



Survival Analysis of Colorectal Cancer Patients Using Exponentiated Weibull Distribution

Moein Yoosefi,¹ Ahmad Reza Baghestani,^{2,*} Naghmeh Khadembashi,³ Mohamad Amin Pourhoseingholi,⁴ Alireza Akbarzadeh Baghban,⁵ and Azin Khosrovirad¹

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

²Physiotherapy Research Center, Department of Biostatistics, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

³Department of English Language, School of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

⁴Gastroenterology and Liver Diseases Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran

⁵Proteomics Research Center, Department of Basic Sciences, School of Rehabilitation, Shahid Beheshti University of Medical Sciences, Tehran, Iran

*Corresponding author: Ahmad Reza Baghestani, Physiotherapy Research Center, Department of Biostatistics, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. E-mail: baghestani.ar@gmail.com

Received 2016 September 05; Revised 2017 April 11; Accepted 2018 February 25.

Abstract

Background: Colorectal cancer is a highly prevalent cancer across the world, which is associated with significant mortality. Therefore, special attention must be paid to the prevention and care of this fatal disease. The present study aimed to evaluate the effective methods to prolong the survival of colorectal cancer patients and determine the influential factors in their longevity using the exponentiated Weibull statistical model.

Methods: Medical records of 446 colorectal cancer patients referring to Taleghani Hospital in Tehran, Iran during 1985 - 2013 were reviewed. The exponentiated Weibull model is an extended version of the Weibull model, which is used to investigate the survival of colorectal cancer patients and assess the influential factors in their survival and longevity. We also compared the model with the Weibull-exponential model for better practice.

Results: Mean survival rate of the patients was 4.52 ± 0.182 years. According to the exponentiated Weibull model, age at diagnosis was the only significant influential factor in the survival of colorectal cancer patients ($P=0.002$). Other factors such as gender, tumor size, family history of colorectal cancer, tumor site, and body mass index had no significant effects on the survival of these patients. The proposed model was considered superior to the Weibull-exponential model due to the lower Akaike's information criterion.

Conclusions: According to the results, age of colorectal cancer patients at the time of diagnosis is the most important influential factor in increasing their survival and reducing the mortality rate. Therefore, it is recommended that proper care be provided for the elderly patients with colorectal cancer, especially at the onset of the disease.

Keywords: Colorectal Cancer, Survival Analysis, Exponentiated Weibull Statistical Model

1. Background

Cancer is a life-threatening, prevalent disease across the world with the mortality rate of 65% in underdeveloped countries (1). Studies show that in 2012 alone, approximately 14.1 million people were afflicted with this fatal disease (1). Lung, breast, colorectal, and stomach cancer are among the most hazardous diseases, comprising 40% of all cancers (2).

Colorectal cancer has recently been a major concern among researchers and physicians due to its prevalence and significance in terms of mortality and disease burden. Irregular growth of cancer cells in the walls of the colon and rectum causes polyps, some of which are benign, while two-thirds of these polyps may give rise to colorectal can-

cer (3). In Iran, Moradi et al. (2009) reported the median survival of colorectal cancer to be 3.5 years (4). In Texas (United States), Scott Kopetz et al. (2009) estimated the time-dependent increasing median survival of colorectal cancer patients to be 2.5 years (5).

Considering the prevalence rate of the disease, colorectal cancer is reported to be the third most common cancer among men and the second most common cancer among women in different countries (6). It is also the fifth and third most frequent cancer among Iranian men and women, respectively (7). Unfortunately, the prevalence of colorectal cancer has been on the rise in Iran (8).

Various influential factors have been determined in the development and exacerbation of colorectal cancer, including genetics, lifestyle, gender, family history, tumor

size, tumor site, body mass index (BMI), and patient's age at the time of diagnosis. In this regard, controlling of many other factors could affect the longevity of these patients. With this background, it is of utmost importance to further investigate this disease to decrease its prevalence. Evidence suggests that early diagnosis and changes in dietary habits along with controlling the associated influential factors could prolong the life of colorectal cancer patients (9).

