



OPEN

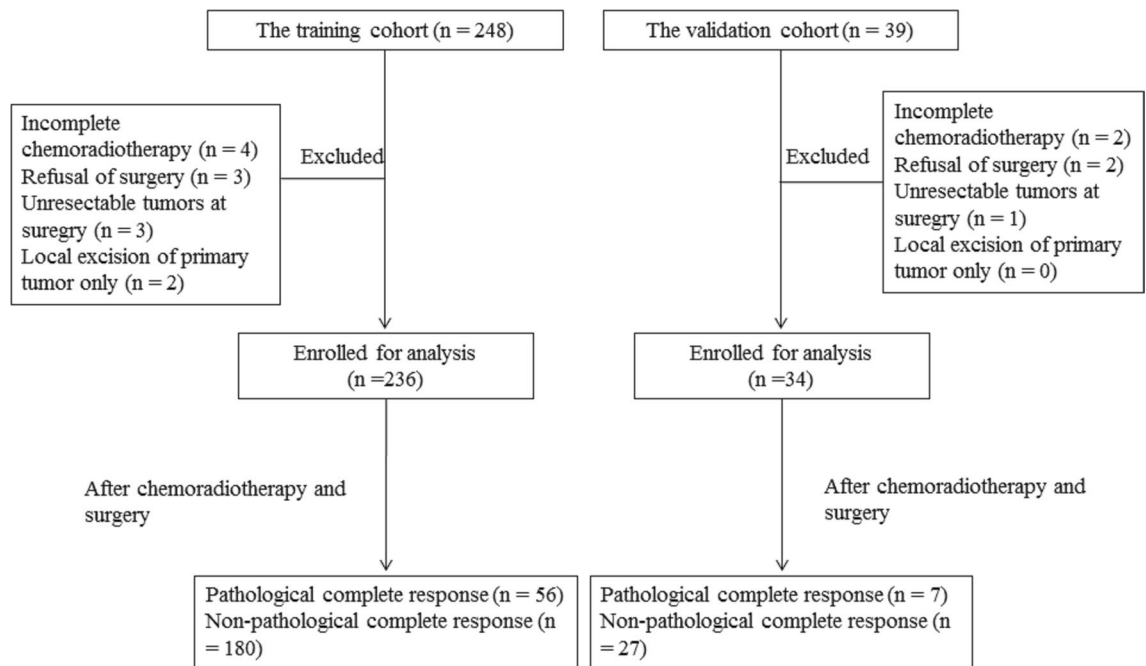
# Machine learning for predicting pathological complete response in patients with locally advanced rectal cancer after neoadjuvant chemoradiotherapy

Chun-Ming Huang<sup>1,2,3,4</sup>, Ming-Yii Huang<sup>1,2,3,5</sup>, Ching-Wen Huang<sup>6,7</sup>, Hsiang-Lin Tsai<sup>6,7</sup>, Wei-Chih Su<sup>6,7</sup>, Wei-Chiao Chang<sup>8,9</sup>, Jaw-Yuan Wang<sup>4,5,6,7,8,9</sup>✉ & Hon-Yi Shi<sup>10,11,12,13</sup>✉

For patients with locally advanced rectal cancer (LARC), achieving a pathological complete response (pCR) after neoadjuvant chemoradiotherapy (CRT) provides them with the optimal prognosis. However, no reliable prediction model is presently available. We evaluated the performance of an artificial neural network (ANN) model in pCR prediction in patients with LARC. Predictive accuracy was compared between the ANN, *k*-nearest neighbor (KNN), support vector machine (SVM), naïve Bayes classifier (NBC), and multiple logistic regression (MLR) models. Data from two hundred seventy patients with LARC were used to compare the efficacy of the forecasting models. We trained the model with an estimation data set and evaluated model performance with a validation data set. The ANN model significantly outperformed the KNN, SVM, NBC, and MLR models in pCR prediction. Our results revealed that the post-CRT carcinoembryonic antigen is the most influential pCR predictor, followed by intervals between CRT and surgery, chemotherapy regimens, clinical nodal stage, and clinical tumor stage. The ANN model was a more accurate pCR predictor than other conventional prediction models. The predictors of pCR can be used to identify which patients with LARC can benefit from watch-and-wait approaches.

Neoadjuvant chemoradiotherapy (CRT) has benefited patients with locally advanced rectal cancer (LARC) with specific respect to improvements in local control, disease-free survival, and sphincter preservation rates<sup>1–3</sup>. However, the patterns of tumor regression after neoadjuvant CRT vary widely, ranging from a pathological complete response (pCR) to disease progression. Patients with a pCR have the most favorable survival and tumor control, but only 10–30% of patients with LARC achieve a pCR to neoadjuvant CRT<sup>4–7</sup>. Furthermore, mounting evidence has demonstrated that in patients who achieve a pCR, radical surgery can be omitted without compromising

<sup>1</sup>Department of Radiation Oncology, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan. <sup>2</sup>Department of Radiation Oncology, Faculty of Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan. <sup>3</sup>Department of Radiation Oncology, Kaohsiung Municipal Ta-Tung Hospital, Kaohsiung Medical University, Kaohsiung, Taiwan. <sup>4</sup>Graduate Institute of Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan. <sup>5</sup>Center for Cancer Research, Kaohsiung Medical University, Kaohsiung, Taiwan. <sup>6</sup>Division of Colorectal Surgery, Department of Surgery, Kaohsiung Medical University Hospital, Kaohsiung Medical University, No. 100, Tzyou 1st Road, Kaohsiung 807, Taiwan. <sup>7</sup>Department of Surgery, Faculty of Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan. <sup>8</sup>Graduate Institute of Clinical Medicine, College of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan. <sup>9</sup>School of Pharmacy, Taipei Medical University, Taipei, Taiwan. <sup>10</sup>Department of Healthcare Administration and Medical Informatics, Kaohsiung Medical University, 100, Shih-Chuan 1st Road, 80708 Kaohsiung, Taiwan. <sup>11</sup>Department of Business Management, National Sun Yat-Sen University, Kaohsiung, Taiwan. <sup>12</sup>Department of Medical Research, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan. <sup>13</sup>Department of Medical Research, China Medical University Hospital, China Medical University, Taichung, Taiwan. ✉email: cy614112@ms14.hinet.net; Jawyuanwang@gmail.com; hshi@kmu.edu.tw



**Figure 1.** Flow chart of patient selection for the training and validation cohorts.

tumor control<sup>8,9</sup>. Therefore, the identification of useful predictors of a pCR in patients with LARC after neoadjuvant CRT is vital.

