Machine learning in predicting gastric cancer survival

Presenting a novel decision support system model

Ann. Ital. Chir., 2023 94, 6: 631-638
pii: S0003469X23038708

Ezgi Altinsoy*, Batuhan Bakirarar**, Serdar Culcu***

**Department of Genaral Surgery, Manisa State Hospital, Manisa, Turkey*
***Department of Biostatistics, Ankara University Faculty of Medicine, Ankara, Turkey*
****Department of Surgical Oncology, Ankara University Faculty of Medicine, Ankara, Turkey*

Machine learning in predicting gastric cancer survival: Presenting a novel decision support system model

BACKGROUND: *Gastric cancer is the 4th most frequent cause of cancer-related deaths, with a 5-year survival rate of less than 40%. In recent years, many artificial intelligence applications have been used in the field of gastric cancer through their effective computing and learning ability. In this study, we aim to develop a software that can accurately detect overall survival in gastric cancer cases with the help of artificial intelligence and machine learning.*
METHODS: *The study included 34417 patients' data diagnosed with gastric cancer between 2010 and 2015. The main hypothesis in the study was overall survival (OS) in years, defined from the date of diagnosis to the date of death or, for living patients, the last control date. In addition to survival, other variables selected for the analyzes were age at diagnosis, race, gender, behavior, primary site, grade, histology, T stage, N stage, M stage and size of the tumor, vital status, and follow-up time (months).*
RESULTS: *The median overall survival of the patients was found to be 15.00±0.20 years. Median life expectancy was found to be 21.00±0.85 years for those younger than 50 years of age, 20.00±0.43 years for those aged 50-69 years, and 10.00±0.22 years for those aged 70 and over. Especially artificial intelligence techniques such as machine learning and deep learning lead to remarkable developments in the field of gastric cancer.*
CONCLUSION: *With the ability to compute and learn we think that use of artificial intelligence will be revolutionary in gastric cancer in terms of diagnosis and prognosis.*

KEY WORDS: Artificial intelligence, Gastric cancer, Survival

Introduction

Gastric cancer (GC) is a malignant tumor that arises from the cells that make up the stomach wall. According to the latest data, approximately 1.2 million new gastric cancer cases are detected and around 865,000 people die from gastric cancer each year. The ratio of these numbers in all cancer types is 5.7% and 8.2%, respectively. It is predicted that one out of every 78 women and one out of every 33 men will be diagnosed with gastric cancer in their lifetime [1,2]. Treatment methods include surgery, chemotherapy, and molecular targeted therapy. Among the treatment options, surgery is still the main treatment for early-stage gastric cancer. The prognosis is very good in patients who are caught at an early stage and undergo radical resection [3]. However, in patients with advanced gastric cancer, the risk of recurrence after surgery is as high as 50-70% in patients who have only surgery [4]. In such patients, chemotherapy is one of the primary treatment options, but its effectiveness is still not at the desired level. Median survival in patients receiving chemotherapy is between 6-13 months [5]. In general, the 5-year survival rate of gastric cancer is below 40% [6].

Pervenuto in Redazione Luglio 2022. Accettato per la pubblicazione Settembre 2022
Correspondence to: Serdar Culcu, Ankara University Faculty of Medicine, Department of Surgical Oncology, Ankara, Turkey (e-mail: serdarculcu@gmail.com)

ABBREVIATIONS

AI: Artificial Intelligence
AJCC: American Joint Committee on Cancer
GC: Gastric Cancer
MCC: Matthews Correlation Coefficient
ML: Machine Learning
OS: Overall Survival
PRC: Precision-Recall Curve
SEER: Surveillance, Epidemiology, and End Results
SVM: Support Vector Machine

Since the symptoms of the disease are atypical in the early period and very aggressive behavior in the later period, advanced screening, imaging, treatment, prognosis, and other technologies are required to prolong survival and reduce recurrence in gastric cancer [7].

