ORIGINAL ARTICLE

Journal of Cachexia, Sarcopenia and Muscle 2021; **12**: 1144–1152
Published online 17 June 2021 in Wiley Online Library (wileyonlinelibrary.com) **DOI:** 10.1002/jcsm.12747

Machine learning model for predicting excessive muscle loss during neoadjuvant chemoradiotherapy in oesophageal cancer

Han Gyul Yoon[1], Dongryul Oh[1]*, Jae Myoung Noh[1], Won Kyung Cho[1], Jong-Mu Sun[2], Hong Kwan Kim[3], Jae Ill Zo[3], Young Mog Shim[3] & Kyunga Kim[4]

[1]*Department of Radiation Oncology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea;* [2]*Department of Internal Medicine, Division of Hematology-Oncology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea;* [3]*Department of Thoracic and Cardiovascular Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea;* [4]*Statistics and Data Center, Research Institute for Future Medicine, Samsung Medical Center, Seoul, Republic of Korea*

Abstract

Background Excessive skeletal muscle loss during neoadjuvant concurrent chemoradiotherapy (NACRT) is significantly related to survival outcomes of oesophageal cancer. However, the conventional method for measuring skeletal muscle mass requires computed tomography (CT) images, and the calculation process is labour-intensive. In this study, we built machine-learning models to predict excessive skeletal muscle loss, using only body mass index data and blood laboratory test results.

Methods We randomly split the data of 232 male patients treated with NACRT for oesophageal cancer into the training (70%) and test (30%) sets for 1000 iterations. The naive random over sampling method was applied to each training set to adjust for class imbalance, and we used seven different machine-learning algorithms to predict excessive skeletal muscle loss. We used five input variables, namely, relative change percentage in body mass index, albumin, prognostic nutritional index, neutrophil-to-lymphocyte ratio, and platelet-to-lymphocyte ratio over 50 days. According to our previous study results, which used the maximal χ^2 method, 10.0% decrease of skeletal muscle index over 50 days was determined as the cut-off value to define the excessive skeletal muscle loss.

Results The five input variables were significantly different between the excessive and the non-excessive muscle loss group (all $P < 0.001$). None of the clinicopathologic variables differed significantly between the two groups. The ensemble model of logistic regression and support vector classifier showed the highest area under the curve value among all the other models [area under the curve = 0.808, 95% confidence interval (CI): 0.708–0.894]. The sensitivity and specificity of the ensemble model were 73.7% (95% CI: 52.6%–89.5%) and 74.5% (95% CI: 62.7%–86.3%), respectively.

Conclusions Machine learning model using the ensemble of logistic regression and support vector classifier most effectively predicted the excessive muscle loss following NACRT in patients with oesophageal cancer. This model can easily screen the patients with excessive muscle loss who need an active intervention or timely care following NACRT.

Keywords Machine learning; Oesophageal cancer; Nutrition; Skeletal muscle loss; Sarcopenia

Received: 18 January 2021; Revised: 12 May 2021; Accepted: 8 June 2021
**Correspondence to: Dongryul Oh, Department of Radiation Oncology, Samsung Medical Center, Sungkyunkwan University School of Medicine, 81 Irwon-ro, Gangnam-gu, Seoul 06351, Republic of Korea. Tel: +82-2-3410-2612; Fax: +82-2-3410-2619. E-mail: dongryul.oh@samsung.com*

Introduction

Recent studies investigating the interactions between tumours, nutrition, and inflammation have revealed that the skeletal muscle can affect anti-tumour immunity and treatment response in cancer therapy.[1] Molecules such as myokines, which are cytokines produced and released by skeletal muscle, can mediate anti-inflammatory reactions and anti-tumour mechanisms.[2–4] Accordingly, researchers have recently focused on the correlation between skeletal muscle degradation and prognosis of patients with cancer. Through a series of studies, it was found that excessive skeletal muscle wasting may lead to poor survival outcomes in many types of cancer.[5–10] Skeletal muscle loss and cancer cachexia are also known to be relevant to several medical issues such as prolonged hospitalization, disability, postoperative infections, and other treatment-related toxicities.[11–13]

Our previous study investigated the relationship between the amount of skeletal muscle loss and the survival outcomes of patients with oesophageal cancer, who received neoadjuvant chemoradiotherapy (NACRT) followed by surgery.[14] We calculated skeletal muscle index (SMI) through computed tomography (CT)-based analysis, one of the most common methods for skeletal muscle mass evaluation. We found that an SMI change of $\leq -10.0\%$/50 days during NACRT ('excessive skeletal muscle loss') was significantly related to the overall survival as well as recurrence-free survival of the patients. Excessive skeletal muscle loss was also associated with the decrease in nutritional markers such as albumin and prognostic nutritional index (PNI), and with the increase in inflammatory markers such as neutrophil-to-lymphocyte ratio (NLR) and platelet-to-lymphocyte ratio (PLR) during NACRT.

But one of the limitations of most studies on skeletal muscle mass, including our previous study, was that CT images were necessary to calculate the skeletal muscle amount. Also, the quantitative measurement of skeletal muscle mass using CT images is a labour-intensive and time-consuming process. The axial images at the reference level (e.g. L3 vertebra) are required to be transferred to certain software, and reference points are manually designated on the images to discriminate fat and muscle tissues and, finally, to calculate the skeletal mass. These complicated processes impair the clinical applicability of the above-mentioned study results to the evaluation and management of skeletal muscle loss in cancer patients. Appropriate management of the body composition of cancer patients is also quite challenging, mainly because it is difficult for physicians to perform monitoring and evaluation of the patients adequately and regularly.