Few studies have compared the artificial neural network (ANN), *k*-nearest neighbor (KNN), support vector machine (SVM), naïve Bayes Classifier (NBC), and multiple logistic regression (MLR) models with respect to internal validity (reproducibility). Validity is a crucial performance metric<sup>10,11</sup>. However, numerous predictive models yield insufficiently reliable predictions of pCR occurrence in patients with LARC after neoadjuvant CRT.

One of the most frequently applied methods for multivariate analysis is regression analysis; in this type of analysis, linear correlations between dependent and independent variables are assumed. Studies have demonstrated that biomedical variables usually vary nonlinearly<sup>12–16</sup>. The KNN model is a simple classification algorithm with straightforward implementation<sup>14</sup>. The KNN model predicts new samples by using training samples; the process entails a majority vote on the outcome of points that are *k*-nearest to the new sample. The SVM model is a supervised learning model associated with learning algorithms that analyse information used for regression analysis and classification<sup>13</sup>. An SVM model constructs multidimensional hyperplanes that separate the 2 classes while maximizing the margin between the 2 classes; it uses kernel functions and can discriminate between nonlinearly separable classes. An NBC model can be used to efficiently develop classification tools for various health domains and transform complex clinical problems into clear, precise, and predictive models<sup>16</sup>. An ANN model has three layers: input, hidden, and output layers. Every layer has nodes connected by links from one layer to the next<sup>12,15</sup>. Nodes in the input layer represent predictors, whereas those in the output layer are considered outcome variables. A common application of neural networks is the multilayer backpropagation learning algorithm, which models nonlinear systems. Although the interpretation of neural networks is more complicated than that of other statistical models, the ANN model has been used in various medical fields.

Although considerable improvements in outcome prediction models have been achieved, pCR prediction models continue to have major limitations<sup>17,18</sup>. For example, many studies have identified effective pCR predictors, but most related variables have exhibited insufficient sensitivity and specificity<sup>19–21</sup>. Therefore, in our study, ANN, KNN, SVM, NBC, and MLR models were used to identify the most powerful predictors of pCR in patients with LARC after neoadjuvant CRT. Thus, the primary purpose of this study was to validate the accuracy of the ANN model for pCR prediction in patients with LARC following neoadjuvant CRT. The secondary purpose was to investigate the predictive performance of various forecasting models.

## Methods

**Patients.** This study identified patients with a LARC diagnosis who were undergoing neoadjuvant CRT at any period between January 2011 and December 2017 at Kaohsiung Medical University Hospital. In total, 248 consecutive patients satisfied the inclusion criteria, which were pathologically proven adenocarcinomas, tumors located within 12 cm of the anal verge, clinical stage II and III rectal tumors (T3 to 4 or N+), and the delivery of neoadjuvant CRT. We excluded twelve patients because they had incomplete neoadjuvant CRT (*n* = 4), rejection of resection (*n* = 3), unresectable tumors after CRT (*n* = 3), or only primary tumor excision (*n* = 2). The remaining 236 patients were enrolled for analysis as the training cohort. For the validation cohort, 34 patients with LARC were recruited at Kaohsiung Medical University Hospital between January 2018 and September 2018. The same inclusion and exclusion criteria were used for the training and validation cohorts (Fig. 1). Pretreatment clinical staging was determined through computed tomography (CT) of the abdomen and chest, pelvic magnetic

resonance imaging (MRI), and a physical examination. Participants' serum carcinoembryonic antigen (CEA) levels and routine laboratory test results were analyzed.

**Treatment.** All participants underwent neoadjuvant CRT. Radiotherapy was delivered from 45 to 50.4 Gy, 1.8 to 2.0 Gy per fraction. Three-dimensional conformal radiotherapy was administered to 45 patients, and intensity-modulated radiotherapy was administered to 191 patients. Chemotherapy was administered concurrently with radiotherapy. Participants underwent 1 of the following 2 chemotherapeutic regimens. The first was the fluoropyrimidine-based regimen ( $n=95$ ), which consisted of six courses of capecitabine (850 mg/m<sup>2</sup> twice daily for 14 days) followed by 7 days of rest after each course. The second was a biweekly schedule of FOLFOX, which included oxaliplatin (85 mg/m<sup>2</sup>) on day 1, in addition to folinic acid (400 mg/m<sup>2</sup>) and a 46-h infusion of 5-fluorouracil (2,800 mg/m<sup>2</sup>) repeated every 2 weeks during radiotherapy; patients continued to receive three to four cycles of consolidation chemotherapy with biweekly FOLFOX after completion of radiotherapy ( $n=141$ )<sup>7</sup>.

All patients in the current study underwent total mesorectal excision after completing neoadjuvant CRT. The surgical procedures included low anterior resection with colorectal or coloanal anastomosis ( $n=207$ ) and abdominoperineal resection ( $n=29$ ).

**Evaluation and follow-up.** Two experienced pathologists evaluated tumor responses to neoadjuvant CRT. A pCR was defined as the absence of malignant cells in primary tumors and nodes (ypT0N0) in the resected specimen following neoadjuvant CRT.

Acute side effects were assessed at each visit during neoadjuvant CRT according to the Common Terminology Criteria for Adverse Events, version 4.03. We defined anemia as a hemoglobin level of < 10 g/dL. Approximately 6–10 weeks after completing CRT, measurements were conducted before surgery, specifically through pelvic MRI, abdominal and chest CT, a CEA test, and colonoscopy. After treatment completion, patients were required to visit the hospital every 3 months during the initial 2 years and then once every 6 months.

