

RESEARCH COMMUNICATION

Multivariate Analysis of Prognostic Factors in Gastric Cancer Patients Using Additive Hazards Regression Models

Saman Maroufizadeh¹, Ebrahim Hajizadeh^{1*}, Ahmad Reza Baghestani², Seyed Reza Fatemi³

Abstract

Background and Objectives: Gastric cancer is the second leading cause of cancer death worldwide and is the most common type of cancer in Iran. The objective of this research was to apply additive hazards models to the study of survival of patients with gastric cancer and to compare with results obtained using the Cox model. **Methods:** We retrospectively studied 213 patients with gastric cancer who were registered in one referral cancer registry center in Tehran, Iran. Age at diagnosis, sex, presence of metastasis, tumor size, histology type, lymph node metastasis, and pathologic stages were entered into analysis using the Cox model and additive hazard models. To visualize a covariate effect over time, the estimated cumulative regression function by the Aalen's model was examined. **Results:** The five-year survival rate and the median life expectancy in the studied patients were 14.6% and 29.6 months, respectively. Multivariate Cox and Additive hazards models analysis identified age at diagnosis, tumor size and pathologic stage as independent prognostic factors for the survival of patients with gastric cancer. Moreover, pathologic stage had a late or delayed effect according to the Aalen's plot. Other clinicopathological characteristics were not statistically significant. **Conclusion:** Since Cox and Aalen models give different aspects of the association between risk factors and the study outcome, it seems desirable to use them together to give a more comprehensive understanding of data. Our results also suggest that early detection of patients at younger age and in primary stages is important to increase survival of patients with gastric cancer.

Keywords: Gastric cancer - additive hazards models - Cox model - prognostic factor - survival

Asian Pacific J Cancer Prev, 12, 901-907

Introduction

Gastric cancer (GC) is the fourth most common cancer and the second leading cause of cancer-related death in the world. A 2005 analysis of the global incidence and cancer-related mortality revealed that 934,000 new cases of GC were diagnosed and approximately 700,000 patients died from this disease in 2002 (Parkin et al., 2005). Approximately 60% of all cases occur in developing countries with the highest rate in eastern Asia (Jemal et al., 2006). Although its incidence has been decreasing in the West for the last 20–30 years, incidence of GC has remained high in some Eastern countries (e.g., China, Korea, and Japan) (Chen et al., 2008; Yang, 2006). In Iran, the incidence is around 7300 cases per year, which is the most common cancer in men. Mortality from GC is also the first cause of death due to cancer in both sexes in Iran (Movahedi et al., 2009). Given to the low rate of 5-year survival of patients with GC, identification and control of risk factors remain the most effective means of prevention (Gonzalez et al., 2002).

In survival analysis major interests are to compare the failure time distribution function or to assess covariate

effects on survival via appropriate hazards regression models. The Cox proportional hazards model is the most widely used model in survival analysis, offering researchers great flexibility in the analysis of time to event data (Cox, 1972). Most published prognostic studies of GC used the Cox proportional hazards (PH) model (Binquet et al., 2009). In 1990s, Altman et al. reported that, although the vast majority of multivariable analyses of survival in different cancers used the Cox's PH model, only about 5% of these studies tested the underlying PH hypothesis (Altman et al., 1995). Other authors seemingly accepted a priori this crucial assumption, which postulates that the impact of a prognostic factor on survival remains constant over the entire follow-up (Binquet et al., 2009).

Nevertheless, the PH assumption may be often incorrect. Indeed, changes over time in the effects of some patient characteristics were reported in several flexible analyses of different cancers (Quantin et al., 1999; Rachet et al., 1998; Hess, 1994; Gray, 1992; Giorgi et al., 2003). For example, one may question if the relative risk associated with higher gastric cancer stage at the time of initial diagnosis remains equally elevated during the entire follow-up period, as implied by the PH model,

¹Department of Biostatistics, Faculty of Medical Science, Tarbiat Modares University, ²Islamic Azad University, South Tehran Branch, ³Research Center for Gastroenterology and Liver Disease, Shahid Beheshti University of Medical Sciences, Tehran, Iran *For correspondence: hajizadeh@modares.ac.ir

or if it decreases with increasing time since diagnosis. This type of information could greatly help physicians for adapting treatments over time to the actual risk of individual patients (Binquet et al., 2009).

