JRHS
Journal of Research in Health Sciences
journal homepage: www.umsha.ac.ir/jrhs

Original Article

Determination of Affected Risk Factors on Time to Recurrence and Death in Patients with Postoperative Gastric Cancer using Copula Function

Ghodratollah Roshanaei (PhD)[a], Anoshirvan Kazemnejad (PhD)[b*], Sanambar Sadighi (PhD)[c]

[a] *Research Center for Health Sciences and Department of Biostatistics and Epidemiology, Hamadan University of Medical Sciences, Hamadan, Iran*
[b] *Department of Biostatistics, Faculty of Medical Sciences, Tarbiat Modares University, Tehran, Iran*
[c] *Department of Medical Oncology, Cancer Research Center, Cancer Institute. Imam Khomeini Hospital, Tehran, Iran*

ARTICLE INFORMATION

Article history:
Received: *01 May 2013*
Revised: *11 July 2013*
Accepted: *30 July 2013*
Available online: *09 August 2013*

Keywords:
Parametric survival model
Recurrence time
Gastric cancer
Semi-competing risk
Dependent event
Copula function

**** Correspondence***
Anoshirvan Kazemnejad (PhD)
Tel:*+982182883875*
Fax: *+982182884555*
E-mail: *kazem_an@modares.ac.ir*

ABSTRACT

Background: In survival studies when the event times are dependent, performing of the analysis by using of methods based on independent assumption, leads to biased. In this paper, using copula function and considering the dependence structure between the event times, a parametric joint distribution has made fitting to the events, and the effective factors on each of these events would be determined.

Methods: This retrospective cohort study was conducted from March 2003 to March 2007. The data collected from 256 patients with gastric cancer who underwent surgery and that the event time of the two outcomes of death and recurrence for them was recorded. Akaike Information Criterion (AIC) was used to determine of suitable parametric models. Moreover, applying copula function with regard to the relationships between the events, the effect of the risk factors of each of the two outcomes was determined. The data analysis was done using R2.12.1 software.

Results: According to the AIC criterion, the Weibull distribution had the best fitting in both of the event times. The median times for recurrence and survival of the patients were estimated 20.2 and 28.1 months respectively. Furthermore, with a fitting of Weibull distribution to the two event times using Clayton copula function, the variables of gender, tumor size and tumor pathological stage on survival, and tumor size and tumor pathological stage on recurrence were significant (P<0.001).

Conclusions: Applying copula function for determining specific risk factors of the semi-competing events produces suitable results opposite the common methods which are based on independent assumption of the events.

Citation: Roshanaei G, Kazemnejad A, Sadighi S. Determination of Affected Risk Factors on Time to Recurrence and Death in Patients with Postoperative Gastric Cancer using Copula Function. J Res Health Sci. 2014;14(1):52-56.

Introduction

Cohort and clinical trial studies that include patients' follow-up are usually a part of common and valid studies in medical researches. In such studies, the researchers may focus on several outcomes which with the occurrence of one of the events the study ends, which is the events, are mutually independent. These kinds of events are called Competing Risks[1]. In the study of such events, the time and type of the first event is recorded and common models such as Proportional Hazards (PH) are employed. The analyses in these methods are based on the assumption of non-informative Censoring[2-5]. In contrast, if the events occur so that some of the events become terminal (such as death or dropout) and some events become non-terminal (such as relapse or disease progression), these kinds of events are called semi-competing Risks [6-7]. In semi-competing paradigm, the terminal events cause non-terminal events to be dependently censored. In such studies using of usual models will have biased[2-4,8-9]. To model the joint distribution of these events, the random effect models by Wang (2001), Liu (2004), and Huang et al. (2004) have been suggested[10-12]. Liu (2004) has also presented a common gamma frailty model for the relapse event and PH model for the terminal events[11].

A semi-parametric model is suggested for marginal distribution in which the dependent structure between relapse and death was regarded by gamma frailty copula1. Moreover this model is extended to other parametric copulas[13].

Jiang et al. (2005) presented a self-consistency estimator for non-terminal event which modeled the dependency between the events with Clayton copula. They suggested a method for estimation of correlation parameter and a non-parametric method for the marginal distributions[14-15].

