

## RESEARCH ARTICLE

# Estimation of Hazard Function and its Associated Factors in Gastric Cancer Patients using Wavelet and Kernel Smoothing Methods

Azadeh Ahmadi<sup>1</sup>, Masoud Roudbari<sup>1,2\*</sup>, Mahmood Reza Gohari<sup>3</sup>, Bistoon Hosseini<sup>4</sup>

### Abstract

**Background and Objectives:** Increase of mortality rates of gastric cancer in Iran and the world in recent years reveal necessity of studies on this disease. Here, hazard function for gastric cancer patients was estimated using Wavelet and Kernel methods and some related factors were assessed. **Materials and Methods:** Ninety-five gastric cancer patients in Fayazbakhsh Hospital between 1996 and 2003 were studied. The effects of age of patients, gender, stage of disease and treatment method on patient's lifetime were assessed. For data analyses, survival analyses using Wavelet method and Log-rank test in R software were used. **Results:** Nearly 25.3% of patients were female. Fourteen percent had surgery treatment and the rest had treatment without surgery. Three fourths died and the rest were censored. Almost 9.5% of patients were in early stages of the disease, 53.7% in locally advance stage and 36.8% in metastatic stage. Hazard function estimation with the wavelet method showed significant difference for stages of disease ( $P < 0.001$ ) and did not reveal any significant difference for age, gender and treatment method. **Conclusion:** Only stage of disease had effects on hazard and most patients were diagnosed in late stages of disease, which is possibly one of the most reasons for high hazard rate and low survival. Therefore, it seems to be necessary a public education about symptoms of disease by media and regular tests and screening for early diagnosis.

**Keywords:** Hazard - gastric cancer - wavelet - kernel - methods

*Asian Pacific J Cancer Prev*, 13 (11), 5643-5646

### Introduction

The increase in mortality rate of gastric cancer (Joossens and Geboers, 1981; Boring, 1994; Parkin et al., 1999; Pisani et al., 1999; Leung and Sung, 2003) reveal necessity of studies on this disease. On the other hand, to obtaining pattern for rate of occurred death, the hazard curve is more obvious than survival curve (Klein and Moeschberger, 2003). The common methods for estimation of hazard function (Antoniadis, 1989; O'Sullivan, 1988; Kooperberg and stone, 1992) have time-consuming and complex calculations or programming. Recently Wavelet estimation of hazard function is considered that fast speed calculations is its excellent advantage (Antoniadis et al., 1999).

Therefore, the aim of this study is to estimate the hazard function of gastric cancer patients in Fayazbakhsh hospital in Tehran, using wavelet and kernel smoothing methods, to compare.

### Materials and Methods

Ninety-five recourse patients to Fayazbakhsh hospital, between 1997-2003 years, with gastric cancer were considered as a prospective study. These were all patient with definite diagnosis, and with complete records, therefore there was no sampling. The last follow up of the patient was done with a telephone call in 2009. The proposed event for these patients was death from Gastric Cancer, so those who were still alive at the end of study, or with no information from their situation after some specific time, considered as right censor. In this study the patients' lifetime for their age (<60 year, 60-70 years and >70 years), gender, stages of disease, type of treatment (surgery, non-surgery) were assessed. For definition of stages of disease, staging system and TNM method was used. According to this system, the patients were divided into following three categories: a) Early stage Gastric Cancer, involvement of the mucosa and sub-mucosa

<sup>1</sup>Department of Mathematics and Statistics, <sup>3</sup>School of Health Management and Information Sciences, Hospital Management Research Center, Tehran University of Medical Sciences, <sup>2</sup>Department of Mathematics and Statistics, School of Health Management and Information Sciences, the Anti-Microbial Resistance Research Center, <sup>4</sup>University of Applied Science and Technology, Tehran, IR Iran \*For correspondence: m-roudbari@tums.ac.ir

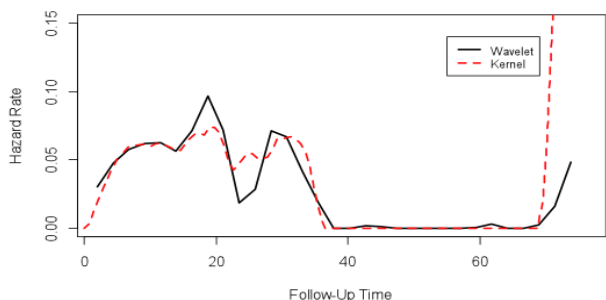
without regional lymph node involvement. b) Locally Advanced Gastric Cancer. c) Metastatic Gastric Cancer with any distant metastases.