To overcome these limitations, we designed several machine learning models to predict excessive skeletal muscle loss using only body mass index (BMI) data and blood laboratory test results. The purpose of this study was to evaluate the performances of the machine learning models and to find the best and robust model to predict excessive skeletal muscle loss. Because we did not know the characteristics of our data distribution and the appropriate model for it, we generated multiple models to find the best model that fits well. Through the study, we aimed to explore whether the machine learning models have the clinical potential to be used as convenient nutritional screening and monitoring tools in real-world clinics.

Materials and methods

Study design and data collection

We retrospectively reviewed the medical records of patients with oesophageal cancer who underwent NACRT followed by surgery. Sixteen of the 248 male patients included in the previous study were excluded because post-radiotherapy (RT) blood tests were not performed prior to the surgery. The final analysis using the machine learning models was performed with 232 patients. All the included patients received NACRT according to our institutional protocol.[14] Along with 5 weeks of RT, two cycles of the same doses of 5-fluorouracil and cisplatin were delivered to the patients 3 weeks apart. This study was approved by our institutional review board (IRB # 2020-12-044) and was performed in accordance with the guidelines of the Declaration of Helsinki.

We obtained pre-RT and post-RT BMI data and the laboratory test results of albumin, PNI, NLR, and PLR. PNI was calculated as [10 × albumin (g/dL) + 0.005 × absolute lymphocyte count (/μL)].[15] Pre- and post-RT values were measured on the day nearest to the pre- and post-RT ^{18}F-fluorodeoxyglucose positron emission tomography-CT (PET-CT) scan dates (*Supporting information*, *Figure S1*). We then calculated the relative ratio of the change amount during RT to the pre-RT value. This value was further divided by the number of interval days (days) and multiplied by 50, which represents the relative change percentage over 50 days [ΔBMI (%/50 days), Δalbumin (%/50 days), ΔPNI (%/50 days), ΔNLR (%/50 days), and ΔPLR (%/50 days)].

Skeletal muscle mass assessment

For the assessment of skeletal muscle mass, we used the CT images of two consecutive PET-CT scans captured before and after RT. Using the in-house software based on MATLAB (version R2014a, MathWorks Inc., Natick, MA, USA), we measured the cross-sectional area (cm^2) of the skeletal muscle on the axial images at the L3 vertebra level. The Hounsfield unit (HU) range to calculate the skeletal muscle cross-sectional area was set as −29 to 150 HU.[16,17] The obtained skeletal muscle area was divided by the height squared (m^2), and this was defined as the SMI. The SMI change during NACRT was then expressed as a percentage

DOI: 10.1002/jcsm.12747

relative to pre-RT SMI [ΔSMI (%)]. Considering the variation in the time interval between pre-RT and post-RT PET-CT scans among patients, we divided ΔSMI (%) by the number of scan interval days (days) and multiplied the result by 50 [ΔSMI (%/50 days)]. According to our previous study results, which used Maxstat, a maximal χ^2 method in R version 3.5.3 (R Development Core Team, Vienna, Austria, http://www.r-project.org), ΔSMI (%/50 days) ≤ -10.0%/50 days was defined as 'excessive' skeletal muscle loss.[14]

Building process and statistical analysis of the machine-learning models

All procedures for building the machine learning-based models were performed using the Python Scikit-learn library (version 0.23.1), LightGBM library (version 2.3.1), imbalanced-learn library (version 0.8.0), and umap-learn library (version 0.4.6). The following five variables were determined as the candidates for the input parameters: ΔBMI (%/50 days), Δalbumin (%/50 days), ΔPNI (%/50 days), ΔNLR (%/50 days), and ΔPLR (%/50 days). Candidate variables with a P value <0.1 in the independent t test on the whole data set were selected as the input features of our prediction models. All the selected variables were then standard-scaled through the 'StandardScaler' function in the Scikit-learn library. Seven machine learning models—k-nearest neighbour, naive Bayes, random forest (RF), light gradient boosting, logistic regression (LR), support vector classifier (SVC), and an ensemble model of LR and SVC using the soft-voting method—were used to predict excessive skeletal muscle loss. In the soft voting method, the probability values of several models are averaged, and the class with the highest probability value is predicted as the answer.[18] The models were built using 'KNeighborsClassifier', 'BernoulliNB', 'RandomForestClassifier', 'LogisticRegression', and 'SVC' function in the Scikit-learn library,[19] and 'LGBMClassifier' function in the LightGBM library. L2 regularization was applied to the LR and SVC models, and radial basis function kernel was used in the SVC. 'VotingClassifier' function in the Scikit-learn library was used to perform soft-voting method.

To compensate for the data distribution bias according to a single partition of the training and test sets, we adopted the concept similar to the 'bootstrap bias-corrected cross-validation (BBC-CV)'.[20,21] This is a kind of internal validation method that enables to estimate more unbiased performance of models trained with a given data set. After the performance estimation, the final model hyperparameters are determined by the cross-validation of the whole given data set, which can be used in testing the future external data sets. The processes are as follows. (i) Randomly split the data into the training (70%) and hold-out test (30%) sets, using the 'train_test_split' function in the Scikit-learn library (In the BBC-CV method, the training set is generated by bootstrapping, which is different from the method of this study). During this process, the proportion of patients with excessive skeletal muscle loss in both sets was maintained at a constant value. (ii) Perform hyperparameter tuning with the training set, using 10-fold cross-validation and grid search methods.[22] The candidate hyperparameters of the models are listed in *Table S1*. Among the candidate hyperparameters, those that maximized area under the curve (AUC) of the receiver operating characteristic (ROC) curve were selected. (iii) A naive random oversampling technique was applied to the training set to adjust for class imbalance, using the 'RandomOverSampler' function in the 'imbalanced-learn' Python package.[23] (iv) Train the models with the selected configurations and estimate their performance on the test set. (v) Repeat the (i)–(iv) for 1000 times. In each iteration, the random seed number was set by subtracting 1 from the number of iterations (e.g. random seed value = 999 in the 1000th iteration). With the 1,000 pairs of training and test sets created, an 'empirical distribution' of the statistics of interest was formulated (*Figure 1*). In the distribution, the lower 2.5% and upper 2.5% values were set as the boundary values of the 95% confidence interval (CI) of each statistic, respectively. Through this process, we calculated the 95% CI of accuracy, AUC of the ROC curve, sensitivity, and specificity for all the machine learning models. Difference of the two distributions of statistics was considered significant if the 95% interval of the difference did not include a zero.