Today, TNM classification is the most ideal prognostic marker for early-stage gastric cancer, but TNM classification may be insufficient to determine survival in advanced gastric cancers [8,9]. It is not uncommon for gastric cancer patients with the same clinical stage and the same treatment to have different clinical courses. Therefore, in order to accurately determine individual treatment options, there is a need to define prognostic markers in gastric cancer and to find additional auxiliary methods. At this point, artificial intelligence (AI) draws attention in the field of gastric cancer with its effective computing power and learning capacity. In the light of new technological developments used in the field of medicine, large-scale data of cancer patients can be collected and made available to researchers. Recently, with the help of large datasets and in-depth learning, researchers are able to better predict disease risk using medical data and machine learning (ML). ML can translate various metrics and classifications into relevant predictive models. Thanks to ML, diagnostic and drug geneticists can find the complexity of the disease, apply treatment, and customize medical options for each patient [10]. These techniques determine the behavior patterns of different types of cancers using historical data and can also effectively predict the outcome and survival of a particular cancer type [7].

In the light of these studies, we thought that demographic and clinicopathological features and survival times obtained with the large database we used could be effective in determining the prognosis of gastric cancer through ML and we have developed a software that can accurately detect overall survival in gastric cancer cases.

Material and Methods

The study included 34417 patients diagnosed with gastric cancer between 2010 and 2015, and all patient data were analyzed for the study. January 2010 was chosen as the starting point for the study. December 2015 was chosen as the end date of the study to show the effect of the developments in operation and treatment techniques after 2010. The data in the study were taken from the Surveillance, Epidemiology, and End Results (SEER) database. These data, published by the National Cancer Center Institute, are a compilation of databases of 18 SEER cancer registries in the USA. The SEER program is used to summarize data from patients' medical records, and it is estimated that more than 95% of all cancer cases are detected and included in this database in areas under surveillance [14]. The duration of follow-up is calculated in months using the date of diagnosis and whichever occurs first, 1) date of death, 2) date last known to be alive, 3) December 2015 (the follow-up cutoff date used in our analysis). Since all patient data were obtained with the permission of SEER without including personal patient information, there is no need to obtain ethical committee approval from any committee within the scope of this research.

The main hypothesis in the study was overall survival (OS) in years (censored observations), defined from the date of diagnosis to the date of death or, for living patients, the last control date. In addition to survival, other variables selected for the analyzes were age, race, gender, behavior, primary site, grade, histology, T stage, N stage, M stage and tumor size at the time of diagnosis, vital status, and follow-up time (months). Surgical methods, radiotherapy, and chemotherapy techniques were not included in the study because of missing data. Then, integrating these results and clinicopathologic features, we used multilayer perceptron, bagging, logistic regression, and Naive Bayes to develop a prognostic classifier.

In this study, in addition to the classical machine learning methods, we created a hybrid model consisting of a combination of existing methods. Such hybrid models have been preferred more in recent years, as they are a combination of machine learning methods and use the strongest aspects of these methods.

Statistical Analysis

SPSS 11.5 and Weka 3.7 programs were used in the analysis of the data. Mean±standard deviation and median (minimum-maximum) were used as descriptors for quantitative variables, and the number of patients (percentage) for qualitative variables. Chi-square test was used to examine the relationship between two qualitative variables. Survival analyzes on qualitative variables were performed using the Kaplan-Meier method, and significant differences between groups were determined using the log-rank test. The statistical significance level was taken as 0.05.

Classification methods of Logistic Regression, Naive

Bayes, AdaBoostM1, Bagging and J48 were used in the WEKA program. The data set was evaluated using the 10-fold Cross Validation test option. Accuracy, F-Measure, Precision, Matthews correlation coefficient (MCC), Precision-Recall Curve (PRC Area) and ROC Area were used as data mining performance criteria.

Results

General descriptors of the variables in the data set are given in Table 1. According to this, 10.7% of the patients were younger than 50 years old, 42.2% were in the 50-69 age range, 47.1% were 70 years old or older. While 71.2% of the patients were Caucasian, 13.9% were Black and 14.9% were from other races. In addition, 60.9% of the patients were male and 39.1% were female.

While 0.8% of the patients were in the in situ period, 99.2% of them were detected in the invasive period. The table shows the primary site, grade, histological behavior, and TNM stages of the tumor. Tumor sizes of the patients are also grouped in the table, and vital status and follow-up periods of the patients are also given (Table I).

Table II shows the survival analysis results of the patients. The median overall survival of the patients was found to be 15.00±0.20 years. In addition, all variables in the table were found to be statistically significant risk factors for survival. Median life expectancy was found to be 21.00±0.85 years for those younger than 50 years of age, 20.00±0.43 years for those aged 50-69 years, and 10.00±0.22 years for those aged 70 and over. When evaluated in terms of race, the median life expectancy was 14.00±0.21 years for the Caucasian race, and 15.00±0.21 years and 18.00±0.62 years for the Black and other races, respectively. In the study, the median life expectancy of men was found to be lower than that of women, and the median life expectancy of patients with malignant behavioral type was found to be lower than those in situ.