**Statistical analysis.** In the current study, we used individual patients who received neoadjuvant CRT with subsequent surgery as the unit of analysis. First, we used univariate logistic regression to select significant risk factors related to pCR. In the forecasting models, the dependent variable was the probability of pCR, and the independent variables were the significant risk factors.

Second, the data set was randomly segmented into training and testing sets, comprising 70% and 30% of the whole data set, respectively. From a probabilistic perspective related to forecasting models, this randomisation was a form of statistical sampling (e.g. Monte Carlo sampling). We used the training set to construct the forecasting models. The independent variables fitted to the forecasting models were the significant risk factors, and the dependent variable was the outcome (pCR probability). Upon completing training, the forecasting model was exposed to the testing set, and the model outputs were calculated for each testing set. Additionally, for cross-validation, data from 34 new patients were used to construct the validation set for the prediction of pCR in patients with LARC after neoadjuvant CRT.

Third, the performance indices including sensitivity, 1-specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, and area under the receiver operating characteristic curve (AUROC) were employed to evaluate the accuracy of the models. Bootstrapping with 1,000 replications was also performed to further amplify the training, testing, and validation data sets to reduce variability in assessments of model performance.

Finally, a global sensitivity analysis was conducted to evaluate the relative significance of input variables in the prediction models; these variables were ranked by their importance. The network error ratio, the sum of squared residuals, represented the global sensitivity of the input variables against the output variables. In general, a variable sensitivity ratio (VSR) of  $\leq 1$  demonstrates that the variable decreased predictive performance and should be removed. STATISTICA 13.0 (StatSoft, Inc., Tulsa, OK, USA) was used for statistical analyses.

**Ethics approval statement.** The study protocol was established according to the ethical principles of the Declaration of Helsinki and was approved by the Institutional Review Board of Kaohsiung Medical University Hospital (KMUHIRB-E (II)-20190280). Each patient provided written informed consent.

## Results

**Patient characteristics.** Two hundred seventy patients with LARC were enrolled for analysis. The training and validation cohorts had 236 and 34 patients, respectively (Table 1). The median post-CRT CEA level was 2.2 ng/mL (range 0.48 to 197.5). Accordingly, the cut-off value of post-CRT CEA level was 2 ng/mL. In the training and validation cohorts, respectively 23.7% and 20.6% of patients achieved pCR following CRT ( $P=0.162$ ).

**Study characteristics.** Table 2 presents the training data set's pCR odds ratio (OR). The univariate analysis indicated that pCR occurrence in patients with LARC after neoadjuvant CRT was significantly associated with gender, age, tumor location, type of chemotherapy, clinical tumor stage, clinical nodal stage, tumor-node-metastasis stage, tumor grade, post-CRT CEA level, anemia, diarrhea, urinary symptoms, dermatitis, leukopenia, radiation therapy dose, and the radiation to surgery interval ( $P<0.01$ ). As a result, the significant variables were further analyzed in the forecasting models.

**Comparisons between these forecasting models.** The differences in patient attributes, clinical attributes, and pCR occurrence between the training and testing data sets were insignificant (data not shown). Con-

Variables	The training cohort mean $\pm$ SD/N (%)	The validation cohort mean $\pm$ SD/N (%)
Number of patients	236	34
<b>Patient attributes</b>		
Gender		
Female	82 (34.7)	12 (35.3)
Male	154 (65.3)	22 (64.7)
Age	62.1 $\pm$ 11.5	62.8 $\pm$ 12.2
<b>Clinical attributes</b>		
Chemotherapy		
Fluoropyrimidine	95 (40.3)	15 (44.1)
FOLFOX	141 (59.7)	19 (55.9)
Tumor location		
Low/middle	141 (59.7)	20 (58.8)
Upper	95 (40.3)	14 (41.2)
Clinical T stage		
T2	13 (5.5)	2 (5.9)
T3	184 (78)	27 (79.4)
T4	39 (16.5)	5 (14.7)
Clinical N stage		
N0	36 (15.3)	6 (17.7)
N1	145 (61.4)	20 (58.8)
N2	55 (23.3)	8 (23.5)
TNM stage		
II	36 (15.3)	6 (17.6)
III	200 (84.7)	28 (82.4)
Tumor grade		
Well differentiation	16 (6.8)	2 (5.8)
Moderate differentiation	212 (89.8)	31 (91.3)
Poor differentiation	8 (3.4)	1 (2.9)
Pre-CRT CEA (ng/mL)		
$\leq 5$	144 (61)	20 (58.8)
$> 5$	92 (39.0)	14 (41.2)
Anemia		
Hb (g/dL) $\leq 10$	76 (32.2)	10 (29.4)
Hb (g/dL) $> 10$	160 (67.8)	24 (70.6)
Diarrhea		
Grade 0–1	102 (43.2)	14 (41.2)
Grade 2–3	134 (56.8)	20 (58.8)
Urinary symptoms		
Grade 0–1	218 (92.4)	31 (91.2)
Grade 2–3	18 (7.6)	3 (8.8)
Dermatitis		
Grade 0–1	166 (70.3)	24 (70.6)
Grade 2–3	70 (29.7)	10 (29.4)
Leukopenia		
WBC $\leq 3,000$ (/uL)	65 (27.5)	11 (32.4)
WBC $> 3,000$ (/uL)	171 (72.5)	23 (67.6)
RT dose (cGy)		
5,040	11 (4.7)	1 (2.9)
5,000	181 (76.7)	27 (79.5)
4,500	44 (18.6)	6 (17.6)
RT-surgery interval		
$\leq 8$ weeks	81 (34.3)	14 (41.2)
$> 8$ weeks	155 (65.7)	20 (58.8)
Post-CRT CEA (ng/mL)		
$\leq 2$	90 (38.1)	13 (38.2)
$> 2$	146 (61.9)	21 (61.8)
Continued		

Variables	The training cohort mean $\pm$ SD/N (%)	The validation cohort mean $\pm$ SD/N (%)
Treatment response		
pCR	56 (23.7)	7 (20.6)
Non-pCR	180 (76.3)	27 (79.4)

**Table 1.** Patient characteristics. *CEA* carcinoembryonic antigen, *CRT* chemoradiotherapy, *FOLFOX* fluorouracil, leucovorin, and oxaliplatin, *Hb* hemoglobin, *pCR* pathological complete response, *SD* standard deviation, *RT* radiation therapy, *WBC* white blood cell.