The semi-parametric inference for the association parameter in the semi-competing risk data is introduced16. Pang et al. (2007) used the marginal association of the non-terminal event by the accelerated failure time. They also employed the time dependent copula for modeling the dependency structure[17]. The estimation of the non-terminal event and the association parameter is obtained using Archimedean copula 18. A regression analysis on the semi-competing risks data is

made. In this research the estimation of the marginal distribution of the non-terminal event in the presence of death dependently censoring will be discussed using copula function approach, and its specific risk factors will be determined[19].

Roshanaei et al. (2012) designed a parametric model to assess the risk factors for recurrence of patients with gastric cancer in the presence of informative censoring using a semi-competing risk approach. They showed that semi-competing risk methods perform well in determining risk factors for recurrence[20]. Some studies applied semi-competing risk and copula paradigm in cancer data [14,16,18].

Because the analysis of data with usual method with considering independent assumption between events would be led to bias, then in this paper we applied copula function by considering the dependence structure between the event and recurrence times and determined the effective factors on each of these event times.

Methods

In a retrospective cohort study, 256 underwent-surgery patients with gastric cancer in the Cancer Institute of Iran, from March 2003 to March 2007, were selected for this study. The mean (SD) of the follow-up time has been 20 (17.5) months. The two focused outcomes were relapse (as the non-terminal event) and death (as the terminal event) whose times of occurrence were determined after diagnosis in month.

All demographic and extra information was obtained from the patients' files. The disease stages were determined by CT scan and Endosonography before surgery and by pathological reports after the surgery. As these two events are dependent on one another, using the copula function method, a joint distribution has fitted to these two events.

In addition to determining an appropriate parametric distribution from the common parametric distribution family (Exponential, Weibull, Log-normal, log-logistic), the effect of each risk factors of each event was assessed. At first, the effect of variables of gender, age (at the timeof the disease diagnosis), degree of the tumor differentiation, place of tumor, stage of disease, margin, type of treatment, metastasis and tumor size as a univariate was investigated by parametric survival models. The significant variables in this stage were put in the multivariate model. In multivariate investigation, considering the events as semi-competing risks and using copula function (regarding Clayton copula function with Weibull marginal distributions for both of the two events of relapse and death), the simultaneous impact of variables on the time of relapse and death was studied and utilizing this model, the specific risk factors of both events was determined. The data analysis was done by R2.12.1 software including copula packages.

Results

Among the 256 patients with postoperative chemotherapy GC, 192 patients (75%) were men. Of 256 patients, 88 patients died during follow-up and 73 patients had recurrence after treatment. The mean (SD) age at diagnosis was 57.4(11.9)) years and the median was 59.4. Table 1 shows the demographic characteristics of these patients. According to AIC criterion and using copula function method with regard to the correlation between the events, Weibull distribution was determined as the best parametric distribution. According to the chosen model, the correlation coefficient between these two events was 0.45. Moreover, according to this model, the median of relapse time and survival time were 20.2 and 28.1 months respectively. The results of the fitted models are drawn up in Table 2.

Table 1: Demographic characteristics of prognostic factors on gastric cancer patients

Variables	Number	Percent
Gender		
Male	192	75.0
Female	64	25.0
Age at diagnosis (yr)		
≤60	131	51.2
>60	125	48.8
Grade of tumor		
Well differentiated	36	14.1
Moderately differentiated	100	39.0
Poorly differentiated	120	46.9
Tumor site		
Cardia	71	30.3
Body	65	27.8
Other	98	41.9
Radiotherapy		
Negative	201	78.5
Positive	55	21.5
Pathologic stage		
II	58	23.8
III	127	52.0
IV	59	24.2
Tumor size (mm)		
<25	129	64.8
25-45	31	15.6
>45	39	19.6
Margin		
Negative	154	71.6
Positive	61	28.4
Recurrence		
Negative	183	71.5
Positive	73	28.5
Death		
Negative	168	65.6
Positive	88	34.4

Table 2: Akaike Information Criteria (AIC) values for assessment of different marginal distributions of time to relapse and death of gastric cancer patients

Marginal's for time to relapse and death	Exponential	Weibull	Log-Logistic
Spearman correlation (95% CI)	0.6 (0.3, 0.7)	0.5 (0.3, 0.6)	0.5 (0.3, 0.7)
Median time to relapse (95% CI)	28.0 (22.5, 34.9)	20.2 (16.3, 26.1)	21.3 (14.9, 28.7)
median survival time (95% CI)	34.2 (28.1, 41.8)	28.1 (19.7, 34.5)	29.0 (20.2, 41.1)
AIC	796.2	771.8	775.3

The results of univariate analysis of risk factors effect on relapse and survival times evinced that the effect of gender, age at diagnosis, pathologic stage, and the size of tumor on the relapse and survival times was significant (Table 3).

