TURKISH JOURNAL of ONCOLOGY

Turk J Oncol 2021;36(2):184–90
doi: 10.5505/tjo.2021.2566

ORIGINAL ARTICLE

Is There Any Advantage of Machine Learning to Multivariate Regression Analysis for Predicting Disease-Related Deaths in Patients with Gastric Cancer? Reevaluation of Retrospective Data

Umut KAYA,[1] Gökhan YAPRAK,[2] Atınç YILMAZ,[3] Alaattin ÖZEN[4]

[1]Department of Computer Engineering, Istanbul Ayvansaray University Faculty of Engineering, İstanbul-*Turkey*
[2]Department of Radiation Oncology, University of Health Sciences, Kartal Dr. Lutfi Kırdar Training and Research Hospital, İstanbul-*Turkey*
[3]Department of Computer Engineering, Beykent University Faculty of Engineering and Architecture, İstanbul-*Turkey*
[4]Department of Radiation Oncology, Eskisehir Osmangazi University Faculty of Medicine, Eskisehir-*Turkey*

OBJECTIVE
The problem in gastric cancer patients is multifactorial and it is not possible to use one factor alone to predict disease survival. Machine learning (ML) algorithms have become popular in the medical field, recently. We aimed to evaluate the power of ML algorithms for predicting deaths due to gastric cancer.

METHODS
We reevaluated the retrospective data published. Seven different ML algorithms (logistic regression [LR], artificial neural networks/multilayer perceptron, gradient boosted trees, support vector machine, random forest, naive Bayes, and probabilistic neural network) tried to predict disease-related deaths using the significant variables effective on disease-specific survival (DSS) obtained from univariate analysis.

RESULTS
Median follow-up time was 34 months (4-156 months), and the death with disease occurred in 194 (86.6%) patients in the follow-up period. The median DSS was 22 (4-139) months. Using the significant variables effective on DSS obtained from univariate analysis, the highest accuracy rate (99%) was the best in the LR, and only one patient was classified incorrectly.

CONCLUSION
We can successfully predict the treatment outcomes such as disease-related deaths in gastric cancer patients using ML algorithms.

Keywords: Disease-related death; gastric cancer; machine learning algortihms.
Copyright © 2021, Turkish Society for Radiation Oncology

Introduction

Gastric cancer, the common cause of cancer-related deaths, is the sixth most common cancer worldwide. [1] Although the TNM stage is the most valuable prognostic factor, lymphovascular space invasion (LVSI), grade, surgery type, and performance score are other factors that can affect the survival of the patient.[2-4] Gastrectomy plus regional lymph node dissection remains the primary treatment of the surgically remov-

Received: January 14, 2021
Accepted: February 07, 2021
Online: March 23, 2021

Accessible online at:
www.onkder.org

OPEN ACCESS This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Öğr. Gör. Umut KAYA
İstanbul Ayvansaray Üniversitesi Mühendislik Fakültesi,
Bilgisayar Mühendisliği Bölümü
İstanbul-Turkey
E-mail: umutkaya@ayvansaray.edu.tr

able disease, but surgery without any pre-operative or post-operative treatment provides a 5-year overall survival (OS) rate of 20-30%.[5,6] The survival benefit of adjuvant treatment options when compared with surgery alone in potentially operable patients has been shown in several randomized trials.[2,4,5] Adjuvant chemoradiotherapy (ChRT) as the result of SWOG 9008/INT-0116 trial and perioperative (pre-operative plus post-operative) chemotherapy (ChT) as the result of the MAGIC trial are the treatment options that have been used until today.[2,6] While the most significant prognostic factors are tumor spread, tumor size, and lymph node metastasis status, tumor grade, histology, surgical margin, tumor localization, and performance status are also using in the treatment decision and estimating prognosis.[2-4,7]

The problem in patients diagnosed with gastric cancer is also multi-factorial as in many different areas in the universe that means many variables contribute to treatment results. We cannot use one factor alone to predict disease survival, as disease, patient, and treatment-related factors are the relationship to cancer patients' survival. In this way, the multivariate analysis tool aims to find patterns and relationships between several variables simultaneously and, multivariate analysis lets us predict the effects of a change in one variable will have on other variables. The multivariate analysis is capable of providing a more accurate depiction and understanding of the behavior of data that are highly correlated with each other. Multivariate analysis techniques are complex and a statistical program is necessary for performing this analysis. One of the significant limitations of multivariate analysis is that statistical modeling outputs are not always easy for clinicians to interpret. Furthermore, to obtain meaningful results for multivariate techniques, a large sample of data is necessary.

