



# Prediction of Response to Neoadjuvant Chemoradiotherapy with Machine Learning in Rectal Cancer: A Pilot Study

Melek YAKAR,<sup>1</sup> Durmuş ETİZ,<sup>1</sup> Bartu BADAĞ,<sup>2</sup> Özer ÇELİK,<sup>3</sup> Deniz KÜTRİ,<sup>1</sup> Alaattin ÖZEN,<sup>1</sup>  
 Evrim YILMAZ<sup>4</sup>

<sup>1</sup>Department of Radiation Oncology, Eskisehir Osmangazi University Faculty of Medicine, Eskisehir,-Turkey

<sup>2</sup>Department of General Surgery, Eskisehir Osmangazi University Faculty of Medicine, Eskisehir-Turkey

<sup>3</sup>Department of Mathematics and Computer, Eskisehir Osmangazi University Faculty of Arts and Sciences, Eskisehir-Turkey

<sup>4</sup>Department of Medical Pathology, Eskisehir Osmangazi University Faculty of Medicine, Eskisehir-Turkey

## OBJECTIVE

In locally advanced rectal cancer, trimodality therapy comprising chemoradiotherapy, total mesorectal excision, and chemotherapy (CT) are accepted as standard treatment. However, standard “one-size-fits-all” therapy based on the TNM staging system may not be suitable for every patient. In cases with a good response, less invasive surgical treatments, such as sphincter-sparing local excision or the watch-and-wait approach may be more appropriate due to their lower recurrence rates. Therefore, it is very important to predict these cases and plan treatment accordingly to ensure effective personalized treatment. Machine learning can successfully predict these cases. Aim: The aim of the study was to predict the response to neoadjuvant chemoradiotherapy with machine learning in locally advanced rectal cancer.

## METHODS

The study included 125 rectal cancer cases who underwent neoadjuvant radiotherapy (RT)±CT between 2010 and 2020, and the cases with a good response (grade 0-1) according to the Modified Ryan classification were predicted using machine learning. A total of 26 variables were evaluated. After determining key variables, the dataset was divided into training/test sets at 80%/20%. Logistic regression, artificial neural network-multilayer perceptron classifier, XGBoost, support vector classification, random forest, and Gaussian Naive Bayes algorithms used to establish a prediction model. In the prediction of the group with a good response, 173 cases were created and evaluated with the synthetic minority oversampling technique method.

## RESULTS

Of the 125 cases, 15 had a complete response and 33 had a good response (Modified Ryan grades 0 and 1). Six algorithms were tested in terms of their ability to predict a good response. Key variables for this prediction were found to be tumor localization, RT break time, age, gender, Karnofsky Performance Scale score, body mass index, pre- and post-treatment carcinoembryonic antigen levels, pre-treatment hemoglobin and neutrophil-to-lymphocyte ratio and platelet-to-lymphocyte ratio, radiological T and N stages, perineural and lymphatic invasion, tumor grade, radiological metastatic lymph node region, RT dose and technique, and presence and scheme of concurrent CT. The algorithm that showed the best performance was determined as logistic regression with an accuracy rate of 84% (CI: 0.69-0.98), sensitivity of 83%, and specificity of 85%.

Received: July 15, 2021

Accepted: August 03, 2021

Online: September 16, 2021

Accessible online at:

www.onkder.org

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



Dr. Melek YAKAR

Eskisehir Osmangazi Üniversitesi Tıp Fakültesi,

Radyasyon Onkolojisi,

Eskişehir-Turkey

E-mail: mcakcay@ogu.edu.tr

## CONCLUSION

It is very important to predict the cases with a good response and plan treatment accordingly to ensure effective personalized treatment. Machine learning can successfully predict these cases.

**Keywords:** Artificial intelligence; machine learning; neoadjuvant chemoradiotherapy; prediction of treatment response; rectal cancer.

Copyright © 2021, Turkish Society for Radiation Oncology

## Introduction

In locally advanced rectal cancer, trimodality therapy comprising chemoradiotherapy (CRT), total mesorectal excision (TME), and chemotherapy (CT) are accepted as standard treatment.[1] Conventionally, the treatment algorithm is established according to the clinical and pathological TNM staging system.[1] However, standard “one-size-fits-all” therapy based on the TNM staging system may not be suitable for every patient. Treatment response and overall survival may not be similar in groups of patients receiving the same treatment at the same stage. Identifying patients at high risk of recurrence and disease-related death will also be valuable in guiding treatment. Therefore, in this complex and heterogeneous disease group, it is important to evaluate the prognosis in a personalized manner and plan the treatment accordingly.

