinussen, T. and Scheike, T. (2006). *Dynamic regression models for survival data*. New York: Springer.

Martinussen, T., Scheike, T.H., Skovgaard, I. (2002). Efficient estimation of fixed and time-varying covariate effects in multiplicative intensity models. *Scand. J. Statist.* **28**, 57–74.

McKeague, I.W., Utikal, K.J. (1991). Goodness-of-fit tests for additive hazards and proportional hazards models. *Scandinavian Journal of Statistics* **18**, 177–195.

Olsen, J., Bolumar, F., Boldsen, J., Bisanti, L. (1997). Does Moderate Alcohol Intake Reduce Fecundability? A European Multicenter Study on Infertility and Subfecundity. *Alcoholism* **21**, 206–212.

O'Quigley, J. (2008). *Proportional hazards regression.* New York, Springer.

Pierre, J.M., Hans, C. and Stijnen, T. (1998). A goodness-of-Fit Test for Cox's Proportional Hazards Model Based on Martingale Residuals. *Biometrics* **54**, 1517-1526.

Scheike, T.H. (2004). *Time-varying effects in Survival Analysis.* Handbooks of Statistic **23**, 0169-7161.

Schoenfeld, D. (1982). Partial residuals for the proportional hazards regression model. *Biometrika* **69**, 239-241.

Sharma, S. and Chaudhury, R. (1970). Studies on mating II. The effect of ethanol on sperm transport and ovulation in successfully mated rabbits. *Indian Journal of Medical Sciences **58***: 501.

Silva, P.D., Cool, J.L and Olson, K.L. (1999). Impact of lifestyle choices on female infertility. *Journal of Reproductive Medicine* **44**:288–296.

Slama, R., Kold-Jensen, T., Scheike, T., Ducot, B., Spira, A. and Keiding, N. (2004). How would a decline in sperm concentration over time influence the probability of pregnancy? *Epidemiology (Cambridge, Mass)* **4**:458-65.

Suonio, S., Saarikoski, S., Kauhanen, O., Metsapelto, A., Terho, J., and Vohlonen, I. (1990). Smoking does affect fecundity. *European Journal of* Obstetrics *Gynecology and Reproductive Biology* **34**, 89-95.

Therneau, T.M. and Grambsch, P.M. (2000). *Modeling Survival Data: Extending the Cox Model*. Springer-Verlag, New York.

Appendix

R – Codes

```
project<-read.table("C:\\Documents and Settings\\DKN\\My Documents\\Summer
project\\newdata.txt",header=T)
TTP=project$TTP
K.GRAVID=project$K_GRAVID
K.ALK=project$K_ALK
K.COF=project$K_COF
M.COF=project$M_COF
K.RYG=project$K_RYG
M.RYG=project$M_RYG
M.ALK=project$M_ALK
MKRYG=project$MKRYG
M.ZKON0=project$M_ZKON0

# 3.1 Data exploration
# survival functions by KM

library(survival)
library(timereg)
fit11<-survfit(Surv(TTP, K.GRAVID)~K.RYG)
plot (fit11,mark.time=F,lty=c(1,2),ylab="prob.non pregnant",xlab="Days elapsed")
legend(5,0.3, c("Female non-smokers","Female smokers"), lty=c(1,2))

fit12<-survfit(Surv(TTP, K.GRAVID)~M.RYG)
plot (fit12,mark.time=F,lty=c(1,2),ylab="prob. non pregnant",xlab="Days elapsed")
legend(5,0.3, c("Male non-smokers","Male smokers"), lty=c(1,2))

# Log rank test

survdiff(Surv(TTP, K.GRAVID)~M.RYG)
survdiff(Surv(TTP, K.GRAVID)~K.RYG)

# Univariate Cox regression

fit13<-coxph(Surv(TTP, K.GRAVID)~M.COF, project)
fit14<-coxph(Surv(TTP, K.GRAVID)~K.COF, project)
fit15<-coxph(Surv(TTP, K.GRAVID)~K.ALK, project)
fit16<-coxph(Surv(TTP, K.GRAVID)~M.ALK, project)
fit17<-coxph(Surv(TTP, K.GRAVID)~M.ZKON0, project)
```