For data analysis, survival analysis with wavelet method and for the comparison of hazard function in different levels of each variable, Log rank test was used. The wavelet smoothing method, which was used in this study, has widespread applications in functions approximation, and one of its most important applications in statistics is the estimation of hazard function in survival analysis. The simplest form of wavelet analysis has base on Harr wavelet and for introducing wavelets people starts with Harr wavelets. In this method, a series of discrete available observations of the hazard function are used to approximate the function graph (Nason, 2008).

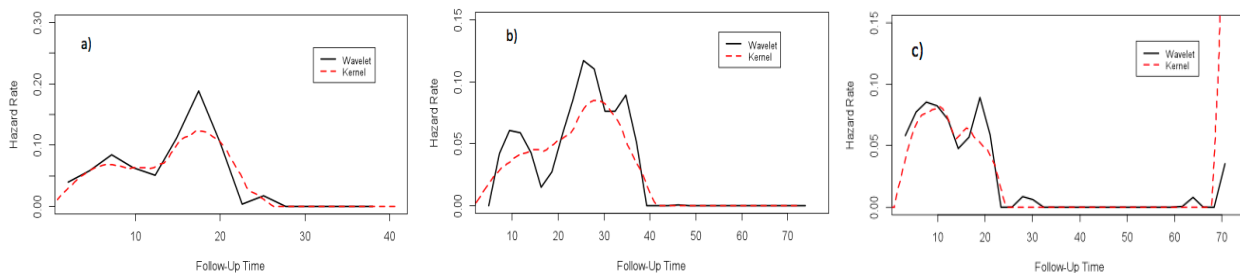
For data analysis the wavethresh and muhaz packages in R software were used. The study was done with the hospital and patients' permission, and their data kept confidential and were analyzed using anonymous codes. The patients were allowed for withdrawal the study anytime they wish.

**Results**

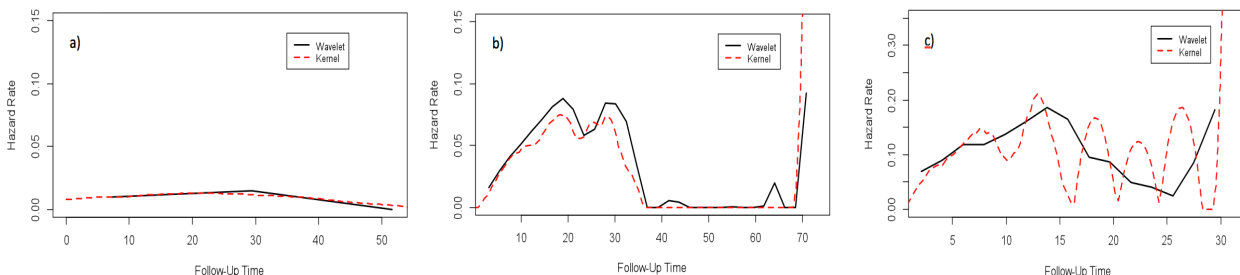
Twenty-four (25.3%) of patients were female and 71



**Figure 1. The Comparison of Estimated Hazard Function Using Smoothing Methods of Wavelet and Kernel**



**Figure 2. The Estimation of Hazard Function Using Wavelet and Kernel Methods for the Age Groups. a) 60 and less, b) between 60 to 70 years and c) more than 70 years**



**Figure 3. The Estimation of Hazard Function using Wavelet and Kernel Methods for Stages of Disease. a) Elementary, b) Locally advanced and c) Metastatic**