When survival is evaluated in terms of primary site types, the lowest median survival time is found in the group classified as N/A, followed by cardia, pylorus, antrum, fundus, and corpus, respectively. Survival statistics for grade, histology, TNM stage, and tumor size are also given in Table II.

Gain II. Ratio Attribute Eval and Information Gain Attribute Eval attribute selection methods in WEKA were used. With using these methods, the importance of the variables and the values it added to the data set were examined. The variables, which were determined to be insignificant by the two methods and considered to be not important as clinical information, were excluded from the data set. A total of 11 variables (10 independent variables and 1 dependent variable) remained in the data set. These variables are M Stage, Histology, Behavior, Grade, N stage, T stage, Age, Tumor Size, Primary Site, Gender and Status. Percentages of variable importance according to dependent variable Status was given in (Fig. 1).

Table I - *Descriptives.*

Variables		
Age, n (%)	<50	3671 (10.7)
	50-69	14534 (42.2)
	70	16212 (47.1)
Race, n (%)	Caucasian	24325 (71.2)
	Black	4738 (13.9)
	Other	5096 (14.9)
Gender, n (%)	Male	20944 (60.9)
	Female	13473 (39.1)
Behavior, n (%)	In situ	292 (0.8)
	Malignant	34125 (99.2)
Primary Site, n (%)	N/A	12362 (35.9)
	Antrum	5644 (16.4)
	Cardia	10032 (29.2)
	Corpus	3902 (11.3)
	Fundus	1732 (5.0)
	Pylorus	745 (2.2)
Grade, n (%)	Grade 1	2760 (10.6)
	Grade 2	6887 (26.6)
	Grade 3	15633 (60.3)
	Grade 4	645 (2.5)
Histology, n (%)	Adenomatous	29416 (85.5)
	Squamous	1502 (4.4)
	Soft Tissue	2580 (7.5)
	Other	919 (2.6)
T Stage, n (%)	T0	66 (0.3)
	T1a	2499 (12.4)
	T1b	1694 (8.4)
	T2	3508 (17.4)
	T3	7283 (36.0)
	T4a	2369 (11.7)
	T4b	2329 (11.5)
	Tis	458 (2.3)
N Stage, n (%)	N0	17149 (58.3)
	N1	7536 (25.6)
	N2	2441 (8.3)
	N3	2304 (7.8)
M Stage, n (%)	M0	23349 (68.1)
	M1	10957 (31.9)
Tumor Size, n (%)	<1 cm	1918 (9.8)
	1-3 cm	4514 (23.2)
	3-5 cm	6623 (33.9)
	>5 cm	6466 (33.1)
Vital Status, n (%)	Alive	12197 (35.4)
	Dead	22220 (64.6)
Follow-up Time (months)	Mean±SD	18.45±17.80
	Median (Min.-Max.)	13.00 (1.00-60.00)