Machine learning (ML) has become popular in the health sector recently. Although there is no consensus on which algorithm is the best, applications related to ML are studied in several trials that include patients with cancer.[8] Many ML algorithms are capable of learning from the provided data by investigators and also the accuracy and efficiency of models to make decisions improve with subsequent training as new data are provided. Although the most significant advantage of ML is the ability to automate various decision-making tasks, the most painful and difficult point of ML is the acquisition of data and the cost of collecting data.

In our study, we aimed to evaluate the power of ML algorithms for predicting deaths due to stomach cancer. Thus, we used statistically significant parameters that we obtained from univariate analysis for disease-specific survival (DSS).

Materials and Methods

In this study, we reevaluated the retrospective data published as prognostic factors for survival in patients with gastric cancer: Single-center experience was reported by Yaprak et al.[9] We aimed to work on these data because we had ready-made data sets of a large number of patients diagnosed with gastric cancer. We excluded patients with Stage 4 and who had missing data.

Patient Characteristics

The patient characteristics are summarized in Table 1. In our study, the median age was 57 (range; 22-87), and 66.5% of the patients were male. Total gastrectomy was performed in 168 (50.3%) patients, and subtotal gastrectomy was performed in 166 (49.7%) patients. 258 (77.8%) patients had positive lymph node disease and 76 (22.8%) patients confirmed as a node-negative disease based on pathologic examination. According to staging, 50 (15%) patients were Stage I, 94 (28.2%) patients were Stage II, and 190 (56.8%) patients were Stage III. Perineural invasion (PNI) was identified in 203 (60.8%) patients, and LVSI was identified in 238 (71.3%) patients. 41 patients had Grade 1 (12.3%), 107 patients had Grade 2 (32.0%), and 186 patients had Grade 3 (55.7%) disease.

Treatment and Relapse Patterns in Follow-up

Two hundred and twelve patients (63.5%) were considered eligible for adjuvant ChRT. The RT treatment was administered as 45 Gy/25 fractions in 172 (81.1%) patients and 50.4 Gy/28 fractions in 27 (12.7%) patients. Thirteen (6.1%) patients could not complete 45 Gy due to toxicity. Two-dimensional technique and three-dimensional conformal technique was used in 52 (24.5%) and 160 (75.5%) patients, respectively. All patients received bolus or infusional 5-FU as one cycle before RT and one cycle after RT. Used concomitant ChT schemes were bolus fluorouracil and leukovorin, or infusional fluorouracil, or oral capecitabine. The characteristics of the received treatments are summarized in Table 2.

ML

In our study, to predict the DSS, we used seven different ML algorithms, such as logistic regression (LR), artificial neural networks/multilayer perceptron (ANN/

Turk J Oncol 2021;36(2):184–90
doi: 10.5505/tjo.2021.2566

Table 1 Demographic and clinicopathologic characteristics of the patients

Characteristics	n	%
Gender		
Male	222	66.5
Female	112	33.5
Age (year), mean±SD	57.1±11.5	
<70	281	84.1
≥70	53	15.9
Tumor location		
Gastroesophageal junction	11	3.3
Fundus, cardia	80	24.0
Corpus	97	29.0
Antrum pylorus	146	43.7
Surgical resection type		
Total gastrectomy	168	50.3
Subtotal gastrectomy	166	49.7
Tumor size		
<5 cm	139	58.4
≥5 cm	195	41.6
Lymphovascular invasion		
Yes	238	71.3
No	96	28.7
Perineural invasion		
Yes	203	60.8
No	131	39.2
Grade		
I	41	12.3
II	107	32.0
III	186	55.7
Surgical margin		
Negative	294	88.0
Positive	40	12.0
T stage		
T1	34	10.2
T2	64	19.2
T3	105	31.4
T4	131	39.2
N stage		
N0	76	22.9
N1	68	20.4
N2	90	26.9
N3a	74	22.2
N3b	26	7.8
TNM stage		
IA	10	3.0
IB	40	12.0
IIA	44	13.2
IIB	50	15.0
IIIA	70	21
IIIB	62	18.6
IIIC	58	17.4