Variables	OR	95% C.I	P-value
<b>Gender</b>			
Male vs. female	3.53	2.41–5.17	<0.001
<b>Age</b>			
	1.02	1.01–1.02	<0.001
<b>Chemotherapy</b>			
FOLFOX vs. fluoropyrimidine	2.53	1.75–3.64	<0.001
<b>Tumor location</b>			
Upper vs. low/middle	4.28	2.56–7.15	<0.001
<b>Clinical T stage</b>			
T2 vs. T3	2.92	2.09–4.06	<0.001
T2 vs. T4	6.80	2.66–17.4	<0.001
<b>Clinical N stage</b>			
N0 vs. N1	3.68	2.47–5.47	<0.001
N0 vs. N2	4.00	2.07–7.75	<0.001
<b>TNM stage</b>			
II vs. III	3.76	2.68–5.29	<0.001
<b>Tumor grade</b>			
Well differentiation vs. moderate differentiation	3.00	2.20–4.09	<0.001
Well differentiation vs. poor differentiation	3.68	2.40–6.97	<0.001
<b>Pre-CRT CEA (ng/mL)</b>			
$\leq 5$ vs. $>5$	4.75	2.77–8.14	<0.001
<b>Anemia</b>			
Grade 0–1 vs. grade 2–3	3.32	2.30–4.80	<0.001
<b>Diarrhea</b>			
Grade 0–1 vs. grade 2–3	2.62	1.80–3.83	<0.001
<b>Urinary symptoms</b>			
Grade 0–1 vs. grade 2–3	8.00	1.84–34.79	0.006
<b>Dermatitis</b>			
Grade 0–1 vs. grade 2–3	3.67	2.07–6.49	<0.001
<b>Leukopenia</b>			
Grade 0–1 vs. grade 2–3	2.89	2.05–4.07	<0.001
<b>RT-dose (cGy)</b>			
5,000 vs. 4,500	2.69	1.94–3.74	<0.001
5,040 vs. 4,500	7.80	3.07–19.79	<0.001
<b>RT-surgery interval</b>			
$>8$ wk vs. $\leq 8$ wk	2.44	1.73–3.46	<0.001
<b>Post-CRT CEA (ng/mL)</b>			
$\leq 2$ vs. $>2$	1.58	0.86–2.88	<0.001

**Table 2.** The univariate analysis of logistic regression model using selected risk factors related to pathological complete response (N = 236). *CEA* carcinoembryonic antigen, *CI* confidence interval, *CRT* chemoradiotherapy, *FOLFOX* fluorouracil, leucovorin, and oxaliplatin, *Hb* hemoglobin, *OR* odds ratio, *pCR* pathological complete response, *RT* radiation therapy, *WBC* white blood cell.

sequently, samples from these two data sets could be compared to improve the reliability of the validation data sets. ANN-based approaches provide three-layer networks and the relative weights of neurons used for pCR prediction. The ANN 16-4-1 model contains 16 input neurons, 4 hidden neurons, 1 bias neuron in the hidden

	Sensitivity	1-Specificity	PPV	NPV	Accuracy	AUROC
<b>Training dataset (n = 165)</b>						
ANN	0.93	0.84	0.87	0.90	0.87	0.79
KNN	0.81	0.64	0.86	0.64	0.78	0.72
SVM	0.91	0.57	0.85	0.57	0.64	0.73
NBC	0.91	0.49	0.75	0.87	0.75	0.50
MLR	0.90	0.47	0.83	0.39	0.80	0.79
<b>Testing dataset (n = 71)</b>						
ANN	0.94	0.87	0.89	0.88	0.86	0.81
KNN	0.89	0.49	0.87	0.46	0.84	0.72
SVM	0.90	0.82	0.85	0.71	0.85	0.74
NBC	0.90	0.85	0.82	0.75	0.78	0.51
MLR	0.84	0.61	0.88	0.69	0.85	0.77

**Table 3.** Comparison of 1,000 pairs of prediction models for predicting pathological complete response. *ANN* artificial neural network, *KNN* K nearest neighbor, *SVM* support vector machines, *NBC* Naive Bayes classifier, *MLR* multiple logistic regression, *PPV* positive predictive value, *NPV* negative predictive value, *AUROC* area under the receiver operating characteristic.

	Sensitivity	1-Specificity	PPV	NPV	Accuracy	AUROC
ANN	0.94	0.80	0.89	0.87	0.88	0.84
KNN	0.80	0.67	0.87	0.60	0.80	0.74
SVM	0.91	0.76	0.86	0.72	0.71	0.76
NBC	0.90	0.53	0.79	0.84	0.80	0.63
MLR	0.88	0.79	0.84	0.49	0.83	0.77

**Table 4.** Comparative performance indices of prediction models when using 34 new validation datasets to predict pathological complete response. *ANN* artificial neural network, *KNN* K nearest neighbor, *SVM* support vector machines, *NBC* Naive Bayes classifier, *MLR* multiple logistic regression, *PPV* positive predictive value, *NPV* negative predictive value, *AUROC* area under the receiver operating characteristic.

	Rank 1st	Rank 2nd	Rank 3rd	Rank 4th	Rank 5th
Variables	Post-CRT CEA	RT-surgery interval	Chemotherapy regimen	Clinical N stage	Clinical T stage
VSR	1.57	1.50	1.45	1.37	1.32

**Table 5.** Global sensitivity analysis of the ANN model in predicting pathological complete response. *ANN* artificial neural network, *CEA* carcinoembryonic antigen, *CRT* chemoradiotherapy, *RT* radiation therapy, *VSR* variable sensitivity ratio.

layer, and 1 output neuron. Table 3 presents comparisons between the training and testing data sets indicating that the ANN model outperformed the KNN, SVM, NBC, and MLR models with respect to sensitivity, 1-specificity, PPV, NPV, accuracy, and AUROC. For cross-validation, data from 34 newly enrolled patients were used to construct the validation data set for pCR prediction; the ANN model remained the most accurate (Table 4).