For a comparison of variables between the two groups, the χ^2 test or Fisher's exact test was employed for categorical variables, whereas the independent t test was used for continuous variables. Before performing the independent t test for each continuous variable, we examined the normality of the variables using the Shapiro–Wilk test. All statistical analyses in this study were conducted using the SPSS software package (version 27.0, IBM Corporation, Armonk, NY, USA).

Results

Patient characteristics

The clinicopathologic characteristics of all patients are summarized in *Table 1*. The age of the excessive muscle loss group tended to be higher, although this was not statistically significant (62.26 ± 7.46 vs. 64.45 ± 8.18, $P = 0.052$). None of the other clinicopathologic variables differed significantly between the two groups. The histology of all patients was squamous cell carcinoma. *Table 2* demonstrates the body components and laboratory test results of the excessive and non-excessive muscle loss groups. There was no significant difference in the pre-RT laboratory test results between the two groups, and the baseline pre-RT SMI and BMI were rather higher in the excessive muscle loss group. Conversely,

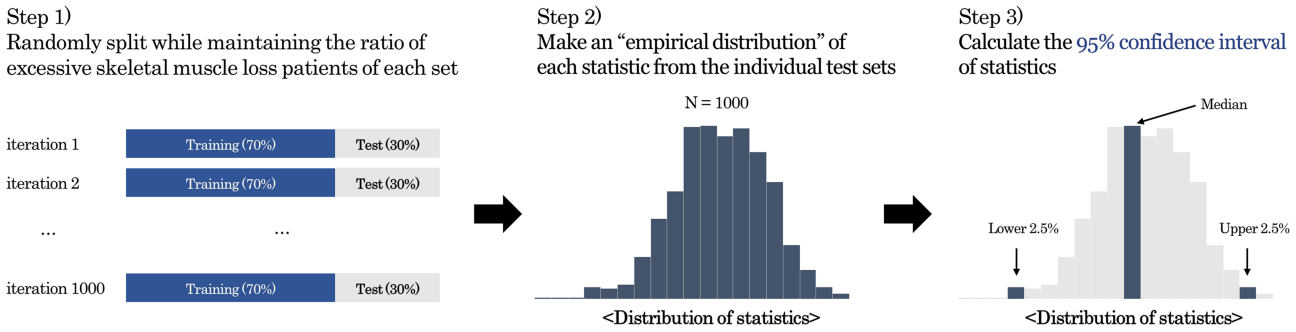

Figure 1 Process of calculating the 95% confidence interval of statistics.

Table 1 Comparison of clinicopathologic characteristics between the excessive muscle loss group and the non-excessive muscle loss group

	Non-excessive muscle loss group (n = 168)	Excessive muscle loss group (n = 64)	P value
Age (years)	62.26 ± 7.46	64.45 ± 8.18	0.052
ECOG performance status			0.578[b]
0–1	164 (97.6%)	64 (100.0%)	
2	4 (2.4%)	0 (0.0%)	
Current smoking			0.150
No	69 (41.1%)	33 (51.6%)	
Yes	99 (58.9%)	31 (48.4%)	
Location			0.415
Upper	44 (26.2%)	16 (25.0%)	
Middle	83 (49.4%)	27 (42.2%)	
Lower	41 (24.4%)	21 (32.8%)	
cT stage			0.353
cT1–2	44 (26.2%)	13 (20.3%)	
cT3–4	124 (73.8%)	51 (79.7%)	
cN stage			0.787
cN0–1	103 (61.3%)	38 (59.4%)	
cN2–3	65 (38.7%)	26 (40.6%)	
ypT stage			0.570
ypT0/Tis	77 (45.8%)	32 (50.0%)	
ypT1–4	91 (54.2%)	32 (50.0%)	
ypN stage			0.557
ypN0	86 (51.2%)	30 (46.9%)	
ypN+	82 (48.8%)	34 (53.1%)	
ypCR			0.570
No	117 (69.6%)	47 (73.4%)	
Yes	51 (30.4%)	17 (26.6%)	
Resection margin			1.000[b]
R0	161 (95.8%)	61 (95.3%)	
R1–2	7 (4.2%)	3 (4.7%)	

CR, complete response; ECOG, Eastern Cooperative Oncology Group.
[a]Excessive muscle group was defined as ΔSMI (%/50 days) < −10 (%/50 days).
[b]Calculated by Fisher's exact test.