Artificial intelligence (AI) is a branch of computer science that tries to imitate human-like intelligence in machines using computer software and algorithms without direct human stimuli to perform certain tasks. [2,3] Machine learning is a subunit of AI that utilizes data-driven algorithms that learn to imitate human behavior based on previous examples or experience.[4] Deep learning (DL) is an ML technique that uses deep neural networks (NNs) to create a model. The growth and sharing of data, increases in computing power, and developments in AI have initiated a transformation in oncology. Advances in radiation oncology and accumulated big data of an increasing number of cases have resulted in the production of a significant amount of data. There are a number of individual differences that are responsible for each patient’s disease or associated with their response to treatment and clinical outcome. The concept of personalized treatment is based on determining and using these factors for each patient. [5] The integration of such a large and heterogeneous amount of data and creation of accurate models may sometimes present with certain difficulties for the human brain and involve subjective individual differences. With machine learning, appropriate algorithms

can be created and the most suitable personalized treatment for each patient can be determined at the initial stage of treatment.

In locally advanced rectal cancer, neoadjuvant CRT (n-CRT) improves local control, disease-free survival, and sphincter preservation rates.[6] However, after n-CRT, tumor regression patterns have a wide spectrum ranging from pathological complete response (pCR) to disease progression. Although cases with pCR have the best survival and tumor control, pCR is achieved using n-CRT in only 10-30% of cases with locally advanced rectal cancer.[7] Some studies have shown that cases with pCR have low recurrence rates, and therefore less invasive alternative surgical treatments, such as sphincter-sparing local excision and the watch-and-wait approach may be more appropriate for this patient population.[8-11] For this reason, identifying patients that are likely to achieve a complete or almost complete response is very important to effectively personalize treatment, refer selected cases directly to surgery without waiting, and prevent unnecessary excessive treatment/toxicity.

This study aimed to predict response to n-CRT among 125 cases who underwent this treatment at the Department of Radiation Oncology of Eskisehir Osmangazi University Faculty of Medicine between 2010 and 2020.

## Materials and Methods

### Patients Characteristics

Between 2010 and 2020, 127 rectal cancer cases who underwent n-CRT at the Department of Radiation Oncology of Eskisehir Osmangazi University were retrospectively evaluated. The study included patients with a histopathological diagnosis of rectal cancer, stage T3-4N0-2 or T1-4N1-2, and a Karnofsky Performance Scale (KPS) score of  $\geq 70$ . Staging was performed with pelvic magnetic resonance imaging (MRI) and Positron emission tomography-computed tomography. TNM staging was undertaken according to the American Joint Committee on Cancer staging system,

eighth edition.[12] After the diagnosis, all the patients were evaluated in the Oncology Council of ESOGUTE, and the treatment decision was made in a multidisciplinary manner.

### Treatment Characteristics

Radiotherapy (RT) was applied to all cases as neoadjuvant. Considering the patient's KPS score, age, and comorbidities, an evaluation for concomitant CT was made and the CT scheme was determined. As concomitant CT, continuous 5-FU (5-Fluorouracil 225 mg/m<sup>2</sup>) or capecitabine (825 mg/m<sup>2</sup> 5 days a week for 5 weeks during RT) was used. During the treatment, the cases were evaluated at least once a week in the outpatient clinic based on complete blood count and blood biochemistry and examination findings. A close follow-up of toxicity and weight was undertaken.

All the patients were immobilized in the supine position, with their arms up. Computed tomography was performed with the Somatom Definition AS<sup>®</sup> Device with a 5 mm slice interval. Pelvic MRI fusion was used to contour gross tumor volume (GTV<sub>t</sub>). In the presence of pathological lymph nodes, MRI and positron emission tomography-computed tomography fusion were used for GTV<sub>n</sub>. For the clinical target volume (CTV)<sub>high</sub>, the area of margin was selected as 2 cm for GTV<sub>t</sub> and 0.5 cm for GTV<sub>n</sub>, and mesorectum and presacral areas were included. CTV<sub>standard</sub> was obtained by adding elective nodal areas to CTV<sub>high</sub>. Regional lymphatics included obturator, presacral, internal iliac, and external iliac lymph nodes (for T4 tumors only). The target volume (PTV) margin planned according to the RT technique was determined as 0.5-1 cm. The bladder, small intestines, and femoral heads were contoured as organs at risk (OARs).

The RT dose was planned as 45-54 (1.8-2 Gy/day). RT was applied with the Varian Trilogy<sup>®</sup>/TrueBeam<sup>®</sup>/Elekta Precise<sup>®</sup> device accompanied by three-dimensional (3D) conformal RT and volumetric modulated arc therapy (VMAT). 3D conformal RT was applied to 73 cases and VMAT technique was applied to 52 cases.