T stage: T1: Tumour invades lamina propria or muscularis mucosae (1a) or invades submucosa (1b), T2: Tumour invades muscularis propria, T3: Tumour penetrates subserosal tissue without invasion of visceral peritoneum or adjacent structures, T4: Tumour penetrates serosa (visceral peritoneum) (4a), Tumour directly invades adjacent organs or structures (4b); N stage: N0 No regional lymph node metastases, N1: Metastases in 1-2 regional lymph nodes, N2: Metastases in 3-6 regional lymph nodes, N3a: Metastases in 7-15 regional lymph nodes, N3b: Metastasis in 16 or more regional lymph nodes; TNM stage: IA: T1N0, IB: T2N0, T1N1, IIA: T3N0, T2N1, T1N2, IIB: T4aN0, T3N1, T2N2, T1N3, IIIA: T4aN1, T3N2, T2N3, IIIB: T4bN0-1, T4aN2, T3N3, IIIC: T4bN2-3

Table 2 Treatment characteristics of the patients

Characteristics	n	%
Adjuvant radiotherapy		
Yes	212	63.5
No	122	36.5
Radiation technique		
Two-dimensional radiotherapy	52	24.5
Three-dimensional radiotherapy	160	75.5
Radiation dose		
<45 Gy	13	6.1
45 Gy	172	81.1
50.4 Gy	27	12.7
Concomitant chemotherapy		
Yes	212	36.5
No	122	63.5
Concomitant chemotherapy regimes		
Bolus 5- fluorouracil	102	48.1
Infusional 5- fluorouracil	78	36.8
Oral Capecitabine	32	15.1

MLP), gradient boosted trees (GBT), support vector machine (SVM), random forest (RF), naive Bayes (NB), and probabilistic neural network (PNN). We used the parameters obtained from the univariate analysis results for predicting DSS using ML algorithms.

LR algorithm generates a curve between 0 and 1 value and makes probability estimation. The algorithm uses the natural logarithm of the probabilities of the target variable while constructing the model.[10] The ANN/MLP algorithm is created by imitating the way nerve cells in the human brain, known as neurons, carries information. While performing the learning process with experience, this algorithm tries to find the relationship between data and create a meaningful pattern between them.[11] The GBT algorithm, which dominates data sets, is an algorithm that is created using gradient supported decision trees, which is preferred due to its speed and performance and is used in the solution of classification and regression problems.[12] The SVM algorithm is an algorithm that can make binary or multiple classifications on the data set and can generalize on data whose distribution is unknown and can predict new data thanks to these data.[13] RF algorithm is a classification algorithm that can work with missing data and show high accuracy when used in large data sets. Since different data and variables are used in each tree, no overfitting problem has been encountered in the algorithm.[14] The NB algorithm is an algorithm that can increase the classification accuracy and is used to process con-

tinuous values frequently.[15] PNN algorithm is an algorithm based on Bayes rule and class probability estimation to minimize the possibility of misclassification. PNN, an algorithm that is used frequently in classification and pattern recognition problems, was created using feed-forward neural networks. This algorithm approximates the parent probability distribution function of the classes using the Parzen window and a non-parametric model. PNN uses the parental probability distribution function of each class to estimate the class probability of new input and adds the class with the highest odds according to the Bayesian approach as the new input.[16]

From the retrospective gastric cancer data we have, we reevaluated patient, disease, and treatment characteristics, such as age, stage, tumor diameter, LVSI, PNI, grade, surgery status, surgery type, radiation technique, and concomitant ChT status. We decided the dataset into two groups for algorithm training and testing the accuracy of prediction. Patients distributed between these two groups in a ratio of 70-30%. The models were constructed using the training set and validated using the testing set.