Table 2 Comparison of body components and laboratory test results between the excessive muscle loss group and the non-excessive muscle loss group

	Non-excessive muscle loss group[a] (n = 168)	Excessive muscle loss group[a] (n = 64)	P value
SMI			
Pre-RT	48.87 ± 7.77	51.62 ± 8.03	0.017
Post-RT	46.21 ± 7.40	42.23 ± 7.51	<0.001
ΔRT (%/50 days)	−3.67 ± 4.36	−13.39 ± 3.82	<0.001
BMI			
Pre-RT	22.65 ± 2.81	23.53 ± 2.96	0.037
Post-RT	21.72 ± 2.87	21.31 ± 3.05	0.341
ΔRT (%/50 days)	−4.75 ± 4.95	−10.04 ± 6.00	<0.001
Albumin (g/dL)			
Pre-RT	4.27 ± 0.37	4.29 ± 0.41	0.727
Post-RT	3.92 ± 0.35	3.66 ± 0.36	<0.001
ΔRT (%/50 days)	−6.90 ± 8.41	−13.01 ± 9.32	<0.001
PNI			
Pre-RT	53.16 ± 5.47	53.63 ± 5.10	0.552
Post-RT	45.02 ± 4.81	42.15 ± 5.86	<0.001
ΔRT (%/50 days)	−13.45 ± 10.26	−19.77 ± 12.88	<0.001
NLR			
Pre-RT	2.43 ± 1.23	2.60 ± 1.29	0.374
Post-RT	3.07 ± 3.86	5.85 ± 10.02	0.002
ΔRT (%/50 days)	48.38 ± 215.72	259.98 ± 915.26	0.005
PLR			
Pre-RT	125.56 ± 43.28	125.66 ± 71.32	0.990
Post-RT	201.09 ± 154.01	254.19 ± 195.33	0.031
ΔRT (%/50 days)	64.38 ± 118.34	140.75 ± 204.11	<0.001

BMI, body mass index; NLR, neutrophil-to-lymphocyte ratio; PLR, platelet-to-lymphocyte ratio; PNI, prognostic nutritional index; RT, radiation therapy; SMI, skeletal muscle index.
[a]All variables are described as mean ± standard deviation.

significant differences were observed in the post-RT values. Patients with excessive muscle loss had a lower post-RT SMI (46.21 ± 7.40 vs. 42.23 ± 7.51, $P < 0.001$), albumin (3.92 ± 0.35 vs. 3.66 ± 0.36, $P < 0.001$), and PNI (45.02 ± 4.81 vs. 42.15 ± 5.86, $P < 0.001$), and a higher NLR (3.07 ± 3.86 vs. 5.85 ± 10.02, $P = 0.002$) and PLR (201.09 ± 154.01 vs. 254.19 ± 195.33, $P = 0.031$). Furthermore, the relative change percentage of the variables over 50 days was significantly different between the two groups: ΔBMI (%/50 days) (−4.75 ± 4.95 vs. −10.04 ± 6.00, $P < 0.001$), Δalbumin (%/50 days) (−6.90 ± 8.41 vs. −13.01 ± 9.32, $P < 0.001$), ΔPNI (%/50 days) (−13.45 ± 10.26 vs. −19.77 ± 12.88, $P < 0.001$), ΔNLR (%/50 days) (48.38 ± 215.72 vs. 259.98 ± 915.26, $P = 0.005$), and ΔPLR (%/50 days) (64.38 ± 118.34 vs. 140.75 ± 204.11, $P < 0.001$).

Visualization of the data distribution and performance of the different models

According to our selection criteria, ΔBMI (%/50 days), Δalbumin (%/50 days), ΔPNI (%/50 days), ΔNLR (%/50 days), and ΔPLR (%/50 days) were all selected as the input parameters. Because our data space was a five-dimensional space composed of the five variables, its visualization was impossible. Therefore, we visualized the data distribution using a dimension reduction technique called the uniform manifold approximation and projection (UMAP) algorithm.[24] We reduced the spatial distribution of data from five to three dimensions using the UMAP algorithm, and the reduced data space was visualized as shown in *Figure 2*, wherein, it can be observed that the samples of the excessive and non-excessive muscle loss groups are clustered to some extent. It can be assumed that the two groups may be well distinguished by finding a specific boundary plane.

Table 3 summarizes the predictive performances of the different machine learning models. The ensemble model of LR and SVC showed the highest AUC of 0.808 (95% CI: 0.708–0.894), accuracy of 74.3% (95% CI: 64.3%–82.9%), and sensitivity of 73.7% (95% CI: 52.6%–89.5%) among the seven prediction models, with a specificity of 74.5% (95% CI: 62.7%–86.3%). Following the ensemble model, the LR (AUC = 0.804, 95% CI: 0.701–0.896) and SVC (AUC = 0.797, 95% CI: 0.692–0.889) models showed the second and third highest AUC values, respectively. The AUC difference between the ensemble model and the LR model was not statistically significant (median = 0.003, 95% CI: −0.030 to 0.043). For specificity, the RF (specificity = 80.4%, 95% CI: 68.6%–90.2%) and light gradient boosting (specificity = 78.4%, 95% CI: 66.7%–88.2%) models represented the highest and second highest values, respectively. However, the AUCs of the two models were relatively low, at 0.728 (95% CI: 0.621–0.821) and 0.681 (95% CI: 0.563–0.781), respectively. Each model's distributions of the statistics are shown as boxplots in the *Figure S2*.