Survival studies have indicated the time response variable to be at stake until the onset of the event. Moreover, the defining characteristic of survival studies is that the response variable could be censored; in other words, there may be no definitive event time for some patients. Finding the proper model for estimating the survival and verifying the prognosis of the covariates is the main concern in survival analysis (10). Matching the features of survival data is of paramount importance in the modeling of these data. Some of the available models for survival analysis include non-parametric methods (Kaplan-Meier), semiparametric analysis (Cox's proportional hazard model), and parametric models (exponential, Weibull, and lognormal). Considering their flexibility and variety in function and performance, parametric models are of particular interest to many researchers.

Exponentiated Weibull statistical model is an extended version of the Weibull model, which is a flexible, robust model in fitting survival data and can be effectively used based on the data structure (11). To estimate the survival of chronic patients and the influential factors in this regard, widely used survival analysis models are often reviewed. One of the most important functions of survival studies is the hazard function. In parametric models, the hazard function associated with the data may have various forms (monotonic, decreasing, increasing, unimodal, and bathtub shape). The main advantage of the exponentiated Weibull model over the Weibull model is the coverage of unimodal hazard function in the related data.

Recently, several studies have exploited the Weibull model (12, 13). The current research is the first to use the exponentiated Weibull model in order to assess the survival rate of colorectal cancer patients and effective measures in the presence of covariates and censored survival time. The present study aimed to investigate the survival of colorectal cancer patients using the exponentiated Weibull model and determine the effects of the aforementioned covariates on the survival time of these patients.

2. Methods

Medical records of 446 patients referring to Taleghani hospital in Tehran, Iran during 1985 - 2012 were reviewed

in the present study. The certification of ethics committee regarding this article is provided by Shahid Beheshti University of Medical Sciences by ethic number: IR.SBMU.REC.1396.182.

In these patients, the survival time was the interval between cancer diagnosis and death (year). Variables such as gender, family history, tumor size, tumor site, BMI, and age at diagnosis were investigated as the covariates of the model. In order to impute the missing values of the tumor size variable, we used the logistic regression imputation method (14). As 10% of the tumor size observation was missing, some variables were considered independent to tumor size for accuracy, with the aid of the logistic regression to impute the missing values of this variable. Mice library in the R software was used for imputation.

In the present study, the exponentiated Weibull model was used to evaluate the survival of the patients with colorectal cancer. This model is in fact a generalized format of the Weibull model, which has two shape parameters (α and β) and a scale parameter (λ) (15). The exponentiated Weibull function is as follow:

$$F(t) = [1 - \exp(-(\lambda t)^\alpha)]^{\beta-1} \quad (1)$$

$$f(t) = \alpha \beta \lambda [1 - \exp(-(\lambda t)^\alpha)]^{\beta-1} \exp(-(\lambda t)^\alpha) (\lambda t)^{\alpha-1} \quad (2)$$

For the specified amount of $\beta = 1$, the exponentiated Weibull was converted into the Weibull model; it was anticipated that the flexibility of the model would increase by adding it to a shape parameter.

In the current research, we used the Akaike information criterion (AIC) to compare the exponentiated Weibull model to the former model. AIC is an approach to select the appropriate model by considering the number of the parameters and maximum likelihood function; accordingly, a proper model has smaller AIC. It is notable that in the present study, the comparison was carried out in the presence of the covariates, including gender, family history of colorectal cancer, tumor size, tumor site, BMI, and age at diagnosis. Landa is used as the link function to insert a covariate in the model, which is defined as

$$\lambda = \exp(b_0 + \sum_i b_i x_i) \quad (3)$$

where b_0 is the intercept, b_i represents the coefficient, and x_i denotes the covariates.

In the present study, the R software was used for the exponentiated Weibull model and its comparison with the exponential and Weibull models.

3. Results

In total, 446 colorectal cancer patients who referred to Taleghani hospital in Tehran were evaluated in this study.

Mean age at diagnosis and BMI of the patients were 51.53 ± 13.9 years and 25.07 ± 4.39 kg/m 2 , respectively. Approximately 55% of the colorectal cancer patients were male, and 45% were female. In 87% of the patients, the tumor size was less than one centimeter. Tumor sites were divided into three main categories, including the colon (28%), rectum (31%), and unknown location (41%). Family history of colorectal cancer was reported in 34% of the patients. Moreover, a number of the patients were excluded before the end of the study for different reasons (69%); in other words, they were censored and assumed to be independent of the event.

