

Evaluation of parametric models by the prediction error in colorectal cancer survival analysis

Ahmad Reza Baghestani¹, Mahmood Reza Gohari², Arezoo Orooji¹, Mohamad Amin Pourhoseingholi³, Mohammad Reza Zali³

¹Department of Biostatistics, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

²Department of Biostatistics, Iran University of Medical Sciences, Tehran, Iran.

³Gastroenterology and Liver Diseases Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ABSTRACT

Aim: The aim of this study is to determine the factors influencing predicted survival time for patients with colorectal cancer (CRC) using parametric models and select the best model by predicting error's technique.

Background: Survival models are statistical techniques to estimate or predict the overall time up to specific events. Prediction is important in medical science and the accuracy of prediction is determined by a measurement, generally based on loss functions, called prediction error.

Patients and methods: A total of 600 colorectal cancer patients who admitted to the Cancer Registry Center of Gastroenterology and Liver Disease Research Center, Taleghani Hospital, Tehran, were followed at least for 5 years and have completed selected information for this study. Body Mass Index (BMI), Sex, family history of CRC, tumor site, stage of disease and histology of tumor included in the analysis. The survival time was compared by the Log-rank test and multivariate analysis was carried out using parametric models including Log normal, Weibull and Log logistic regression. For selecting the best model, the prediction error by apparent loss was used.

Results: Log rank test showed a better survival for females, BMI more than 25, patients with early stage at diagnosis and patients with colon tumor site. Prediction error by apparent loss was estimated and indicated that Weibull model was the best one for multivariate analysis. BMI and Stage were independent prognostic factors, according to Weibull model.

Conclusion: In this study, according to prediction error Weibull regression showed a better fit. Prediction error would be a criterion to select the best model with the ability to make predictions of prognostic factors in survival analysis.

Keywords: Parametric model, Prediction error, Apparent loss, Colorectal cancer.

(Please cite as: Baghestani AR, Gohari MR, Orooji A, Pourhoseingholi MA. Evaluation of parametric models by the prediction error in colorectal cancer survival analysis. *Gastroenterol Hepatol Bed Bench* 2015;8(3):183-187).

Introduction

Colorectal cancer is the third most common cancer and cause of cancer death worldwide (1). The incidence and mortality of colorectal cancer are rising rapidly in Asian countries (2-4).

Gastrointestinal cancers are the most important causes of mortality in Iran, which is located in the Middle East, Asia, and the burden of these cancers are increasing (5, 6) and CRC is not an exception (7). The prognosis of CRC is relatively good in terms of survival. Over the past 60 years, numerous claims have been made of variables being related to survival of colorectal cancer. Several studies have also considered independent prognostic factors,

Received: 14 January 2015 Accepted: 4 March 2015

Reprint or Correspondence: Mohamad Amin Pourhoseingholi, PhD, Gastroenterology and Liver Diseases Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

E-mail: amin_pgh@yahoo.com

including age at diagnosis (8), sex (9), stage (10), histological grade (8) etc.

Survival models or failure time models are statistical techniques to estimate the overall time up to specific events and find the related factors or predict the outcome. Prediction is important in medical science, because doctors need to estimate the survival of patients to choose the best treatment and it helps one to know about the disease condition in the future (11). The accuracy of prediction is determined by a measurement, generally based on loss functions, called prediction error. Recently, this technique developed to estimate the prediction error in survival analysis, in order to find the best model for analyzing the prognosis factors (12, 13).

The aim of this study is to determine the factors influencing predicted survival time for patients with colorectal cancer using parametric models and select the best model by predicting error's technique.

Patients and Methods

The data belongs to registered patients with colorectal carcinoma who admitted to the Cancer

Registry Center of Gastroenterology and Liver Disease Research Center, Taleghani Hospital, Shahid Beheshti University of Medical Sciences; Tehran, Iran, in the period between 2002 to 2007. All patients were followed until January 1, 2007(as failure time) from their diagnosis by telephone contact (14, 15). Each patient was informed by a consent form for documenting his/her information in the Cancer Registry Center. The data of 600 patients who were followed at least for 5 years and have completed information selected for this study. Body Mass Index (BMI), Sex, family history of CRC, tumor site (colon, rectum), stage of disease (early, advanced) and histology of tumor (Mucinous, others) included in the analysis. The survival time was compared by the Log-rank test and multivariate analysis was carried out using parametric models including Log normal, Weibull and Log logistic regression. For selecting the best model, the prediction error by the apparent loss (12) was used in which smaller error indicates a better model. $P < 0.05$ was considered as statistically significant and all analysis carried out using R software (16).