SD: Standard Deviation, Min.: Minimum, Max.: Maximum

TABLE II - *Kaplan-Meier analysis results.*

Variables		Survival			Duration of life		
		1 year (%)	3 years (%)	5 years (%)	Average±SD	Median±SD	p value
General		53.9	33.3	26.2	25.66±0.14	15.00±0.20	-
Age	<50	60.8	39.0	32.6	29.67±0.43	21.00±0.85	<0.001
	50-69	60.6	39.0	32.2	29.19±0.22	20.00±0.43	
	70	46.1	26.9	19.1	21.57±0.19	10.00±0.22	
Race	Caucasian	52.8	32.2	25.0	25.03±0.16	14.00±0.21	<0.001
	Black	53.5	33.5	26.0	25.66±0.37	15.00±0.54	
	Other	57.9	37.1	30.3	27.86±0.37	18.00±0.62	
Gender	Male	52.5	30.9	23.0	24.42±0.17	14.00±0.21	<0.001
	Female	55.9	37.2	31.3	27.61±0.23	17.00±0.40	
Behavior	In situ	82.6	69.8	57.9	44.62±1.37	-	<0.001
	Malignant	53.6	33.0	25.9	25.49±0.14	15.00±0.19	
Primary Site	N/A	49.6	32.4	26.5	24.56±0.24	12.00±0.29	<0.001
	Antrum	55.4	35.2	27.3	26.52±0.34	16.00±0.56	
	Cardia	52.8	27.6	19.7	23.38±0.24	14.00±0.26	
	Corpus	64.1	44.4	35.9	31.70±0.42	27.00±1.20	
	Fundus	61.8	43.7	37.6	30.95±0.64	24.00±1.86	
	Pylorus	53.1	31.8	25.3	24.76±0.93	14.00±1.41	
Grade	1	85.4	74.3	66.4	47.09±0.43	-	<0.001
	2	61.1	38.6	28.4	28.80±0.31	20.00±0.60	
	3	46.0	22.5	16.5	20.20±0.18	11.00±0.18	
	4	39.5	18.9	10.9	18.19±1.02	8.00±0.87	
Histology	Adenomatous	52.1	30.4	23.3	24.24±0.15	14.00±0.18	<0.001
	Squamous	35.3	19.5	15.3	16.81±0.58	6.00±0.44	
	Soft Tissue	90.0	79.0	68.7	49.64±0.40	-	
	Other	28.0	16.4	12.9	14.08±0.88	3.00±0.30	
T Stage	T0	90.7	80.2	70.2	51.14±1.37	-	<0.001
	T1a	68.8	53.5	44.4	35.91±0.53	45.00±2.94	
	T1b	80.5	64.6	52.4	42.26±0.59	-	
	T2	74.4	55.4	45.5	37.92±0.43	55.00±-	
	T3	65.8	37.0	27.5	29.28±0.29	21.00±0.48	
	T4a	52.9	22.4	13.3	21.51±0.29	21.00±0.48	
	T4b	28.9	8.3	5.0	11.65±0.33	6.00±0.23	
	Tis	88.2	77.6	66.5	48.58±1.00	-	
N Stage	N0	62.5	45.8	37.6	31.90±0.20	27.00±0.77	<0.001
	N1	46.9	22.0	15.9	25.87±0.47	18.00±0.57	
	N2	61.8	30.0	19.2	20.30±0.26	11.00±0.23	
	N3	53.2	19.7	11.9	20.53±0.43	13.00±0.37	
M Stage	M0	67.3	46.0	36.7	32.86±0.17	29.00±0.58	<0.001
	M1	25.7	6.6	3.7	10.46±0.14	5.00±0.10	
Tumor Size	<1 cm	86.1	76.8	68.6	48.16±0.51	-	<0.001
	1-3 cm	75.4	55.3	45.0	37.94±0.38	51.00±2.08	
	3-5 cm	61.1	37.5	28.9	28.60±0.31	20.00±0.52	
	>5 cm	59.0	35.2	26.3	27.22±0.31	17.00±0.45	

SD: Standard Deviation

The performance criteria of Machine Learning methods for the 5-year survival prediction model are given in (Table III). Looking at the machine learning results, the Hybrid Model gave best results according to Accuracy, F-measure and MCC performance criteria, which are the most accepted performance criteria in the literature. Considering these three performance criteria, the Hybrid model is followed by Multilayer Perceptron, Logistic Regression, Naive Bayes, Bagging and J48, respectively. Also in the study, decision support system software was made based on the Hybrid Model, which is the method that gives the best results, and the outputs of this software are given in (Fig. 2).