Figure 3 shows the change in sensitivity and specificity according to the probability threshold for excessive skeletal muscle loss. The results are based on the ensemble model of LR and SVC, and the cut-off probability is changed from 0.05 to 0.95 at 0.05 intervals. Among the threshold values included, the threshold value of 0.50 showed the most balanced results in terms of the sensitivity–specificity trade-off relationship. Sensitivity greater than 80% and 90% was achieved when the probability threshold was set to less than 0.40 (sensitivity = 84.2%, 95% CI: 63.2%–100.0%) and 0.30

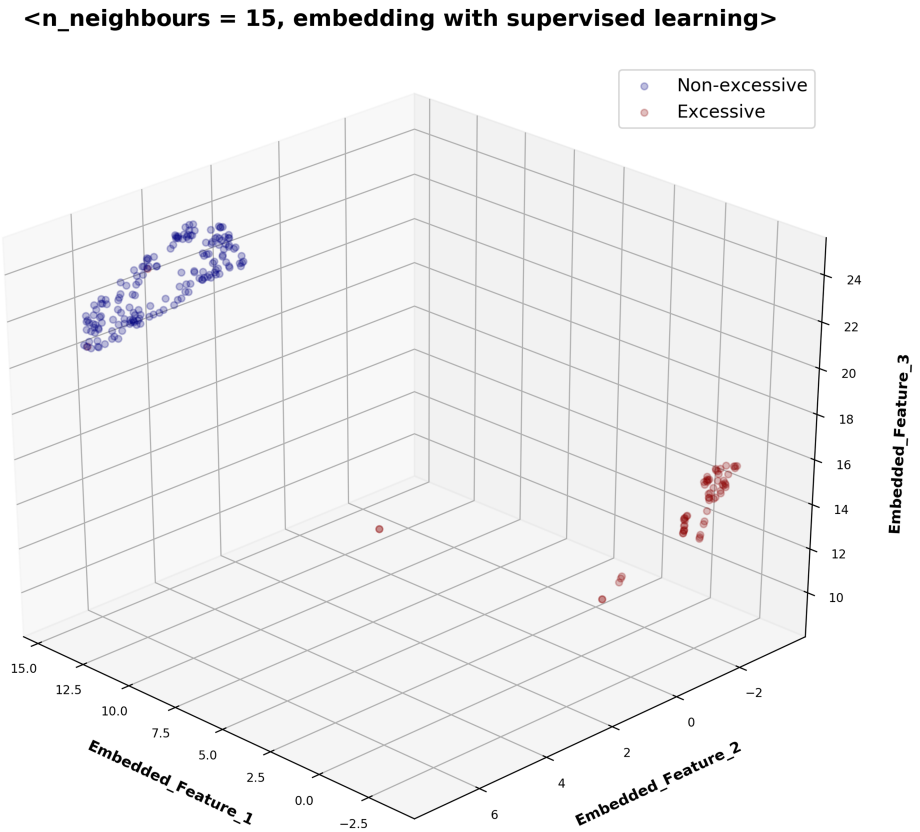

Figure 2 Visualization of the data distribution using the uniform manifold approximation and projection algorithm. The number of neighbours and random state value were set to 15 and 0, respectively.

Table 3 Predictive performance of the machine learning models

Model	AUC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
k-nearest neighbour	0.781 (0.679–0.878)	0.729 (0.629–0.814)	0.684 (0.421–0.895)	0.745 (0.608–0.863)
Bernoulli naive-Bayes	0.754 (0.636–0.852)	0.686 (0.571–0.786)	0.684 (0.474–0.842)	0.686 (0.549–0.824)
Random forest	0.728 (0.621–0.821)	0.714 (0.614–0.800)	0.474 (0.263–0.684)	0.804 (0.686–0.902)
Light gradient boosting	0.684 (0.563–0.781)	0.686 (0.586–0.757)	0.421 (0.211–0.632)	0.784 (0.667–0.882)
Logistic regression	0.804 (0.701–0.896)	0.729 (0.643–0.829)	0.737 (0.526–0.895)	0.745 (0.627–0.863)
Support vector classifier	0.797 (0.692–0.889)	0.729 (0.629–0.814)	0.737 (0.526–0.947)	0.725 (0.569–0.843)
Ensemble (LR + SVC)	0.808 (0.708–0.894)	0.743 (0.643–0.829)	0.737 (0.526–0.895)	0.745 (0.627–0.863)

AUC, area under the curve; CI, confidence interval; LR, logistic regression; SVC, support vector classifier.

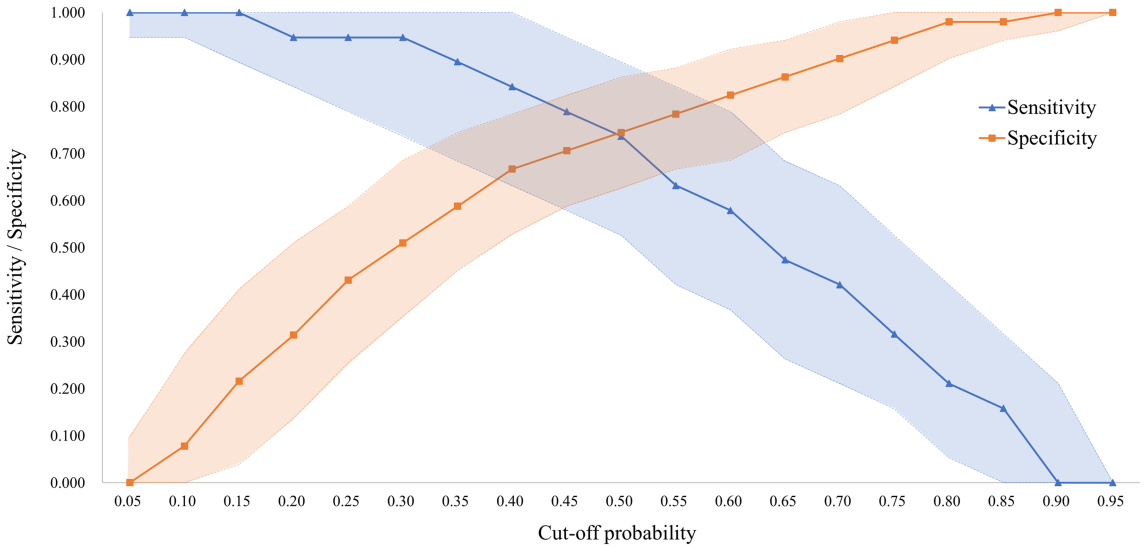

Figure 3 Sensitivity and specificity according to the probability threshold for excessive skeletal muscle loss.