Statistics and Application

The complexity matrix is a matrix created from the information obtained by comparing the actual and predicted data and applying the classification process to these data. The complexity matrix is used to determine the classification performance of the methods used.[17,18] The accuracy rate method is used to determine how it performs the classification process accurately. This method is calculated by dividing the number of true-positive and true-negative samples in the samples by the total number. The error rate is calculated by proportioning the number of false-positive and false-negative samples to the total number.[19]

We defined DSS as a period from the date of diagnosis to the date of cancer-related death or the last follow-up date. The Kaplan-Meier method was used for survival analysis. A Cox proportional hazard model was utilized for multivariate analysis to determine independent prognostic factors. All the tests were two-sided and, $p<0.05$ was considered to be statistically significant.

Results

Median follow-up time was 34 months (4-156 months), and in the follow-up, locoregional relapse, and/or distant relapse occurred in 204 (61.1) patients. Locoregional relapse alone occurred in 48 (23.5%) patients and, distant relapse alone occurred in 145 (71.1%) patients. The death occurred in 224 (67.1) patients and, death with disease occurred in 194 (86.6%) patients. The median OS was 34.5 (4-156) months and DSS was 22 (4-139) months.

The univariate analysis showed that age (<70 vs. ≥70 years, p=0.042), tumor diameter (<5 vs. ≥5 cm, p=0.006), T stage (p<0.001), N stage (p<0.001), stage (p<0.001), LVSI (p=0.005), grade (p<0.001), adjuvant RT dose (<45 Gy vs. ≥45 Gy, p=0.023), and relapse situation (p<0.001) were affecting factors on DSS.

According to the univariate analysis results, two different multivariate analysis models were described. In the first model age, tumor diameter, T stage, N stage, LVSI, grade, adjuvant RT dose, and relapse, and in the second model age, tumor diameter, TNM Stage, LVSI, grade, adjuvant RT dose, and relapse situation were included in the study. As a result of multivariate analysis, the independent prognostic factor was the N stage (p=004) and TNM stage (p<0.001) in two different models, respectively.

The results for the prediction of DSS obtained from seven different ML algorithms using parameters obtained from the univariate analysis are shown in Table 3. The accuracy rate was the best in the LR algorithm, and also incorrectly classified patient count was only 1.

Table 3 Recall, precision, sensitivity, specificity, f-measure, and accuracy of machine learning algorithms

Algorithms	Recall	Precision	Sensitivity	Specificity	F-Measure	Accuracy
Naive Bayes	0.983-0.905	0.935-0.974	0.983-0.905	0.905-0.983	0.959-0.938	0.95
Gradient boosted trees	0.967-0.950	0.967-0.950	0.967-0.950	0.950-0.967	0.967-0.950	0.96
Logistic regression	1.000-0.977	0.98-1.000	1.000-0.977	0.977-1.000	0.991-0.989	0.99
Multi-layer perceptron	0.984-0.974	0.984-0.974	0.984-0.974	0.974-0.984	0.984-0.974	0.98
Probabilistic neural network	0.946-0.867	0.898-0.929	0.946-0.867	0.867-0.946	0.922-0.897	0.91
Random forest	0.981-0.979	0.981-0.979	0.981-0.979	0.979-0.981	0.981-0.979	0.98
Support vector machine	0.982-0.933	0.948-0.977	0.982-0.933	0.933-0.982	0.965-0.955	0.96

Turk J Oncol 2021;36(2):184–90
doi: 10.5505/tjo.2021.2566

Discussion

The mortality rate and incidence of gastric cancer differ throughout the world.[20] Surgery is a curative treatment. Despite the improvements in surgical techniques, surgery alone without any pre- or post-operative treatment option provides a reasonable OS rate. Randomized studies demonstrated the OS between 20 and 30% following surgery alone in patients with operable gastric cancer.[3,21] The survival rate varies according to the T and N stage, such as 85-90% in patients with T1 tumors and 15-20% in patients with T4 and node-positive patients. Furthermore, locoregional recurrence rates are a serious concern in resected patients.[22] Given the information above, a multi-modal approach is necessary to improve surgical results.