According to the findings, the survival rate of colorectal cancer was 0.93, 0.85, 0.79, 0.74, and 0.70 years in the first five survival times. In addition, the median survival time of the patients who died due to colorectal cancer and those who were censored was 2.44 and 4.91 years, respectively (total: 3.94 years).

As mentioned earlier, the main advantage of the exponentiated Weibull model over the Weibull model is the coverage of the unimodal hazard function in the related data. In the present study, the shape of hazard function in the exponentiated Weibull model was unimodal, whereas the Weibull model did not cover the unimodal shape in the hazard function (Figure 1).

According to the AIC values, the exponential, Weibull, and exponentiated Weibull models were expressed in the presence of the covariates, revealing the AIC for the exponentiated Weibull model to be lower than the other two models, which confirms the superiority of this model.

In the current research, the estimation curve of the survival function was drawn for the three parametric models in the presence of the covariates, and the estimated survival time in the exponentiated Weibull model was observed to be higher comparatively (Figure 2).

According to the results, the age of colorectal cancer patients at the time of diagnosis was the only influential factor in their survival, which significantly affected all the three models ($P < 0.001$) (Table 1).

To interpret and compare the applied models, patients aged 40 - 41 years were investigated. According to the results the hazard ratio in these patients was the constant over the survival time in the Weibull model, while this finding differed in the exponentiated Weibull model as the hazard ratio changed over the survival time (Figure 3). Similar to the Weibull model, the exponential model constant hazard ratio function was ignored (Figure 3).

During the study, interpreting the effect of age at diagnosis on the survival time was different in the exponentiated Weibull model as the interpretation depended on the time of the study. Effect of the patient's age on the survival time showed the highest value at the beginning of the

study (Figure 2). However, factors such as gender, family history, tumor size, tumor site, and BMI had no significant effects on the survival of the patients with colorectal cancer.

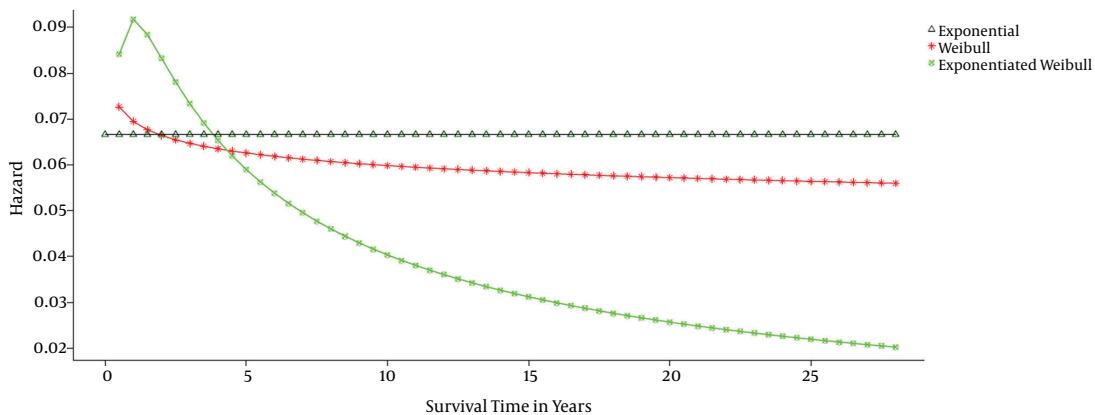
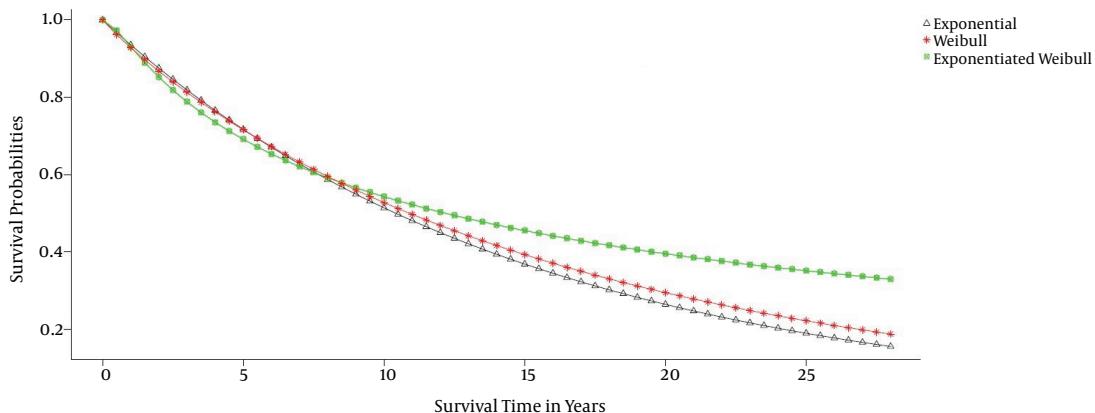
4. Discussion