The median of relapse and death times, with independent assumption and employing Kaplan-Meier method, were 41 and 32 months respectively. Then using Weibull parametric model (as the best chosen model in Table 2) the effect of contributing factors on each of the relapse and death events was measured separately and the results are shown in Table 4.

Finally, for a simultaneous study of these factors on the relapse and death times with regard to dependency structure

between these two events, the copula function with Weibull distribution for the marginal of both times of relapse and death, the effect of risk factors on events was investigated simultaneously. The contributing factors affecting each of these events are shown in Table 5.

Table 3: Evaluating results of the risk factors on relapse and death of gastric cancer patients using log-rank

Variables	Recurrence (non-terminal event) Median (month)	*P* value	Death (terminal event) Median (month)	*P* value
Gender		0.041		0.018
Male	31		27	
Female	42		38	
Age at diagnosis (yr)		0.028		0.038
≤60	42		33	
>60	27		25	
Tumor grade		0.381		0.641
Well differentiated	23		35	
Moderately differentiated	34		32	
Poorly differentiated	37		27	
Tumor site		0.178		0.704
Cardia	30		24	
Body	34		29	
Other	37		31	
Radiotherapy		0.721		0.207
Negative	32		27	
Positive	38		32	
Pathologic stage		0.025		0.002
II	-		35	
III	31		27	
IV	42		25	
Tumor size (mm)		0.033		0.040
<25	58		33	
25-45	32		27	
>45	21		25	
Margin		0.621		0.470
Negative	35		32	
Positive	42		29	

Table 4: The results of the effect of contributing factors on relapse and death time of the gastric cancer patients with Weibull model (with the independent assumption of events)

Risk factor	Relative Risk (95% CI) Recurrence (Non-terminal event)	Death (Terminal event)
Gender		
Male	1.00	1.00
Female	0.61 (0.30, 1.25)	0.46 (0.23, 0.92)
Age at diagnosis (yr)		
<60	1.00	1.00
>60	1.24 (0.69, 2.12)	1.41 (0.83, 2.30)
Tumor size (mm)		
<25	1.00	1.00
25-45	1.89 (0.97, 3.68)	1.25 (0.60, 2.82)
>45	1.13 (0.50, 2.32)	1.32 (0.74, 2.44)
Pathologic stage		
II	1.00	1.00
III	2.06 (0.95, 4.02)	3.33 (1.53, 7.31)
IV	1.35 (0.36, 5.08)	5.00 (1.93, 12.92)
Median (month)	20.21 (16.30, 26.12)	28.10 (19.72, 34.53)

Discussion

Although the full surgery of the main tumor is the only way of effective treatment of gastric cancer, the recurrence of the disease locally or with distance metastasis is plausible, especially, when the disease passes through sub-mucous membrane, the possibility of expansion to lymph glands and relapse increases. Some factors may affect the relapse event of the disease[21-22]. All the patients in this study used chemotherapy after the surgery.

Table 5: Evaluating results of risk factors on relapse and death in gastric cancer patients using copula approach

Prognostic factors	Relative Risk (95% CI) Relapse (non-terminal event)	Death (terminal event)
Gender		
Female	1.00	1.00
Male	1.91 (0.91, 3.85)	2.44 (1.23, 4.79)
Age at diagnosis (yr)		
≤60	1.00	1.00
>60	1.02 (0.61, 1.70)	1.18 (0.74, 1.92)
Tumor size (mm)		
<25	1.00	1.00
25-45	1.73 (1.08, 2.83)	1.23 (1.11, 1.44)
>45	2.07 (1.13, 4.12)	1.41(1.07, 1.75)
Pathologic stage		
II	1.00	1.00
III	2.23 (1.18, 4.09)	4.42 (2.14, 8.92)
IV	2.58 (1.12, 5.94)	4.68 (1.81, 12.13)
Spearman correlation	0.45 (0.33, 0.57)	