(74.7%) patients were male. The mean age was 63.2 year in females, 63.5 year in males and in total participant, it was 63.4 year. Twenty-nine (30.5%) patients had the age of 60 or less, 39 (41.1%) with 60-70 and 27 (28.4%) with the age of more than 70. Thirteen (14%) patients had surgical treatment and 82 (86%) patients had non-surgical treatment. Seventy-one (74.7%) patients died until the end of the study and 24 (25.3%) were censored. Nine (9.5%) patients were at the early stage with the mean age of 56.4 year, 51 (53.7%) patients in local spreading with the mean age of 63.5 year, and 35 (36.8%) patients were in Metastatic stage with the mean age of 65 year.

Figure 1 shows the estimate of hazard function using both wavelet and kernel smoothing methods. As it can be seen in wavelet method, the death rate for gastric cancer in more than 35 or less than 75 month, follow up after diagnosis, is almost zero. Also, the most death took place before month 35 of diagnosis. The same trend can be observed in kernel method.

Figure 2 shows the estimation of hazard function using wavelet and kernel methods for patients' three age groups. The graphs are almost the same as it can be seen, and there was no significant different between the hazard functions, using log rank test ( $p=0.14$ ). Using log rank test there was no significant different between hazard functions in males and females ( $p=0.296$ ).

Figure 3 shows the hazard function for three stages of disease using both wavelet and kernel methods. As it can be seen from hazard functions, the functions are different in different stages of disease. Using log rank test, there was a significant different between hazard functions among stages of disease ( $p<0.001$ ).

In addition, there was no significant different between hazard functions for the type of treatment using log rank test ( $p=0.997$ ), so surgical and non-surgical treatments have no significant effect different in the hazard function. Therefore, the only variable which has significant effect on hazard function in this study is stages of disease.

## Discussion

The variable of patients' lifetime in different age, gender, stages of disease and type of treatment were assessed in this study. The hazard function showed significant different for the stages of disease. The death rate for disease in time more than 35 month and less than 75 month was almost zero. The most death was before month 35. This is probably due to the recourse of patients at the advanced stages of disease, because only nine patients recourse at the elementary stage and the rest recourse were in local spreading or metastatic stage, then their life time reduced and do not remain at the study for a long time. Therefore, most of death occurred before month 35 and patients which were remained until this time, are probably alive at the end of the study except a little patient which died at the end of study. On the other hand, the death rate after month 35 is almost zero and patients which were remained until this time, have small probability for die. At the end of the Figure 1 the probability of death event has increased due to remain fewer patients with hazard exposure.

Although in some studies, it proved that age is a significant factor in lifetime (Matley et al., 1988; Mitsudomi et al., 1989; Arveux et al., 1992; Maehara et al., 1995; Haugstvedt et al., 1993; Bucchi et al., 2004; Enzinger and Mayer, 2004; Saidi et al., 2004) there was no significant relationship between age and the lifetime in this study but, in some other studies there was shown that there is not a significant relationship between age and the lifetime (Khedmat et al., 2007; Biglarian et al., 2009; Roshanaei et al., 2010).

The cause of not being relationship between age and lifetime is probably due to the method of changing the age to three categories. When there are old patients in lowest age group (less than 60 years), they cause reducing the lifetime of their group. Therefore, finding significant relationship between first and second categories becomes more difficult. If the age categories were divided into four new categories such as (less than 50 years, 50-59 years, 60-69 year and 70 year and more), we would probably be able to find the significant relationship between age and the life time.

The next non-significant variable on lifetime was gender. The result of the study is different from those who proved that there is significant relationship between sex and the lifetime (Bako et al., 1985; Curtis et al., 1985; Jonasson et al., 1994), but the study results is the same as which proved that there is not significant relationship between sex and the lifetime (Damhuis and Tilanus, 1995; Sigon et al., 2003; Liu et al., 2004).

One of the causes that gender has non-significant relationship with the lifetime, may be due to the disease diagnosis in all patients which was in advanced stages of disease for every gender. In other word, the ratio of patients in advanced stages of disease is the same in both gender and it was not possible to find any significant relationship between gender and the lifetime.