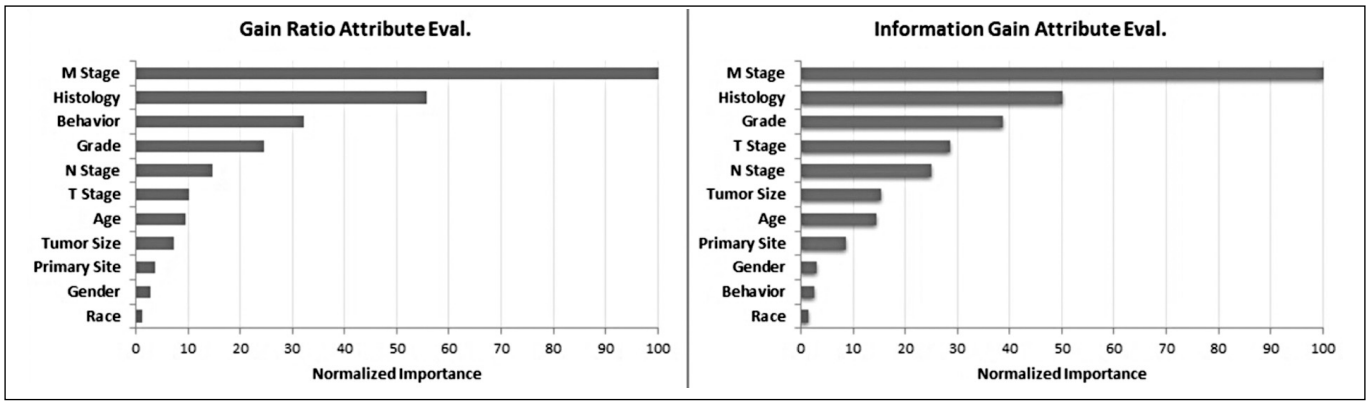

Fig. 1: Variable importance according to Status variable.

Fig. 2: Outputs of the Software Created Based on the Hybrid Model.

Discussion

Despite a wide variet y of treatment options in gastric cancer, recurrence rates are between 14% and 60% [11,12]. Gastric cancer is a very heterogeneous disease clinically and clinical outcomes are very variable even in patients at the same stage, however, prognosis is very important to determine appropriate treatment alternatives [13].

In our study, we first analyzed the demographic characteristics of gastric cancer cases in our database.

We found that the results we obtained were compatible with the literature in terms of age, race, and gender [14]. Also, the analyzes of our clinicopathological data were similar with the literature [15].

Increasing and shared data, increasing computational power, and advances in machine learning (ML) are transforming healthcare. Demographic, pathological, and physiological characteristics and even social relationships have an impact on the prognosis of gastric cancer patients. However, conventional methods such as TNM classification may fail to analyze the complicated relationships between these characteristics. Today, the 8th edition of TNM Malignant Tumors classification (TNM), developed by the American Joint Committee on Cancer (AJCC), is used for gastric cancer staging. Graziosi et al. compared the 8th edition recommended by AJCC with the previous 7th edition and suggested that the results for stage 3a are not clear, although it improved the survival differences between stages 3b and 3c [16]. In a study conducted in another large center, it

TABLE III - *Performance results of Machine Learning methods for 5-year survival.*

Methods		Performance Criteria					
		Accuracy	F-measure	Precision	MCC	PRC Area	ROC Area
Logistic Regression	Live	0.571	0.640	0.728	0.484	0.739	0.828
	Dead	0.883	0.833	0.789	0.484	0.885	0.828
	Overall	0.772	0.765	0.767	0.484	0.833	0.828
Naive Bayes	Live	0.588	0.637	0.695	0.466	0.713	0.804
	Dead	0.859	0.824	0.791	0.466	0.863	0.804
	Overall	0.763	0.763	0.757	0.466	0.810	0.804
Multilayer Perceptron	Live	0.608	0.654	0.709	0.490	0.742	0.828
	Dead	0.863	0.831	0.800	0.490	0.880	0.828
	Overall	0.773	0.768	0.768	0.490	0.831	0.828
Bagging	Live	0.546	0.619	0.713	0.458	0.722	0.815
	Dead	0.880	0.826	0.779	0.458	0.876	0.815
	Overall	0.761	0.753	0.756	0.458	0.821	0.815
J48	Live	0.532	0.610	0.715	0.451	0.690	0.800
	Dead	0.884	0.826	0.775	0.451	0.858	0.800
	Overall	0.759	0.749	0.753	0.451	0.799	0.800
Hybrid Model	Live	0.766	0.807	0.853	0.712	0.803	0.847
	Dead	0.927	0.902	0.878	0.712	0.918	0.847
	Overall	0.870	0.868	0.869	0.712	0.877	0.847

MCC: Matthews correlation coefficient, PRC: DOIPrecision-Recall Curve

was suggested that in cases where less than 30 lymph nodes were removed, the 8th edition was not a good prognostic marker in stage 3 gastric cancer and could not provide an improvement compared to the 7th edition [17].

Advanced computation and integration ability of artificial intelligence is used to increase the survival of gastric cancer patients. In recent years, artificial intelligence has been used in the prediction of survival , recurrence risk, and metastasis [18-26]. Jiang et al applied support vector machine (SVM) in their survival analysis and developed a classification that predicts prognosis. As a result, it was revealed that this classification has higher predictive power than the TNM classification in predictions of overall survival and disease-free survival. Lu et al [19] created a novel multimodal hypergraph learning framework by combining demographic data, pathological markers, and physiological characteristics of 939 cases to improve the accuracy of survival prediction. Results of that study showed that this proposed new method was more accurate in estimating overall survival than the classical method and SVM.