 # 3.2 Cox proportional hazards Model

```
fit21<-coxph(Surv(TTP, K.GRAVID)~K.ALK+K.RYG+M.ZKON0+M.COF, project)
summary(fit21)
```

 # 3.2.1 Goodness of fit procedures for the cox model
#3.2.1.1 Test for proportional Hazards assumption

```
diag<-coxph(Surv(TTP, K.GRAVID)~K.ALK+K.RYG+M.ZKON0+M.COF, project)
time.test<-cox.zph(diag,transform="log")
print(time.test)
par(mfrow = c(2, 2))
plot(time.test)
```

#3.2.1.2 Cox Snell residual

```
mr=residuals(diag,type="martingale")
csr= K.GRAVID-mr
r.surv=survfit(Surv(csr,K.GRAVID)~1)
plot(r.surv$time,-log(r.surv$surv),xlab="cox-snell residual rd", ylab="cumulative
hazard on csr")
lines(c(0,3),c(0,3))
```

#3.2.1.3 Cumulative martingale residuals.

```
fit23<-cox.aalen(Surv(TTP, K.GRAVID)~prop(K.ALK)
+prop(K.RYG)+prop(M.ZKON0)+prop(M.COF),max.time=195,residuals=1,n.sim=1000,project)
resids<-cum.residuals(fit23,project,cum.resid=1)
summary(resids)
par(mfrow = c(1, 3))
plot(resids,score=2)

logM.ZKON0<-log(1+M.ZKON0)

fit24<-cox.aalen(Surv(TTP,
K.GRAVID)~prop(K.ALK)+prop(K.RYG)+prop(logM.ZKON0)+prop(M.COF),max.time=195,residuals
=1,n.sim=1000,project)
resids<-cum.residuals(fit24,project,cum.resid=1)
summary(resids)
par(mfrow = c(1, 1))
plot(resids,score=2, specific.comp=2)
```

 # 3.3 Extended Cox model with time varying regression effects

```
fit21<-timecox(Surv(TTP, K.GRAVID)~ K.ALK+K.RYG+M.ZKON0+M.COF, max.time=195,project)
par(mfrow = c(3, 3))
plot(fit21,ylab="Cummulative coefficient",xlab="Time(days)")
summary(fit21)
```

 # 3.3.2 Goodness of fit

```
fit10<-timecox(Surv(TTP, K.GRAVID)~K.ALK+M.RYG+M.ZKON0+M.COF,project,max.time=195
,residuals=1,n.sim=1000)
resids<-cum.residuals(fit10,project,cum.resid=1)
par(mfrow = c(1, 3))
plot(resids,score=2)
summary(resids)

fit11<-timecox(Surv(TTP,
K.GRAVID)~K.ALK+M.RYG+logM.ZKON0+M.COF,project,max.time=195,residuals=1,n.sim=1000)
resids<-cum.residuals(fit11,project,cum.resid=1)
par(mfrow = c(1, 3))
plot(resids,score=2)
summary(resids)
```

##3.4 Additive models

```
fit31<-aalen(Surv(TTP, K.GRAVID)~K.ALK+K.RYG+M.ZKON0+M.COF, project,max.time=195)
summary(fit31)
par(mfrow = c(3, 3))
plot(fit31, xlab="Time(days)")
```

 # 3.4.1 Goodness of fit
Cumulative martingale residuals

```
fit12<-aalen(Surv(TTP,
K.GRAVID)~K.ALK+M.RYG+M.ZKON0+M.COF,project,max.time=195,residuals=1,n.sim=1000)
resids<-cum.residuals(fit12,project,cum.resid=1)
par(mfrow = c(1, 3))
plot(resids,score=2)
summary(resids)

fit121<-aalen(Surv(TTP,
K.GRAVID)~K.ALK+M.RYG+logM.ZKON0+M.COF,project,max.time=195,residuals=1,n.sim=1000)
resids2<-cum.residuals(fit121,project,cum.resid=1)
par(mfrow = c(1, 3))
plot(resids2,score=2)
summary(resids2)
```