### Evaluation of Treatment Response

At 4 to 6 weeks after treatment, response to n-CRT was evaluated with pelvic MRI, surgery was planned 6 to 12 weeks later. The modified Ryan classification was used for response evaluation after n-CRT.[13]

### ML

In this study, for the evaluation of response to n-CRT, logistic regression, artificial NN (ANN)-multilayer perceptron (MLP) classifier, XGBoost, support vector

classification (SVC), random forest (RF), and gaussian naive bayes (GNB) algorithms were used.

MLP is a feed-forward class of ANN. The term MLP is sometimes loosely used to refer to any feed-forward ANN and can also refer to specific networks consisting of more than one sensor layer (with threshold activation). Multilayer sensors are called "vanilla" NNs in the spoken language, especially when they have a single hidden layer.[14] MLP consists of at least three node layers: An input layer, a hidden layer, and an output layer.

XGBoost is an optimized distributed gradient boosting library designed as a highly efficient, flexible and portable tool. XGB provides parallel tree enhancement (also known as GBDT and GBM) that quickly and accurately solves many data science problems. The same code runs in a large distributed environment (e.g., Hadoop, Sun Grid Engine, and Message Passing Interface) and can solve problems beyond billions of examples. The most important features of the algorithm are its high predictive power, its ability to prevent over-learning and manage empty data, and perform these tasks in a quick manner.[15]

SVC involves clustering the data set according to some criteria to organize the data in a more meaningful way. There are many ways to achieve this goal. Clustering can proceed by performing grouping according to a certain parametric model or based on a measure of distance or similarity, as in hierarchical clustering. A natural way of setting cluster boundaries is the use of "valleys" in regions where there is very little data in the data area; i.e., in the probability distribution of the data.[16]

RFC method has multiple estimation trees and combines each tree to depend on the randomly selected vector value equally distributed among all the trees in the forest. Thus, in RFC, a random  $\theta_k$  vector independent of the previous random vectors and distributed across all trees is selected, and each tree is grown using a training set and a random  $\theta_k$  vector, resulting in an ensemble of trees.[17]

GNB classifier is one of the top 10 algorithms in data mining. GNB is a useful classifier widely used in many applications, such as text categorization, spam filtering, and data flow classification. Bayesian classifiers operate based on the Bayes rule and probability theorems.[18] A density distribution is drawn for the Gauss model of each class. The line shows the decision limit corresponding to the curve where a new point has an equal probability of being part of each class.

Logistic regression model simply uses a logistic function to model a binary dependent variable, but it is fundamentally much more complex. In regression

analysis, logistic (or logit) regression estimates the parameters of a logistic model. This is a method of classifying the relationship between multiple independent variables and dependent variables. In logistic regression analysis, the probability of a dependent variable is estimated with two values. In addition, the variables in the model are continuous, which makes this technique favorable for use in classifying observations.[19]

A total of 26 variables were evaluated: Age, gender, KPS score, history of comorbidities, body mass index (BMI), tumor grade, radiological T, N, and TNM stages, radiological metastatic lymph node region, tumor localization (cm), tumor localization (lower/middle/upper), pre-treatment carcinoembryonic antigen (CEA) level, post-treatment CEA level, pre-treatment hemoglobin, neutrophil-to-lymphocyte ratio (NLR) and platelet-to-lymphocyte ratio (PLR), RT dose (Gy), RT technique, RT break time, presence and scheme of concomitant CT, time from RT to surgery, presence of lymphatic, vascular, and perineural invasion. Key variables were selected by the permutation-based feature selection method, which is based on the evaluation of the significance of each feature separately. This technique measures changes in prediction quality (based on the coefficient of decrease in determination score) after processing in a single feature vector. The rate of decrease in the coefficient of determination shows how important a feature is.[20]

The dataset was divided into training and test sets at 80% and 20%, respectively. Models were created using the training set and verified using the test set. The optimal model was selected according to the receiver operating characteristic curves. Cross validation is a model validation technique that tests what result a statistical analysis will yield in an independent dataset. The main use of this technique is to predict the accuracy of a predictive system in practice. In a prediction problem, the model is usually trained with a “known dataset” (training set) and tested with an unknown dataset (verification or test set), which is also known as supervised learning. The purpose of this test is to measure the generalizability of the trained model to new data and identify problems of overfitting or selection bias.[21] In the current study, five-fold cross validation was performed.

The dataset contained 15 cases evaluated to have a complete response and 33 cases with a good response (modified Ryan classification grades 0 and 1). In an unbalanced dataset, the model predicts in favor of the group with a higher number of samples, which results in overfitting. In statistics, overfitting refers to a pro-

duced analysis aligning too closely to a certain dataset (memorization), which leads to the inability to adapt to new data that are not included in this dataset, and to eliminate the possibility of this problem, a balanced dataset should be used.[22] In the synthetic minority oversampling technique (SMOTE),[23] the class type with an unbalanced data distribution is artificially replicated, and thus balancing is achieved. In the current study, for the evaluation of cases with a complete response, the samples belonging to the minority class type was multiplied using SMOTE. The dataset was divided into training (80%) and test (20%) sets, with 100 cases being included in the training group (26 with and 74 without good response) and 25 (7 with and 18 without good response) in the test group. The training group was oversampled with SMOTE, resulting in 148 cases, of which 74 did not have a good response and 74 had a good response.