In the present study, mean age at diagnosis and BMI of the colorectal cancer patients were 51.53 ± 13.9 years and 25.07 ± 4.39 kg/m 2 , respectively. On average, the evaluated patients survived 4.52 years after the disease diagnosis, while the rate has been reported differently in the previous studies (16, 17). One of the most important objectives of the current research was to investigate the survival of colorectal cancer patient based on the exponentiated Weibull model in the presence of the influential factors and compare the mentioned model with the most common models of survival analysis, such as the exponential and Weibull models.

Parametrical models are generally practical in survival studies, and researchers select them to ensure accurate estimation (9). Parametrical models could be used for the estimation of the patients' survival and speculating the related influential factors without the need for specific defaults. Considering the spread, precision, and flexibility of these models, they have acquired a significant role in the field of medicine (12, 13). Despite the abundant use of parametric models, due attention has not been paid to these models for the survival analysis of colorectal cancer patients as compared to other models (e.g., Kaplan-Meier and Cox's proportional hazard model).

In a study, Baghestani et al. (2015) investigated the survival model of 600 colorectal cancer patients based on the Weibull model (12). In the present study, we used the exponentiated Weibull model, which is an extension of the Weibull model. The suitability of this model has been confirmed for colorectal cancer patients in the presence of age at diagnosis, as well as other factors (gender, family history, tumor size, tumor site, and BMI). According to our findings, this model is more appropriate for the analysis of colorectal cancer data compared to the other two models.

According to the findings of the current research, the only variable with a significant effect on the longevity of colorectal cancer patients across the three models was the age at the time of diagnosis; in other words, elderly patients at diagnosis are at a higher risk of death compared to the other age groups, which becomes more evident in the case of early diagnosis. In this regard, Mehrkhani et al. (2008) conducted a research on 1,090 colorectal cancer patients and concluded that the patients aged less than 65 years had a better chance of survival upon diagnosis compared to those aged more than 65 years (17). Furthermore,

**Figure 1.** Hazard Function in Exponential, Weibull, and Exponentiated Weibull Models**Figure 2.** Survival Trend in Exponential, Weibull, and Exponentiated Weibull Models**Table 1.** Results of Parametric Estimates in Exponential, Weibull and, Exponentiated Weibull Models

Variables	Exponential			Weibull			Exponentiated Weibull		
	Estimate	SE	P Value	Estimate	SE	P Value	Estimate	SE	P Value
Intercept	-3.195	0.905	0.000 ^a	-3.189	7	0.000 ^a	10.856	6.59	0.121
Gender	-0.216	0.182	0.235	-0.229	0.206	0.242	-0.192	0.206	0.354
Tumor size	-0.681	0.319	0.032 ^a	-0.723	0.345	0.036 ^a	-0.624	0.325	0.056
Tumor site (1)	0.231	0.224	0.302	0.243	0.241	0.311	0.176	0.253	0.485
Tumor site (2)	0.056	0.246	0.821	0.051	0.263	0.845	-0.045	0.269	0.866
Family history	-0.042	0.187	0.821	-0.057	0.201	0.774	-0.153	0.215	0.476
BMI	-0.012	0.023	0.591	-0.014	0.025	0.563	-0.009	0.024	0.696
Age at diagnosis	0.029	0.006	0.000 ^a	0.031	0.007	0.000 ^a	0.021	0.007	0.001 ^a
AIC	944			944.7			934.8		

Abbreviations: AIC, Akaike Information Criterion; BMI, Body Mass Index.
^astatistically significant.