Although the Cox model has an unknown baseline hazard, one may assume that the baseline hazard has a particular parametric form, such as Weibull, log-normal, log-logistic, generalized gamma, etc. However, if the assumed parametric baseline is incorrect, the resulting estimates are biased and inconsistent (Meyer, 1990; Bhat, 1996). An alternative approach is frailty models. Frailty models have been used for univariate data to extend parametric models and to understand the effect of the unobserved heterogeneity (Vaupel et al., 1979). Heterogeneity can be accounted for by incorporating additional unobserved random frailty effects into standard survival models (Oakes, 1992; Hougaard, 1995; Lim et al., 2007). Ignoring frailty when it is present can lead to an underestimation of covariate effects and inaccuracy in fitted survival curves (Henderson et al., 1999).

The Cox model can lead to potentially biased conclusions when the proportionality assumption is not satisfied. Parametric models can also lead to biased conclusions when a parametric baseline hazard distribution is not satisfied. An alternative, but less widely used method is the additive hazards model. Additive hazards model has been considered by several authors (Aalen, 1980; Buckley, 1984; Cox and Oakes, 1984; Huffer and McKeague, 1991; Andersen et al., 1993). Unlike the Cox model, the additive hazards model assumes that covariates act in an additive manner on an unknown baseline hazard rate. Numerous authors advocate the additive hazards model in various forms (Buckley, 1984; Aalen, 1989; Huffer and McKeague, 1991; Lin and Ying, 1994; Hougaard, 2000; Martinussen and Scheike, 2002; Klein, 2006).

The primary aim of this paper is to apply the additive hazards models to the study of patients with GC and to compare results obtained by the additive hazards models and the Cox model.

Materials and Methods

Study Population

This is a retrospective study of patients treated from February 2003 through January 2008, between 213 patients whom were admitted to the Taleghani hospital with a diagnosis of GC. The hospital is a referral center for gastrointestinal cancers, and all of the patients were diagnosed by endoscopy and biopsies. The exclusion criteria were the patients who had not completed document at hospital registry or treated out of the time February 2003 to January 2008 and the start point for survival time was the time of diagnosis which extracted from the patient's document. The study protocol was approved by the ethics committee of the Research Center for Gastroenterology and Liver Disease of Shaheed Beheshti Medical University. In the research center, all patients who register with gastrointestinal cancer are monthly followed for survival. The case of patient's death was confirmed by contact with the patient's family by telephone and clinical information was extracted from hospital documents.

The Clinicopathological features analyzed for GC patients were age at diagnosis, sex, pathologic distant metastasis, tumor size, histology type, regional lymph node metastasis, and pathologic stage.

Statistical Analysis

In this paper, Cox proportional hazards model (Cox model) and additive hazards models were used for multivariate analysis. The additive hazards regression models considered in this paper are the Aalen's additive hazards model (Aalen model) and Lin and Ying's additive hazards model (L-Y model). In this section, all models for analysis are reviewed.

Cox Proportional Hazards Model

Currently, the most popular regression method for survival analysis in biomedical studies is the Cox proportional hazards model (Cox model). In this model, the effect of the covariates was to act multiplicatively on some unknown baseline hazard rate. Thus, under the Cox model, the hazard function for the failure time T_i associated with a p -vector of the covariates $Z_i=(z_{i1}, \dots, z_{ip})$ is defined as:

$$\lambda_i(t) = \lambda_0(t) \exp(\beta_1 z_{i1} + \dots + \beta_p z_{ip}) \quad (1)$$

where $\lambda_0(t)$ is an unspecified baseline hazard function and β_i is the regression coefficient, where $k=1, 2, \dots, p$. Estimation of β_i proceeds through partial likelihood such that $\lambda_0(t)$ is not involved in the estimation of β_i . Cox regression is the predominant model in biomedical studies and the original paper proposing model (1) is one of the most cited papers in science, let alone statistics. However, the assumption of proportional hazards in the Cox model is a crucial one that needs to be fulfilled for the results to be meaningful.