In this study the marginal distribution of the events were estimated using the independent assumption methods between the events. The medians of relapse and survival times were 41 and 32 months respectively. Among 73 patients in this study who had relapse, 35 patients (48%) died at the end of the study while from 183 patients without relapse, 53 (29%) died. The results demonstrate that death risk in the patients with relapse has been 2.3 times more than non-relapse patients and there is a significant relation between these two events ($P<0.001$). As the analysis of the correlated events with the common methods (based on independent assumption) causes bias in results[6,14-15], these numbers seem far-fetched. In addition to marginal distribution estimation of these events, the specific risk factors of the events by the common method, due to the relation between the events (Table 4), are not valid either, because the spearman correlation coefficient of the relapse and death times in this study was 0.45 which was statistically ($P<0.001$) significant (Table 5). Determination of effective risk factors on relapse and death at the presence of each other, with considering the relationship between relapse and death events, is the most importance for researchers and it can help physicians to find remedies and solutions to prevent relapse and to predict death.

The simultaneous study of the events was done using Clayton copula function and Weibull distribution for both of the events. The results indicated that the size of tumor had significant effect on both relapse and death so that the patients, who had tumor between 25-45 mm, and more than 45 mm, had the risk of relapse respectively 1.73 and 2.1 times more than the patients who had tumors less than 25 mm. Furthermore, the risk of death among them was 1.23 and 1.4 respectively. In the study by Bando et al. (1999), conducted on the tumor size at two levels, this variable has had a significant effect on the two events and the relative risk for relapse and death has been 4.38 and 1.49 respectively[23]. Some studies indicated that the tumor size has been significant on both events[24-26].

In this study, the pathological stage of tumor was effective on relapse and death and the relative risks of relapse in the patients who were in the 3rd and 4th stages of the disease vis-à-vis stage two were respectively 2.23 and 2.6 times and the relative risks of death were 4.4 and 4.7 times. In the re-

search carried out by Lurje (2010) the stage of tumor had not been effective on survival but had been effective on relapse of the disease[27].

In the present study, the mean and the median of the diagnosis age were 57.4 and 59.4 years respectively, and the results evinced that the effect of this variable was not significant on none of the two events of relapse and death. The results of this study are consistent with Park's et al. (2003) study in which the median age of the patients under the study was 54 years [28]. The results of this study, though, are inconsistent with the Marrillo's et al. (2008) study in which the median age of the patients under the study was 65 years[25,28]. Thus, it could be concluded that diagnosis age has no effect on events and it is because of low age at diagnosis among the patients in the present study. Moreover, the effect of age on relapse and survival had not been significant in the study conducted by Lurje (2010)[27].

Although men had the risk of relapse 1.9 times more than women in this study, but the effect of gender on relapse and survival of the patients were not significant. The relative risk of death for men has been 2.44 times in comparison with women. The results of the study done by Moriguchi et al. (1991) and Lurje et al. (2010) disclose that the gender of the patients has not been significant for none of the events of relapse and death of the patients with gastric cancer[26-27].

Conclusions

Applying copula function for determining specific risk factors of the jointly dependent events produces appropriate results but the common methods which are based on independent assumption of the events is led to bias in estimation of factor effect.

Acknowledgments

The authors thank Cancer Research Centre (CRC) of Cancer Institute of Imam Khomeini Hospital Complex and Research Department of Tehran University of Medical Sciences for data gathering.

Conflict of interest statement

There was no potential conflict of interest in this article.

Funding

There was no financial support for this article.