Statistical analysis was performed and ML algorithms were run using Python software (Python Software Foundation. Python Language Reference, version 3.5. Available from <http://www.python.org>) and Scikit-Learn library. All analyses and operations were undertaken on a computer with the specifications of Intel Core i7-9750 CPU 2.6 GHz 12MB Cache and 16GB 2666 MHz DDR4 RAM memory running 64-bit Windows 10 operating system.

## Results

### Patient, Tumor, and Treatment Characteristics

For the prediction of cases with a good response, 125 cases were used and SMOTE was applied. The median age was 61 (min: 23, max: 85) years, and the male/female ratio was 86/39. Patient and tumor characteristics are given in Table 1. The median RT dose was 50.4 (min: 45, max: 54) Gy. The median RT break time was 2 (min: 0, max: 18) days. Concomitant CT was performed in 106 cases and adjuvant CT in 110 cases. Treatment characteristics are given in Table 2.

### Treatment Response

The modified Ryan classification was used for response prediction, and the number of cases with grades 0, 1, 2, and 3 was 15 (12.0%), 18 (14.4%), 80 (64.0%), and 12 (9.6%), respectively.

### Results of ML

Of the 26 variables, ten were determined as important using the permutation-based feature selection method: Tumor localization, RT break time, age, gender, KPS

Characteristic	Number of patients (%) / (min-max)
Age	61 (23-85)
Gender	
Male	86 (68.8)
Female	39 (31.2)
KPS score	90 (70-100)
BMI (kg/m <sup>2</sup> )	27(15-40)
Tumor grade	
Grad 1	12 (9.6)
Grad 2	106 (84.8)
Grad 3	7 (5.6)
Radiological T stage	
rT2	4 (3.2)
rT3	105 (84.0)
rT4	16 (12.8)
rN stage	
rN0	57 (45.6)
rN1	53 (42.4)
rN2	15 (12.0)
TNM stage	
IIA/IIIB/IIC	45 (36.0)/4(3.2)/6 (4.8)
IIIA/IIIB/IIIC	3 (2.4)/60 (48.0)/3 (2.4)
IVA	4 (3.2)
Tumor localization	
Upper/middle/lower	16 (12.8)/38 (30.4)/84 (67.2)
Pre- /post-treatment CEA	4.0 (0.8-313.0)/2.6 (0.4-78.7)
NLR	5 (0.8-31)
PLR	303 (41-964)
Lymphatic invasion	
Present/absent	17 (13.6)/108 (86.4)
Vascular invasion	
Present/absent	17 (13.6)/108 (86.4)
Perineural invasion	
Present/absent	27 (21.6)/98 (78.4)
Pathological tumor size (cm)	2 (0-6.5)
Lymph nodes removed during surgery, n	14 (2-40)
Metastatic lymph nodes removed during surgery, n	0 (0-12)

KPS: Karnofsky performance scale; BMI: Body mass index; r: Radiological; CEA: Carcinoembryonic antigen; NLR: Neutrophil-to-lymphocyte ratio; PLR: Platelet-to-lymphocyte ratio; TNM: Tumor, lymph node and metastasis staging

score, BMI, pre- and post-treatment CEA levels, pre-treatment hemoglobin, NLR and PLR values, radiological T and N stages, perineural and lymphatic invasion, tumor grade, radiological metastatic lymph node area, RT dose and technique, and presence and scheme of concomitant CT. The feature importance graph is given in Figure 1a. The algorithm that showed the best performance was determined as logistic regression with an