(sensitivity = 94.7%, 95% CI: 73.7%–100.0%), respectively. In contrast, specificity greater than 80% and 90% was achieved when the probability threshold was set to more than 0.60 (specificity = 82.4%, 95% CI: 68.6%–92.2%) and 0.70 (specificity = 90.2%, 95% CI: 78.4%–98.0%), respectively.

Discussion

This study investigated the performances of the machine learning models that were designed for the prediction of excessive skeletal muscle loss in patients with oesophageal cancer. Among the prediction models, the ensemble model of LR and SVC showed the highest AUC, accuracy, and sensitivity. The ensemble model represented an AUC over 0.8 and sensitivity/specificity over 70%. By adjusting the threshold value of the probability for excessive skeletal muscle loss to 0.4 or 0.3, sensitivity of more than 80% or 90% was achieved, respectively.

To the best of our knowledge, this is the first study to evaluate the severity of muscle loss with machine learning models that only used BMI data and blood test results. Eisner *et al*. designed machine learning models to predict the skeletal muscle loss $< -0.75\%/100$ days with urine metabolites in 93 colon and lung cancer, and the highest accuracy obtained was 82.2%.[25] However, the study was performed with urine samples, and ^{1}H nuclear magnetic resonance spectroscopy results were required to obtain the input features of the prediction models. Although very few studies have predicted the amount of muscle loss using machine learning, several studies have attempted to predict the presence of sarcopenia, a condition characterized by the loss of muscle mass and strength,[26] using machine learning models. Cui *et al*. used the SVC and RF models for the prediction of sarcopenia in patients with type 2 diabetes mellitus and achieved the highest AUC of 0.87 ± 0.07.[27] Another study by Kang *et al*. investigated the performance of several machine learning-based sarcopenia prediction models in 4020 post-menopausal women from the Korea National Health and Nutrition Examination Surveys, and the highest AUC among the included models was 0.82 ± 0.06.[28]

Skeletal muscle wasting remains to be one of the critical issues in cancer patients, leading to poor survival outcomes,

treatment response, and multiple medical issues.[1,5–13] Furthermore, several tumour-related and treatment-related factors can aggravate skeletal muscle degeneration. The tumour itself releases molecules and pro-inflammatory cytokines that can impair the protein synthesis and muscle regeneration mechanisms.[29] Neoadjuvant treatment can further exacerbate muscle wasting,[30,31] reduce muscle strength,[32,33] and cause adverse effects such as esophagitis, dysphagia, nausea, vomiting, and anorexia.[34,35] Our previous study has shown that excessive skeletal muscle loss after treatment is a significant factor for prognosis and recurrence rather than the presence of sarcopenia itself prior to treatment.[14] Adequate support and intervention for nutrition and physical activity during treatment may be beneficial to improve their prognosis and quality of life.[36–38] Indeed, in one randomized controlled pilot study, patients who received a walk-and-eat intervention during NACRT showed less decrease in the distance on the 6 min walk test handgrip strength and body weight and showed lower rates of intravenous nutritional support and wheelchair use than those of the control group.[36]

Appropriate management of the body composition of cancer patients is quite challenging, mainly because it is difficult for physicians to perform monitoring and evaluation of the patients adequately and regularly. Moreover, conventional CT-based methods that evaluate skeletal muscle mass are inevitably subject to a labour-intensive and time-consuming process. The machine learning models in our study can easily assess the change in skeletal muscle mass of cancer patients after treatment and screen patients who require a more active intervention or nutritional support before surgery. Using our machine learning models as cost-effective and convenient monitoring tools, physicians will be able to automatically assess the nutritional status of cancer patients and help them receive more timely care. Additionally, a model that can predict significant excessive skeletal muscle loss, with the data obtained during the treatment rather than after the treatment, could be clinically more useful for managing the patients. Future research should consider early prediction of excessive skeletal muscle loss using algorithms for time-series data analysis such as recurrent neural networks.[39]

This study has several limitations. First, the patient data used in this study were retrospectively collected from a single institution. Second, variables that are potentially associated with excessive skeletal muscle loss were not included in this study. Third, we conducted the bootstrapping process similar to BBC-CV with our internal data and performed statistical analysis for feature selection using the whole given data set, which may have biased the outcomes. Even though we estimated the bias-corrected performance of our models through the 1000 iterations of random split, further analyses with external validation sets and multi-centre design are necessary to verify whether the selected features and trained models are appropriate. The hyperparameters selected by 10-fold cross-validation on our entire data set are shown in *Table S2*, which can be used in future external validation. Fourth, the baseline pre-RT SMI and BMI were rather higher in the excessive muscle loss group, and there is a possibility that this may have influenced the study results. Fifth, although the ensemble model showed the highest AUC, it was not significantly different from that of the LR model. This might suggest that the samples of the two classes are grouped to some extent with each other and tended to be linearly separated, which are generally known as conditions that the LR model can operate effectively.[40,41] Considering these points, it is also possible to consider using the LR model alone. The coefficient of LR model obtained from the 10-fold cross-validation of the whole given data set and 95% CIs of the coefficients acquired by the 1000 random splits are summarized in *Table S3*. Finally, the patients included in this study were all male patients diagnosed with squamous cell carcinoma. Although most patients with oesophageal cancer in Korea are male with squamous carcinoma,[42] future studies that include female patients and various histologic types are needed to assess the robustness of our prediction models.