Aalen's Additive Hazards Model

In the Aalen's additive hazards model (Aalen, 1989), the covariates are assumed to impact additively upon a (unknown) baseline hazard, but the effects are not constrained to be constant. Thus, under the Aalen's additive hazards model (Aalen's model), the hazard function for the failure time T_i associated with a p -vector of the covariates $Z_i=(z_{i1}, \dots, z_{ip})$ is defined as: An important one in this framework is Aalen's model:

$$\lambda_i(t) = \lambda_0(t) + \gamma_1(t)z_{i1}(t) + \dots + \gamma_p(t)z_{ip}(t) \quad (2)$$

where $\lambda_0(t)$ is an unspecified baseline hazard function, and coefficient $\gamma_k(t)$ is allowed to vary freely over time, where $k=1, 2, \dots, p$. Aalen shows that if a covariate is independent of all the other covariates in the model, then the regression model with this covariate eliminated is the same as the regression model with this covariate included. Note that this is not true for the Cox model (Aalen, 1989). The additive effect $\gamma_k(t)$ may change in magnitude and even sign with time. As it is not straightforward to estimate $\gamma_0(t)$ non-parametrically, direct estimation of the coefficient $\gamma_k(t)$ is difficult. Aalen and others (Aalen, 1980; Huffer and McKeague, 1991) have developed least square estimation of integrated coefficients.

$$\Gamma_k(t) = \int_0^t \gamma_k(u) du.$$

The usual method of representing the effect $\gamma_k(t)$ is to graph them against time. To define how the effects of covariates changes over the time, cumulative regression function plots estimated by the Aalen's model is examined. The values of $\gamma_k(t)$, the absolute increase in hazard at time t, are not actually observed, but their relative size may be inferred from the slope of the line. These plots are sometimes called Aalen plots, and they are also used to provide an informal assessment of the adequacy of the proportional hazards assumption in the Cox model, although Aalen considered its primary role as an alternative model in its own right (Aalen, 1993). The Aalen's plots are obtained by estimating the instantaneous contributions of covariates to the hazard at each distinct failure time and summing up the resulting estimates. The slope of such plots indicates whether a specific covariate has a time-dependent or constant effect (Mau, 1986). Slope of an estimated cumulative regression function is positive when covariate increase corresponds to hazard increase, and negative when covariate increase corresponds to hazard decrease. A Cumulative sums slope approaches zero when a covariate has no effect on the hazard (Henderson and Milner, 1991). From a practical standpoint, the graphical representation of the cumulative regression functions is attractive, because it provides a direct perception of data and a picture of how effects and the model fit in with change over time. However, some caution is needed when estimate plots are interpreted especially in later periods, when few subjects remain in the risk set (Lim et al., 2009).

Lin and Ying's additive hazards model

Additive hazards model proposed by Lin and Ying is the most closely connected and analogue to the Cox model. In the Lin and Ying's additive hazards model, the covariates are assumed to act additively on a baseline hazard, but the effects are constrained to be constant. Thus, under the L-Y model, the hazard function for the failure time T_i associated with a p-vector of the covariates Z_i is defined as:

$$\lambda_i(t) = \lambda_0(t) + \gamma_1 z_{i1}(t) + \dots + \gamma_p z_{ip}(t) \quad (3)$$

In the additive hazards models (3), the function $\lambda_0(t)$ is an unspecified baseline hazard function. Coefficient γ_k is constant additive effects, where $k=1, 2, \dots, p$. Lin and Ying propose a heuristic estimation method based on a estimating equation due to the Cox's partial likelihood. Their method successfully treats the baseline hazard as nuisance and removed them from estimating the regression coefficients. Using the counting process and martingale approach, they obtained closed-form estimators for the regression parameters of (3), and the cumulative baseline hazard function $\Lambda_0(t)$ (Lin and Ying, 1994).

Statistical analyses were performed using the computer program SAS 9.1 (Howell and Klein, 1997; Schaubel and Wei, 2007). A P-value of less than 0.05 was considered

Results

The male-to-female ratio among the 213 patients was 2.61:1 and the mean age at diagnosis was 58.6±12.8 years (range: 29-85 years). The mean and median of overall survival time were 31.2 and 29.6 month respectively. The overall survival rate for our series was 79.0% after one year, 35.1% after three, and 14.6% after five years. Of the patients, 22 (10.3%) had pathologic distant metastasis, 158 (74.2%) had tumor size greater than 35mm, 133 (62.4%) diagnosed with advanced stage of GC, 162 (76.1%) with