Thanks to the development of machine learning techniques, there has been an improvement of 15-20% in the prediction of treatment outcomes and survival in cancer patients [26]. There are machine learning methods used not only for survival but also for risk assessment [27,28].

To develop this survival prediction model, we retrieved 34417 patients' data which includes age, race, gender, tumor behavior, primary site, grade, histology, T stage, N stage, M stage, tumor size, vital status, and follow-up time (months) information from the database. Surgical methods and radiotherapy and chemotherapy techniques could not be included in the analysis due to missing data. The difficulties we encountered during this study are that the TNM staging has changed frequently in the last 2 decades and some of the data has not been fully entered by the centers. Therefore, we had to exclude some patient data from the computation. In addition, since the treatment models were not fully notated, we could not analyze the relationship between treatment methods and survival and use them in our software.

Then, by analyzing the relationship between these variables and survival time through machine learning, we developed a decision support system software that predicts 5-year overall survival.

For example, according to the decision support system, a male patient younger than 50 years who has a Grade 1: T0N0M0 in situ gastric soft tissue sarcoma in the fundus; has an 83.8% 5-year survival rate. In another example, a male patient between 50-69 years old who has a Grade 4: T4bN3M1 malignant gastric adenocarcinoma in the fundus has a 9.3% 5-year survival rate.

Considering its computational power and learning ability, it is predicted that artificial intelligence will play a role in many areas related to gastric cancer. Analyzing this mixed data is very difficult for clinicians. An AI model analyzes this huge amount of mixed data, sig-

nificantly reducing the clinician's workload. However, due to some safety and ethical concerns, the predictions made by AI need to be developed and reviewed by professional clinicians. Therefore, artificial intelligence techniques cannot completely replace clinicians in the future, but the partnership of human and artificial intelligence will provide ideal usefulness [29].

A flexible AI model needs a large amount of well-annotated data for training, validation, and testing, while small samples are likely to cause measurement errors [30]. With the advancement of medical screening modalities such as endoscopy and pathology, there has been a wealth of data streams that can assist clinicians in clinical diagnosis and decision making. Access to high-quality data is essential for the development and improvement of artificial intelligence. To access these quality data sets, large-scale open databases are needed. In addition, existing data sources should be used effectively.

We believe that the future of medicine cannot be separated from technology and artificial intelligence. The program we created must be tested, validated, and made clinically available with further research and more up-to-date data to be added. The use of these algorithms, which are not yet routinely used in daily practice, should be increased and encouraged, primarily for the benefit of patients and, of course, clinicians.

Conclusion

Especially artificial intelligence techniques such as machine learning and deep learning lead to remarkable developments in the field of gastric cancer. Many studies have also revealed that the performance of artificial intelligence is more reliable than standard statistical methods. Despite the proven success of artificial intelligence in the field of medicine in terms of diagnosis and prognosis, there are still obstacles to be overcome before it can be used in clinical practice. Among these, the scarcity of well-organized and labeled data can be cited. However, with the ability to compute and learn; we think that use of artificial intelligence will be revolutionary in gastric cancer in terms of diagnosis and prognosis.

Riassunto

Il cancro gastrico è la quarta causa più frequente di decessi per cancro, con un tasso di sopravvivenza a 5 anni inferiore al 40%. Negli ultimi anni, molte applicazioni di intelligenza artificiale sono state utilizzate nel campo del cancro gastrico grazie alla loro efficace capacità di elaborazione e apprendimento. In questo studio, miriamo a sviluppare un software in grado di rilevare con precisione la sopravvivenza globale nei casi di cancro gastrico con l'aiuto dell'intelligenza artificiale e dell'apprendimento automatico.

METODI: Lo studio ha incluso 34417 dati di pazienti con diagnosi di cancro gastrico tra il 2010 e il 2015. L'ipotesi principale nello studio era la sopravvivenza globale (OS) in anni, definita dalla data di diagnosi alla data di morte o, per i pazienti vivi, l'ultima data di controllo. Oltre alla sopravvivenza, altre variabili selezionate per le analisi erano l'età alla diagnosi, la razza, il sesso, il comportamento, la sede primaria, il grado, l'istologia, lo stadio T, lo stadio N, lo stadio M e le dimensioni del tumore, lo stato vitale e il follow-up, tempo (mesi).