Jing Li et al. investigated the effect of age on the survival of colorectal cancer patients using the multivariate Cox

and univariate Log Rank test models, concluding that the chance of survival decreases with increasing age (18).

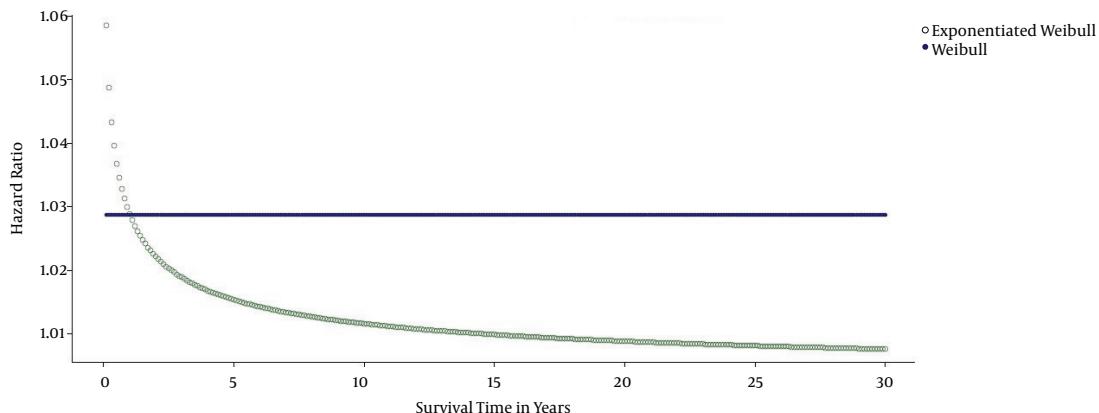


Figure 3. Hazard Ratio in Weibull and Exponentiated Weibull Models

Tumor size was another variable which proved to be of importance in the present study. Although it had no significant effect in the exponentiated Weibull model, it showed a significant effect when the other two models were used, so that the patients with larger tumors had a higher chance of survival compared to those with smaller tumors. This finding is inconsistent with the findings of Zhai et al. (2012), who assessed the effect of tumor size on colorectal cancer patients, claiming that larger tumors were associated with a lower chance of survival (19). With regard to the variable of tumor size, the Weibull and exponential models showed contrary results to our expectation; in the mentioned study, patients with small tumors were at a higher risk of mortality compared to those with large tumors, while the exponentiated Weibull model showed a more significant, logical outcome in this respect. The exponentiated Weibull model showed a quite similar result to that of Park's study, which was performed on 2,230 colorectal cancer patients (20).

In the present study, gender had no significant effect on the survival of colorectal cancer patients. On the other hand, Amando et al. (2013) reported the significant effect of gender on the survival of 3,284 patients (21). Moreover, in the study by Baghestani (2015), the Weibull and lognormal models were applied to determine the survival time of 600 colorectal cancer patients, and gender was observed to have a significant effect on survival (12); this finding was later confirmed by Jing (18). In contrast to the aforementioned studies, the present study revealed no correlation between gender and survival, which is in line with some studies in this regard (22).

BMI, which was an important variable in the current research, has no significant effect on the survival time of colorectal cancer patients. Findings regarding the effect

of BMI on the mortality rate of these patients are contradictory. In 2010, Shibakita stated the significant effect of BMI on the mortality rate of colorectal cancer patients, concluding that patients with the BMI of less than 21 and more than 24 kg/m^2 were at a higher risk of death (23). Furthermore, another study in 2000 revealed that men with the BMI of $25 - 30 \text{ kg/m}^2$ had better longevity (24).

Family history and tumor site were the two factors that proved influential in some studies, while other researchers denoted no significant effects in this regard (12, 18, 20-22). The discrepancy in the results of the studies could be due to the differences in the number of the variables used for survival analysis and sample populations.