Table 1. Clinicopathological Characteristics of the Gastric Cancer Patients

Characteristics	n (%)
Age at diagnosis (mean±SD)	58.6±12.8
Sex	
Female	59 (27.7)
Male	154 (72.3)
Distant metastasis	
Absent	191 (89.7)
Present	22 (10.3)
Tumor size	
<35mm	55 (25.8)
>35mm	158 (74.2)
Histology type	
Adenocarcinoma NOS	162 (76.1)
SCC, MPA, MA	26 (12.2)
Other histology	25 (11.7)
Regional node metastasis	
N1	59 (27.7)
N2	128 (60.1)
N3	26 (12.2)
Pathologic stage	
Early	80 (37.6)
Advanced	133 (62.4)

SCC, signet cell carcinoma; MPA, mucin-producing adenocarcinoma; MA, mucinous adenocarcinoma. N1, Metastasis in 1-6 regional lymph nodes; N2, Metastasis in 7-15 regional lymph nodes; N3, Metastasis in more than 15 regional lymph nodes (According to SEER Summary Staging Manual 2000)

Table 2. Multivariate Analysis of Prognostic Characters for Patients with GC using the Cox Proportional Hazards Model

	Regression coefficient	Standard error	HR (95% CI)	p-value
Age at diagnosis	0.027	0.010	1.027 (1.006-1.048)	0.010
Sex Male vs. Female	0.227	0.279	1.255 (0.727-2.169)	0.415
Pathologic distant metastasis Present vs. Absent	0.322	0.344	1.380 (0.703-2.710)	0.349
Tumor size >35mm vs. <35mm	0.993	0.322	2.700 (1.436-5.079)	0.002*
Histology type SCC, MPA, MA vs. Adenocarcinoma NOS	-0.127	0.491	0.881 (0.336-2.307)	0.796
Others vs. Adenocarcinoma NOS	0.232	0.354	1.261 (0.630-2.525)	0.512
Regional lymph node metastasis				
N2 vs. N1	-0.471	0.352	0.624 (0.313-1.244)	0.181
N3 vs. N1	-0.626	0.504	0.535 (0.199-1.435)	0.214
Pathologic stage Advanced vs. Early	0.622	0.343	1.862 (0.951-3.644)	0.069

HR, hazard ratio; CI, confidence interval

Table 3. Multivariate Analysis of the Prognostic Characters for GC Patients with using the Lin-Ying's Additive Hazards Model

	Regression coefficient	Standard error	p-value
Age at diagnosis	0.017	0.008	0.039*
Sex Male vs. Female	0.035	0.229	0.877
Pathologic distant metastasis Present vs. Absent	0.160	0.275	0.561
Tumor size >35mm vs. <35mm	0.969	0.250	<0.001
Histology type SCC, MPA, MA vs. Adenocarcinoma NOS	-0.050	0.439	0.909
Other vs. Adenocarcinoma NOS	0.360	0.291	0.220
Regional lymph node metastasis N2 vs. N1	-0.419	0.353	0.239
N3 vs. N1	-0.602	0.447	0.182
Pathologic stage Advanced vs. Early	0.724	0.353	0.043

Table 4. Multivariate Analysis of the Prognostic Characters for GC Patients with using the Aalen's Additive Hazards Model

	chi-square	p-value
Age at diagnosis	5.153	0.023
Sex Male vs. Female	0.749	0.386
Pathologic distant metastasis Present vs. Absent	0.800	0.371
Tumor size >35mm vs. <35mm	12.405	<0.001
Histology type SCC, MPA, MA vs. Adenocarcinoma NOS	0.025	0.874
Other vs. Adenocarcinoma NOS	0.067	0.795
Regional lymph node metastasis N2 vs. N1	1.656	0.198
N3 vs. N1	1.428	0.232
Pathologic stage Advanced vs. Early	4.478	0.034

histology type of adenocarcinoma NOS, and 26 (12.2%) in N3 level of regional lymph nodes metastasis (Table 1).

The results of the Cox and L-Y models are given in Tables 2 and 3, respectively. The Cox and L-Y models approximately give similar results. Two covariates showed significant impact on the gastric cancer patients' data in both hazards models: age at diagnosis and tumor size. We found that pathologic stage was significant under the L-Y model (P=0.043), but was marginally significant under the Cox model (P=0.069). Neither the Cox model nor L-Y model found sex, pathologic distant metastasis, histology type, and regional lymph node metastasis as a prognostic factor.

The results of the Aalen's models are given in Table 4. The multivariate analysis using the Aalen's model identified that age at diagnosis (P=0.023), tumor size (P<0.001) and pathologic stage (P=0.034) were independent prognostic factors for the survival of patients with GC. Other clinicopathological characteristics were not statistically significant (P>0.05).