The stage has been the most commonly used and most-effective factor for predicting the prognosis in patients with gastric cancer.[23] Similarly, in our study, we found the N and the AJCC Stage is negative prognostic factors on DSS. However, receiving different treatments, such as a combination of surgery, ChT, and RT, will cause many more significant risk factors influencing the DSS. Therefore, more-comprehensive prognosis models, such as nomograms, have been implemented, including demographics and other significant clinical parameters except the stage.[24] The constructed nomograms for this purpose over time include several independent prognostic factors. First, Zhong et al.[25] presented a nomogram for predicting the 10-year DSS for patients with gastric cancer. Age is an important prognostic factor for DSS in many studies, and low survival has been demonstrated in elderly patients.[26,27] According to these studies, we have shown that advanced age is a worse prognostic factor in our assessment.

There is currently no consensus on the optimal algorithm to predict treatment results by ML. Several studies in the literature used clinical, radiological, tissue, and blood genomics for predicting survival by ML in several cancer types.[28-33] However, to the best of our knowledge, there are no studies that evaluate using ML algorithms for predicting DSS in patients with gastric cancer to date. Using our retrospective data, in the present study, we aimed to compare seven ML methods commonly used in the literature as LR, ANN/MLP, GBoosted, SVM, RF, NB, and PNN. Although all algorithms had a high accuracy rate of >90%; in our study, the best algorithm with the highest accuracy to predict DSS was the LR algorithm.

LR algorithm has high success in classification problems where dependent variables are not continuous.[10] Our findings suggest that the issue in the data set we used in our study is more suitable for the LR method. ANN/MLP algorithm can work with missing data and is also successful in solving both regression and classification problems. When the literature is examined, the findings suggest that the ANN/MLP algorithm achieves better results in large data sets.[11] The failure of the ANN/MLP algorithm in solving the problem with the data set we used in our study may be related to the number of patients. Because of the outliers in the data set used in the study, the GBT algorithm showed a lower classification success compared to the other methods because it was overly adapted to the outliers in the classification. Since there are many tree structures in the model, it is more expensive concerning computation time and requires more memory.[12] SVM algorithm is a simple and practical, high performance. and useful algorithm based on estimating the most suitable function to separate data from each other. In this algorithm, the number of samples is insignificant, and superiority to the other algorithms is that it can classify data that have not been seen during training without any problems.[13] However, finding the optimal plane to separate the samples is critical for the algorithm, and samples were classified in a multifactorial disease, such as cancer, may not always be separated by a linear line. Since its ability in probabilistic classification is lower than other methods, it has led to lower performance in this problem. RF algorithm has a high success rate that classifies using many decision tree structures. The number of instances to be used in each node and the number of trees to be created must be determined in the algorithm to create a tree structure. Given the class ratios in the data set, the algorithm uses 2/3 of the whole data set as training and 1/3 as test data. RF generally outperforms decision trees and has a lower accuracy than gradient supported trees. Since there are too many tree structures in the algorithm, it is known to obtain slower results in real-time classification problems.[14] Although the RF algorithm showed high performance in our study, it worked slower than other algorithms. NB algorithm is a probabilistic simple classification method based on Bayes' theorem. The algorithm accepts attributes independently of each other, and the examples are all equally important. However, it is not possible given that each datum has the same significance in patients with cancer.[15] In our study, parameters and class information are dependent on each other. Since the NB algorithm considers the relationship between parameters independent from each other, the successful performance has calculated

lower than other methods. PNN algorithm is relatively insensitive to outliers. Outliers in the data set used in our study negatively affected the performance of the model, and it was less successful than other methods. Furthermore, the PNN algorithm is slower in classifying new cases than multi-layer sensor networks and requires more memory space to store the model.[16]

Conclusion

The prediction of prognosis in patients with cancer that underlies critical clinical decisions regarding treatment or monitoring is vital. Although our study has several limitations, we assessed the potential of predictions of disease-related deaths using an ML trained with prognostic parameters. Consequently, we can predict the treatment outcomes using these algorithms that enable learning based on different data types and by providing computers with the ability to detect complicated patterns and make rational decisions based on the data in patients with gastric cancer.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The authors have no conflicts of interest to declare.