References

1. Hougaard P. *Analysis of multivariate survival data*. New York: Springer; 2000.
2. Kleinbaum D, Klein M. *Survival analysis: a self-learning text*. 5th ed. New York: Springer; 2005.
3. Klein J, Moeschberger M. *Survival analysis: techniquesfor censored and truncated data*. New York: Springer; 2003.
4. Klein JP, Bajorunaite R. Inference for Competing Risks. In: Balakrishnan N, Rao CR, editors. *Handbook of statistics*. Vol 23. Netherlands: Elsevier; 2004. pp. 291-311.
5. Satagopan JM, Ben-Porat L, Berwick M, Robson M, Kutler D, Auerbach AD. A note on competing risks in survival data analysis. *Br J Cancer*. 2004;91:1229-1235.
6. Fine J, Jiang H, Chappell R. On semi-competing risks data. *Biometrika*. 2001;88:907-719.
7. Day R, Bryant J, Lefkopoulou M. Adaptation of bivariate frailty models for prediction, with application to biological markers as prognostic indicators. *Biometrika*.1997;84:45-56.
8. Barker P, Henderson R. Small sample bias in the gamma frailty model for univariate survival. *Liftime Data Anal*. 2005;11:256-284.
9. Rondeau V, Commenges D, Joly P. Maximum penalized likelihood estimation in a gamma frailty model. *Liftime Data Anal*. 2003;9:139-153.
10. Wang MC, Qin J, Chiang C. Analyzing recurrent event data with informative censoring. *J Am Statist Ass*. 2001;96:1057-1065.
11. Liu L, Wolfe RA, Huang X. Shared frailty models for recurrent events and a terminal event. *Biometrics*. 2004;60:747-756.
12. Huang CY, Wang MC. Joint modeling and estimation for recurrent event processes and failure time data. *J Am Statist Ass*. 2004;99:1153-1167.
13. Wang, W. Estimating the association parameter for copula models under dependent censoring. *J R Stat Soc Series B Stat Methodo*. 2003;65:257-273.
14. Jiang H, Fine JP, Kosorok R, Chappell R. Pseudo self-consistent estimation of a copula model with informative censoring. *Scand J Stat*. 2005;33:1-20.
15. Jiang H, Fine J, Chappell R: Semiparametric analysis ofsurvival data with left truncation and dependent right censoring. *Biometrics*. 2005;61:567-575.
16. Ghosh D. Semi-parametric inferences for association with semi-competing risks data. *Stat Med*. 2006; 30;25(12):2059-2070.
17. Peng L, Fine J. Regression modeling of semi-competing risks data. *Biometrics*. 2007;63:96-108.
18. Lakhal L, Rivest LP, Abdous B. Estimating survival and association in a semicompeting risks model. *Biometrics*. 2008;64:180-188.
19. Hsieh JJ, Wang W, Ding AA. Regression analysis based on semicompeting risks data. *J R Stat Soc Series B Stat Methodo*. 2008;70:3-20.
20. Roshanaei G, Kazemnejad A, Sadighi S.Assessment of risk factors affecting recurrence of patients with gastric cancer in the presence of informative censoring in Iran. *Asian Pacific J Cancer Prev*. 2012;12:2443-2446.
21. Siewert JR, Bohcher K, Roder JD, Busch R. Prognostic relevance of systemic lymph node dissection in gastric carcinoma. *Br J Surg*. 1993;80:1015-1018.
22. Sadighi S, Mohagheghi MA, Haddad P, Omranipoor R, Moosavi Jarrahi AR, Meemari F, et al. Life expectancy with perioperative chemotherapy and chemoradiotherapy for locally advanced gastric adenocarcinoma. *Tehran University Medical Journal*. 2008;66(9):664-669.
23. Bando E, Yonemura Y, Takeshita Y, Taniguchi K, Yasui T, Yoshimitsu Y, et al. Intraoperative lavage for cytological examination in 1297 patients with gastric carcinoma. *Am J Surg*. 1999;178:256-262.
24. Marashi A, Moghimi M, Salehian MT, Peirovi H, Sodagari F. Local recurrence following total or subtotal gastrectomy for adenocarcinoma of the antrum. *Gastroenterology and Hepatology From Bed to Bench*. 2008;1(3):133-138.
25. Marrelli D, Stefano AD, Manzoni G, Morgagni P, Leo A.D, Roviello F. Prediction of recurrence after radical surgery for

gastric cancer: a scoring system obtained from a prospective multicenter study. *Ann Surg Oncol*. 2005;21:247-255.

26. Moriguchi S, Odakal T, Hayashi Y, Nose Y, Maehara Y, Korenaga D, et al. Death due to recurrence following curative resection of early gastric cancer depends on age of the patient. *Br J Cancer*.1991;64:555-558.

27. Lurje G, Husain H, Power DG, Yang D, Groshen S, Pohl A, et al. Genetic variations in angiogenesis pathway genes associated with clinical outcome in localized gastric adenocarcinoma. *Ann Oncol*. 2010;21:78-86.

28. Park SH, Kim DY, Heo JS, Lim DH, Park CK, Lee KW, et al. Postoperative chemoradiotherapy for gastric cancer. *Ann Oncol*. 2003;14:1373-1377.