Figure 1 shows the excess risk due to age at diagnosis and a 95% point-wise confidence interval. In Figure 1,

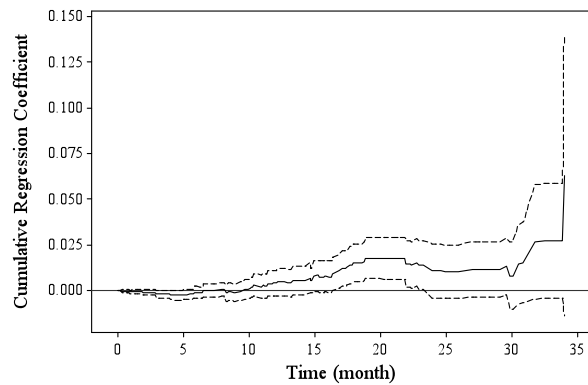


Figure 1. Estimate of Cumulative Effect of Age at Diagnosis and a 95% Point-wise Confidence Interval

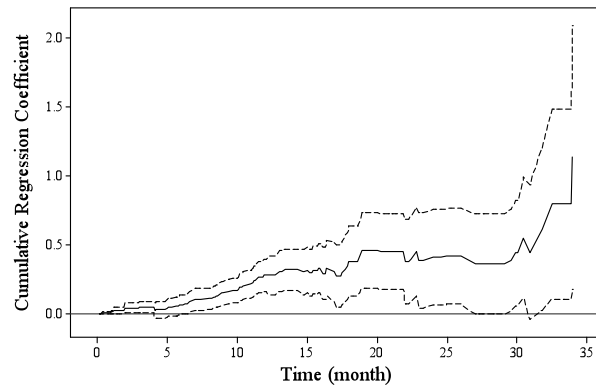


Figure 2. Estimate of Cumulative Excess Risk of Tumor Size >35mm as Compared to <35mm and a 95% Point-wise Confidence Interval

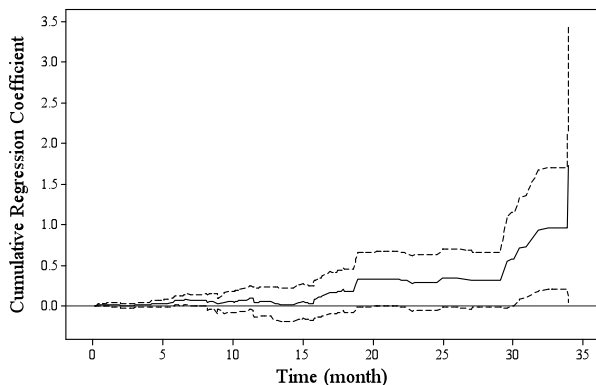


Figure 3. Estimate of Cumulative Excess Risk of Advanced as Compared to Early Stage and a 95% Point-wise Confidence Interval

the estimated cumulative regression coefficient increases nearly linearly over the entire 35-month interval. There is a slight downwards bump in the plot between 22 and 30 months, but the plot continues to increase linearly after 30 months. This bump is likely an artifact of chance variability. Overall, the plot suggests that there is an increase in the hazard rate with increasing age that remains in effect over the entire time period. Figure 2 shows the plot of estimate of cumulative excess risk of tumor size greater than 35mm as compared to tumor size less than 35mm and a 95% point-wise confidence interval. This plot is nearly linear, with a positive slope over the entire 35 month. This plot suggests that the effect of the tumor size does not change over time, and that the size of tumor

increases the hazard over the entire time period. Figure 3 shows the plot of estimate of cumulative excess risk of advanced stage as compared to early stage and a 95% point-wise confidence interval. The plot is nearly linear with a slight positive slope over the first 18 months. However, in this time period the zero line is contained within the lower 95 percent confidence band, which suggests that the covariate may not provide a significant additive increase to the hazard rate during the first 18 month of follow-up. We note that the plot is linear with a much steeper slope after 18 months. This suggests that pathologic stage has a late or delayed effect. Other plots show no consistent trend in any time interval, and the zero line is contained within 95 percent confidence interval (other plots not shown here).