Ethics Committee Approval: The study protocol was approved by the University of Health Science, Dr. Lütfi Kırdar Training and Research Hospital Clinical Research Ethics Committee. (Number: 2018/514/136/1, Date: 28/08/2018).

Financial Support: The authors declared that this study has received no financial support.

Authorship contributions: Concept – U.K., G.Y.; Design – U.K., G.Y.; Supervision – A.Y., A.Ö.; Funding – None; Materials – G.Y.; Data collection and/or processing – G.Y.; Data analysis and/or interpretation – U.K., A.Y.; Literature search – U.K., A.Ö.; Writing – U.K.; Critical review – A.Y., G.Y., A.Ö.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68(6):394–424.
2. Macdonald JS, Smalley SR, Benedetti J, Hundahl SA, Estes NC, Stemmermann GN, et al. Chemoradiotherapy after surgery compared with surgery alone for adenocarcinoma of the stomach or gastroesophageal junction. N Engl J Med 2001;345(10):725–30.
3. Cunningham D, Allum WH, Stenning SP, Thompson JN, Van de Velde CJ, Nicolson M, et al. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. N Engl J Med 2006;355(1):11–20.
4. Zhang ZX, Gu XZ, Yin WB, Huang GJ, Zhang DW, Zhang RG. Randomized clinical trial on the combination of preoperative irradiation and surgery in the treatment of adenocarcinoma of gastric cardia (AGC)--report on 370 Patients. Int J Radiat Oncol Biol Phys 1998;42(5):929–34.
5. Neri B, Cini G, Andreoli F, Boffi B, Francesconi D, Mazzanti R, et al. Randomized trial of adjuvant chemotherapy versus control after curative resection for gastric cancer: 5-year follow-up. Br J Cancer 2001;84(7):878–80.
6. Ychou M, Boige V, Pignon JP, Conroy T, Bouché O, Lebreton G, et al. Perioperative chemotherapy compared with surgery alone for resectable gastroesophageal adenocarcinoma: an FNCLCC and FFCD multicenter phase iii trial. J Clin Oncol 2011;29(13):1715–21.
7. National Comprehensive Cancer Network. Gastic Cancer (Version 3.2020). Available at: https://www.nccn.org/professionals/physician_gls/pdf/gastric.pdf. Accessed, Aug 14, 2020.
8. Deist TM, Dankers FJ, Valdes G, Wijsman R, Hsu IC, Oberije C, et al. Machine learning algorithms for outcome prediction in (chemo)radiotherapy: an empirical comparison of classifiers. Med Phys 2018;45(7):3449–59.
9. Yaprak G, Tataroglu D, Dogan B, Pekyurek M. Prognostic factors for survival in patients with gastric cancer: single-centre experience. North Clin Istanb 2020;7(2):146–52.
10. Tolles J, Meurer WJ. Logistic regression relating patient characteristics to outcomes. JAMA 2016;316(5):533–4.
11. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction. New York: Springer; 2009.
12. Mason L, Baxter J, Bartlett PL, Frean M. Boosting algorithms as gradient descent. In: Solla SA, Leen TK, Müller K, editors. Advances in Neural Information Processing Systems 12. Cambridge, Massachusetts: MIT Press; 1999.
13. Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20(3):273–97.
14. Breiman L. Random forests. Mach Learn 2001;45(1):5–32.
15. Lou W, Wang X, Chen F, Chen Y, Jiang B, Zhang H. Sequence based prediction of DNA-binding proteins based on hybrid feature selection using random forest and Gaussian Naïve Bayes. PLoS One 2014;9(1):e86703.
16. Mohebali B, Tahmassebi A, Meyer-Baese A, Gandomi AH. Probabilistic neural networks: a brief overview of theory, implementation, and application. Amsterdam:

Turk J Oncol 2021;36(2):184–90
doi: 10.5505/tjo.2021.2566

Elsevier; 2020.
17. Stehman SV. Selecting and interpreting measures of thematic classification accuracy. Remote sens environ 1997;62(1):77–89.
18. Powers DM. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. J Mach Learn Technol 2011;2(1):37–63.
19. Chicco D, Jurman G. The advantages of the Matthews Correlation Coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Gen 2020;21(1):6.
20. Avital I, Pisters PW, Kelsen DP, Willet CG. Cancer of the stomach. In: Devita VT, Lawrence T, Rosenberg SA, editors. Cancer: principles and practise of oncology. 9th ed. Philadelphia, PA: Lippincott Williams and Wilkins; 2011.
21. Cuschieri A, Weeden S, Fielding J, Bancewicz J, Craven J, Joypaul V, et al. Patient survival after D1 and D2 resections for gastric cancer: longterm results of the MRC randomized surgical trial. Surgical Co-operative Group. Br J Cancer 1999;79(9-10):1522–30.
22. Gunderson LL. Gastric cancer-patterns of relapse after surgical resection. Semin Radiat Oncol 2002;12(2):150–61.
23. Wang PL, Xiao FT, Gong BC, Liu FN, Xu HM. A nomogram for predicting overall survival of gastric cancer patients with insufficient lymph nodes examined. J Gastrointest Surg 2017;21(6):947–56.
24. Dikken JL, Coit DG, Baser RE, Gönen M, Goodman KA, Brennan MF, et al. Nomogram for predicting the survival of gastric adenocarcinoma patients who receive surgery and chemotherapy. BMC Cancer 2020;20(1):10.
25. Zhong Q, Chen QY, Li P, Xie JW, Wang JB, Lin JX, et al. Prediction of conditional probability of survival after surgery for gastric cancer: a study based on eastern and western large data sets. Surgery 2018;163(6):1307–16.
26. Kim Y, Spolverato G, Ejaz A, Squires MH, Poultsides G, Fields RC, et al. A nomogram to predict overall survival and disease-free survival after curative resection of gastric adenocarcinoma. Ann Surg Oncol 2015;22(6):1828–-35.
27. Shang-Guan XC, Chen QY, Li P, Xie JW, Wang JB, Lin JX, et al. Preoperative lymph node size is helpful to predict the prognosis of patients with stage iii gastric cancer after radical resection. Surg Oncol 2018;27(1):54–60.
28. Savareh BA, Aghdaie HA, Behmanesh A, Bashiri A, Sadeghi A, Zali M, et al. A machine learning approach identified a diagnostic model for pancreatic cancer through using circulating microRNA signatures. Pancreatology 2020;20(6):1195–204.
29. Li R, Shinde A, Liu A, Glaser S, Lyou Y, Yuh B, et al. Machine learning-based interpretation and visualization of nonlinear interactions in prostate cancer survival. JCO Clin Cancer Inform 2020;4:637–46.
30. Bhambhvani HP, Zamora A, Shkolyar E, Prado K, Greenberg DR, Kasman AM, et al. Development of robust artificial neural networks for prediction of 5-year survival in bladder cancer. Urol Oncol 2020;2020:30213–1.
31. Wulczyn E, Steiner DF, Xu Z, Sadhwani A, Wang H, Flament-Auvigne I, et al. Deep learning-based survival prediction for multiple cancer types using histopathology images. PLoS One 2020;15(6):e0233678.
32. Alongi P, Laudicella R, Stefano A, Caobelli F, Comelli A, Vento A, et al. Choline PET/CT features to predict survival outcome in high risk prostate cancer restaging: a preliminary machine-learning radiomics study. Q J Nucl Med Mol Imaging 2020;2020:03227–6.
33. Lu L, Wang D, Wang L, Linning E, Guo P, Li Z, et al. A quantitative imaging biomarker for predicting disease-free-survival-associated histologic subgroups in lung adenocarcinoma. Eur Radiol 2020;30(7):3614–23.