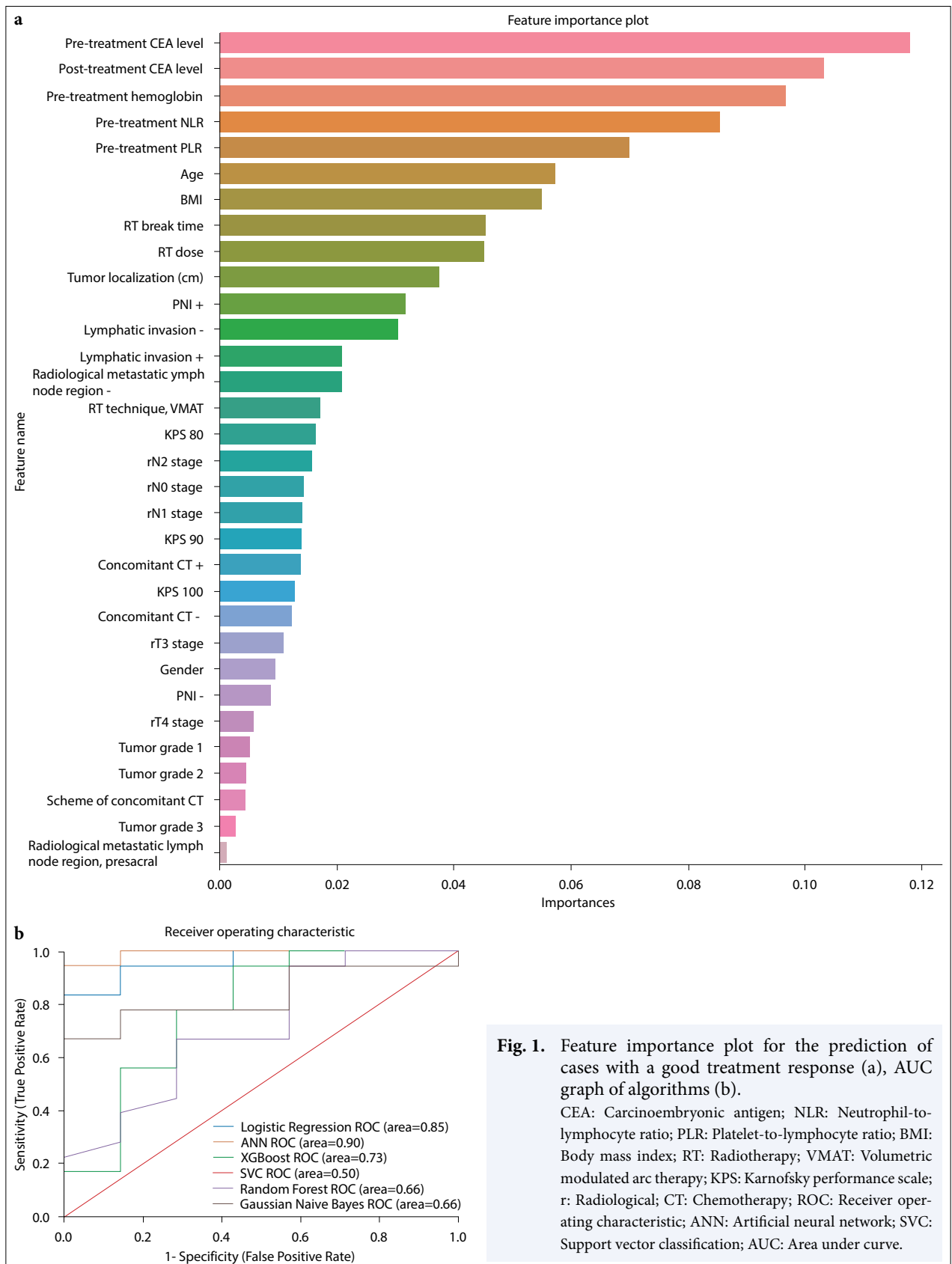
Characteristic	Number of patients (%) / (min-max)
RT dose (Gy)	50.4 (25-54)
RT break time (days)	2 (0-18)
Concomitant CT	
Present	106 (84.8)
Absent	19 (15.2)
Concomitant CT scheme	
Capecitabine	84 (66.1)
Continuous 5-FU	22 (17.3)
Time from RT to surgery (days)	61 (27-70)
Resection	
R <sub>0</sub>	119 (95.2)
R <sub>1</sub>	4 (3.2)
R <sub>2</sub>	2 (1.6)
Adjuvant CT	
Present	110 (88.0)
Absent	15 (12.0)

Gy: Gray; CT: Chemotherapy; 5-FU: 5-Fluorouracil; RT: Radiotherapy

accuracy rate of 84% (confidence interval: 0.69-0.98) and area under the curve (AUC) value of 0.84. After SMOTE was applied to the dataset of 125 cases, 148 cases were used for training and 25 for testing. In the test stage, the logistic regression algorithm accurately predicted six of the seven cases with a good response, and 15 of the 18 cases without a good response. The AUC graph of the algorithms is given in Figure 1b. The accuracy rates of the algorithms are given in Table 3, and the confusion matrix of the logistic regression algorithm is presented in Table 4.

## Discussion

n-CRT followed by TME is currently the standard treatment for locally advanced rectal cancer. Following n-CRT, approximately 10-30% of patients achieve pCR.[7] For these patients no longer having invasive cancer, the question is whether they need TME because this invasive surgical method is associated with significant complications and morbidity.[8,24] Several studies have shown that patients with pCR have low local recurrence rates, and thus less invasive, alternative surgical therapies, such as sphincter-sparing local excision or the watch-and-wait approach are gaining more popularity among these patients.[8] There is a need to confirm pCR by surgery, and if these cases can be predicted in advance, less aggressive surgery can be performed. Similarly, less invasive surgical methods



**Fig. 1.** Feature importance plot for the prediction of cases with a good treatment response (a), AUC graph of algorithms (b).