RISULTATI: la sopravvivenza globale mediana dei pazienti è risultata essere di 15,00±0,20 anni. L'aspettativa di vita mediana è risultata essere di 21,00 ± 0,85 anni per quelli di età inferiore ai 50 anni, 20,00 ± 0,43 anni per quelli di età compresa tra 50 e 69 anni e 10,00 ± 0,22 anni per quelli di età pari o superiore a 70 anni. Soprattutto le tecniche di intelligenza artificiale come l'apprendimento automatico e il deep learning portano a notevoli sviluppi nel campo del cancro gastrico.

CONCLUSIONE: Con la capacità di calcolare e apprendere, pensiamo che l'uso dell'intelligenza artificiale sarà rivoluzionario nel cancro gastrico in termini di diagnosi e prognosi.

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 [published correction appears in CA Cancer.* J Clin, 2020; 70(4):313, CA Cancer, J Clin, 2018; 68(6):394-24, doi: 10.3322/caac.21492.

2. Fitzmaurice C, Abate D, Abbasi N, et al: *Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2017: A systematic analysis for the global burden of disease study.* JAMA Oncology, 2019; 5(12):1749-768, doi: 10.1001/jamaoncol.2019.2996.

3. Ajani J. A, Lee J, Sano T, Janjigian YY, Fan D, Song S: *Gastric adenocarcinoma.* Nature Reviews Disease Primers, 2017; 3(1), doi: 10.1038/nrdp.2017.36.17036.

4. Wagner AD, Syn NL, Moehler M, et al: *Chemotherapy for advanced gastric cancer*. The Cochrane Database of Systematic Reviews, 2017; 8, doi: 10.1002/14651858.CD004064.pub4.Cd 004064.

5. Japanese Gastric Cancer Association: *Japanese gastric cancer treatment guidelines 2014 (ver. 4).* Gastric Cancer, 2017; 20(1):1-19, doi: 10.1007/s10120-016-0622-4.

6. Jemal A, Ward EM, Johnson CJ, et al: *Annual report to the nation on the status of cancer, 1975-2014, featuring survival.* J Natl Cancer Inst, 2017; 109(9):djx030, doi: 10.1093/jnci/djx030.

7. Hanahan D, Weinberg RA: *Hallmarks of cancer: The next generation*. Cell, 2011; 144(5):646-74, doi: 10.1016/j.cell.2011.02.013.

8. Noh SH, Park SR, Yang HK, et al: *Adjuvant capecitabine plus oxaliplatin for gastric cancer after D2 gastrectomy (Classic): 5-: year follow-up of an open-label, randomized phase 3 trial.* Lancet Oncol, 2014; 15(12):1389-96, doi: 10.1016/S1470-2045(14)70473-5.

9. Nishida T: *Adjuvant therapy for gastric cancer after D2 gastrectomy*. Lancet, 2012; 379(9813):291-92, doi: 10.1016/S0140-6736(11)61928-4.

10. Zhou C, Wang Y, Ji MH, Tong J, Yang JJ, Xia H: *Predicting peritoneal metastasis of gastric cancer patients based on machine learning*. Cancer Control, 2020; 27(1):1073274820968900, doi: 10.1177/1073274820968900.

11. National Cancer Institute, SEER*Stat software, version 8.3.6.1 (seer.cancer.gov/seerstat): *Surveillance research program.*

12. Seyfried F, von Rahden BH, Miras AD, et al: *Incidence, time course and independent risk factors for metachronous peritoneal carcinomatosis of gastric origin: A longitudinal experience from a prospectively collected database of 1108 patients*. BMC Cancer, 2015; 15:73, Published, 2015; 19, doi: 10.1186/s12885-015-1081-8.

13. Razzak M: *Genetics: New molecular classification of gastric adenocarcinoma proposed by the cancer genome atlas*. Nat Rev Clin Oncol, 2014; 11(9):499, doi: 10.1038/nrclinonc.2014.138.

14. Karimi P, Islami F, Anandasabapathy S, Freedman ND, Kamangar F: *Gastric cancer: Descriptive epidemiology, risk factors, screening, and prevention*. Cancer Epidemiol Biomarkers Prev, 2014; 23(5):700-13, doi: 10.1158/1055-9965, EPI-13-1057.