Despite these limitations, the machine learning models of the present study effectively predicted the excessive muscle loss following NACRT in oesophageal cancer. These models can be used to screen patients requiring an active intervention after NACRT. More research is needed to address the clinical usefulness of machine learning models that detect severe skeletal muscle loss in large populations of patients afflicted with different cancer types.

Acknowledgements

The authors of this manuscript certify that they comply with the ethical guidelines for authorship and publishing in the *Journal of Cachexia, Sarcopenia and Muscle*.[43]

Online supplementary material

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Calculation process of the skeletal muscle index and its change before and after radiotherapy. CCRT, Concurrent chemoradiotherapy; RT, Radiotherapy; CT, Computed tomograpy; PET-CT, ^{18}F-fluorodeoxyglucose positron emission tomography-computed tomography; SMI, Skeletal muscle index.

Figure S2. Boxplots for the distributions of performance statistics. AUC, Area Under the Curve; KNN, K-nearest neighbour; NB, Naive Bayes; RF, Random Forest; LGBM, Light Gradient Boosting Model; LR, Logistic Regression; SVC, Support Vector Classifier.

Table S1. Candidate hyperparameters of the machine learning models

Table S2. The hyperparameters of the machine learning models selected by cross-validation on our whole dataset

Table S3. The coefficients of the multivariable logistic regression model

Conflict of interest

The authors do not have any conflict of interest to disclose.

References

1. Zitvogel L, Pietrocola F, Kroemer G. Nutrition, inflammation and cancer. *Nat Immunol* 2017;**18**:843–850.
2. Roy P, Chowdhury S, Roy HK. Exercise-induced myokines as emerging therapeutic agents in colorectal cancer prevention and treatment. *Future Oncol* 2018;**14**: 309–312.
3. Steensberg A, Fischer CP, Keller C, Moller K, Pedersen BK. IL-6 enhances plasma IL-1ra, IL-10, and cortisol in humans. *Am J Physiol Endocrinol Metab* 2003;**285**:E433–E437.
4. Lucia A, Ramirez M. Muscling in on cancer. *N Engl J Med* 2016;**375**:892–894.
5. Basile D, Parnofiello A, Vitale MG, Cortiula F, Gerratana L, Fanotto V, et al. The IMPACT study: early loss of skeletal muscle mass in advanced pancreatic cancer patients. *J Cachexia Sarcopenia Muscle* 2019;**10**:368–377.
6. Reisinger KW, Bosmans JW, Uittenbogaart M, Alsoumali A, Poeze M, Sosef MN, et al. Loss of skeletal muscle mass during neoadjuvant chemoradiotherapy predicts postoperative mortality in esophageal cancer surgery. *Ann Surg Oncol* 2015;**22**: 4445–4452.
7. Daly LE, Ni Bhuachalla EB, Power DG, Cushen SJ, James K, Ryan AM. Loss of skeletal muscle during systemic chemotherapy is prognostic of poor survival in patients with foregut cancer. *J Cachexia Sarcopenia Muscle* 2018;**9**:315–325.
8. Takamori S, Tagawa T, Toyokawa G, Shimokawa M, Kinoshita F, Kozuma Y, et al. Prognostic impact of postoperative skeletal muscle decrease in non-small cell lung cancer. *Ann Thorac Surg* 2020;**109**: 914–920.
9. Chang KV, Chen JD, Wu WT, Huang KC, Hsu CT, Han DS. Association between loss of skeletal muscle mass and mortality and tumor recurrence in hepatocellular carcinoma: a systematic review and meta-analysis. *Liver Cancer* 2018;**7**:90–103.
10. Lee J, Chang CL, Lin JB, Wu MH, Sun FJ, Jan YT, et al. Skeletal muscle loss is an imaging biomarker of outcome after definitive chemoradiotherapy for locally advanced cervical cancer. *Clin Cancer Res* 2018;**24**: 5028–5036.
11. Motoori M, Fujitani K, Sugimura K, Miyata H, Nakatsuka R, Nishizawa Y, et al. Skeletal muscle loss during neoadjuvant chemotherapy is an independent risk factor for postoperative infectious complications in patients with advanced esophageal cancer. *Oncology* 2018;**95**:281–287.
12. Kurk S, Peeters P, Stellato R, Dorresteijn B, de Jong P, Jourdan M, et al. Skeletal muscle mass loss and dose-limiting toxicities in metastatic colorectal cancer patients. *J Cachexia Sarcopenia Muscle* 2019;**10**: 803–813.
13. Naito T, Okayama T, Aoyama T, Ohashi T, Masuda Y, Kimura M, et al. Unfavorable impact of cancer cachexia on activity of daily living and need for inpatient care in elderly patients with advanced non-small-cell lung cancer in Japan: a prospective longitudinal observational study. *BMC Cancer* 2017;**17**:800.
14. Yoon HG, Oh D, Ahn YC, Noh JM, Pyo H, Cho WK, et al. Prognostic impact of sarcopenia and skeletal muscle loss during neoadjuvant chemoradiotherapy in esophageal cancer. *Cancers (Basel)* 2020;**12**.
15. Onodera T, Goseki N, Kosaki G. Prognostic nutritional index in gastrointestinal surgery of malnourished cancer patients. *Nihon Geka Gakkai Zasshi* 1984;**85**:1001–1005.
16. Mitsiopoulos N, Baumgartner RN, Heymsfield SB, Lyons W, Gallagher D, Ross R. Cadaver validation of skeletal muscle measurement by magnetic resonance imaging and computerized tomography. *J Appl Physiol (1985)* 1998;**85**:115–122.
17. Aubrey J, Esfandiari N, Baracos V, Buteau F, Frenette J, Putman C, et al. Measurement of skeletal muscle radiation attenuation and basis of its biological variation. *Acta Physiol (Oxf)* 2014;**210**:489–497.
18. Swamynathan M. *Mastering Machine Learning with Python in Six Steps: A Practical Implementation Guide to Predictive Data Analytics Using Python*. Apress; 2019.
19. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. *J Mac Lear Res* 2011;**12**:2825–2830.
20. Tsamardinos I, Greasidou E, Borboudakis GJML. Bootstrapping the out-of-sample predictions for efficient and accurate cross-validation. *Mac Lea* 2018;**107**: 1895–1922.
21. Widera P, Welsing PM, Ladel C, Loughlin J, Lafeber FP, Dop FP, et al. Multi-classifier prediction of knee osteoarthritis progression from incomplete imbalanced longitudinal data. *Sci Rep* 2020;**10**:1–15.
22. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *IJCAI (Intern Joint Conf Articial Intell)* 1995;**14**:1137–1145.
23. Lemaître G, Nogueira F, Aridas CKJTJoMLR. Imbalanced-learn: a Python toolbox to tackle the curse of imbalanced datasets in machine learning. *J Mac Lear Res* 2017;**18**:559–563.
24. McInnes L, Healy J, Melville J. UMAP: uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426. 2018.
25. Eisner R, Stretch C, Eastman T, Xia J, Hau D, Damaraju S, et al. Learning to predict cancer-associated skeletal muscle wasting from ^{1}H-NMR profiles of urinary metabolites. *Metabolomics* 2010;**7**:25–34.
26. Evans WJ. What is sarcopenia? *J Gerontol A Biol Sci Med Sci* 1995;50 Spec No;5–8.
27. Cui M, Gang X, Gao F, Wang G, Xiao X, Li Z, et al. Risk assessment of sarcopenia in patients with type 2 diabetes mellitus using data mining methods. *Front Endoc (Lausanne)* 2020;**11**:123.
28. Kang YJ, Yoo JI, Ha YC. Sarcopenia feature selection and risk prediction using machine learning: A cross-sectional study. *Med (Baltimore)* 2019;**98**:e17699.
29. Fukushima H, Fujii Y, Koga F. Metabolic and molecular basis of sarcopenia: implications in the management of urothelial carcinoma. *Int J Mol Sci* 2019;**20**.
30. Fearon K, Arends J, Baracos V. Understanding the mechanisms and treatment options in cancer cachexia. *Nat Rev Clin Oncol* 2013;**10**:90–99.
31. Fanzani A, Zanola A, Rovetta F, Rossi S, Aleo MF. Cisplatin triggers atrophy of skeletal C2C12 myotubes via impairment of Akt signalling pathway and subsequent increment activity of proteasome and autophagy systems. *Toxicol Appl Pharmacol* 2011;**250**:312–321.
32. Guinan EM, Doyle SL, Bennett AE, O'Neill L, Gannon J, Elliott JA, et al. Sarcopenia during neoadjuvant therapy for oesophageal cancer: characterising the impact on muscle strength and physical performance. *Support Care Cancer* 2018;**26**:1569–1576.
33. Jack S, West MA, Raw D, Marwood S, Ambler G, Cope TM, et al. The effect of neoadjuvant chemotherapy on physical fitness and survival in patients undergoing oesophagogastric cancer surgery. *Eur J Surg Oncol* 2014;**40**:1313–1320.
34. Chowhan NM. Injurious effects of radiation on the esophagus. *Am J Gastroenterol* 1990;**85**:115–120.
35. Duan XF, Tang P, Yu ZT. Neoadjuvant chemoradiotherapy for resectable esophageal cancer: an in-depth study of randomized controlled trials and literature review. *Cancer Biol Med* 2014;**11**:191–201.