CEA: Carcinoembryonic antigen; NLR: Neutrophil-to-lymphocyte ratio; PLR: Platelet-to-lymphocyte ratio; BMI: Body mass index; RT: Radiotherapy; VMAT: Volumetric modulated arc therapy; KPS: Karnofsky performance scale; r: Radiological; CT: Chemotherapy; ROC: Receiver operating characteristic; ANN: Artificial neural network; SVC: Support vector classification; AUC: Area under curve.

**Table 3** Algorithms used in the prediction of complete or good response

Algorithm performance						
Algorithm	LR	ANN	XGB	SVC	RF	GNB
Accuracy	0.84	0.72	0.60	0.72	0.72	0.68
ROC AUC	0.84	0.76	0.54	0.76	0.63	0.47
Precision	0.93	0.92	0.75	0.92	0.78	0.70
Sensitivity/recall	0.83	0.66	0.66	0.66	0.83	0.94
Specificity	0.85	0.85	0.42	0.85	0.42	0.30
Confidence interval	0.69-0.98	0.54-0.89	0.40-0.79	0.54-0.89	0.54-0.89	0.49-0.68

LR: Logistic regression; ANN: Artificial neural network; XGB: XGBoost; SVC: Support vector classification; RF: Random forest; GNB: Gaussian naive Bayes; ROC: Receiver operating characteristic; AUC: Area under the curve

**Table 4** Confusion matrix of the best-performing algorithm

Logistic regression	Actual class	
	Good response (+)	Good response (-)
Predicted class		
Good response (+)	6	1
Good response (-)	3	15

can be used in cases that have a good response to treatment. This possible prediction can significantly reduce health-care costs in the treatment of rectal cancer.

With the early detection and prediction of their response to treatment and following a personalized treatment approach, patients can be divided into different prognostic groups. Of the patients with locally advanced rectal cancer that has undergone n-CRT and surgery, 45% will require permanent colostomy. Identifying those with a complete or good clinical response and good response before surgery will allow for the optimization of the surgical approach with “organ-sparing” procedures, resulting in a reduction in surgical morbidity. In addition, among patients diagnosed with locally advanced rectal cancer, the early detection of those with a poor response to n-CRT will offer the opportunity to directly move on to surgery, thus avoiding morbidities associated with n-CRT or intensified treatment regimen.[25]

In a study conducted with 696 patients with Stage I-III rectal cancer, Cai et al.[26] determined tumor size and pre-treatment CEA to be poor prognostic factors. Bacha et al.[27] evaluating 44 patients with locally advanced rectal cancer, accepted age as a factor affecting response to treatment. In the current study, the CEA level and age were accepted as key variables.

In a study conducted with 248 patients diagnosed with locally advanced rectal cancer who underwent n-CRT, Huang et al.[28] performed pCR prediction using patient and treatment characteristics. The authors obtained the highest accuracy rate from the ANN algorithm at 88%. Key variables were accepted as post-treatment CEA value, time from RT to surgery, CT scheme, and clinical N and T stages. In the current study, post-treatment CEA level, time from RT to surgery, CT scheme, radiological N and T stages were found to be key variables, and the best-performing algorithm was determined as logistic regression with an accuracy rate of 84% in predicting a good response.

Imaging methods have also been used for pCR prediction.[19,29-32] Shayesteh et al.[25] included 98 cases diagnosed with rectal cancer in their sample and performed MRI 1 week before CRT to extract radiomics such as density, shape, and tissue features. The authors used 53 cases for training and 45 for validation. They used the SVM, BN, NN, and K nearest neighbor (KNN) algorithms both individually and together to evaluate their ability to predict response to n-CRT using AUC. When the algorithms were evaluated separately, the best result was obtained from the BN algorithm with the AUC value and accuracy rate of 0.75 and 80.9%, respectively. When the algorithms (SVM, NN, BN, and KNN) were evaluated together, the AUC and accuracy values were 0.97 and 92.8%, respectively. The authors suggested that the prediction process could be improved when algorithms were used as hybrid. In a study including 95 patients diagnosed with T2-4N0-1 rectal cancer, radiomics were obtained from the computed tomography images taken before CRT (1683 radiomic features per case) together with clinical and treatment data, and response prediction was made with AI.[32] In the creation of prediction models, the deep NN

(DNN) and SVM algorithms were combined with radiomics while only TNM staging was added to linear regression (LR). pCR was achieved in a total of 23 cases. The accuracy rates of the DNN, SVM, and LR algorithms were reported as 80.0%, 71.5%, and 69.5%, respectively.