The stages of disease have significant effect on patients' lifetime ( $P < 0.001$ ), such as many other studies

(Msika et al., 2000; Barchielli et al., 2001; Thong-Ngam et al., 2001; Tian et al., 2004; Khedmat et al., 2007; Biglarian et al., 2009; Roshanaei et al., 2010) and this proved the necessity of disease diagnosis at early stages of disease.

In almost all studies, there is a significant relationship between disease stages and the patients' lifetime. It seems the cause of this event is the growth of gastric cancer in the body and existence of metastasis to other body parts due to late diagnosis of the disease. Therefore, with spread of the disease its treatment is more difficult and has more mortality.

The type of treatment has not significant relationship with patients' lifetime which is similar to a recent study (Biglarian et al., 2009). In Khedmat study (Khedmat et al., 2007) he proved that patients which received no treatment had more mortality than patient with surgical treatment. The cause of no finding significant relationship for type of treatment may be due to the short number of patients (13 patients) with surgical treatment. This can be due to the large number of patients in advanced stages of disease which it has metastasis to other parts, so surgery treatment is less applicable.

It can be concluded that the advanced stages of disease has relationship with lifetime. Almost 90% of patients refer to doctors in local spreading or metastatic stage and this is the most important cause of death hazard and low lifetime. The delay in disease diagnosis causes spread of disease in other parts of body (metastasis). In this condition, the treatment has less success and therefore, these patients have low lifetime.

It is important to educate the element signs of disease to the public and performing regular experiments for early disease diagnosis. In addition, it is possible to divide people into different parts according to probable factors affecting the disease, which people with high probability of getting the disease, can be separated from others. Using this method and applying consultation and regular experiments, it is possible to consider people with more risk of disease. With early diagnosis of the disease, it is feasible to treat the patients and reduce the disease mortality.

We used the estimation of hazard function using wavelet method for right censored data. It is possible to perform the same research with left and interval censored and right and left truncated data. In addition, it is feasible to use wavelet method with other methods of estimating hazard function such as penalized likelihood and the methods which are relying on multivariate regression.

Also, more than age, gender, stages of disease, and type of treatment which used in the study, it is possible to consider weight, using smoke and alcohol, nutrition, family history of cancer, BMI, the location of tumor etc., for finding more factors affecting gastric cancer.

## Acknowledgements

This study was part of a master thesis supported by Tehran University of Medical Sciences, grant number 14656.