Discussion

In this study we demonstrated the additive hazards regression models and showed the differences in estimates obtained by the additive hazards models and the Cox model using the GC patients' data. The Cox and L-Y models approximately revealed similar results. Although declining in incidence, GC remains an important cancer problem. The five-year survival rate in this study was 14.6% which is lower than that found in many other countries such as United States (37.0%) (Schwarz and Zagala-Nevarez, 2002), France (30.0%) (Triboulet et al., 2001), China (29.6%) (Ding et al., 2004) and Switzerland (22.0%) (Adachi et al., 2003). This may be explained by the fact that Iranian patients generally seek medical attention when disease has reached an advanced stage. Therefore, diagnosis is made when the chance of a full cure is slim.

As we expected life expectancy significantly decreased with age. A study performed in the United States also showed that older age groups have a shortened life expectancy in comparison to young (Saidi et al., 2004). This fact has been verified by studies performed in Japan and Italy as well (Otsuji et al., 1999; Bucchi et al., 2004). Some studies reported better survival rate for females (Ries et al., 1992). In our results, sex had no effect on survival rate. Liu et al indicated that there was no association between gender and survival of patients with early GC (Liu et al., 2010). A study carried out on 2773 patients by Rotterdam cancer registry reported similar results which rate for male and female were similar (Damhuis and Tilanus, 1995). Pathologic distant metastasis is another important prognostic factor of GC (Shiraishi et al., 2000); however in our results no association was observed according to the Cox and additive hazards models.

Tumor size was another significant factor where affected the survival probability of patients in both models. This finding was similar to studies which pointed a higher hazard ratio of death for patients with large tumors (Coburn et al., 2006; Li et al., 2009). It is generally accepted that lymph node metastasis is one of the most important prognostic factors (Kim et al., 1995; Yokota et al., 2004; Liu et al., 2010), however in our study no association was observed according to all models. Histology type did not seem to be significant in our study.

The Cox and L-Y models are difficult to compare directly because the coefficients of the Cox model act in multiplicative way on unknown baseline hazard, while in the L-Y model they act in additive way on unknown baseline hazard or represent coefficient function for added risks. In our study, the Cox and L-Y models similarly identify significant covariates on the GC patients' data; however, interpretations of coefficients from the models are different. The additive hazards models describe the association between the covariates and failure time in terms of the risk difference or excess risk rather than the risk ratio. From a public health point of view, the risk difference can be more important than the risk ratio in understanding an association between a risk factor and disease occurrence (Lin and Ying, 1997).

The additive models have limitations. The Aalen's model may provide more in depth information on the effect of a prognostic factor over time. However, for Aalen's model, one has to visualize all covariate effects over time, and a simple interpretation of the effects is not possible. This makes Aalen's model less appealing in real applications than other models. As Lin and Ying noted, a theoretical limitation of the additive model is the linear predictors in the model to constrain to be positive (Lin and Ying, 1997). These reasons, together with the relative lack of statistical software, are probably the deciding factors in the relatively minimal use of additive hazards models. In general, the choice between the Cox and additive models will normally be an empirical matter. Although in theory either model can provide adequate fit to a given same data set, the more parsimonious one will undoubtedly be more appealing to clinical investigators. An overall conclusion is that the Cox and additive hazards models give different aspects of the association between risk factors and the study outcome. It seems desirable to use them, not as alternatives to each other but as complementary methods, together to give a more comprehensive understanding of data.

Also, our study showed that age at diagnosis, tumor size and pathologic stage were associated factors for survival time in patients with GC. According to these results the early detection of patients at younger age and in primary stages may be important to increase survival of patients with GC.

Acknowledgements

This research extracted from the Master's thesis in biostatistics field with the financial support of faculty of Medical Science of Tarbiat Modares University. Authors would like to thank the Cancer Registry Center of the Research Center for Gastroenterology and Liver Diseases, Shaheed Beheshti Medical University, for data gathering and collaboration in the study.

References

- Aalen OO (1980). A model for nonparametric regression analysis of counting process. In: Klonecki N, Kosek A, Rosinski J, Eds. *Survival Analysis: State of the Art*, 2nd ed. New York: Springer, 1-25.