In the literature, there are very few studies that predict good response to n-CRT based on patient, tumor and treatment characteristics, and such prediction evaluations have mostly been undertaken using imaging methods and radiomics. However, response rates are also related to patient and treatment characteristics. In the current study, a prediction model was created using not only patient and treatment but also tumor characteristics. An accurate classification of cases with a good response could help determine less invasive therapeutic strategies, such as sphincter-sparing surgery, mucosectomy, or the wait-watch approach. In addition, the prediction of cases that do not respond to n-CRT would allow for these patients to be referred to more effective treatments and thus significantly reduce unnecessary health expenses.

The limitations of the study are the small number of cases, the inclusion of metastatic (single liver metastasis) cases and being a single-center study. Prediction software obtained in such studies has not yet entered into routine treatment use, and it is not clear which health authorities can give their ethical approval. The strengths of the study are the inclusion of patient, tumor, and treatment characteristics in the algorithm. In addition, this study is important in terms of forming the basis for the decisions to be taken about the patient in the next oncology councils.

## Conclusion

In recent years, the increasing interest in AI in all fields of science has also led to the development of innovative tools in oncology. The development of prediction tools with a wide variety of variables and models help plan personalized treatments. Using such prediction models, rectal cancer groups that will respond well to n-CRT can be identified to use less invasive methods while surgical treatments can be applied to cases predicted to be unresponsive to n-CRT to improve their oncological outcomes.

**Peer-review:** Externally peer-reviewed.

**Conflict of Interest:** All authors declared no conflict of interest.

**Ethics Committee Approval:** The study was approved by the Eskisehir Osmangazi University Non-Invasive Clinical Research Ethics Committee (No: 25, Date: 30/03/2021).

**Financial Support:** None declared.

**Authorship contributions:** Concept – D.E., M.Y., Ö.Ç., B.B.; Design – D.E., M.Y., Ö.Ç., B.B.; Supervision – D.E., M.Y., A.Ö., B.B., E.Y.; Funding – D.E., M.Y., A.Ö., B.B., E.Y.; Materials – D.E., M.Y., B.B., D.K., E.Y.; Data collection and/or processing – D.E., M.Y., Ö.Ç., D.K., E.Y.; Data analysis and/or interpretation – D.E., M.Y., Ö.Ç., D.K.; Literature search – D.E., M.Y., B.B., D.K.; Writing – D.E., M.Y.; Critical review – D.E., M.Y., Ö.Ç., A.Ö., E.Y.

## References

1. National Comprehensive Cancer Network. NCCN Guidelines. Rectal Cancer, Version 1. Available at: National Comprehensive Cancer Network; 2021. Available from: <http://www.nccn.org>. Accessed Sep 1, 2021.
2. Meyer P, Noblet V, Mazzara C, Lallement A. Survey on deep learning for radiotherapy. *Comput Biol Med* 2018;98:126–46.
3. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521(7553):436–44.
4. Jarrett D, Stride E, Vallis K, Gooding MJ. Applications and limitations of machine learning in radiation oncology. *Br J Radiol* 2019;92(1100):20190001.
5. Bibault JE, Giraud P, Burgun A. Big data and machine learning in radiation oncology: State of the art and future prospects. *Cancer Lett* 2016;382(1):110–7.
6. Sauer R, Liersch T, Merkel S, Fietkau R, Hohenberger W, Hess C, et al. Preoperative versus postoperative chemoradiotherapy for locally advanced rectal cancer: Results of the German CAO/ARO/AIO-94 randomized phase III trial after a median follow-up of 11 years. *J Clin Oncol* 2012;30(16):1926–33.
7. Gérard JP, Azria D, Gourgou-Bourgade S, Martel-Laffay I, Hennequin C, Etienne PL, et al. Comparison of two neoadjuvant chemoradiotherapy regimens for locally advanced rectal cancer: Results of the phase III trial ACCORD 12/0405-prodige 2. *J Clin Oncol* 2010;28(10):1638–44.
8. Borschitz T, Wachtlin D, Möhler M, Schmidberger H, Junginger T. Neoadjuvant chemoradiation and local excision for T2-3 rectal cancer. *Ann Surg Oncol* 2008;15(3):712–20.
9. Renehan AG, Malcomson L, Emsley R, Gollins S, Maw A, Myint AS, et al. Watch-and-wait approach versus surgical resection after chemoradiotherapy for patients with rectal cancer (the OnCoRe project): A propensity-score matched cohort analysis. *Lancet Oncol* 2016;17(2):174–83.