## References

- Antoniadis A (1989). A penalty method for nonparametric estimation of the intensity of a counting process. *Ann Inst Math Statist*, **41**, 781-807.
- Antoniadis A, Gregoire G, Nason G (1999). Density and hazard rate estimation for right-censored data by using wavelet methods. *J R Statist Soc B*, 61.
- Arveux P, Faivre J, Boutron MC, et al (1992). Prognosis of gastric carcinoma after curative surgery. A population-based study using multivariate crude and relative survival analysis. *Dig Dis Sci*, **37**, 757-63.
- Bako G, Ferenczi L, Hanson J, Hill GH, Dewar R (1985). Factors influencing the survival of patients with cancer of the stomach. *Clin Invest Med*, **8**, 158-64.
- Barchielli A, Amorosi A, Balzi D, Crocetti E, Nesi G (2001). Long-term prognosis of gastric cancer in a European country: a population-based study in Florence. 10-year survival of cases diagnosed in 1985-1987. *Eur J Cancer*, **37**, 1674-80.
- Biglarian A, Hajizadeh E, Kazemnejad A, Zali MR (2009). Survival analysis of gastric cancer patients using Cox model: a five years study. *Tehran University Med J*, **67**, 317-25.
- Boring CC, Squires TS, Tong T, Montgomery S (1994). Cancer statistics. *CA Cancer J Clin*, **44**, 7-26.
- Bucchi L, Nanni O, Ravaioli A, et al (2004). Cancer mortality in a cohort of male agricultural workers from northern Italy. *J Occup Environ Med*, **46**, 249-56.
- Curtis RE, Kennedy BJ, Myers MH, Hankey BF (1985). Evaluation of AJC stomach cancer staging using the SEER population. *Semin Oncol*, **12**, 21-31.
- Damhuis RA, Tilanus HW (1995). The influence of age on resection rates and postoperative mortality in 2773 patients with gastric cancer. *Eur J Cancer*, **31**, 928-31.
- Enzinger PC, Mayer RJ (2004). Gastrointestinal cancer in older patients. *Semin Oncol*, **31**, 206-19.
- Haugstvedt TK, Viste A, Eide GE, Søreide O (1993). Norwegian multicentre study of survival and prognostic factors in patients undergoing curative resection for gastric carcinoma. The norwegian stomach cancer trial. *Br J Surg*, **80**, 475-8.
- Joossens JV, Geboers J (1981). Nutrition and gastric cancer. *Nutr Cancer*, **2**, 250-61.
- Jónasson L, Hallgrímsson J, Sigvaldason H, Olafsdóttir G, Tulinius H (1994). Gastric cancer in Iceland: a retrospective study of resected gastric specimens in a high-risk country during 30 years (1960-1989). *Int J Cancer*, **57**, 793-8.
- Khedmat H, Panahian M, Amini M, et al (2007). Survival rate of stomach cancer among patients hospitalized in Baghiatollah hospital. *Teb Nezami*, **9**, 167-77.
- Klein JP, Moeschberger ML (2003). Survival analysis, techniques for censored and truncated data. 2<sup>nd</sup> ed. New York: Springer-Verlag.
- Kooperberg C, Stone CJ (1992). Log-spline density estimation for censored data. *J Comput Graph Statist*, **1**, 301-28.
- Leung W, Ng E, Sung J (2003). Tumors of the Stomach. In: Yammada T, eds. *Textbook of Gastroenterology: Lippincott Williams & Wilkins*, 1416-31.
- Liu JF, Wang QZ, Hou J (2004). Surgical treatment for cancer of the esophagus and gastric cardio in Hebei, China. *Br J Surg*, **91**, 90-8.
- Matley PJ, Dent DM, Madden MV, Price SK (1988). Gastric carcinoma in young adults. *Ann Surg*, **208**, 593-6.
- Maehara Y, Oshiro T, Oiwa H, et al (1995). Gastric carcinoma in patients over 70 years of age. *Br J Surg*, **82**, 102-5.
- Mitsudomi T, Matsusaka T, Wakasugi K, et al (1989). A clinicopathological study of gastric cancer with special reference to age of the patients; an analysis of 1630 cases. *World J Surg*, **13**, 225-31.
- Msika S, Benhamiche AM, Rat P, Faivre J (2000). Long-term prognosis of gastric cancer in the population of Cote-d'Or. *Gastroenterol Clin Biol*, **24**, 649-55.
- Nason GP (2008). Wavelet Methods in Statistics with R. New York: Springer.
- O'Sullivan F (1988). Fast computation of fully automated log-density and log-hazard estimators. *SIAM J Sci Statist Comput*, **9**, 363-79.
- Parkin DM, Pisani P, Ferlay J (1999). Global cancer statistics. *CA Cancer J Clin*, **49**, 33-64.
- Pisani P, Parkin DM, Bray F, Ferlay J (1999). Estimates of the worldwide mortality from 25 cancers in 1990. *Int J Cancer*, **83**, 18-29.
- Roshanaei G, Kazemnejad A, Sedighi S (2010). Postoperative survival estimation of gastric cancer patients in cancer institute of tehran, imam khomeini hospital and its relative factors. *Hamadan University Med J*, **17**, 13-8.
- 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.
- Sigon R, Canzonieri V, Rossi C (2003). Early gastric cancer: a single institution experience on 60 cases. *Suppl Tumori*, **2**, 23-6.
- Tian J, Wang XD, Chen ZC (2004). Survival of patients with stomach cancer in Changle city of China. *World J Gastroenterol*, **10**, 1543-6.
- Thong-Ngam D, Tangkijvanich P, Mahachai V, Kullavanijaya P (2001). Current status of gastric cancer in Thai patients. *J Med Assoc Thai*, **84**, 475-82.