Table1. The result of log rank test for Univariate analysis

Prognostic Factors	Number of death	Number of patients	Mean (SE)	Test statistic	P-value
BMI					
Less than 25	113(31.2)	362(60.3)	71.565(5.055)	28.55	<0.0001
More than 25	38(16)	238(39.7)	128.032(9.719)		
Tumor site					
Colon	96(23)	418(69.7)	109.166(6.99)	4.936	0.026
Rectum	55(30.2)	182(30.3)	77.32(6.83)		
Family history					
Yes	88(25)	345(57.5)	74.557(6.523)	2.034	0.154
No	63(24.7)	255(42.5)	115.178(6.812)		
sex					
Male	57(22.3)	256(42.7)	111.503(9.348)	3.87	0.049
Female	94(27.3)	344(57.3)	88.114(8.334)		
Stage					
Early	57(18.8)	303(50.5)	117.436(7.768)	25.537	<0.0001
Advance	94(31.6)	297(49.5)	68.118(4.938)		
Histology					
Mucinous	18(28.1)	64(10.7)	105.113(14.006)	0.766	0.381
Others	133(24.8)	536(89.3)	99.799(6.369)		

Results

Among 600 patients, 344 were men (57.3%) and 256 were women (42.7%). Among 151 patients who died, 62.3% were men. The mean of survival for patients was 105.08 months (95% CI: 950.5-115.1) and the median was 94.5 months (95% CI: 58.6-130.4). Log rank test showed a better survival for females, BMI more than 25, patients with early stage at diagnosis and patients with colon tumor site (Table 1).

All factors included in parametric models (Log normal, Weibull and Log logistic censored regression) and prediction error by apparent loss was estimated for each model consequently, which resulted as the 1.46 for Log logistic, 1.49 for Log normal and 1.28 for Weibull. Therefore, Weibull model was the best model among these parametric models (Table 2) and revealed that BMI and stage of disease were independent prognostic factors of CRC survival. The relative risk of death for patients in the advanced stage of disease is 2.27 times more compared to patients is in the early first stage and patients with low BMI (less than 25) were at higher risk of death, compared to those with BMI more than 25. The other variables were not significant.

Table2. The results of Weibull Censored Regression for Multivariate analysis

	RR ¹	SE	P-value
BMI	2.2	0.16	<0.001
Tumor site ¹	1.17	0.14	0.2
Family history	1.12	0.14	0.4
Sex	1.13	0.14	0.3
Stage	1.96	0.14	<0.001
Histology ¹	1.08	0.21	0.7

Relative Risk, 2.Colon is as the reference group, 3. Others were as the reference group.

Discussion

In this study, BMI and stage of the disease were prognostic factors of CRC survival, according to parametric regression model and the apparent loss prediction error indicated that Weibull model was the best option among

parametric models to analyze the survival of CRC patients who admitted in Taleghani hospital.

However, in the Log rank test, sex was a significant factor, multivariate model showed no relation between sex and survival of CRC patients. A population study on about 165,000 CRC patients in Germany reported a better survival for women (17).

According to the histology type of tumor, Log rank test and Weibull model showed no difference in survival. This result is consistent to some studies (18, 19).

The Log rank analysis revealed a better survival of colon cancer, compared to rectal's. However, this result was not significant in multivariate analysis. Other studies reported a better survival for colon cancer (20, 21). People with *rectal cancer tend to be older* and may have other serious health issues. Therefore, it would be the reason of different survival.

Family history of CRC was another risk factor in our analysis. Although individuals with a family history of colorectal cancer are diagnosed more often than the general population (22), the study suggests that survival from colorectal cancer may not be worse (23) and the result of this analysis in both univariate and multivariate confirmed that.

BMI was a prognostic factor of CRC survival in both Log rank and Weibull analysis and the patients with higher BMI had a better survival. A similar study suggested that underweight and obese women with colon cancer were at increased risk of death (24).

In multivariate and univariate analysis, the effect of the cancer's stage was significant on survival time. A similar study indicated that patients whose cancer is in the early stage have a better survival time (25).

In this study, we used a parametric model to analyze the survival rate of patients with CRC and select the best-appropriated model using prediction error. Parametric models are more flexible than Cox semi parametric model (26-28).

Besides, prediction error would be criteria to select the best model with the ability to make predictions of prognostic factors in survival analysis.

Acknowledgment

This paper is a result of Ms Thesis in Biostatistics, Shahid Beheshti University of Medical Sciences.