- Aalen OO (1989). A linear regression model for the analysis of life times. *Stat Med*, 8, 907–925.
- Aalen OO (1993). Further results on the non-parametric linear regression model in survival analysis. *Stat Med*, 12, 1569–1588.
- Adachi Y, Tsuchihashi J, Shiraishi N, Yasuda K, Etoh T, Kitano S (2003). AFP-producing gastric carcinoma: multivariate analysis of prognostic factors in 270 patients. *Oncology*, 65, 95–101.
- Altman D, DeStavola B, Love S, Stepniwska K (1995). Review of survival analyses published in cancer journals. *Br J Cancer*, 72, 511–518.
- Andersen PK, Borgan O, Gill RD, Keiding N (1993). *Statistical Models Based on Counting Processes*. New York: Springer-Verlag.
- Bhat CR (1996). A hazard-based duration model of shopping activity with nonparametric baseline specification and nonparametric control for unobserved heterogeneity. *Transportation Research Part B: Methodological*, 30, 189–207.
- Binquet C, Abrahamowicz M, Astruc K, Faivre J, Bonithon-Kopp C, Quantin C (2009). Flexible statistical models provided new insights into the role of quantitative prognostic factors for mortality in gastric cancer. *J Clin Epidemiol*, 62, 232–240.
- Bradburn MJ, Clark TG, Love SB, Altman DG (2003). *Survival Analysis Part II: Multivariate data analysis – an introduction to concepts and methods*. *Br J Cancer*, 89, 431–436.
- Bucchi L, Nanni O, Ravaioli A, Falcini F, Ricci R, Buiatti E, et al (2004). Cancer mortality in a cohort of male agricultural workers from northern Italy. *J Occup Environ Med*, 46, 249–256.
- Buckley JD (1984). Additive and multiplicative models for relative survival rates. *Biometrics*, 40, 51–62.
- Chen XZ, Jiang K, Hu JK, et al (2008). Cost-effectiveness analysis of chemotherapy for advanced gastric cancer in China. *World J Gastroenterol*, 14, 2715–2722.
- Coburn NG, Swallow CJ, Kiss A, Law C (2006). Significant regional variation in adequacy of lymph node assessment and survival in gastric cancer. *Cancer*, 107, 2143–2215.
- Cox DR (1972). Regression models and life tables (with discussion). *J R Statist Soc*, B34, 187–220.
- Damhuis RA, Tilanus HW (1995). The influence of age on resection rates and postoperative mortality in 2773 patients with gastric cancer. *Eur J Cancer*, 31A, 928–931.
- Ding YB, Chen GY, Xia JG, et al (2004). Correlation of tumor positive ratio and number of perigastric lymph nodes with prognosis of patients with surgically removed gastric carcinoma. *World J Gastroenterol*, 10, 182–185.
- Giorgi R, Abrahamowicz M, Quantin C, et al (2003). A relative survival regression model using B-spline functions to model non-proportional hazards. *Stat Med*, 22, 2767–2784.
- Gonzalez CA, Sala N, Capella G (2002). Genetic susceptibility and gastric cancer risk. *Int J Cancer*, 100, 249–260.
- Gray R (1992). Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. *J Am Stat Assoc*, 87, 942–951.
- Henderson R, Milner A (1991). Aalen plots under proportional hazards. *Appl Statist*, 40, 401–409.
- Henderson R, Oman P (1999). Effect of frailty on marginal regression estimates in survival analysis. *J R Statist Soc*, B61, 367–379.
- Hess K (1994). Assessing time-by-covariate interactions in proportional hazards regression models using cubic spline function. *Stat Med*, 13, 1045–1062.
- Hougaard P (1995). Frailty models for survival data. *Lifetime Data Analysis*, 1, 255–273.
- Hougaard P (2000). *Analysis of Multivariate Survival Data*. New York: Springer.
- Howell A, Klein JP (1997). A SAS macro for the additive regression hazards model. *Proc ASA 1997*, 282–7.
- Jemal A, Siegel R, Ward E, et al (2006). *Cancer statistics, 2006*. *CA Cancer J Clin*, 56, 106–130.
- Kim JP, Hur YS, Yang HK (1995). Lymph node metastasis as a significant prognostic factor in early gastric cancer: analysis of 1,136 early gastric cancers. *Ann Surg Oncol*, 2, 308–313.
- Klein JP (2006). Modelling competing risks in cancer studies. *Stat Med*, 25, 1015–1034.