36. Xu YJ, Cheng JC, Lee JM, Huang PM, Huang GH, Chen CC. A walk-and-eat intervention improves outcomes for patients with esophageal cancer undergoing neoadjuvant chemoradiotherapy. *Oncologist* 2015;**20**:1216–1222.
37. Birnstein E, Schattner M. Nutritional support in esophagogastric cancers. *Surg Oncol Clin N Am* 2017;**26**:325–333.
38. Christensen JF, Simonsen C, Banck-Petersen A, Thorsen-Streit S, Herrstedt A, Djurhuus SS, et al. Safety and feasibility of preoperative exercise training during neoadjuvant treatment before surgery for adenocarcinoma of the gastro-oesophageal junction. *BJS Open* 2019;**3**:74–84.
39. Petneházi G. Recurrent neural networks for time series forecasting. arXiv preprint arXiv:1901.00069. 2019.
40. Wang J. Research on machine learning and its algorithm. *DEStech Transac Comp Sci Engin* 2017.
41. Patel M, Sharma P. Cherian AKJJoCR. Bully identification with machine learning algorithms. *J Crit Rev* 2020;**7**:417–425.
42. Shin A, Won YJ, Jung HK, Kong HJ, Jung KW, Oh CM, et al. Trends in incidence and survival of esophageal cancer in Korea: analysis of the Korea Central Cancer Registry Database. *J Gastroenterol Hepatol* 2018;**33**:1961–1968.
43. von Haehling S, Morley JE, Coats AJS, Anker SD. Ethical guidelines for publishing in the *Journal of Cachexia, Sarcopenia and Muscle*: update 2019. *J Cachexia Sarcopenia Muscle* 2019;**10**:1143–1145.