15. Terry MB, Gaudet MM, Gammon MD: *The epidemiology of gastric cancer*. Semin Radiat Oncol, 2002; 12(2):111-27, doi: 10.1053/srao.30814.

16. Graziosi L, Marino E, Donini A: *Survival comparison in gastric cancer patients between 7th and 8th edition of the AJCC TNM staging system: The first western single-center experience*. Eur J Surg Oncol, 2019; 45(6):1105-108, doi: 10.1016/j.ejso.2018.12.010.

17. Lu J, Zheng CH, Cao LL, et al: *Validation of the american joint commission on cancer (8th edition) changes for patients with stage III gastric cancer: Survival analysis of a large series from a specialized eastern* center. Cancer Med, 2017; 6(10):2179-87, doi: 10.1002/cam4.1118.

18. Jiang Y, Xie J, Han Z, et al: *Immunomarker support vector machine classifier for prediction of gastric cancer survival and adjuvant chemotherapeutic benefit*. Clin Cancer Res, 2018; 24(22):5574-84, doi: 10.1158/1078-0432.CCR-18-0848.

19. Lu F, Chen ZK, Yuan X, Li Q, Du ZD, Luo L, Zhang FY: *MMHG: Multi-modal Hypergraph Learning for Overall Survival After D2 Gastrectomy for Gastric Cancer. DASC/PiCom/DataCom/CyberSci Tech 2017: Proceedings of the 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing, and Cyber Science and Technology Congress. 2017; Nov 6-10, Orlando FL, USA, California.* IEEE Computer Society, 2017; 164-9.

20. Korhani Kangi A, Bahrampour: *Predicting the survival of gastric cancer patients using: artificial and bayesian neural networks.* Asian Pac J Cancer Prev, 2018; 19(2):487-90, Published, 2018; 26, doi: 10.22034/APJCP.2018.19.2.487.

21. Zhang W, Fang M, Dong D, et al: *Development and validation of a CT-based radiomic nomogram for preoperative prediction of early recurrence in advanced gastric cancer.* Radiother Oncol, 2020; 145:13-20, doi: 10.1016/j.radonc.2019.11.023.

22. Liu B, Tan J, Wang X, Liu X: *Identification of recurrent risk-related genes and establishment of support vector machine prediction model for gastric cancer*. Neoplasma, 2018; 65(3):360-366, doi: 10.4149/neo_2018_170507N326.

23. Bollschweiler EH, Mönig SP, Hensler K, Baldus SE, Maruyama K, Hölscher AH: *Artificial neural network for prediction of lymph node metastases in gastric cancer: A phase II diagnostic study.* Ann Surg Oncol, 2004; 11(5):506-511, doi: 10.1245/ASO.2004.04.018.

24. Hensler K, Waschulzik T, Mönig SP, Maruyama K, Hölscher AH, Bollschweiler E: *Quality-assured efficient engineering of feedforward neural networks (QUEEN): Pretherapeutic estimation of lymph node status in patients with gastric carcinoma*. Methods Inf Med, 2005; 44(5):647-54.

25. Jagric T, Potrc S, Jagric T: *Prediction of liver metastases after gastric cancer resection with the use of learning vector quantization neural networks.* Dig Dis Sci, 2010; 55(11):3252-61, doi: 10.1007/s10620-010-1155-z.

26. Cruz JA, Wishart DS: *Applications of machine learning in cancer prediction and prognosis*. Cancer Inform, 2007; 2:59-77, Published, 2007; 11.

27. Exarchos KP, Goletsis Y, Fotiadis D: *Multiparametric decision support system for the prediction of oral cancer reoccurrence*. IEEE Trans Inf Technol Biomed, 2012; 16(6):1127-34, doi: 10.1109/TITB.2011.2165076.

28. Park K, Ali A, Kim D, An Y, Kim M, Shin H: *Robust predictive model for evaluating breast cancer survivability*. Engl Appl Artif Intell, 2013; 26:2194-205.

29. Niu PH, Zhao LL, Wu HL, Zhao DB, Chen YT: *Artificial intelligence in gastric cancer: Application and future perspectives*. World J Gastroenterol, 2020; 26(36):5408-419, doi: 10.3748/wjg.v26.i36.5408.

30. Loken E, Gelman A: *Measurement error and the replication crisis*. Science, 2017; 355(6325):584-85, doi: 10.1126/science.aal3618.