- Klein JP, Moeschberger ML (2003). *Survival Analysis: Techniques for Censored and Truncated Data (Second Edition)*. New York: Springer-Verlag.
- Li C, Oh SJ, Kim S, et al (2009). Risk factors of survival and surgical treatment for advanced gastric cancer with large tumor size. *J Gastrointest Surg*, 13, 881–885.
- Lim HJ, Liu L, Meltzer-Lange M (2007). Comparison of Methods for analyzing recurrent event data: application to Emergency department visits of pediatric firearm victims. *Accid Anal Prev*, 39, 290–299.
- Lim HJ, Zhang X (2009). Semi-parametric additive risk models: Application to injury duration study. *Accid Anal Prev*, 41, 211–216.
- Lin DY, Ying Z (1994). Semiparametric analysis of the additive risk model. *Biometrika*, 81, 61–71.
- Lin DY, Ying Z (1997). *Additive Regression Models for Survival Data*. In *Proceedings of the First Seattle Symposium in Biostatistics: Survival Analysis*. DY Lin and TR Fleming, eds. New York: Springer, 185–198.
- Liu C, Zhang R, Lu Y, et al (2010). Prognostic role of lymphatic vessel invasion in early gastric cancer: A retrospective study of 188 cases. *Surg Oncol*, 19, 4–10.
- Martinussen T, Scheike TH (2002). A flexible additive multiplicative hazard model. *Biometrika*, 89, 283–298.
- Mau J (1986). On a graphical method for the detection of time-dependent effects of covariate in survival data. *Appl Statist*, 35, 245–255.
- Meyer BD (1990). Unemployment insurance and unemployment spells. *Econometrica*, 58, 775–782.
- Movahedi M, Afsharfard A, Moradi A, et al (2009). Survival rate of gastric cancer in Iran. *JRMS*, 14, 367–373.
- Oakes D (1992). Frailty models for multiple event times. In: Klein JP, Goel PK, Eds. *Survival Analysis: State of the Art*. Dordrecht: Kluwer Academic Publisher.
- Otsuji E, Yamaguchi T, Sawai K, et al (1999). Regional lymph node metastasis as a predictor of peritoneal carcinomatosis in patients with borrmann type IV gastric carcinoma. *Am J Gastroenterol*, 94, 434–437.
- Parkin DM, Bray F, Ferlay J, Pisani P (2005). *Global cancer statistics, 2002*. *CA Cancer J Clin*, 55, 74–108.
- Quantin C, Abrahamowicz M, Moreau T, Bartlett G, MacKenzie T, Tazi M, et al (1999). Variation over time of the effects of prognostic factors in a population based study of colon cancer: comparison of statistical models. *Am J Epidemiol*, 150, 188–200.
- Rachet B, Sasco A, Abrahamowicz M, Benyamine D (1998). Prognostic factors in nasopharyngeal cancer: accounting for time-dependence of relative risks. *Jpn J Med*, 28, 772–780.
- Ries LAG, Eisner MP, Kosary CL, Hankey BF, Miller BA, Clegg L, et al (1992). *Cancer Statistics Review: 1973-1989*. NIH Pub. No. 92-2789. Bethesda, National Cancer Institute, 23, 1–9.
- Saidi RF, Bell JL, Dudrick PS (2004). Surgical resection for gastric cancer in elderly patients: is there a difference in outcome? *J Surg Res*, 118, 15–20.
- Schaubel DE, Wei G (2007). *Fitting Semiparametric Additive*

- Hazards Models using Standard Statistical Software. *Biometrical Journal*, 49, 719–730.
- Schwarz RE, Zagala-Nevarez K (2002). Recurrence patterns after radical gastrectomy for gastric cancer: Prognostic factors and implications for postoperative adjuvant therapy. *Ann Surg Oncol*, 9, 394–400.
- Shiraishi N, Inomata M, Osawa N, et al (2000). Early and late recurrence after gastrectomy for gastric carcinoma: Univariate and multivariate analyses. *Cancer*, 89, 255–261.
- Triboulet JP, Fabre S, Castel B, Tournel H (2001). Adenocarcinomas of the distal esophagus and cardia: Surgical management. *Cancer Radither*, 5, 90–97 (In French).
- Vaupel JW, Manton KG, Stallard E (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16, 439–454.
- Yang L (2006). Incidence and mortality of gastric cancer in China. *World J Gastroenterol*, 12, 17–20.
- Yokota T, Ishiyama S, Saito T, et al (2004). Lymph node metastasis as a significant prognostic factor in gastric cancer: a multiple logistic regression analysis. *Scand J Gastroenterol*, 39, 380–384.