**Significant predictors in the ANN model.** We used the training data sets to compute the VSR for the ANN model. The global sensitivity analysis demonstrated that the most sensitive variable for predicting pCR occurrence in patients with LARC after neoadjuvant CRT was post-CRT CEA levels (VSR = 1.57), followed by intervals between radiation and surgery (VSR = 1.50), types of chemotherapy (VSR = 1.45), clinical nodal stages (VSR = 1.37), and clinical tumor stages (VSR = 1.32) (Table 5). All VSR values in the current study exceeded 1, indicating that the network operated better when we considered all variables.

## Discussion

We used performance indices to compare the forecasting models with respect to their accuracy in predicting pCR occurrence in patients with LARC after neoadjuvant CRT. Overall, the ANN model exhibited higher accuracy than did the KNN, SVM, NBC, and MLR models. When we used actual validation data sets to compare performance among forecasting models based on pCR occurrence, the ANN model significantly outperformed

the KNN, SVM, NBC, and MLR models, which were constructed using the same limited number of clinical parameters.

Recent studies have consistently demonstrated the ANN model's superiority relative to the KNN, SVM, NBC, or MLR models<sup>22–24</sup>. Furthermore, statistical analyses have proven the advantages of the ANN model<sup>23</sup>. In particular, the high fault tolerance of ANN models facilitates accurate and appropriate processing of incomplete or noise-added inputs. In addition, nonnormally distributed and highly correlated data can be used to develop nonlinear and linear ANN models, with extensive application in medical big data analysis. Clinical studies have commonly used ANN models for prognosis prediction<sup>11,22,24</sup>. This study's comparison of various models indicated that the ANN model had the best performance in terms of expanding the set of predictive variables; this facilitates evaluation of the effectiveness of research methods and enables comprehensive prediction of pCR occurrence. For other cancer types, the established model can be used to predict clinical outcomes or events.

Machine learning has been widely applied for predicting responses to cancer therapy. Bibault et al. used deep learning combined with clinical and radiomic features to predict pCR in patients with LARC following neoadjuvant CRT. They demonstrated that the deep neural network achieved higher accuracy than the linear regression and SVM models did<sup>25</sup>. Metser et al. evaluated the correlation between radiomic features and pCR by using machine learning algorithms and revealed that the classifier trained on pretreatment positron emission tomography scans had an accuracy of 92.8% in predicting pCR to CRT in patients with LARC<sup>26</sup>. Furthermore, machine learning for treatment response prediction has been used for patients with cancer of the head and neck, breast, lung, and prostate<sup>15,22,27,28</sup>. Many studies have demonstrated the favorable performance of machine learning for treatment response prediction related to different cancer types. Our results supported the high accuracy of the ANN model in predicting the efficacy of CRT for LARC.

In the current study, the ANN model exhibited higher accuracy than did the MLR model, a traditional and widely used statistical model in medicine. Growing evidence indicates that overall, machine learning models have higher accuracy in predicting oncologic outcomes than do logistic regression models. According to Faradmal et al., ANN model accuracy was higher than that of the logistic regression model for predicting breast cancer recurrence<sup>29</sup>. Similarly, Alabi et al. compared an ANN model with a logistic regression model based on their prediction of locoregional recurrence in patients with early oral tongue carcinoma, and the ANN model was superior<sup>28</sup>. In the aforementioned studies, machine learning methods exhibited superior accuracy than traditional methods.

A global sensitivity analysis was performed to evaluate the value of significant predictors affecting pCR occurrence. We determined post-CRT CEA level to be the most important predictor of pCR occurrence in patients with LARC after neoadjuvant CRT. CEA level has been commonly evaluated in colorectal cancer-related predictions. Several studies have highlighted the predictive value of post-CRT CEA levels in patients with LARC treated with neoadjuvant CRT. Peng et al. revealed that a post-CRT CEA level of  $\leq 2$  ng/mL was an independent predictor of pCR (OR 1.579; 95% confidence interval [CI] 1.02 to 2.43;  $P = 0.038$ )<sup>30</sup>. Yang et al. identified a post-CRT CEA level of  $\leq 2.61$  ng/mL as being significantly associated with pCR (OR 0.605; 95% CI 0.41 to 0.89;  $P = 0.011$ ) and improved overall survival<sup>31</sup>. Kleiman et al. reported a significant correlation between decreased post-CRT CEA levels and pCR occurrence (OR 1.74; 95% CI 1.06 to 3.81)<sup>32</sup>. A possible explanation might be that decreased post-CRT CEA levels indicate prominent effects of CRT and consequently favorable tumor regression. However, the literature on the exact mechanism remains scarce.

Because radiation-induced necrosis requires time to develop, a prolonged interval between radiation and surgery potentially increases pCR occurrence. In the current study, a radiation-surgery interval  $> 8$  weeks was associated with high pCR rates. The association between longer intervals and pCR occurrence has been studied in several retrospective cohorts, with inconsistent findings. Kalady et al. and Probst et al. have demonstrated that intervals  $> 8$  weeks are associated with increased pCR occurrence<sup>33,34</sup>, but Stein et al. and Sun et al. have reported the opposite result<sup>35,36</sup>. In our previous study, we demonstrated that a longer CRT-surgery interval was associated with increased pCR rates<sup>7</sup>. Several randomized trials have been published to resolve this inconsistency. Two randomized trials by Akgun et al. and Terzi et al. have demonstrated that pCR rates are higher for long intervals ( $> 8$  weeks) than for short intervals, although both intervals have exhibited similar rates in postoperative mortality and morbidity<sup>37,38</sup>. However, the GRECCAR-6 trial revealed no significant difference between long (11 weeks) and short intervals (7 weeks) concerning pCR occurrence, although greater complications and difficulties in surgery were observed for participants with an 11-week interval<sup>39</sup>. More data are required to determine which interval best increases pCR occurrence.