The main strength of the present study is that it is the first to examine the appropriateness of the exponentiated Weibull model for evaluating the survival rate of colorectal cancer patients in the presence of covariates. With the inclusion of the unimodal shape in the hazard function and minimum AIC in the exponentiated Weibull model, this model proved to be more appropriate compared to the exponential and Weibull models.

In the current research, the tumor size was not documented in 10% of the patients; therefore, using the logistic regression, we estimated the missing values by finding an appropriate cutoff point and used the maximum sensitivity and specificity to impute the missing values of the tumor size.

In the exponential and Weibull models, analysis of the covariates was performed based on the hazard ratio, which was fixed across time for these models. However, the exponentiated Weibull model has a different analysis compared to the two mentioned models. Contrary to the previous model, it has a non-fixed risk ratio across time; therefore, in this model, the covariates analysis across time is differ-

ent.

In conclusion, it is recommended that the effects of other covariates (e.g., genetics, lifestyle, and dietary habits) on survival time be investigated in further research. The exponentiated Weibull model could be used to assess the risk in the healed patients formerly afflicted by colorectal cancer.

Acknowledgments

This article was extracted from an MSc thesis performed at the department of biostatistics, faculty of paramedical Sciences, Shahid Beheshti University of Medical Sciences in Tehran, Iran. Hereby, we extend our gratitude to the staff of the faculty of paramedical Sciences.

Footnotes

Authors' Contribution: Study design, Ahmad Reza Baghestani and Moein Yoosefi; statistical analysis and interpretation of data, Ahmad Reza Baghestani and Moein Yoosefi; drafting of the manuscript, Moein Yoosefi, Ahmad Reza Baghestani, Naghmeh Khadembashi, Mohamad Amin Pourhoseingholi, Alireza Akbarzadeh Baghban, Azin Khosrovirad.

Conflicts of Interests: None declared.

Financial Disclosure: This study was financially supported by Shahid Beheshti University of Medical Sciences, Tehran, Iran