References

- Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin* 2011; 61:69-90.
- Moghimidehkhordi B, Safaee A. An overview of colorectal cancer survival rates and prognosis in Asia. *World J Gastrointest Oncol* 2012;4:71-5.
- Pourhoseingholi MA. Increased burden of colorectal cancer in Asia. *World J Gastrointest Oncol* 2012;4:68-70.
- Pourhoseingholi MA, Vahedi M, Baghestani AR. Burden of gastrointestinal cancer in Asia; an overview. *Gastroenterol Hepatol Bed Bench* 2015;8:19-27.
- Pourhoseingholi MA, Vahedi M, Moghimidehkhordi B, Pourhoseingholi A, Ghafarnejad F, Maserat E, et al. Burden of hospitalization for gastrointestinal tract cancer patients - Results from a cross-sectional study in Tehran. *Asian Pac J Cancer Prev* 2009;10:107-10.
- Pourhoseingholi MA, Fazeli Z, Ashtari S, Bavand-Pour FS. Mortality trends of gastrointestinal cancers in Iranian population. *Gastroenterol Hepatol Bed Bench* 2013;6:S52-7.
- Pourhoseingholi MA, Faghihzadeh S, Hajizadeh E, Abadi A, Zali MR. Bayesian estimation of colorectal cancer mortality in the presence of misclassification in Iran. *Asian Pac J Cancer Prev* 2009;10:691-4.
- Weiss JM, Pfau PR, O'Connor ES, King J, LoConte N, Kennedy G, et al. Mortality by stage for right-versus left-sided colon cancer: analysis of surveillance, epidemiology, and end results - Medicare data. *J Clin Oncol* 2011;29:4401-9.
- Coleman M, Quaresma M, Berrino F, Lutz JM, De Angelis R, Capocaccia R, et al. Cancer survival in five continents: a worldwide population-based study (CONCORD). *Lancet Oncol* 2008;9:730-56.
- Nedrebø B, Søreide K, Eriksen M, Dørum L, Kvaløy J, Søreide J, et al. Survival effect of implementing national treatment strategies for curatively resected colonic and rectal cancer. *J Surg* 2011; 98:716-23.
- Henderson R, Keiding N. Individual survival time prediction using statistical models. *J Med Ethics* 2005; 31:703-706.
- Gerds TA, Schumacher M. Efron-type measures of prediction error for survival analysis. *Biometrics* 2007; 63:1283-87.
- Lawless JF, Yuan Y. Estimation of prediction error for survival models. *Stat Med* 2010; 29:262-74.
- Azadeh S, Moghimidehkhordi B, Fatem SR, Pourhoseingholi MA, Ghiasi S, Zali MR. Colorectal cancer in Iran: an epidemiological study. *Asian Pac J Cancer Prev* 2008;9:123-26.
- Moghimidehkhordi B, Safaee A, Zali MR. Prognostic factors in 1,138 Iranian colorectal cancer patients. *Int J Colorectal Dis* 2008;23:683-88.
- Hothorn T, Everitt BS. *A Handbook of Statistical Analyses Using R*. 3rd ed. Boca Raton, Florida, USA: Chapman & Hall/CRC Press; 2014.
- Majek O, Gondos A, Jansen L, Emrich K, Holleczeck B, Katalinic A, et al. Sex differences in colorectal cancer survival: population-based analysis of 164,996 colorectal cancer patients in Germany. *PLoS One* 2013;8:e68077.
- Nilsson KR, Berenholtz SM, Dorman T, Garrett P, Kaufman HS, Pronovost PJ. Preoperative predictors of blood transfusion in colorectal cancer surgery. *J Gastrointest Surg* 2002; 6:753-62.
- Liang H, Wang XN, Wang BG, Pan Y, Liu N, Wang DC, et al. Prognostic factors of young patients with colon cancer after surgery. *World J Gastroenterol* 2006;12:1458-62.
- Wray CM, Ziogas A, Hinojosa MW, LeH, Stamos MJ, Zell JA. Tumor subsite location within the colon is prognostic for survival after colon cancer diagnosis. *DisColon Rectum* 2009;52:1359-66.
- Meguid RA, Slidell MB, Wolfgang CL, Ahuja N. Is there a difference in survival between right-versus left-sided coloncancers? *Ann SurgOncol* 2008;15:2388-94.
- Safaee A, Moghimidehkhordi B, Pourhoseingholi MA, Vahedi M, Maserat E, Ghiasi S, et al. Risk of colorectal cancer in relatives: A case control study. *Indian J Cancer* 2010;47:27-30.
- Kirchhoff AC, Newcomb PA, Trentham-Dietz A, Nichols HB, Hampton JM. Family history and

colorectal cancer survival in women. *Fam Cancer* 2008;7:287-92.

24. Doria-Rose VP, Newcomb PA, Morimoto LM, Hampton JM, Trentham-Dietz A. Body mass index and the risk of death following the diagnosis of colorectal cancer in postmenopausal women (United States). *Cancer Causes Control* 2006;17:63-70.

25. Maringe C, Walters S, Rachet B, Butler J, Fields T, Finan P, et al. Stage at diagnosis and colorectal cancer survival in six high-income countries: a population-based study of patients diagnosed during 2000-2007. *Acta Oncol* 2013;52:919-32.

26. Pourhoseingholi MA, Hajizadeh E, Moghimi Dehkordi B, Safaee A, Abadi A, Zali MR. Comparing Cox regression and parametric models for survival of patients with gastric carcinoma. *Asian Pac J Cancer Prev* 2007;8:412-6.

27. Moghimi-Dehkordi B, Safaee A, Pourhoseingholi MA, Fatemi R, Tabeie Z, Zali MR. Statistical comparison of survival models for analysis of cancer data. *Asian Pac J Cancer Prev* 2008;9:417-20.

28. Pourhoseingholi MA, Moghimi-Dehkordi B, Safaee A, Hajizadeh E, Solhpour A, Zali MR. Prognostic factors in gastric cancer using log-normal censored regression model. *Indian J Med Res* 2009;129:262-67.