To enhance response to CRT, several chemotherapeutic drugs were added to standard fluoropyrimidine-based CRT. Two randomized trials have reported an increase in pCR after the addition of oxaliplatin to CRT<sup>5,6</sup>, but other trials have revealed no such increase<sup>4,40,41</sup>. To resolve this inconsistency, Yang et al. reviewed the published randomized trials and demonstrated that the addition of oxaliplatin to CRT significantly increased pCR rates (risk ratio 1.24; 95% CI 1.02 to 1.51;  $P = 0.03$ )<sup>42</sup>. Our previous study revealed that FOLFOX-based CRT resulted in improved pCR rates relative to fluoropyrimidine-based CRT<sup>7</sup>. In the current study, we also determined that FOLFOX-based CRT constituted an independent predictor of pCR in machine-learning prediction models.

In agreement with our results, several studies have demonstrated that having clinically node-negative rectal cancer is independently associated with an increase in pCR occurrence<sup>43–45</sup>. Our previous study reviewed 236 patients with LARC undergoing neoadjuvant CRT with subsequent surgery. According to the results, pCR rates in clinically node-negative diseases were three times higher than in node-positive diseases (OR 3.2; 95% CI 1.27 to 8.41;  $P = 0.013$ )<sup>46</sup>. Based on these studies, clinical node positivity may indicate more advanced disease, which results in poor response to CRT. Therefore, watch-and-wait treatment is likely to be suitable for patients with clinically node-negative rectal cancer.

In this study, the ANN model identified clinical T4 as an independent predictor for the absence of pCR. This finding is consistent with those of other studies demonstrating that an advanced tumor stage is associated with

unfavorable tumor regression<sup>43–45</sup>. Despite contradictory findings on the association between clinical tumor stage and pCR occurrence<sup>33,47</sup>, clinical experience suggests that a highly advanced tumor stage is associated with highly aggressive tumor behavior, indicating lower sensitivity to CRT.

In addition to improving the analysis of variance in the correlation between clinical parameters and pCR occurrence, predictive models have broad clinical applications. The methods used in this study can be applied to investigate the effectiveness of other treatment methods, and the quality of care can thus be improved. Because the proposed ANN model exhibited high accuracy in predicting pCRs, the model can help clinicians identify which patients can benefit from watch-and-wait treatment after neoadjuvant CRT. More studies are required to confirm the reliability of the ANN model and to clarify whether it can be used to effectively predict clinical outcomes and optimize cancer treatment.

This study had some limitations. First, MRI features were not assessed: comparisons are of limited validity because of incompleteness in MRI data. Second, the focus on pCR as the endpoint of this prediction model potentially limits the overall clinical utility of the ANN model to a small subset of patients who have a high likelihood of achieving pCR. Third, we only ran forecasting models to predict pCR in patients with LARC after neoadjuvant CRT. Because of the robust magnitude and statistical significance of the effects in the current study, we contend these limitations did not compromise the validity of the results.

## Conclusions

Relative to the KNN, SVM, NBC, and MLR models, this study's ANN model was more accurate in predicting pCR in patients with LARC after neoadjuvant CRT, at higher overall performance indices. Those giving preoperative consultations can use this study's predictors to educate candidates on choices of LARC surgery in terms of health outcomes and the expected prognosis. These findings can serve as a vital and empirical foundation for improving the quality of life of patients with LARC due to the omission of radical surgery.

## Data availability

All data generated or analyzed during this study are included in this published article.

Received: 1 April 2020; Accepted: 8 July 2020

Published online: 28 July 2020

## References

- Sauer, R. *et al.* Preoperative versus postoperative chemoradiotherapy for rectal cancer. *N. Engl. J. Med.* **351**, 1731–1740. <https://doi.org/10.1056/NEJMoa040694> (2004).
- Roh, M. S. *et al.* Preoperative multimodality therapy improves disease-free survival in patients with carcinoma of the rectum: NSABP R-03. *J. Clin. Oncol.* **27**, 5124–5130. <https://doi.org/10.1200/jco.2009.22.0467> (2009).
- Sebag-Montefiore, D. *et al.* Preoperative radiotherapy versus selective postoperative chemoradiotherapy in patients with rectal cancer (MRC CR07 and NCIC-CTG C016): A multicentre, randomised trial. *Lancet* **373**, 811–820. [https://doi.org/10.1016/s0140-6736\(09\)60484-0](https://doi.org/10.1016/s0140-6736(09)60484-0) (2009).
- Gerard, J. P. *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.* **28**, 1638–1644. [https://doi.org/10.1200/jco.2009.25.8376\(2010\)](https://doi.org/10.1200/jco.2009.25.8376(2010)) (2010).
- Rodel, C. *et al.* Preoperative chemoradiotherapy and postoperative chemotherapy with fluorouracil and oxaliplatin versus fluorouracil alone in locally advanced rectal cancer: Initial results of the German CAO/ARO/AIO-04 randomised phase 3 trial. *Lancet Oncol.* **13**, 679–687. [https://doi.org/10.1016/s1470-2045\(12\)70187-0](https://doi.org/10.1016/s1470-2045(12)70187-0) (2012).
- Deng, Y. *et al.* Modified FOLFOX6 with or without radiation versus fluorouracil and leucovorin with radiation in neoadjuvant treatment of locally advanced rectal cancer: initial results of the Chinese FOWARC multicenter, open-label, randomized three-arm phase III trial. *J. Clin. Oncol.* **34**, 3300–3307. <https://doi.org/10.1200/jco.2016.66.6198> (2016).
- Huang, C. M. *et al.* An observational study of extending FOLFOX chemotherapy, lengthening the interval between radiotherapy and surgery, and enhancing pathological complete response rates in rectal cancer patients following preoperative chemoradiotherapy. *Ther. Adv. Gastroenterol.* **9**, 702–712. <https://doi.org/10.1177/1756283x16656690> (2016).
- Habr-Gama, A. *et al.* Local recurrence after complete clinical response and watch and wait in rectal cancer after neoadjuvant chemoradiation: Impact of salvage therapy on local disease control. *Int. J. Radiat. Oncol. Biol. Phys.* **88**, 822–828. <https://doi.org/10.1016/j.ijrobp.2013.12.012> (2014).
- Renehan, A. G. *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.* **17**, 174–183. [https://doi.org/10.1016/s1470-2045\(15\)00467-2](https://doi.org/10.1016/s1470-2045(15)00467-2) (2016).
- Parreco, J. & Rattan, R. Machine learning models for prediction of reinjury after penetrating trauma. *JAMA Surg.* **153**, 184–186. <https://doi.org/10.1001/jamasurg.2017.3116> (2018).
- Grewal, J. K. *et al.* Application of a neural network whole transcriptome-based pan-cancer method for diagnosis of primary and metastatic cancers. *JAMA Netw. Open* **2**, e192597. <https://doi.org/10.1001/jamanetworkopen.2019.2597> (2019).
- Burke, H. B. *et al.* Artificial neural networks improve the accuracy of cancer survival prediction. *Cancer* **79**, 857–862 (1997).
- Burges, C. J. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.* **2**, 121–167 (1998).
- Anand, S. S. *et al.* An evaluation of intelligent prognostic systems for colorectal cancer. *Artif. Intell. Med.* **15**, 193–214 (1999).
- Ronco, A. L. Use of artificial neural networks in modeling associations of discriminant factors: Towards an intelligent selective breast cancer screening. *Artif. Intell. Med.* **16**, 299–309 (1999).
- Santafe, G., Lozano, J. A. & Larranaga, P. Bayesian model averaging of naive Bayes for clustering. *IEEE Trans. Syst. Man Cybern. B Cybern.* **36**, 1149–1161 (2006).
- Ryan, J. E. *et al.* Predicting pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer: A systematic review. *Colorectal Dis.* **18**, 234–246. <https://doi.org/10.1111/codi.13207> (2016).
- Ryan, E. J. *et al.* Meta-analysis of the effect of extending the interval after long-course chemoradiotherapy before surgery in locally advanced rectal cancer. *Br. J. Surg.* **106**, 1298–1310. <https://doi.org/10.1002/bjs.11220> (2019).
- Dayde, D., Tanaka, I., Jain, R., Tai, M. C. & Taguchi, A. Predictive and prognostic molecular biomarkers for response to neoadjuvant chemoradiation in rectal cancer. *Int. J. Mol. Sci.* <https://doi.org/10.3390/ijms18030573> (2017).
- Zhang, C. *et al.* Morphologic predictors of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Oncotarget* **9**, 4862–4874. <https://doi.org/10.18632/oncotarget.23419> (2018).