References

- International Agency for Research on Cancer. *Latest world cancer statistics Global cancer burden rises to 14.1 million new cases in 2012: Marked increase in breast cancers must be addressed*. World Health Organization; 2013.
- International Agency for Research on Cancer. *World cancer fact sheet*. Geneva, Switzerland: World Health Organization; 2012.
- Levin B, Lieberman DA, McFarland B, Andrews KS, Brooks D, Bond J, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *Gastroenterology*. 2008;134(5):1570–95. doi: [10.1053/j.gastro.2008.02.002](https://doi.org/10.1053/j.gastro.2008.02.002). [PubMed: 18384785].
- Moradi A, Khayamzadeh M, Guya M, Mirzaei HR, Salmanian R, Rakhsha A, et al. Survival of colorectal cancer in Iran. *Asian Pac J Cancer Prev*. 2009;10(4):583–6. [PubMed: 19827873].
- Kopetz S, Chang GJ, Overman MJ, Eng C, Sargent DJ, Larson DW, et al. Improved survival in metastatic colorectal cancer is associated with adoption of hepatic resection and improved chemotherapy. *J Clin Oncol*. 2009;27(22):3677–83. doi: [10.1200/JCO.2008.20.5278](https://doi.org/10.1200/JCO.2008.20.5278). [PubMed: 19470929].
- Haggag FA, Boushey RP. Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. *Clin Colon Rectal Surg*. 2009;22(4):191–7. doi: [10.1055/s-0029-1242458](https://doi.org/10.1055/s-0029-1242458). [PubMed: 21037809].
- Mohammadianpanah M. Colorectal cancer incidence: Does Iran follow the west?. *Ann Colorectal Res*. 2015;3(1). doi: [10.5812/acr.28045](https://doi.org/10.5812/acr.28045).
- Rezaianzadeh A, Safarpour A, Marzban M, Marzban A. A systematic review of the incidence of colorectal cancer in Iran. *Ann Colorectal Res*. 2014;2(3):5.
- Stewart B, Wild C. *World Cancer report 2014*. Geneva: IARC Nonserial Publication World Health Organization; 2014.
- Kleinbaum DG, Klein M. *Survival analysis: a self-learning text*. Springer Science Business Media; 2006.
- Wahed AS, Luong TM, Jeong JH. A new generalization of Weibull distribution with application to a breast cancer data set. *Stat Med*. 2009;28(16):2077–94. doi: [10.1002/sim.3598](https://doi.org/10.1002/sim.3598). [PubMed: 19424958].
- Baghestani AR, Gohari MR, Orooji A, Pourhoseingholi MA, Zali MR. Evaluation of parametric models by the prediction error in colorectal cancer survival analysis. *Gastroenterol Hepatol Bed Bench*. 2015;8(3):183–7. [PubMed: 26328040].
- Baghestani AR, Hajizadeh E, Fatemi SR. Parametric model to analyse the survival of gastric cancer in the presence of interval censoring. *Tumori*. 2010;96(3):433–7. doi: [10.1177/030089161009600309](https://doi.org/10.1177/030089161009600309). [PubMed: 20845804].
- Royston P. Multiple imputation of missing values. *Stata J*. 2004;4(3):227–41.
- Mudholkar GS, Srivastava DK. Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Trans Rel*. 1993;42(2):299–302. doi: [10.1109/24.229504](https://doi.org/10.1109/24.229504).
- Kavlakoglu B, Ustun I, Oksuz O, Pekcici R, Ergocen S, Oral S. Surgical treatment of liver metastases from colorectal cancer: experience of a single institution. *Arch Iran Med*. 2011;14(2):120–5. [PubMed: 21361719].
- Mehrkhani F, Nasiri S, Donboli K, Meysamie A, Hedayat A. Prognostic factors in survival of colorectal cancer patients after surgery. *Colorectal Dis*. 2009;11(2):157–61. doi: [10.1111/j.1463-1318.2008.01556.x](https://doi.org/10.1111/j.1463-1318.2008.01556.x). [PubMed: 18462239].
- Li J, Wang Z, Yuan X, Xu L, Tong J. The prognostic significance of age in operated and non-operated colorectal cancer. *BMC Cancer*. 2015;15:83. doi: [10.1186/s12885-015-1071-x](https://doi.org/10.1186/s12885-015-1071-x). [PubMed: 25885448].
- Zhai ZW, Gu J. [Influence of tumor size on the prognosis in patients with colon cancer]. *Zhonghua Wei Chang Wai Ke Za Zhi*. 2012;15(5):495–8. [PubMed: 22648846].
- Park YJ, Park KJ, Park JG, Lee KU, Choe KJ, Kim JP. Prognostic factors in 2230 Korean colorectal cancer patients: analysis of consecutively operated cases. *World J Surg*. 1999;23(7):721–6. doi: [10.1007/PL00012376](https://doi.org/10.1007/PL00012376). [PubMed: 10390594].
- Phipps AI, Lindor NM, Jenkins MA, Baron JA, Win AK, Gallinger S, et al. Colon and rectal cancer survival by tumor location and microsatellite instability: the Colon Cancer Family Registry. *Dis Colon Rectum*. 2013;56(8):937–44. doi: [10.1097/DCR.0b013e31828f9a57](https://doi.org/10.1097/DCR.0b013e31828f9a57). [PubMed: 23838861].
- Josa V, Krzystanek M, Eklund AC, Salamon F, Zarand A, Szallasi Z, et al. Relationship of postoperative thrombocytosis and survival of patients with colorectal cancer. *Int J Surg*. 2015;18:1–6. doi: [10.1016/j.ijsu.2015.03.005](https://doi.org/10.1016/j.ijsu.2015.03.005). [PubMed: 25843227].
- Shibakita M, Yoshimura H, Tachibana M, Ueda S, Nagasue N. Body mass index influences long-term outcome in patients with colorectal cancer. *Hepatogastroenterology*. 2010;57(97):62–9. [PubMed: 20422873].
- Murphy TK, Calle EE, Rodriguez C, Kahn HS, Thun MJ. Body mass index and colon cancer mortality in a large prospective study. *Am J Epidemiol*. 2000;152(9):847–54. doi: [10.1093/aje/152.9.847](https://doi.org/10.1093/aje/152.9.847). [PubMed: 11085396].