10. Ludwig KA. Sphincter-sparing resection for rectal cancer. *Clin Colon Rectal Surg* 2007;20(3):203–12.
11. Marijnen CA. Organ preservation in rectal cancer: Have all questions been answered? *Lancet Oncol* 2015;16(1):e13–22.
12. Weiser MR. AJCC 8<sup>th</sup> edition: Colorectal cancer. *Ann Surg Oncol* 2018;25(6):1454–5.
13. Ryan R, Gibbons D, Hyland JM, Treanor D, White A, Mulcahy HE, et al. Pathological response following long-course neoadjuvant chemoradiotherapy for locally advanced rectal cancer. *Histopathology* 2005;47(2):141–6.
14. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer; 2009.
15. Chen T, Guestrin C. Xgboost: A Scalable Tree Boosting System. In: *KDD 16: Proceedings of the 22<sup>nd</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; 2016. p. 785–94.
16. Ben-Hur A, Horn D, Siegelmann HT, Vapnik V. Support vector clustering. *J Mac Learn Res* 2001;2:125–37.
17. Masetic Z, Subasi A. Congestive heart failure detection using random forest classifier. *Comput Methods Programs Biomed* 2016;130:54–64.
18. Jahromi AH, Taheri M. A non-parametric Mixture of Gaussian Naive Bayes Classifiers Based on Local Independent Features. In: *2017 Artificial Intelligence and Signal Processing Conference (AISP)*. IEEE; 2017. p. 209–12.
19. Celik O, Osmanoglu UO. Comparing to techniques used in customer churn analysis. *J Multidiscip Res Dev* 2019;4(1):30–8.
20. Galkin F, Aliper A, Putin E, Kuznetsov I, Gladyshev VN, Zhavoronkov A. Human microbiome aging clocks based on deep learning and tandem of permutation feature importance and accumulated local effects. *EBioMedicine* 2019;45:1–2.
21. Kohavi R. A Study of Cross-validation and Bootstrap for Accuracy Estimation and Model Selection. *Proceedings of the 14<sup>th</sup> International Joint Conference on Artificial Intelligence*. San Mateo, CA: Morgan Kaufmann; 1995. p. 1137–43.
22. Sun Y, Kamel MS, Wong AK, Wang Y. Cost-sensitive boosting for classification of imbalanced data. *Pattern Recognit* 2007;40(1):3358–78.
23. Han H, Wang WY, Mao BH. Borderline-SMOTE: A New Over-sampling Method in Imbalanced Data Sets Learning. In: Huang DS, Zhang XP, Huang GB, editors. *Advances in Intelligent Computing ICIC 2005*. Lecture Notes in Computer Science; Berlin: Springer, 2014. p. 878–87.
24. Sanghera P, Wong DW, McConkey CC, Geh JI, Hartley A. Chemoradiotherapy for rectal cancer: An updated analysis of factors affecting pathological response. *Clin Oncol (R Coll Radiol)* 2008;20(2):176–83.
25. Shayesteh SP, Alikhassi A, Esfahani AF, Miraie M, Geramifar P, Bitarafan-Rajabi A, et al. Neo-adjuvant chemoradiotherapy response prediction using MRI based ensemble learning method in rectal cancer patients. *Phys Med* 2019;62:111–9.
26. Cai D, Huang ZH, Yu HC, Wang XL, Bai LL, Tang GN, et al. Prognostic value of preoperative carcinoembryonic antigen/tumor size in rectal cancer. *World J Gastroenterol* 2019;25(33):4945–58.
27. Bacha D, Talbi G, Ben Slama S, Lahmar A, Gharbi L, Bouraoui S. Predictive factors for histological response to neo-adjuvant treatment in rectal cancers. *Tunis Med* 2019;97(11):1284–90.
28. Huang CM, Huang MY, Huang CW, Tsai HL, Su WC, Chang WC, et al. Machine learning for predicting pathological complete response in patients with locally advanced rectal cancer after neoadjuvant chemoradiotherapy. *Sci Rep* 2020;10:12555.
29. Shi L, Zhang Y, Nie K, Sun X, Niu T, Yue N, et al. Machine learning for prediction of chemoradiation therapy response in rectal cancer using pre-treatment and mid-radiation multi-parametric MRI. *Magn Reson Imaging* 2019;61:33–40.
30. Nie K, Shi L, Chen Q, Hu X, Jabbour SK, Yue N, et al. Rectal cancer: Assessment of neoadjuvant chemoradiation outcome based on radiomics of multiparametric MRI. *Clin Cancer Res* 2016;22(21):5256–64.
31. Fu J, Zhong X, Li N, Van Dams R, Lewis J, Sung K, et al. Deep learning-based radiomic features for improving neoadjuvant chemoradiation response prediction in locally advanced rectal cancer. *Phys Med Biol* 2020;65(7):075001.
32. Bibault JE, Giraud P, Housset M, Durdux C, Taieb J, Berger A, et al. Deep learning and radiomics predict complete response after neo-adjuvant chemoradiation for locally advanced rectal cancer. *Sci Rep* 2018;8(1):16914.