21. Giannini, V. *et al.* Predicting locally advanced rectal cancer response to neoadjuvant therapy with (18)F-FDG PET and MRI radiomics features. *Eur. J. Nucl. Med. Mol. Imaging* **46**, 878–888. <https://doi.org/10.1007/s00259-018-4250-6> (2019).
22. Kang, J., Schwartz, R., Flickinger, J. & Beriwal, S. Machine learning approaches for predicting radiation therapy outcomes: A clinician's perspective. *Int. J. Radiat. Oncol. Biol. Phys.* **93**, 1127–1135. <https://doi.org/10.1016/j.ijrobp.2015.07.2286> (2015).
23. Alanazi, H. O., Abdullah, A. H. & Qureshi, K. N. A critical review for developing accurate and dynamic predictive models using machine learning methods in medicine and health care. *J. Med. Syst.* **41**, 69. <https://doi.org/10.1007/s10916-017-0715-6> (2017).
24. Hernandez-Suarez, D. F. *et al.* Machine learning prediction models for in-hospital mortality after transcatheter aortic valve replacement. *JACC Cardiovasc. Interv.* **12**, 1328–1338. <https://doi.org/10.1016/j.jcin.2019.06.013> (2019).
25. Bibault, J.-E. *et al.* Deep learning and radiomics predict complete response after neo-adjuvant chemoradiation for locally advanced rectal cancer. *Sci. Rep.* **8**, 12611. <https://doi.org/10.1038/s41598-018-30657-6> (2018).
26. Metser, U. *et al.* Multiparametric PET-MR assessment of response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer: PET, MR, PET-MR and tumor texture analysis: A pilot study. *Adv. Mol. Imaging* **05**, 49–60. <https://doi.org/10.4236/ami.2015.53005> (2015).
27. Gulliford, S. L., Webb, S., Rowbottom, C. G., Corne, D. W. & Dearnaley, D. P. Use of artificial neural networks to predict biological outcomes for patients receiving radical radiotherapy of the prostate. *Radiother. Oncol.* **71**, 3–12. <https://doi.org/10.1016/j.radonc.2003.03.001> (2004).
28. Alabi, R. O. *et al.* Machine learning application for prediction of locoregional recurrences in early oral tongue cancer: A web-based prognostic tool. *Virchows Arch.* **475**, 489–497. <https://doi.org/10.1007/s00428-019-02642-5> (2019).
29. Faradmal, J., Soltanian, A. R., Roshanaei, G., Khodabakhshi, R. & Kasaeian, A. Comparison of the performance of log-logistic regression and artificial neural networks for predicting breast cancer relapse. *Asian Pac. J. Cancer Prev.* **15**, 5883–5888. <https://doi.org/10.7314/apjcp.2014.15.14.5883> (2014).
30. Peng, J. *et al.* Clinical factors of post-chemoradiotherapy as valuable indicators for pathological complete response in locally advanced rectal cancer. *Clinics (Sao Paulo)* **71**, 449–454. [https://doi.org/10.6061/clinics/2016\(08\)07](https://doi.org/10.6061/clinics/2016(08)07) (2016).
31. Yang, K. L. *et al.* Carcinoembryonic antigen (CEA) level, CEA ratio, and treatment outcome of rectal cancer patients receiving pre-operative chemoradiation and surgery. *Radiat. Oncol.* **8**, 43. <https://doi.org/10.1186/1748-717x-8-43> (2013).
32. Kleiman, A. *et al.* Normalization of CEA levels post-neoadjuvant therapy is a strong predictor of pathologic complete response in rectal cancer. *J. Gastrointest. Surg.* **19**, 1106–1112. <https://doi.org/10.1007/s11605-015-2814-3> (2015).
33. Kalady, M. F. *et al.* Predictive factors of pathologic complete response after neoadjuvant chemoradiation for rectal cancer. *Ann. Surg.* **250**, 582–589. <https://doi.org/10.1097/SLA.0b013e3181b91e63> (2009).
34. Probst, C. P. *et al.* Extended intervals after neoadjuvant therapy in locally advanced rectal cancer: The key to improved tumor response and potential organ preservation. *J. Am. Coll. Surg.* **221**, 430–440. <https://doi.org/10.1016/j.jamcollsurg.2015.04.010> (2015).
35. Stein, D. E. *et al.* Longer time interval between completion of neoadjuvant chemoradiation and surgical resection does not improve downstaging of rectal carcinoma. *Dis. Colon Rectum* **46**, 448–453. <https://doi.org/10.1007/s10350-004-6579-0> (2003).
36. Sun, Z. *et al.* Optimal timing to surgery after neoadjuvant chemoradiotherapy for locally advanced rectal cancer. *J. Am. Coll. Surg.* **222**, 367–374. <https://doi.org/10.1016/j.jamcollsurg.2015.12.017> (2016).
37. Akgun, E. *et al.* Randomized clinical trial of short or long interval between neoadjuvant chemoradiotherapy and surgery for rectal cancer. *Br. J. Surg.* **105**, 1417–1425. <https://doi.org/10.1002/bjs.10984> (2018).
38. Terzi, C. *et al.* Randomized controlled trial of 8 weeks vs 12 weeks interval between neoadjuvant chemoradiotherapy and surgery for locally advanced rectal cancer. *Colorectal Dis.* <https://doi.org/10.1111/codi.14867> (2019).
39. Lefevre, J. H. *et al.* Effect of interval (7 or 11 weeks) between neoadjuvant radiochemotherapy and surgery on complete pathologic response in rectal cancer: A multicenter, randomized, controlled trial (GRECCAR-6). *J. Clin. Oncol.* **34**, 3773–3780. <https://doi.org/10.1200/jco.2016.67.6049> (2016).
40. Aschele, C. *et al.* Primary tumor response to preoperative chemoradiation with or without oxaliplatin in locally advanced rectal cancer: Pathologic results of the STAR-01 randomized phase III trial. *J. Clin. Oncol.* **29**, 2773–2780. <https://doi.org/10.1200/jco.2010.34.4911> (2011).
41. Allegra, C. J. *et al.* Neoadjuvant 5-FU or capecitabine plus radiation with or without oxaliplatin in rectal cancer patients: A phase III randomized clinical trial. *J. Natl. Cancer Inst.* <https://doi.org/10.1093/jnci/djv248> (2015).
42. Yang, Y. J. *et al.* Fluorouracil-based neoadjuvant chemoradiotherapy with or without oxaliplatin for treatment of locally advanced rectal cancer: An updated systematic review and meta-analysis. *Oncotarget* **7**, 45513–45524. <https://doi.org/10.18632/oncotarget.9995> (2016).
43. Garland, M. L., Vather, R., Bunkley, N., Pearse, M. & Bissett, I. P. Clinical tumour size and nodal status predict pathologic complete response following neoadjuvant chemoradiotherapy for rectal cancer. *Int. J. Colorectal Dis.* **29**, 301–307. <https://doi.org/10.1007/s00384-013-1821-7> (2014).
44. Bitterman, D. S. *et al.* Predictors of complete response and disease recurrence following chemoradiation for rectal cancer. *Front. Oncol.* **5**, 286. <https://doi.org/10.3389/fonc.2015.00286> (2015).
45. Al-Sukhni, E., Attwood, K., Mattson, D. M., Gabriel, E. & Nurkin, S. J. Predictors of pathologic complete response following neoadjuvant chemoradiotherapy for rectal cancer. *Ann. Surg. Oncol.* **23**, 1177–1186. <https://doi.org/10.1245/s10434-015-5017-y> (2016).
46. Huang, C.-M. *et al.* Predictive value of FOLFOX-based regimen, long interval, hemoglobin levels and clinical negative nodal status, and postchemoradiotherapy CEA levels for pathological complete response in patients with locally advanced rectal cancer after neoadjuvant chemoradiotherapy. *J. Oncol.* **2020**, 1–9. <https://doi.org/10.1155/2020/9437684> (2020).
47. Zeng, W. G. *et al.* Clinical parameters predicting pathologic complete response following neoadjuvant chemoradiotherapy for rectal cancer. *Chin. J. Cancer* **34**, 468–474. <https://doi.org/10.1186/s40880-015-0033-7> (2015).

## Acknowledgements

This work was supported by grants through funding from the Ministry of Science and Technology (MOST108-2321-B-037-001, MOST107-2321-B-037-003, MOST107-2314-B-037-116, MOST107-2314-B-037-022-MY2, MOST107-2314-B-037-023-MY2) and the Ministry of Health and Welfare (MOHW107-TDU-B-212-123006, MOHW107-TDU-B-212-114026B, MOHW108-TDU-B-212-133006, MOHW108-TDU-B-212-124026) funded by Health and welfare surcharge of tobacco products, and the Kaohsiung Medical University Hospital (KMTTH-108-042, KMUH108-8R34, KMUH108-8R35, KMUH108-8M33, KMUH108-8M35, KMUH108-8M36, KMUHS10801, KMUHSA10804, KMUHS10807, KMUH-DK109005~3) and Center for Cancer Research, Kaohsiung Medical University (KMU-TC108A04). In addition, this study was supported by the Grant of Taiwan Precision Medicine Initiative, Academia Sinica, Taiwan, R.O.C.

### Author contributions

J.Y.W. and H.Y.S. conceived the concept of this study and supervised the study. C.M.H. wrote and drafted the manuscript. H.Y.S., H.L.T. and W.C.C. conducted the statistical analysis and interpreted the results of analysis. C.W.H., M.Y.H. and W.C.S. collected the clinical information. All authors read and approved the final manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to J.-Y.W. or H.-Y.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020